Bernoulli **18**(4), 2012, 1249–1266
DOI: 10.3150/12-BEJ378

Bootstrap confidence intervals for isotonic estimators in a stereological problem

BODHISATTVA SEN[1] and MICHAEL WOODROOFE[2]

[1]*Department of Statistics, Columbia University, New York, NY 10027, USA.*
E-mail: bodhi@stat.columbia.edu
[2]*Department of Statistics, University of Michigan, Ann Arbor, MI 48109, USA.*
E-mail: michaelw@umich.edu

Let $\mathbf{X} = (X_1, X_2, X_3)$ be a spherically symmetric random vector of which only (X_1, X_2) can be observed. We focus attention on estimating F, the distribution function of the squared radius $Z := X_1^2 + X_2^2 + X_3^2$, from a random sample of (X_1, X_2). Such a problem arises in astronomy where (X_1, X_2, X_3) denotes the three dimensional position of a star in a galaxy but we can only observe the projected stellar positions (X_1, X_2). We consider isotonic estimators of F and derive their limit distributions. The results are nonstandard with a rate of convergence $\sqrt{n/\log n}$. The isotonized estimators of F have exactly half the limiting variance when compared to naive estimators, which do not incorporate the shape constraint. We consider the problem of constructing point-wise confidence intervals for F, state sufficient conditions for the consistency of a bootstrap procedure, and show that the conditions are met by the conventional bootstrap method (generating samples from the empirical distribution function).

Keywords: asymptotic normality; consistency of bootstrap; globular cluster; nonstandard problem; shape restricted estimation; spherically symmetric distribution

1. Introduction

Stereology is the study of three-dimensional properties of objects or matter usually observed two-dimensionally. We consider such a problem, which arises in astronomy. Suppose that the position $\mathbf{X} := (X_1, X_2, X_3)$ of a star within a given galaxy has a spherically symmetric distribution and that we observe the projected stellar positions, that is, (X_1, X_2) (with a proper choice of co-ordinates); and consider the problem of estimating the distribution function F of the squared distance $Z := X_1^2 + X_2^2 + X_3^2$ of a star to the center of the galaxy from a random sample of (X_1, X_2). In this paper, we study the statistical properties of three estimators of F. We show that enforcing *known* shape restrictions (monotonicity) in the estimation procedure leads to estimators with lower asymptotic variance (exactly by one-half in this case). We also consider the problem of constructing point-wise confidence intervals (CIs) around F, and show that the conventional bootstrap method can be used to construct valid CIs. Our treatment is similar in flavor to Groeneboom and Jongbloed's [6] study of the Wicksell's [13] "Corpuscle Problem."

Suppose that $\mathbf{X}$ has a density of the form $\rho(x_1^2 + x_2^2 + x_3^2)$. Then $Y := X_1^2 + X_2^2 \sim G$ and Z have densities

$$g(y) = \pi \int_y^\infty \frac{\rho(z)}{\sqrt{z - y}}\, \mathrm{d}z \tag{1.1}$$

and $f(z) = 2\pi\sqrt{z}\rho(z)$. The reader may recognize (1.1) as Abel's transformation. It may be inverted as follows. Let

$$V(y) = \int_y^\infty \frac{g(u)}{\sqrt{u-y}}\,\mathrm{d}u.$$

Then

$$V(y) = \pi \int_y^\infty \left[\int_u^\infty \frac{\rho(z)\,\mathrm{d}z}{\sqrt{z-u}}\right] \frac{\mathrm{d}u}{\sqrt{u-y}} = \pi^2 \int_y^\infty \rho(z)\,\mathrm{d}z \tag{1.2}$$

so that $\rho(z) = -V'(z)/\pi^2$ at continuity points. Observe that V is a *nonincreasing* function. The quantity of interest, F, can be related to V and, therefore, to the distribution of (X_1, X_2) by

$$F(x) = \int_0^x 2\pi\sqrt{u}\rho(u)\,\mathrm{d}u = 1 + \frac{2}{\pi}\int_x^\infty \sqrt{z}\,\mathrm{d}V(z), \tag{1.3}$$

where the last equality follows from $\int_0^\infty 2\pi\sqrt{u}\rho(u)\,\mathrm{d}u = 1$. Relationship (1.3) will be used extensively in the sequel. Let

$$U(x) := \int_0^x V(t)\,\mathrm{d}t$$

for $x > 0$. Then U is concave since V is nonincreasing. Concavity can also be seen from

$$U(x) = 2\int_0^\infty \left\{\sqrt{u} - \sqrt{(u-x)_+}\right\} g(u)\,\mathrm{d}u,$$

where $y_+ = \max\{y, 0\}$. Let $J(t) := \int_t^\infty \sqrt{z-t}\,\mathrm{d}V(z)$. Then

$$\begin{aligned} G(t) &= \pi \int_0^\infty \int_0^{t\wedge z} \frac{\rho(z)}{\sqrt{z-y}}\,\mathrm{d}y\,\mathrm{d}z \\ &= 2\pi \int_0^\infty \left\{\sqrt{z} - \sqrt{(z-t)_+}\right\}\rho(z)\,\mathrm{d}z = 1 + \frac{2}{\pi}J(t), \end{aligned} \tag{1.4}$$

where the last step follows from $\int_0^\infty 2\pi\sqrt{z}\rho(z)\,\mathrm{d}z = 1$ and $J(t) = -\pi^2\int_t^\infty \sqrt{z-t}\,\rho(z)\,\mathrm{d}z$ (using (1.2)).

Now suppose that we observe an i.i.d. sample $\{(X_{i1}, X_{i2})\}_{i=1}^n$ having the same distribution as (X_1, X_2). Letting $Y_i = X_{i1}^2 + X_{i2}^2$, a natural (unbiased) "naive" estimator of V is

$$V_n^\#(y) := \int_y^\infty \frac{\mathrm{d}G_n^\#(u)}{\sqrt{u-y}} = \frac{1}{n}\sum_{i=1}^n \frac{\mathbf{1}\{Y_i > y\}}{\sqrt{Y_i - y}},$$

where $G_n^\#$ is the empirical distribution function (EDF) of the Y_i's. Then $V_n^\#(y)$ is an unbiased estimator of $V(y)$ for each fixed y; but $V_n^\#$ has infinite discontinuities at the data points Y_i and

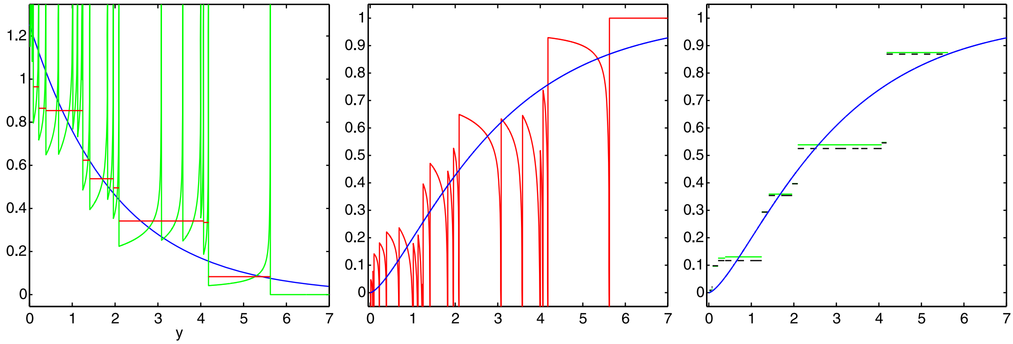

Figure 1. Left panel: Plots of $V_n^\#$ (green), $\tilde{V}_n^\#$ (red, piece-wise constant) and V (blue, smooth); middle panel: $F_n^\#$ (red), F (blue, smooth); right panel: $\tilde{F}_n^\#$ (green, piece-wise constant), $\check{F}_n$ (black, piece-wise constant) and F (blue, smooth) from a sample with $n = 20$ data points.

is, therefore, not monotonic when viewed as a function of y. See Figure 1. We call $V_n^\#$ the *naive estimator*. The naive estimator can be improved by requiring monotonicity. If $V_n^\#$ were square integrable, this could be accomplished by minimizing the integral of $(W - V_n^\#)^2$ over all nonincreasing functions W, or equivalently,

$$\int_0^\infty W^2(y)\,\mathrm{d}y - 2\int_0^\infty W(y)V_n^\#(y)\,\mathrm{d}y. \tag{1.5}$$

The function $V_n^\#$ is not square integrable, but it is integrable, so (1.5) is well defined. Let $\tilde{V}_n^\#$ be the nonincreasing function W that minimizes (1.5). Existence and uniqueness can be shown along the lines of Theorem 1.2.1 of Robertson *et al.* [9], replacing the sums by integrals. Groeneboom and Jongbloed [6] derived the limit distributions of $V_n^\#$ and $\tilde{V}_n^\#$: Let $x_0 > 0$ and

$$\varepsilon_n := \sqrt{n^{-1}\log n}, \tag{1.6}$$

then under appropriate conditions,

$$\frac{V_n^\#(x_0) - V(x_0)}{\varepsilon_n} \Rightarrow N(0, g(x_0)), \tag{1.7}$$

$$\frac{\tilde{V}_n^\#(x_0) - V(x_0)}{\varepsilon_n} \Rightarrow N\left(0, \frac{1}{2}g(x_0)\right), \tag{1.8}$$

where $\Rightarrow$ denotes weak convergence.

We can define two estimators of F, F_n and $\tilde{F}_n^\#$, by replacing V from the right-hand side of (1.3) with $V_n^\#$ and $\tilde{V}_n^\#$, respectively. Observe that $F_n^\#$ is not even nondecreasing; $\tilde{F}_n^\#$ is nondecreasing, and $\max\{\tilde{F}_n^\#, 0\}$ (as $\tilde{F}_n^\# \le 1$), is a valid distribution function and a more appealing estimator of F (see Figure 1).

Yet another estimator of F can be obtained by isotonizing $F_n^{\#}$ over all nondecreasing functions. Let $\check{F}_n$ be the nondecreasing function that is closest to $F_n^{\#}$, in the sense that it minimizes (1.5) with $V_n^{\#}$ replaced by $F_n^{\#}$. It is not difficult to see that then $\max\{0, \min(\check{F}_n, 1)\}$ is a valid distribution function. Figure 1 shows the graphs of the estimators $V_n^{\#}$, $\tilde{V}_n^{\#}$, $F_n^{\#}$, $\tilde{F}_n^{\#}$, and $\check{F}_n$ obtained from simulated data with $n = 20$.

It will be shown later that for $x_0 > 0$,

$$\frac{F_n^{\#}(x_0) - F(x_0)}{\varepsilon_n} \Rightarrow N\left(0, \frac{4}{\pi^2} x_0 g(x_0)\right), \tag{1.9}$$

$$\frac{\tilde{F}_n^{\#}(x_0) - F(x_0)}{\varepsilon_n} \Rightarrow N\left(0, \frac{2}{\pi^2} x_0 g(x_0)\right) \quad \text{and} \tag{1.10}$$

$$\frac{\check{F}_n(x_0) - F(x_0)}{\varepsilon_n} \Rightarrow N\left(0, \frac{2}{\pi^2} x_0 g(x_0)\right) \tag{1.11}$$

under modest conditions. As above the isotonized estimators have exactly half limiting variances of corresponding naive estimators.

Construction of confidence intervals for $F(x_0)$ using these limiting distributions is still complicated as they require the estimation of the nuisance parameter $g(x_0)$. Bootstrap intervals avoid this problem and are generally reliable and accurate in problems with $\sqrt{n}$ convergence rate (see Bickel and Freedman [4], Singh [12], Shao and Tu [11] and its references). However, conventional bootstrap estimators are inconsistent for some shape restricted estimators – dramatically so for the Grenander estimator. See Kosorok [8], Abrevaya and Huang [1] and Sen *et al.* [10] and its references. So, it is not a priori clear whether bootstrap methods are consistent in the present context. We show that they are.

In Section 2, we prove uniform versions of (1.7), (1.8), (1.9), (1.10) and (1.11). These are used in Section 3 to establish the consistency of bootstrap methods in approximating the sampling distribution of the various estimators of V and F, while generating samples from the EDF. Using data on the globular cluster M62 we illustrate the isotonized estimators of F along with the corresponding bootstrap based point-wise CIs in Section 4. Section A, the Appendix, gives the details of some of the arguments in the proofs of the main results.

2. Uniform convergence

In this section, we prove central limit theorems for estimates of V and F when we have a triangular array of random variables whose row-distributions satisfy certain regularity conditions. This generalization will also help us analyze the asymptotic properties of the bootstrap estimators (to be introduced in Section 3). Note that conditional on the data, bootstrap samples can be embedded in a triangular array of random variables, with the nth row being generated from a distribution (built from the first n data points) that approximates the data-generating mechanism.

Suppose that we have i.i.d. triangular data $\{Y_{n,i}\}_{i=1}^{n}$ having distribution function G_n. We consider a special construction of $Y_{n,i}$, namely, let $Y_{n,i} = G_n^{-1}(T_i)$, where $G_n^{-1}(u) = \inf\{x\colon\ G_n(x) \geq$

$u\}$ and $T_1, T_2, \dots$ are i.i.d. Uniform(0, 1) random variables. Let V_n and U_n be defined as

$$V_n(y) = \int_y^\infty \frac{\mathrm{d}G_n(u)}{\sqrt{u-y}} \quad \text{and} \quad U_n(x) = \int_0^x V_n(y)\,\mathrm{d}y.$$

Let LCM_I be the operator that maps a function $h:\mathbb{R}\to\mathbb{R}$ into the least concave majorant (LCM) of its restriction to the interval $I \subset \mathbb{R}$. Define $\tilde{V}_n := \mathrm{LCM}_{[0,\infty)}[U_n]'$ where $'$ denotes the right derivative. Let $G_n^\#$ denote the EDF of $Y_{n,1}, Y_{n,2}, \dots, Y_{n,n}$,

$$V_n^\#(y) := \int_y^\infty \frac{\mathrm{d}G_n^\#(u)}{\sqrt{u-y}} = \frac{1}{n} \sum_{i:Y_{n,i}>y} \frac{1}{\sqrt{Y_{n,i}-y}}.$$

Then $V_n^\#$ is a nonmonotonic, unbiased estimate of $V_n(y)$, as above, and we call $V_n^\#$ the naive estimator. The naive estimator can be improved by imposing the monotonicity constraint as in (1.5) to obtain $\tilde{V}_n^\#$. Observe that

$$U_n^\#(x) := \frac{2}{n} \sum_{i=1}^n \left\{\sqrt{Y_{n,i}} - \sqrt{(Y_{n,i}-x)_+}\right\}$$

is an unbiased estimate of $U_n(x)$ for all $x \in [0,\infty)$; $U_n^\#$ is a nondecreasing function; $V_n^\#$ is the derivative of $U_n^\#$ a.e. Let $\tilde{U}_n^\#$ be the LCM of $U_n^\#$. Then $\tilde{V}_n^\#$ is the right-derivative of $\tilde{U}_n^\#$ (see, e.g., Lemma 2 of [6]). Let F_n and $F_n^\#$ be defined by replacing V from the right-hand side of (1.3) with V_n and $V_n^\#$, respectively.

2.1. CLT for estimates of V

Fix $x_0 \in (0,\infty)$ such that $g(x_0) > 0$. We consider two estimates of $V(x_0)$, namely $V_n^\#(x_0)$ and $\tilde{V}_n^\#(x_0)$. To find the limit distribution of $V_n^\#(x_0)$, we assume the following conditions on G_n:

$$V_n(x_0) \to V(x_0), \tag{2.1}$$

$$n\left\{G_n\left(x_0 + \frac{1}{\varepsilon^2 n \log n}\right) - G_n(x_0)\right\} \to 0 \qquad \text{for all } \varepsilon > 0, \tag{2.2}$$

$$\int_{x_0}^{x_0+c_n} \frac{\mathrm{d}G_n(y)}{\sqrt{y-x_0}} = \mathrm{o}(\varepsilon_n), \tag{2.3}$$

$$\frac{1}{\log n} \int_{x_0+c_n}^\infty \frac{\mathrm{d}G_n(y)}{y-x_0} \to g(x_0), \tag{2.4}$$

where $c_n = 1/(\sqrt{n\log n} + V_n(x_0))^2$ and ε_n is defined in (1.6).

Proposition 2.1. *If* (2.1)–(2.4) *hold then* $\varepsilon_n^{-1}\{V_n^\#(x_0) - V_n(x_0)\} \Rightarrow N(0, g(x_0))$.

The proof of the proposition is given in the Appendix. Next, we study the limiting distribution of

$$\Delta_n := \frac{\tilde{V}_n^{\#}(x_0) - \hat{V}_n(x_0)}{\varepsilon_n},$$

where $\hat{V}_n(x_0)$ can be $V_n(x_0)$ or $\tilde{V}_n(x_0)$. Define the stochastic process

$$\mathbb{Z}_n(t) = \varepsilon_n^{-2}\{U_n^{\#}(x_0 + \varepsilon_n t) - U_n^{\#}(x_0) - \hat{V}_n(x_0)\varepsilon_n t\}$$

for $t \in I_n := [-\varepsilon_n^{-1}x_0, \infty)$ and note that $\Delta_n = \mathrm{LCM}_{I_n}[\mathbb{Z}_n]'(0)$, that is, Δ_n is the right-hand slope at 0 of the LCM of the process $\mathbb{Z}_n$. We will study the limiting behavior of the process $\mathbb{Z}_n$ and use continuous mapping arguments to derive the limiting distribution of Δ_n. We consider all stochastic processes as random elements in $C(\mathbb{R})$, the space of continuous functions on $\mathbb{R}$, equipped it with the Borel σ-field and the metric of uniform convergence on compacta. To better understand the limiting behavior of $\mathbb{Z}_n$, we decompose $\mathbb{Z}_n$ into the sum of

$$\mathbb{Z}_{n,1}(t) = \varepsilon_n^{-2}\{(U_n^{\#} - U_n)(x_0 + \varepsilon_n t) - (U_n^{\#} - U_n)(x_0)\} \quad \text{and}$$
$$\mathbb{Z}_{n,2}(t) = \varepsilon_n^{-2}\{U_n(x_0 + \varepsilon_n t) - U_n(x_0) - \hat{V}_n(x_0)\varepsilon_n t\}.$$

Observe that $\mathbb{Z}_{n,2}$ depends only on G_n and not on the $Y_{n,j}$. Let

$$\mathbb{Z}_1(t) = tW \quad \text{and} \quad \mathbb{Z}(t) = \mathbb{Z}_1(t) + \tfrac{1}{2}t^2 V'(x_0)$$

for $t \in \mathbb{R}$, where W is a normal random variable having mean 0 and variance $\frac{1}{2}g(x_0)$. We state some conditions on the behavior of G_n, $\hat{V}_n$ and U_n used to obtain the limiting distribution of Δ_n.

(a) $D_n := \|G_n - G\| = \mathrm{O}(\varepsilon_n)$, where $\|\cdot\|$ refers to the uniform norm, that is, $\|G_n - G\| = \sup_{t\in\mathbb{R}} |G_n(t) - G(t)|$.
(b) $\mathbb{Z}_{n,2}(t) \to \frac{1}{2}t^2 V'(x_0)$ as $n \to \infty$ uniformly on compacta.
(c) For each $\varepsilon > 0$,

$$\left|U_n(x_0 + \beta) - U_n(x_0) - \beta\hat{V}_n(x_0) - \tfrac{1}{2}\beta^2 V'(x_0)\right| \le \varepsilon\beta^2 + \mathrm{o}(\beta^2) + \mathrm{O}(\varepsilon_n^2)$$

for large n, uniformly in β varying over a neighborhood of zero.

Theorem 2.1. *Under condition* (a) *the distribution of* $\mathbb{Z}_{n,1}$ *converges to that of* $\mathbb{Z}_1$. *Further, if* (b) *holds, then the distribution of* $\mathbb{Z}_n$ *converges to that of* $\mathbb{Z}$.

Proof. The covariance of $\mathbb{Z}_{n,1}(s)$ and $\mathbb{Z}_{n,1}(t)$, is needed. To compute it let

$$\phi(y, \eta) := \sqrt{(y - x_0)_+} - \sqrt{(y - x_0 - \eta)_+}$$

for $y, \eta \in \mathbb{R}$ and observe the following two properties:

$$\phi(\cdot, \eta) \le \sqrt{|\eta|}, \tag{P1}$$

$$\int_0^\infty |\phi'(y, \eta)| \, \mathrm{d}y \le 2\sqrt{|\eta|} \tag{P2}$$

of which the second follows from splitting the interval of integration into $[0, x_0 + \eta]$ and $(x_0 + \eta, \infty)$. Observe that

$$\mathbb{Z}_{n,1}(t) = \frac{2}{\varepsilon_n^2} \int \phi(u, \varepsilon_n t) \, \mathrm{d}(G_n^\# - G_n)(u)$$

and

$$\operatorname{Cov}(\mathbb{Z}_{n,1}(s), \mathbb{Z}_{n,1}(t)) = \frac{4}{n\varepsilon_n^4} \operatorname{Cov}(\phi(Y_{n,1}, \varepsilon_n s), \phi(Y_{n,1}, \varepsilon_n t)),$$

where $Y_{n,1} \sim G_n$. We first show that $E[\phi(Y_{n,1}, \varepsilon_n t)] = \mathrm{O}(\varepsilon_n)$, so that $\operatorname{Cov}(\mathbb{Z}_{n,1}(s), \mathbb{Z}_{n,1}(t)) = (1/n\varepsilon_n^4) E[\phi(Y_{n,1}, \varepsilon_n s)\phi(Y_{n,1}, \varepsilon_n t)] + \mathrm{o}(1)$. For this, observe that

$$E[\phi(Y_{n,1}, \varepsilon_n t)] = 2 \int \phi(u, \varepsilon_n t) \, \mathrm{d}(G_n - G)(u) + \{U(x_0 + t\varepsilon_n) - U(x_0)\}. \tag{2.5}$$

The first term is at most

$$\begin{aligned} 2\int_0^\infty (G_n - G)(u)\phi'(u, \varepsilon_n t) \, \mathrm{d}u &\le 2D_n \int_0^\infty |\phi'(u, \varepsilon_n t)| \, \mathrm{d}u \\ &= 2\mathrm{O}(\varepsilon_n) 2\sqrt{|\varepsilon_n t|} = \mathrm{O}(\varepsilon_n^{3/2}), \end{aligned}$$

and the second term in (2.5) is at most $\mathrm{O}(\varepsilon_n)$ by using a one term Taylor expansion. Next, suppose that $s \le t$ and write $E[\phi(Y_{n,1}, \varepsilon_n s)\phi(Y_{n,1}, \varepsilon_n t)]$ as

$$\int \phi(u, \varepsilon_n s)\phi(u, \varepsilon_n t) \, \mathrm{d}(G_n - G)(u) + \int \phi(u, \varepsilon_n s)\phi(u, \varepsilon_n t) \, \mathrm{d}G(u). \tag{2.6}$$

From Lemma 3 of Groeneboom and Jongbloed [6], page 1539,

$$\int \phi(u, \varepsilon_n s)\phi(u, \varepsilon_n t) \, \mathrm{d}G(u) = -\frac{1}{4} g(x_0) st \varepsilon_n^2 \log \varepsilon_n + \mathrm{O}(\varepsilon_n^2). \tag{2.7}$$

Using integration by parts, (P1), and (P2), the first term in (2.6) is at most

$$\begin{aligned} &\left| \int_0^\infty \{\phi'(u, \varepsilon_n s)\phi(u, \varepsilon_n t) + \phi(u, \varepsilon_n s)\phi'(u, \varepsilon_n t)\}(G_n - G)(u) \, \mathrm{d}u \right| \\ &\quad \le 2D_n \sqrt{|\varepsilon_n t|} \left\{ 2\sqrt{|\varepsilon_n s|} \right\} = \mathrm{O}(\varepsilon_n^2). \end{aligned}$$

So,

$$\begin{aligned}\operatorname{Cov}(\mathbb{Z}_{n,1}(s),\mathbb{Z}_{n,1}(t)) &= \frac{4}{n\varepsilon_n^4}\left\{\frac{1}{4}stg(x_0)\varepsilon_n^2\log\left(\frac{1}{\varepsilon_n}\right)+\mathrm{O}(\varepsilon_n^2)\right\}\\ &= \frac{1}{2}g(x_0)st\left\{1-\frac{\log\log n}{\log n}\right\}+\mathrm{O}\left(\frac{1}{\log n}\right).\end{aligned}\tag{2.8}$$

It follows directly from the Lindeberg–Feller central limit theorem for triangular arrays that $\mathbb{Z}_{n,1}(1)\Rightarrow N(0,\frac{1}{2}g(x_0))$; and Chebyshev's inequality implies that $|s\mathbb{Z}_{n,1}(t)-t\mathbb{Z}_{n,1}(s)|=\mathrm{o}_P(1)$ as $n\to\infty$ for all for all fixed $s,t\in\mathbb{R}$. So, the finite dimensional distributions of $\mathbb{Z}_{n,1}$ converges weakly to the finite dimensional distributions of $\mathbb{Z}_1$.

For the the convergence in distribution of $\mathbb{Z}_{n,1}$ to $\mathbb{Z}_1$ in $C(\mathbb{R})$, it suffices to show that for each $M>0$ and sequence of positive numbers $\{\delta_n\}$ converging to zero,

$$E\{\sup|\mathbb{Z}_{n,1}(s)-\mathbb{Z}_{n,1}(t)|\colon |s-t|\le\delta_n,\max(|s|,|t|)\le M\}\to 0.$$

See Theorem 2.3 of Kim and Pollard [7]. Consider the class of functions $\mathcal{C}_R=\{\phi(\cdot,\eta)\colon|\eta|<R\}$ with its natural envelope $\Phi_R(y):=\sqrt{(y-x_0+R)_+}-\sqrt{(y-x_0-R)_+}$. Observe that $\mathcal{C}_R$ are uniformly manageable for its envelope Φ_R and that $\Phi_R\le\sqrt{2R}$. Let δ_n be a sequence of positive numbers converging to zero, $h(y;s,t):=\phi(y,t)-\phi(y,s)=\sqrt{(y-x_0-s)_+}-\sqrt{(y-x_0-t)_+}$ for $y,s,t\in\mathbb{R}$, and $\mathcal{H}_n:=\{h(\cdot;s\varepsilon_n,t\varepsilon_n)\colon\max(|s|,|t|)<M,|s-t|\le\delta_n\}$. The class $\mathcal{H}_n$ has envelope $H_n:=2\Phi_{M\varepsilon_n}$. Observe that

$$\mathbb{Z}_{n,1}(t)=2\varepsilon_n^{-2}(G_n^{\#}-G_n)\phi(\cdot,t\varepsilon_n).$$

So, it suffices to show that $\varepsilon_n^{-2}E[\sup_{h\in\mathcal{H}_n}|(G_n^{\#}-G_n)h|]=\mathrm{o}(1)$. Define $S_n:=G_n^{\#}H_n^2/(n\varepsilon_n^4)$ and $T_n:=\sup_{h\in\mathcal{H}_n}G_n^{\#}h^2$. Then by the maximal inequality of Section 3.1 in Kim and Pollard [7], there is a (single) continuous function $J(\cdot)$ for which $J(0)=0$, $J(1)<\infty$, and

$$\begin{aligned}\frac{1}{\varepsilon_n^2}E\Big[\sup_{h\in\mathcal{H}_n}|G_n^{\#}h-G_nh|\Big] &\le \frac{1}{\varepsilon_n^2\sqrt{n}}E\left[\sqrt{G_n^{\#}H_n^2}\,J\left(\sup_{\mathcal{H}_n}\frac{G_n^{\#}h^2}{G_n^{\#}H_n^2}\right)\right]\\ &= E\left[\sqrt{S_n}\,J\left(\frac{T_n}{n\varepsilon_n^4S_n}\right)\right].\end{aligned}$$

Let $\eta>0$. Splitting according to whether $\{S_n\le\eta\}$ or not, using the fact that $n\varepsilon_n^4S_n\ge T_n$ and invoking the Cauchy–Schwarz inequality for the contribution from $\{S_n>\eta\}$, we may bound the last expected value by

$$\begin{aligned}&E\left[\sqrt{S_n}1\{S_n\le\eta\}J\left(\frac{T_n}{n\varepsilon_n^4S_n}\right)\right]+E\left[\sqrt{S_n}1\{S_n>\eta\}J\left(\frac{T_n}{n\varepsilon_n^4S_n}\right)\right]\\ &\quad\le\sqrt{\eta}J(1)+\sqrt{ES_n}\sqrt{EJ^2\left(\min\left(1,\frac{T_n}{n\varepsilon_n^4\eta}\right)\right)}.\end{aligned}$$

Noting that $\Phi_{M\varepsilon_n} = \phi(\cdot, M\varepsilon_n) - \phi(\cdot, -M\varepsilon_n)$ and using (2.7) and (2.8) with $-s = t = M$, we have

$$ES_n = \frac{1}{n\varepsilon_n^4} E\left[\frac{1}{n}\sum_{i=1}^{n} H_n^2(Y_{n,i})\right] = \frac{G_n H_n^2}{n\varepsilon_n^4} = \mathrm{O}(1). \tag{2.9}$$

So, it suffices to show that $T_n = \mathrm{o}_P(n\varepsilon_n^4)$, which implies $E[J^2(\min(1, T_n/(n\varepsilon_n^4\eta))] \to 0$ (note that $J(1) < \infty$). We will establish the stronger result $ET_n = \mathrm{o}(n\varepsilon_n^4)$. Observe that

$$E\Big[\sup_{\mathcal{H}_n} G_n^{\#} h^2\Big] \le E \sup_{\mathcal{H}_n} G_n h^2 + E\Big[\sup_{\mathcal{H}_n} |G_n^{\#} h^2 - G_n h^2|\Big]$$

and

$$\begin{aligned} G_n h^2 &= G_n[\phi(y, t\varepsilon_n) - \phi(y, s\varepsilon_n)]^2 = -\tfrac{1}{4} g(x_0)(s-t)^2 \varepsilon_n^2 \log\varepsilon_n + \mathrm{O}(\varepsilon_n^2) \\ &= \mathrm{O}(\delta_n^2 n\varepsilon_n^4) + \mathrm{O}(\varepsilon_n^2) = \mathrm{o}(n\varepsilon_n^4) \end{aligned}$$

by (2.7). The maximal inequality applied to the uniformly manageable class $\{h^2\colon h \in \mathcal{H}_n\}$ with envelope H_n^2 bounds the second term by $\tilde{J}(1)\sqrt{G_n H_n^4/n} \le 8M\varepsilon_n/\sqrt{n} = \mathrm{o}(n\varepsilon_n^4)$, where we have used (2.9) and the fact that $H_n^2 \le 8M\varepsilon_n$. That $\mathbb{Z}_n$ converges in distribution to $\mathbb{Z}$ in $C(\mathbb{R})$ follows directly. □

A rigorous proof of the convergence of Δ_n involves a little more than an application of a continuous mapping theorem. The convergence $\mathbb{Z}_n \Rightarrow \mathbb{Z}$ is only in the sense of the metric of uniform convergence on compacta. A concave majorant near the origin might be determined by values of the process long way from the origin; the convergence $\mathbb{Z}_n \Rightarrow \mathbb{Z}$ by itself does not imply the convergence $\mathrm{LCM}_{I_n}[\mathbb{Z}_n] \Rightarrow \mathrm{LCM}_{\mathbb{R}}[\mathbb{Z}]$. We need to show that $\mathrm{LCM}_{I_n}[\mathbb{Z}_n]$ is determined by values of $\mathbb{Z}_n$ for t in an $\mathrm{O}_P(1)$ neighborhood of the origin. Corollary 2.1 shows the convergence of Δ_n, and its proof is given in the Appendix.

Corollary 2.1. *Under conditions* (a)–(c), *the distribution of* Δ_n *converges to that of* $W \stackrel{d}{=} \mathrm{LCM}_{\mathbb{R}}[\mathbb{Z}]'(0)$.

2.2. CLT for estimates of F

We consider three estimates of F, namely $F_n^{\#}$, $\tilde{F}_n^{\#}$ and $\check{F}_n^{\#}$, where $F_n^{\#}$ and $\tilde{F}_n^{\#}$ are obtained by replacing V from the right-hand side of (1.3) with $V_n^{\#}$ and $\tilde{V}_n^{\#}$, respectively; and $\check{F}_n^{\#}$ is the closest (in the sense of minimizing (1.5) with V_n replaced with $F_n^{\#}$) nondecreasing function to $F_n^{\#}$. We start by deriving the limit distribution of $F_n^{\#}$. Let $\sigma^2 := \mathrm{Var}[\sin^{-1}\sqrt{1 \wedge (x_0/Y)}]$ where $Y \sim G$.

Proposition 2.2. *If* $g(x_0) > 0$ *and* $\|G_n - G\| \to 0$ *as* $n \to \infty$, *then*

$$\sqrt{n}\int_{x_0}^{\infty} \frac{V_n^{\#}(u) - V_n(u)}{2\sqrt{u}}\,\mathrm{d}u \Rightarrow N(0, \sigma^2). \tag{2.10}$$

If also (2.1)–(2.4) *hold*, *then*

$$\frac{F_n^{\#}(x_0) - F_n(x_0)}{\varepsilon_n} \Rightarrow N\left(0, \frac{4}{\pi^2} x_0 g(x_0)\right). \tag{2.11}$$

Proof. For (2.10), observe that

$$\int_{x_0}^{\infty} \frac{V_n^{\#}(u)}{2\sqrt{u}}\,\mathrm{d}u = \frac{1}{n}\sum_{i=1}^{n}\int_{x_0}^{\infty} \frac{\mathbf{1}\{Y_{n,i} > u\}}{2\sqrt{u}\sqrt{Y_{n,i} - u}}\,\mathrm{d}u = \frac{\pi}{2} - \frac{1}{n}\sum_{i=1}^{n}\sin^{-1}\sqrt{1 \wedge \frac{x_0}{Y_{n,i}}}$$

after some simplification, and (similarly),

$$\int_{x_0}^{\infty} \frac{V_n(u)}{2\sqrt{u}}\,\mathrm{d}u = \frac{\pi}{2} - \int_0^{\infty} \sin^{-1}\sqrt{1 \wedge \frac{x_0}{y}}\,\mathrm{d}G_n(y).$$

Relation (2.10) now follows from the Lindeberg–Feller CLT. For (2.11), first observe that $F_n^{\#}(x_0) - F_n(x_0)$ may be written as

$$-\frac{2}{\pi}\sqrt{x_0}\{V_n^{\#}(x_0) - V_n(x_0)\} - \frac{2}{\pi}\int_{x_0}^{\infty} \frac{V_n^{\#}(u) - V_n(u)}{2\sqrt{u}}\,\mathrm{d}u.$$

From Proposition 2.1, $\varepsilon_n^{-1}\{V_n^{\#}(x_0) - V_n(x_0)\} \Rightarrow N(0, g(x_0))$. Relation (2.11) follows directly from this and (2.10). □

Applying the proposition with $G_n = G$ verifies (1.9). Next, we derive the limiting distribution of $\tilde{F}_n^{\#}$.

Proposition 2.3. *Suppose that* (a)–(c) *hold with* $\hat{V}_n = \tilde{V}_n$, *then*,

$$\frac{\tilde{F}_n^{\#}(x_0) - \tilde{F}_n(x_0)}{\varepsilon_n} \Rightarrow N\left(0, \frac{2}{\pi^2} x_0 g(x_0)\right).$$

Proof. As above $\tilde{F}_n^{\#}(x_0) - \tilde{F}_n(x_0)$ may be written as

$$\frac{2}{\pi}\sqrt{x_0}\{\tilde{V}_n(x_0) - \tilde{V}_n^{\#}(x_0)\} + \frac{2}{\pi}\int_{x_0}^{\infty} \frac{\tilde{V}_n(u) - \tilde{V}_n^{\#}(u)}{2\sqrt{u}}\,\mathrm{d}u.$$

From Corollary 2.1, $\varepsilon_n^{-1}\{\tilde{V}_n(x_0) - \tilde{V}_n^{\#}(x_0)\} \Rightarrow N(0, \frac{1}{2}g(x_0))$. Integrating by parts, the integral on the last display is a most

$$\begin{aligned}
&\frac{|\tilde{U}_n(x_0) - \tilde{U}_n^{\#}(x_0)|}{2\sqrt{x_0}} + \frac{1}{4}\left|\int_{x_0}^{\infty} \frac{\tilde{U}_n(u) - \tilde{U}_n^{\#}(u)}{u^{3/2}}\,\mathrm{d}u\right| \\
&\quad \le \frac{\|\tilde{U}_n - \tilde{U}_n^{\#}\|}{2\sqrt{x_0}} + \|\tilde{U}_n - \tilde{U}_n^{\#}\|\frac{1}{2\sqrt{x_0}}
\end{aligned}$$

$$= \frac{\|\tilde{U}_n - \tilde{U}_n^{\#}\|}{\sqrt{x_0}} \le \frac{\|U_n - U_n^{\#}\|}{\sqrt{x_0}}$$
$$= \mathrm{O}_P(n^{-1/2}) = \mathrm{o}_P(\varepsilon_n)$$

by Marshall's lemma and maximal inequality 3.1 of Kim and Pollard [7] (to bound $\|U_n - U_n^{\#}\|$). The proposition follows. □

Now let $H_n(x) := \int_0^x F_n(z)\,\mathrm{d}z$ and $H_n^{\#}(x) := \int_0^x F_n^{\#}(z)\,\mathrm{d}z$. Note that $F_n^{\#}$ is the derivative of $H_n^{\#}$ a.e. Let $\check{H}_n^{\#}$ be the greatest convex minorant (GCM) of $H_n^{\#}$. Then $\check{F}_n$ is the right-derivative of $\check{H}_n^{\#}$. We want to study the limit distribution of

$$\Lambda_n := \frac{\check{F}_n(x_0) - \hat{F}_n(x_0)}{\varepsilon_n},$$

where $\hat{F}_n$ can be F_n or $\tilde{F}_n$. Let

$$\mathbb{X}_n(t) := \varepsilon_n^{-2}\{H_n^{\#}(x_0 + \varepsilon_n t) - H_n^{\#}(x_0) - \hat{F}_n(x_0)\varepsilon_n t\}$$

for $t \in I_n := [-\varepsilon_n^{-1} x_0, \infty)$. As before, we decompose $\mathbb{X}_n$ into $\mathbb{X}_{n,1}$ and $\mathbb{X}_{n,2}$ where

$$\mathbb{X}_{n,1}(t) := \varepsilon_n^{-2}\{(H_n^{\#} - H_n)(x_0 + \varepsilon_n t) - (H_n^{\#} - H_n)(x_0)\} \quad \text{and}$$
$$\mathbb{X}_{n,2}(t) := \varepsilon_n^{-2}\{H_n(x_0 + \varepsilon_n t) - H_n(x_0) - \hat{F}_n(x_0)\varepsilon_n t\}.$$

Let GCM_I be the operator that maps the restriction of a function $h : \mathbb{R} \to \mathbb{R}$ to the interval I into its GCM, and observe that $\Lambda_n = \mathrm{GCM}_{I_n}[\mathbb{X}_n]'(0)$. Also let

$$\mathbb{X}_1(t) = tW \quad \text{and} \quad \mathbb{X}(t) = \mathbb{X}_1(t) + \tfrac{1}{2}t^2 f(x_0)$$

for $t \in \mathbb{R}$, where W is a normal random variable having mean 0 and variance $2x_0 g(x_0)/\pi^2$ and f is the density of $Z = X_1^2 + X_2^2 + X_3^2$. The following conditions will be used.

(b′) $\mathbb{X}_{n,2}(t) \to \frac{1}{2}t^2 f(x_0)$ as $n \to \infty$ uniformly on compacta.
(c′) For each $\varepsilon > 0$,

$$\left|H_n(x_0 + \beta) - H_n(x_0) - \beta\hat{F}_n(x_0) - \tfrac{1}{2}\beta^2 f(x_0)\right| \le \varepsilon\beta^2 + \mathrm{o}(\beta^2) + \mathrm{O}(\varepsilon_n^2)$$

for large n, uniformly in β varying over a neighborhood of zero.

Theorem 2.2. *Under condition* (a), *the distribution of* $\mathbb{X}_{n,1}$ *converges to that of* $\mathbb{X}_1$. *Further, if* (b′) *holds, then the distribution of* $\mathbb{X}_n$ *converges to that of* $\mathbb{X}$.

Proof. Using the definitions of $F_n^{\#}$, H_n and $H_n^{\#}$, we may write $H_n^{\#}(x) - H_n(x)$ as

$$-\frac{2}{\pi}\left[\int_0^x \sqrt{z}\{V_n^{\#}(z) - V_n(z)\}\,\mathrm{d}z + \int_0^x \int_z^{\infty} \frac{(V_n^{\#} - V_n)(u)}{2\sqrt{u}}\,\mathrm{d}u\,\mathrm{d}z\right]$$

and

$$
\begin{aligned}
\left|\int_z^\infty \frac{(V_n^\# - V_n)(u)}{2\sqrt{u}}\,\mathrm{d}u\right| &\le \left|\int_z^\infty \int_z^y \frac{\mathrm{d}u}{2\sqrt{y-u}\sqrt{u}}\,\mathrm{d}(G_n^\# - G_n)(y)\right| \\
&= \left|\int_z^\infty \left\{\frac{\pi}{2} - \sin^{-1}\sqrt{\frac{z}{y}}\right\}\mathrm{d}(G_n^\# - G_n)(y)\right| \\
&= \left|\int_z^\infty \frac{\mathrm{d}}{\mathrm{d}y}\left[\sin^{-1}\sqrt{\frac{z}{y}}\right](G_n^\# - G)(y)\,\mathrm{d}y\right| \\
&= \frac{\pi}{2}\|G_n^\# - G_n\| = \mathrm{o}(\varepsilon_n) \qquad \text{a.s.},
\end{aligned}
$$

using the Law of Iterated Logarithms for $\|G_n^\# - G_n\| = \mathrm{o}(\varepsilon_n)$ a.s. Fix a compact set $K = [-M, M]$. Then

$$
\begin{aligned}
\mathbb{X}_{n,1}(t) &= -\frac{2}{\pi}\int_{x_0}^{x_0+\varepsilon_n t} \frac{\sqrt{z}[V_n^\#(z) - V_n(z)]}{\varepsilon_n^2}\,\mathrm{d}z + \mathrm{o}(1) \\
&= -\frac{2}{\pi}\sqrt{x_0}\int_{x_0}^{x_0+\varepsilon_n t} \frac{[V_n^\#(z) - V_n(z)]}{\varepsilon_n^2}\,\mathrm{d}z + \mathrm{o}_P(1) \qquad (2.12)\\
&= -\frac{2}{\pi}\sqrt{x_0}\,\mathbb{Z}_{n,1}(t) + \mathrm{o}_P(1)
\end{aligned}
$$

uniformly on K. Note that (2.12) follows as $|\int_{x_0}^{x_0+\varepsilon_n t}(\sqrt{z} - \sqrt{x_0})[V_n^\#(z) - V_n(z)]/\varepsilon_n^2\,\mathrm{d}z|$ can be bounded, using integration by parts, by

$$
\left|\sqrt{x_0 + \varepsilon_n t} - \sqrt{x_0}\right| \max_{|s|\le M} |\mathbb{Z}_{n,1}(s)| + \max_{|s|\le M} |\mathbb{Z}_{n,1}(s)| \int_{x_0}^{x_0+\varepsilon_n t} \frac{\mathrm{d}z}{2\sqrt{z}} = \mathrm{o}_P(1), \qquad (2.13)
$$

as $\max_{|s|\le M} |\mathbb{Z}_{n,1}(s)| = \mathrm{O}_P(1)$. The theorem now follows. □

Corollary 2.2. *Under conditions* (a), (b′) *and* (c′), *the distribution of* Λ_n *converges to that of* $W \stackrel{d}{=} \mathrm{GCM}_{\mathbb{R}}[\mathbb{X}]'(0)$.

The proof is very similar to that Corollary 2.1 with the LCMs changed to GCMs. The modifications are outlined in the Appendix.

3. Consistency of the bootstrap

We begin with a brief discussion on the bootstrap. Suppose we have i.i.d. random variables (vectors) $T_1, T_2, \ldots, T_n$ having an unknown distribution function λ defined on a probability space $(\Omega, \mathcal{A}, P)$ and we seek to estimate the sampling distribution of the random variable $R_n(\mathbf{T}_n, \lambda)$, based on the observed data $\mathbf{T}_n = (T_1, T_2, \ldots, T_n)$. Let μ_n be the distribution function of $R_n(\mathbf{T}_n, \lambda)$. The bootstrap methodology can be broken into three simple steps:

Step 1: Construct an estimate λ_n of λ based on the data (for example, the EDF).

Step 2: With λ_n fixed, draw a random sample of size n from λ_n, say $\mathbf{T}_n^* = (T_1^*, T_2^*, \ldots, T_n^*)$ (identically distributed and conditionally independent given $\mathbf{T}_n$). This is called the *bootstrap sample*.

Step 3: Approximate the sampling distribution of $R_n(\mathbf{T}_n, \lambda)$ by the sampling distribution of $R_n^* = R_n(\mathbf{T}_n^*, \lambda_n)$. The sampling distribution of R_n^*, the bootstrap distribution, can be simulated on the computer by drawing a large number of bootstrap samples and computing R_n^* for each sample.

Thus the bootstrap estimator of the sampling distribution function of $R_n(\mathbf{T}_n, \lambda)$ is given by $\mu_n^*(x) = P^*\{R_n^* \le x\}$ where $P^*\{\cdot\}$ is the conditional probability given the data $\mathbf{T}_n$. Let L denote the Levy metric or any other metric metrizing weak convergence of distribution functions. We say that μ_n^* is *(weakly) consistent* if $L(\mu_n, \mu_n^*) \stackrel{P}{\to} 0$. Similarly, μ_n^* is *strongly consistent* if $L(\mu_n, \mu_n^*) \to 0$ a.s. If μ_n has a weak limit μ, for the bootstrap procedure to be consistent, μ_n^* must converge weakly to μ, in probability. In addition, if μ is continuous, we must have

$$\sup_{x \in \mathbb{R}} |\mu_n^*(x) - \mu(x)| \stackrel{P}{\to} 0 \qquad \text{as } n \to \infty.$$

3.1. Bootstrapping $\tilde{V}_n$

Given data $Y_1, Y_2, \ldots, Y_n \sim G$ let $G_n^\#$ denote its EDF. Suppose that we draw conditionally independent and identically distributed random variables $Y_{n,1}^*, Y_{n,2}^*, \ldots, Y_{n,n}^*$ having distribution function $G_n^\#$; and let G_n^* be the EDF of the bootstrap sample. Letting

$$V_n^*(y) := \frac{1}{n} \sum_{i : Y_{n,i}^* > y} \frac{1}{\sqrt{Y_{n,i}^* - y}} = \int \frac{\mathbf{1}_{[y,\infty)}(u)}{\sqrt{u-y}} \, \mathrm{d}G_n^*(u) \quad \text{and}$$

$$\begin{aligned} U_n^*(x) &:= \frac{2}{n} \sum_{i=1}^n \left\{ \sqrt{Y_{n,i}^*} - \sqrt{(Y_{n,i}^* - x)_+} \right\} \\ &= 2 \int \left\{ \sqrt{u} - \sqrt{(u-x)_+} \right\} \mathrm{d}G_n^*(u), \end{aligned}$$

the isotonic estimate of V based on the bootstrap sample is $\tilde{V}_n^* = \mathrm{LCM}_{[0,\infty)}[U_n^*]'$. The bootstrap estimator of the distribution function of $\Delta_n = \varepsilon_n^{-1}\{\tilde{V}_n(x_0) - V(x_0)\}$ is then the conditional distribution function of $\Delta_n^* := \varepsilon_n^{-1}\{\tilde{V}_n^*(x_0) - \tilde{V}_n(x_0)\}$ given the sample $Y_1, \ldots, Y_n$. To find its limit let

$$\mathbb{Z}_n^*(t) = \varepsilon_n^{-2}\{U_n^*(x_0 + \varepsilon_n t) - U_n^*(x_0) - \tilde{V}_n(x_0)\varepsilon_n t\}$$

for $t \in I_n := [-\varepsilon_n^{-1}x_0, \infty)$ and decompose $\mathbb{Z}_n^*$ into $\mathbb{Z}_{n,1}^*$ and $\mathbb{Z}_{n,2}^*$ where

$$\mathbb{Z}_{n,1}^*(t) = \varepsilon_n^{-2}\{(U_n^* - U_n^{\#})(x_0 + \varepsilon_n t) - (U_n^* - U_n^{\#})(x_0)\},$$
$$\mathbb{Z}_{n,2}^*(t) = \varepsilon_n^{-2}\{U_n^{\#}(x_0 + \varepsilon_n t) - U_n^{\#}(x_0) - \tilde{V}_n^{\#}(x_0)\varepsilon_n t\}.$$

Recall that $\mathbb{Z}_1(t) = tW$ and $\mathbb{Z}(t) = \mathbb{Z}_1(t) + \frac{1}{2}t^2 V'(x_0)$ are two processes defined for $t \in \mathbb{R}$, where W is a normal random variable having mean 0 and variance $\frac{1}{2}g(x_0)$. Let $\mathbf{Y} = (Y_1, Y_2, \ldots)$. The following theorem shows that bootstrapping from the EDF $G_n^{\#}$ is weakly consistent.

Theorem 3.1. *Suppose that V is continuously differentiable around x_0, and $g(x_0) \neq 0$. Then*:

(i) *The conditional distribution of the process $\mathbb{Z}_{n,1}^*$, given $\mathbf{Y}$, converges to that of $\mathbb{Z}_1$ a.s.*
(ii) *Unconditionally, $\mathbb{Z}_{n,2}^*(t)$ converges in probability to $\frac{1}{2}t^2 V'(x_0)$, uniformly on compacta.*
(iii) *The conditional distribution of the process $\mathbb{Z}_n^*$, given $\mathbf{Y}$, converges to that of $\mathbb{Z}$, in probability.*
(iv) *The bootstrap procedure is weakly consistent, that is, the conditional distribution of Δ_n^*, given $\mathbf{Y}$, converges to that of W, in probability.*

Proof. Assertion (i) follows directly from Theorem 2.1, applied with $G_n = G_n^{\#}$, $G_n^{\#} = G_n^*$ and $P\{\cdot\} = P^*\{\cdot\} = P\{\cdot|\mathbf{Y}\}$, since condition (a) required for Theorem 2.1 holds a.s. For (ii) and (iii), let

$$\mathbb{Z}_n^0(t) = \varepsilon_n^{-2}\{U_n^{\#}(x_0 + t\varepsilon_n) - U_n^{\#}(x_0) - \varepsilon_n t V(x_0)\}$$

for $t \in I_n$. By Theorem 2.1, applied with $G_n = G$, $V_n = V$ and $U_n = U$ for all n, $\mathbb{Z}_n^0$ converges in distribution to $\mathbb{Z}$. To prove (ii) observe that

$$\mathbb{Z}_{n,2}^*(t) = \mathbb{Z}_n^0(t) - t \cdot \mathrm{LCM}_{I_n}[\mathbb{Z}_n^0]'(0).$$

Unconditionally, using the continuous mapping theorem along with a localization argument as in Corollary 2.1, we obtain $\mathbb{Z}_{n,2}^*(t) \Rightarrow \mathbb{Z}(t) - t \cdot \mathrm{LCM}_{\mathbb{R}}[\mathbb{Z}]'(0) = \frac{1}{2}t^2 V'(x_0)$. As the limiting process is a constant, $\mathbb{Z}_{n,2}^*(t) \xrightarrow{P} \frac{1}{2}t^2 V'(x_0)$. Let $\{n_k\}$ be a subsequence of $\mathbb{N}$. We will show that there exists a further subsequence such that conditional on $\mathbf{Y}$, $\mathbb{Z}_n \Rightarrow \mathbb{Z}$ a.s. along the subsequence. Now, given $\{n_k\}$, there exists a further subsequence $\{n_{k_l}\}$ such that $\mathbb{Z}_{n_{k_l},2}^*(t) \to \frac{1}{2}t^2 V'(x_0)$ uniformly on compacta a.s. Thus, the conditional distribution of $\mathbb{Z}_{n_{k_l}}^*$ given $\mathbf{Y}$, converges to that of $\mathbb{Z}$, for a.e. $\mathbf{Y}$. This completes the proof of (iii).

For (iv), we use Corollary 2.1. Although conditions (a) and (b) hold in probability, condition (c) holds with $\hat{V}_n = \tilde{V}_n$ and the $\mathrm{O}(\varepsilon_n^2)$ term replaced by $\mathrm{O}_P(\varepsilon_n^2)$. Thus we cannot appeal directly to Corollary 2.1. Let $\xi > 0$ and $\eta > 0$ be given. We will show that there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $P\{L(K, K_n^*) > \xi\} < \eta$, where L is the Levy metric (Gnedenko and

Kolmogorov [5], page 33), K is the distribution function of $W \sim N(0, \frac{1}{2}g(x_0))$ and K_n^* is the distribution function of Δ_n^*, conditional on the data. For $\varepsilon > 0$, sufficiently small, let

$$A_n := \{|U_n^{\#}(x_0+\beta) - U_n^{\#}(x_0) - \beta \tilde{V}_n^{\#}(x_0) - \tfrac{1}{2}\beta^2 V'(x_0)| < C\varepsilon_n^2 + \varepsilon\beta^2\},$$

where $C > 0$ is chosen such that $P\{A_n^c\} < \frac{\eta}{2}$. This can be done since (c) holds with $\mathrm{O}(\varepsilon_n^2)$ term replaced by $\mathrm{O}_P(\varepsilon_n^2)$. Further, let $P_n^0(E)(\omega) = P^*(E)(\omega)$, if $\omega \in A_n$ and $P_n^0(E)(\omega) = P(E)$, if $\omega \notin A_n$; and let K_n^0 be the distribution function of Δ_n^* under the probability measure P_n^0. Observe that $K_n^0 = K_n^*$ on A_n and that $L(K, K_n^0) \stackrel{P_n^0}{\to} 0$ by Corollary 2.1 can be applied. Therefore, for all sufficiently large n,

$$\begin{aligned} P\{L(K, K_n^*) > \xi\} &\le P\left\{L(K, K_n^0) > \frac{\xi}{2}\right\} + P\left\{L(K_n^0, K_n^*) > \frac{\xi}{2}\right\} \\ &\le \frac{\eta}{2} + P\left\{L(K_n^0, K_n^*) > \frac{\xi}{2}, A_n^c\right\} \le \eta. \end{aligned}$$

This completes the proof of (iv). $\square$

Remark 3.1. Let $J_n(t) = \int_t^\infty \sqrt{z-t}\,\mathrm{d}V_n^{\#}(z)$, for $t \ge 0$, as in (1.4). Then $G_n^{\#} = 1 + 2J_n(t)/\pi$ after some simplification. So, using (1.4) to generate the bootstrap sample would lead back to $G_n^{\#}$.

3.2. Bootstrapping F_n, $\tilde{F}_n$ and $\check{F}_n$

Bootstrap versions of the three estimators of F under study, F_n^*, $\tilde{F}_n^*$ and $\check{F}_n^*$ say, are defined as in Section 2.2; for example, $F_n^*(x) = 1 + (2/\pi)\int_x^\infty \sqrt{z}\,\mathrm{d}V_n^*(z)$. We approximate the sampling distribution of $\varepsilon_n^{-1}\{F_n^{\#}(x_0) - F(x_0)\}$ by the bootstrap distribution of $\varepsilon_n^{-1}\{F_n^*(x_0) - F_n^{\#}(x_0)\}$. The bootstrap samples are generated from $G_n^{\#}$, the EDF of the Y_i's. By appealing to Proposition 2.2 with $G_n = G_n^{\#}$, it is easy to see that the bootstrap method is weakly consistent as (2.1)–(2.4) hold in probability.

The sampling distribution of $\varepsilon_n^{-1}\{\tilde{F}_n^{\#}(x_0) - F(x_0)\}$ is approximated by that of $\varepsilon_n^{-1}\{\tilde{F}_n^*(x_0) - \tilde{F}_n(x_0)\}$. Using Proposition 2.3, we can establish the consistency of the method. Note that the proof of Theorem 3.1 shows how conditions (a)–(c) are satisfied with $G_n = G_n^{\#}$, $\hat{V}_n = \tilde{V}_n$ required to apply Proposition 2.3.

Recall that $\check{F}_n^*$ is the nondecreasing function closest to F_n^*. Let $H_n^{\#}(x) := \int_0^x F_n^{\#}(z)\,\mathrm{d}z$ and $H_n^*(x) := \int_0^x F_n^*(z)\,\mathrm{d}z$. Next, we show that approximating the distribution of $\Lambda_n = \varepsilon_n^{-1}\{\check{F}_n(x_0) - F(x_0)\}$ by the bootstrap distribution of $\Lambda_n^* := \varepsilon_n^{-1}\{\check{F}_n^*(x_0) - \check{F}_n(x_0)\}$ is consistent. To find the limit of the conditional distribution of Λ_n^*, let

$$\mathbb{X}_n^*(t) = \varepsilon_n^{-2}\{H_n^*(x_0 + \varepsilon_n t) - H_n^*(x_0) - \check{F}_n(x_0)\varepsilon_n t\}$$

for $t \in I_n := [-\varepsilon_n^{-1} x_0, \infty)$ and decompose it into $\mathbb{X}^*_{n,1}$ and $\mathbb{X}^*_{n,2}$, where

$$\mathbb{X}^*_{n,1}(t) = \varepsilon_n^{-2}\{(H_n^* - H_n^\#)(x_0 + \varepsilon_n t) - (H_n^* - H_n^\#)(x_0)\},$$
$$\mathbb{X}^*_{n,2}(t) = \varepsilon_n^{-2}\{H_n^\#(x_0 + \varepsilon_n t) - H_n^\#(x_0) - \check{F}_n(x_0)\varepsilon_n t\}.$$

Recall that $\mathbb{X}_1(t) = tW$ and $\mathbb{X}(t) = \mathbb{X}_1(t) + \frac{1}{2}t^2 f(x_0)$ are two processes defined for $t \in \mathbb{R}$, where W is a normal random variable having mean 0 and variance $\frac{2}{\pi^2} x_0 g(x_0)$.

Theorem 3.2. *Suppose that F is continuously differentiable around x_0, and $g(x_0) \neq 0$. Then*:

(i) *The conditional distribution of the process $\mathbb{X}^*_{n,1}$, given* $\mathbf{Y}$, *converges to that of $\mathbb{X}_1$ a.s.*
(ii) *Unconditionally, $\mathbb{X}^*_{n,2}(t)$ converges in probability to $\frac{1}{2}t^2 f(x_0)$, uniformly on compacta.*
(iii) *The conditional distribution of the process $\mathbb{X}^*_n$, given* $\mathbf{Y}$, *converges to that of $\mathbb{X}$, in probability.*
(iv) *The bootstrap procedure is weakly consistent, that is, the conditional distribution of Λ^*_n, given* $\mathbf{Y}$, *converges to that of W, in probability.*

Proof. The proof is very similar to that of Theorem 3.1. To find the conditional distribution of $\mathbb{X}^*_{n,1}$ given $\mathbf{Y}$, we appeal to Theorem 2.2 with $G_n = G_n^\#$, $G_n^\# = G_n^*$ and $P\{\cdot\} = P^*\{\cdot\} = P\{\cdot|\mathbf{Y}\}$. Note that condition (a) required for Theorem 2.2 holds a.s. We express $\mathbb{X}^*_{n,2}(t)$ as $\mathbb{X}^0_n(t) - t \cdot \mathrm{GCM}_{I_n}[\mathbb{X}^0_n]'(0)$ where

$$\mathbb{X}^0_n(t) = \varepsilon_n^{-2}\{H_n^\#(x_0 + t\varepsilon_n) - H_n^\#(x_0) - F(x_0)\varepsilon_n t\}.$$

Note that unconditionally $\mathbb{X}^0_n$ converges in distribution to $\mathbb{X}$ by an application of Theorem 2.2 with $G_n = G$, $\hat{F}_n = F$ and $H_n = H$ for all n.

Unconditionally, using the continuous mapping theorem along with a localization argument as in Corollary 2.1, we obtain $\mathbb{X}^*_{n,2}(t) \Rightarrow \mathbb{X}(t) - t \cdot \mathrm{GCM}_{\mathbb{R}}[\mathbb{X}]'(0) = \frac{1}{2}t^2 f(x_0)$. As the limiting process is a constant, $\mathbb{X}^*_{n,2}(t) \xrightarrow{P} \frac{1}{2}t^2 f(x_0)$.

An argument using subsequences as in the proof of (iii) of Theorem 3.1 shows that the conditional distribution of the process $\mathbb{X}^*_n$, given $\mathbf{Y}$, converges to that of $\mathbb{X}$, in probability. The last part of the theorem follows along similar lines as in the proof of (iv) of Theorem 3.1. □

4. Data application

A globular cluster (GC) is a spherical collection of stars that orbits a galactic core as a satellite. GCs are very tightly bound by gravity, which gives them their spherical shapes and relatively high stellar densities toward their centers. The study of the inner Galactic GCs is important for several reasons – to understand the morphology of the inner Galaxy, to better constraint the characteristics of the Galactic bulge, etc. Data is available on individual stars in 25 globular clusters located toward the center of the Milky Way (see, e.g., Alonso [2] and Alonso *et al.* [3]). The left panel of Figure 2 shows the projected positions of $n = 2000$ stars in the inner

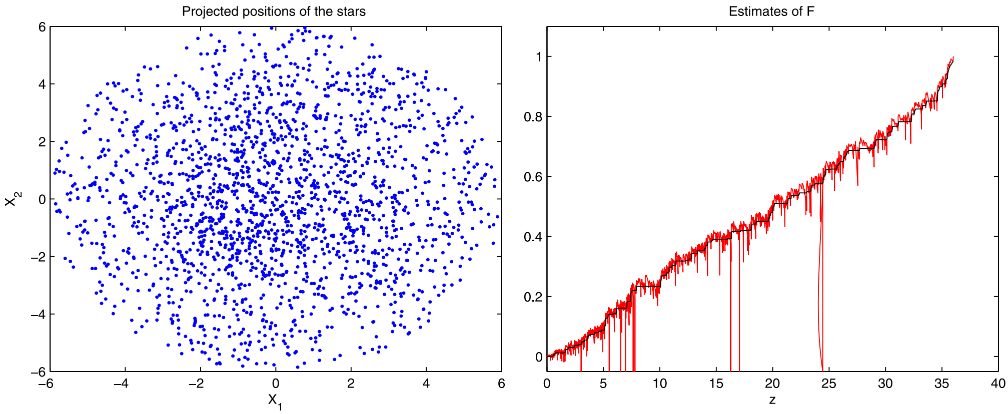

Figure 2. Left panel: Scatter plots of the projected positions of the stars; right panel: $F_n^{\#}$ (red), $\tilde{F}_n^{\#}$ (black, piece-wise constant).

core of the globular cluster M62 (also known as NGC 6266). Interest focuses on estimating the distribution function F of the squared radius. The naive estimator of F, $F_n^{\#}$, is shown in the right panel of Figure 2 along with the isotonized estimator $\tilde{F}_n^{\#}$. The two isotonic estimators $\tilde{F}_n^{\#}$ and $\check{F}_n$ are virtually indistinguishable, and the left panel of Figure 3 shows the difference between the two estimators. Note that both the isotonic estimators have the same pointwise normal limit distribution. The right panel of Figure 3 shows the point-wise bootstrap based 95% CIs for F using the estimator $\tilde{F}_n^{\#}$. A very similar plot is obtained using the estimator $\check{F}_n$.

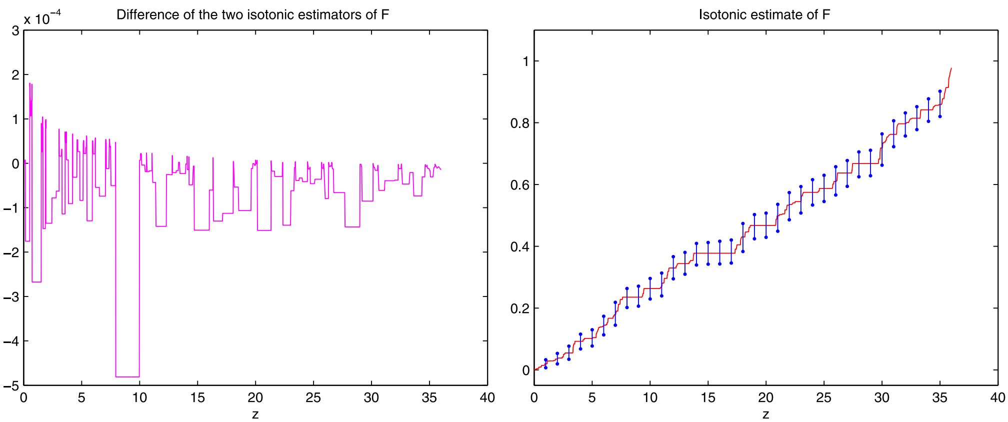

Figure 3. Left panel: $\tilde{F}_n^{\#} - \check{F}_n$; right panel: Bootstrap based 95% pointwise CIs around $\tilde{F}_n^{\#}$.

Acknowledgements

The first author's research was supported by the National Science Foundation, USA.

Supplementary Material

Proofs (DOI: 10.3150/12-BEJ378SUPP; .pdf). The Appendix gives the details of some of the arguments in the proofs of the main results.

References

[1] Abrevaya, J. and Huang, J. (2005). On the bootstrap of the maximum score estimator. *Econometrica* **73** 1175–1204. MR2149245

[2] Alonso, J. (2010). Uncloaking globular clusters in the inner galaxy. Ph.D. thesis. Available at http://hdl.handle.net/2027.42/75831.

[3] Alonso, J., Matio, M. and Sen, B. (2007). Uncloaking globular clusters of the inner galaxy. *Proceedings of the International Astronomical Union: Cambridge Univ. Press* **3** 359–360.

[4] Bickel, P.J. and Freedman, D.A. (1981). Some asymptotic theory for the bootstrap. *Ann. Statist.* **9** 1196–1217. MR0630103

[5] Gnedenko, B.V. and Kolmogorov, A.N. (1968). *Limit Distributions for Sums of Independent Random Variables*. Reading, MA–London–Don Mills, ON: Addison-Wesley. Translated from the Russian, annotated and revised by K.L. Chung. With Appendices by J.L. Doob and P.L. Hsu. Revised edition. MR0233400

[6] Groeneboom, P. and Jongbloed, G. (1995). Isotonic estimation and rates of convergence in Wicksell's problem. *Ann. Statist.* **23** 1518–1542. MR1370294

[7] Kim, J. and Pollard, D. (1990). Cube root asymptotics. *Ann. Statist.* **18** 191–219. MR1041391

[8] Kosorok, M.R. (2008). Bootstrapping in Grenander estimator. In *Beyond Parametrics in Interdisciplinary Research*: *Festschrift in Honor of Professor Pranab K. Sen. Inst. Math. Stat. Collect.* **1** 282–292. Beachwood, OH: IMS. MR2462212

[9] Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Inference*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Chichester: Wiley. MR0961262

[10] Sen, B., Banerjee, M. and Woodroofe, M. (2010). Inconsistency of bootstrap: The Grenander estimator. *Ann. Statist.* **38** 1953–1977. MR2676880

[11] Shao, J. and Tu, D.S. (1995). *The Jackknife and Bootstrap*. *Springer Series in Statistics*. New York: Springer. MR1351010

[12] Singh, K. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9** 1187–1195. MR0630102

[13] Wicksell, S.D. (1925). The corpuscle problem. *Biometrika* **17** 84–99.

Received September 2010 and revised April 2011