Introduction to the Bootstrap

Thomas Wiemann
University of Chicago

TA Discussion # 7
Econ 31720

November 18, 2021

Outline

- Motivation
- Bootstrap CDF Estimation
- Alternative bootstrap confidence intervals
- Bias reduction (asymptotic refinements)
- Failures of the bootstrap

Motivation

Suppose we observe a sample $\{x_i\}_{i=1}^n$, where the CDF of $x_i, \forall i$ is $F_0 \in \mathcal{F}$ (for some family of distributions $\mathcal{F}$). Let T_n be a statistic (i.e., a function of the data only). Let $G_n(\tau, F)$ denote the exact CDF of T_n evaluated at τ when the data is sampled from F – that is, $P_F(T_n \leq \tau)$.

Usually, $G_n(\tau, F)$ is a different function of τ for different distributions F. An exception occurs if $G_n(\cdot, F)$ does not depend on F, in which case T_n is said to be *pivotal*.

As an example of a pivotal statistic, suppose $\mathcal{F}$ is the family of nondegenerate normal distributions with finite variance. Then

$$T_n = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sqrt{s_n^2}} \sim \mathcal{T}(n-1), \tag{1}$$

where $\bar{X}_n$ and s_n^2 are the usual plug-in mean and variance estimators. The t statistic for testing a hypothesis about the mean of a normal population is thus pivotal.

But (unfortunately) pivotal statistics are rare in econometric applications.

Motivation (Contd.)

Suppose that the parameter of interest is $\theta_0 = \mu(F_0)$ for some known function μ (e.g., $\mu(F) = \int_{\mathcal{X}} x dF(x)$) and $\theta_n = \mu(F_n)$ is our estimate.

Define $T_n = \theta_n - \theta_0$ and consider the confidence interval

$$C_n = \left[\theta_n - G_n^{-1}(1 - \frac{\alpha}{2}, F_0),\ \theta_n - G_n^{-1}(\frac{\alpha}{2}, F_0)\right], \tag{2}$$

for some $\alpha \in (0, 1)$. We have

$$\begin{aligned} P(\theta_0 \in C_n) &= P\left(G_n^{-1}(\frac{\alpha}{2}, F_0) \leq T_n \leq G_n^{-1}(1 - \frac{\alpha}{2}, F_0)\right) \\ &= G_n\left(G_n^{-1}\left(1 - \frac{\alpha}{2}, F_0\right), F_0\right) - G_n\left(G_n^{-1}\left(\frac{\alpha}{2}, F_0\right), F_0\right) \\ & 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha, \end{aligned} \tag{3}$$

so that C_n is an exact $1 - \alpha$ confidence interval. Unfortunately, however, it depends on the unknown distribution function $G_n(\cdot, F_0)$ unless T_n is pivotal. We thus need an approximation of C_n.

Motivation (Contd.)

You're probably most used to using asymptotic theory for this purpose.

Let $G_\infty(\cdot, F)$ denote the asymptotic distribution of T_n when the data are sampled from a distribution with CDF F. We say that a statistic T_n is *asymptotically pivotal* when it's distribution function $G_\infty(\cdot, F) \equiv G_\infty(\cdot)$ does not depend on F.

Thus, if n is sufficiently large, $G_n(\cdot, F_0)$ can be estimated by $G_\infty(\cdot)$ without knowledge of F_0.

Many econometric estimators/statistics have asymptotic distributions that are standard normal or chi-square (possibly after appropriate centering and scaling), and so this asymptotic approach is often easy to implement and is widely used in applications.

Motivation (Contd.)

Continuing the example from before, we replace $G_n(\cdot, F_0)$ by $G_\infty(\cdot)$ in (2). Suppose that $\theta_0 = E_{F_0}[X]$ is the parameter of interest and $\mathcal{F}$ is the family of scalar-valued distributions with finite variance. Then $\frac{\sqrt{n}}{s_n}(\bar{X}_n - \theta_0) \xrightarrow{d} \mathcal{N}(0,1)$ by the CLT.

We can then approximate the confidence interval C_n in (2) using

$$C_\infty = \left[\theta_n - \frac{s_n}{\sqrt{n}}\mathcal{Z}_{1-\frac{\alpha}{2}},\, \theta_n + \frac{s_n}{\sqrt{n}}\mathcal{Z}_{1-\frac{\alpha}{2}}\right], \tag{4}$$

where $\mathcal{Z}_\alpha$ is the α quantile of a standard normal.

In some key cases, however, $G_\infty(\cdot)$ can be a poor approximation to $G_n(\cdot, F_0)$ with samples of the sizes encountered in applications, so that C_∞ is a poor approximation to C_n. In other cases (i.e., outside this example), $G_\infty(\cdot)$ may be difficult to implement or estimate (or both).

Motivation (Contd.)

The bootstrap provides an alternative approximation to $G_n(\cdot, F_0)$. Instead of replacing the unknown distribution function G_n with the known function G_∞, the bootstrap replaces the unknown distribution function F_0 with a known estimator F_n. A common choice for F_n being the empirical distribution function (EDF) of the data.

The bootstrap thus provides a way of substituting computation for mathematical analysis. In some contexts, bootstraps are more accurate in finite samples that conventional asymptotic approximations. Such cases are called *asymptotic refinements*. An example is the potential for the bootstrap to reduce the finite-sample bias of an estimator.

The term "bootstrap" was introduced in Efron (1979). In the monograph Efron and Tibshirani (1994), he explained:
"The use of the term bootstrap derives from the phrase to pull oneself up by one's bootstrap, widely thought to be based on one of the eighteenth century Adventures of Baron Munchausen, by Rudolph Erich Raspe. (The Baron had fallen to the bottom of a deep lake. Just when it looked like all was lost, he thought to pick himself up by his own bootstraps.)" (p. 5)

Motivation (Contd.)

There are several helpful text-book style references on the bootstrap. The ones I relied on most heavily for today are:

- Horowitz (2001), which provides excellent intuition about when and why the bootstrap works as well as a good introduction to asymptotic refinements, and
- Chapter 3 of Wasserman (2006), which gives a useful overview of some variants of bootstrap confidence intervals.

You may also find Chapter 29 of DasGupta (2008) and Chapter 23 of van der Vaart (2000) interesting. The first provides a more extensive overview of various kinds of bootstrap procedures. The second provides a brief (but much more rigorous) discussion of some key theoretical results.

And of course there is the monograph Efron and Tibshirani (1994), which provides a discussion of practical considerations of various bootstrap-based estimators (but largely avoids technical details).

Consider the bootstrap estimator $G_n(\cdot, F_n)$ of $G_n(\cdot, F_0)$. The estimator cannot typically be evaluated analytically regardless of the choice of F_n. Fortunately, we can leverage Monte Carlo simulation to estimate $G_n(\tau, F_n)$ to arbitrary accuracy:

Required input: τ a value of interest, F_n a consistent estimate of F_0, an estimate θ_n on the full data;

procedure Bootstrap CDF Estimation
    for $b \in \{1, \dots, B\}$ **do**
        $\{x_i^{(b)}\}_{i=1}^n \sim F_n$ ▷ Sample bootstrap data
        $\theta^*_{n,b} \leftarrow \mathrm{ESTIMATE}\theta(\{x_i^{(b)}\}_{i=1}^n)$
        $T^*_{n,b} \leftarrow \theta^*_{n,b} - \theta_n$ ▷ Compute the bootstrap analogue
    Return: $B^{-1} \sum_{b=1}^B \mathbb{1}\{T^*_{n,b} \leq \tau\}$

We can thus estimate the α quantile of $G_n(\cdot, F_n)$ by the α quantile of $\{T^*_{n,b}\}_b$. The larger we choose B, the more accurate this estimate is.

Bootstrap CDF Estimation (Contd.)

The bootstrap confidence interval is then given by

$$\begin{aligned} C_n^* &= \left[\theta_n - G_n^{-1}(1-\frac{\alpha}{2}, F_n),\ \theta_n - G_n^{-1}(\frac{\alpha}{2}, F_n)\right] \\ &= \left[2\theta_n - \theta^*_{1-\frac{\alpha}{2}},\ 2\theta_n - \theta^*_{\frac{\alpha}{2}}\right], \end{aligned} \quad (5)$$

where θ^*_α denotes the α quantile of the bootstrap coefficients.

Note that because $F_n \neq F_0$ in general, $G_n(\cdot, F_n) \neq G_n(\cdot, F_0)$ unless T_n is pivotal. (Of course, estimation of $G_n(\cdot, F_n)$ to arbitrary accuracy does not imply accurate estimation of $G_n(\cdot, F_0)$.) We thus now turn to the minimal criterion for adequacy of $G_n(\cdot, F_n)$ for $G_n(\cdot, F_0)$: consistency.

Definition 2.1 Horowitz (2001). Let P_n denote the joint probability distribution of the sample $\{x_i\}_{i=1}^n$. The bootstrap estimator $G_n(\cdot, F_n)$ is *consistent* if for each $\epsilon > 0$ and $F_0 \in \mathcal{F}$

$$\lim_{n\to\infty} P_n\left(\sup_\tau \left|G_n(\tau, F_n) - G_\infty(\tau, F_0)\right| > \epsilon\right) = 0. \quad (6)$$

Bootstrap CDF Estimation (Contd.)

The following theorem from Beran and Durcharme (1991) gives conditions under which the bootstrap is consistent. It is given as Theorem 2.1 in Horowitz (2001).

Theorem (Beran and Durcharme, 1991).

$G_n(\cdot, F_n)$ is consistent if for any $\epsilon > 0$ and $F_0 \in \mathcal{F}$:

1. $\lim_{n\to\infty} P_n(\rho(F_n, F_0) > \epsilon) = 0$;
2. $G_\infty(\tau, F)$ is a continuous function of τ for each $F \in \mathcal{F}$; and
3. for any τ and any sequence $(H_n) \in \mathcal{F}$ such that $\lim_{n\to\infty} \rho(H_n, F_0) = 0$, we have $G_n(\tau, H_n) \to G_\infty(\tau, F_0)$.

Here, ρ is a metric on the space $\mathcal{F}$ of permitted distribution functions. The exact metric depends on the particular application under consideration. A metric you may encounter in this setting is the Wasserstein (or Mallows) metric given by $\rho(P, Q) = \inf_{\pi \in \mathcal{M}(P,Q)} E_\pi[\|Y - X\|^2]$ where the infimum is over the joint distributions of (Y, X) whose marginals are P and Q.

Bootstrap CDF Estimation (Contd.)

Horowitz (2001) gives an example for checking the conditions of the previous theorem in the setting of a sample mean estimator (Example 2.1), but it is not as straightforward as one would hope.

The next theorem gives necessary and sufficient conditions for the bootstrap to consistently estimate the distribution of a linear functional of F_0 when F_n is the empirical distribution of the data. It's easier to check the conditions, and many estimators/test statistics in econometrics are asymptotically equivalent to linear functions of some F_0.

Theorem (Mammen, 1992).

Let $\{x_i\}_{i=1}^n$ be a random sample from a population. For a sequence of functions g_n and sequences of numbers t_n and σ_n, define $\bar{g}_n = n^{-1} \sum_{i=1}^n g_n(x_i)$ and $T_n = \sigma_n^{-1}(\bar{g}_n - t_n)$. For the bootstrap sample $\{x_i^*\}_{i=1}^n$, define $\bar{g}_n^* = n^{-1} \sum_{i=1}^n g_n(x_i^*)$ and $T_n^* = \sigma_n^{-1}(\bar{g}_n^* - \bar{g}_n)$. Let $G_n(\tau) = P(T_n \leq \tau)$ and $G_n^*(\tau) = P^*(T_n^* \leq \tau)$, where P^* is the probability induced by bootstrap sampling. Then G_n^* consistently estimates G_n if and only if $T_n \xrightarrow{d} \mathcal{N}(0, 1)$.

Alternative bootstrap confidence intervals

The bootstrap confidence interval given in (5) is one of many possible bootstrap-based confidence intervals.

An even simpler procedure is the *normal bootstrap interval* defined by

$$C^*_{\mathcal{N}} = \left[\theta_n - \widehat{\mathrm{se}}_{boot}\mathcal{Z}_{1-\frac{\alpha}{2}}\; \theta_n + \widehat{\mathrm{se}}_{boot}\mathcal{Z}_{1-\frac{\alpha}{2}}\right], \tag{7}$$

where $\widehat{\mathrm{se}}_{boot}$ is a bootstrap-based standard error estimate constructed using the following procedure:

1: <u>Required input</u>: F_n a consistent estimate of F_0;

2: **procedure** BOOTSTRAP STD. ERROR ESTIMATION
3:     **for** $b \in \{1, \ldots, B\}$ **do**
4:         $\{x_i^{(b)}\}_{i=1}^n \sim F_n$ ▷ Sample bootstrap data
5:         $\theta_n^{(b)} \leftarrow \theta_n(\{x_i^{(b)}\}_{i=1}^n)$ ▷ Compute the bootstrap parameter
6:     <u>Return</u>: $\sqrt{B^{-1}\sum_{b=1}^B \left(\theta_n^{(b)} - B^{-1}\sum_{j=1}^B \theta_n^{(j)}\right)^2}$

Note that the confidence interval $C^*_{\mathcal{N}}$ is not accurate unless $G_n(\tau, F_0)$ is close to being normal.

Alternative bootstrap confidence intervals (Contd.)

Another alternative is the *bootstrap percentile interval*, defined by

$$C_p^* = \left[\theta^*_{\frac{\alpha}{2}},\, \theta^*_{1-\frac{\alpha}{2}}\right], \tag{8}$$

that is, using the $\frac{\alpha}{2}$ and $1-\frac{\alpha}{2}$ quantiles of the bootstrap sample.

Justification of this is as follows. Suppose there exists a monotone transformation $U = m(\theta_n)$ such that $U \sim \mathcal{N}(m(\theta_0), c^2)$. Define $U_b^* = m(\theta^*_{n,b})$. We have

$$\begin{aligned} P(\theta_0 \in C_p^*) &= P(m(\theta^*_{\frac{\alpha}{2}}) \leq m(\theta_0) \leq m(\theta^*_{1-\frac{\alpha}{2}})) \\ &= P(U^*_{\frac{\alpha}{2}} \leq m(\theta_0) \leq U^*_{1-\frac{\alpha}{2}}) \\ &\approx P(U - c\mathcal{Z}_{1-\frac{\alpha}{2}} \leq m(\theta_0) \leq U + c\mathcal{Z}_{1-\frac{\alpha}{2}}) \\ &= P(-\mathcal{Z}_{1-\frac{\alpha}{2}} \leq \frac{U - m(\theta_0)}{c} \leq \mathcal{Z}_{1-\frac{\alpha}{2}}) = 1-\alpha. \end{aligned} \tag{9}$$

Note that we do not need to know this transformation, only that one exists. (There are alo variants where approximate transformations are available. This leads to adjusted percentile methods.)

Alternative bootstrap confidence intervals (Contd.)

There is a different version of the bootstrap confidence interval of (5) that has some advantages. Define

$$Z_n = \frac{\theta_n - \theta_0}{\widehat{\mathrm{se}}_{boot}}, \qquad Z^*_{n,b} = \frac{\theta^*_{n,b} - \theta_n}{\widehat{\mathrm{se}}^*_b}, \tag{10}$$

where $\widehat{\mathrm{se}}^*_b$ is an estimate of the standard error of $\theta^*_{n,b}$ (not θ_n!). Let z^*_α denote the α quantile of $\{Z^*_{n,b}\}_b$. The studentized pivot interval is

$$C^*_{st} = \left[\theta_n - \widehat{\mathrm{se}}_{boot} z^*_{1-\frac{\alpha}{2}}\ \theta_n + \widehat{\mathrm{se}}_{boot} z^*_{1-\frac{\alpha}{2}}\right], \tag{11}$$

We have

$$\begin{aligned} P(\theta_0 \in C^*_{st}) &= P(-z^*_{1-\frac{\alpha}{2}} \leq \frac{\theta_n - \theta}{\widehat{\mathrm{se}}_{boot}} \leq z^*_{1-\frac{\alpha}{2}}) \\ &= P(-z^*_{1-\frac{\alpha}{2}} \leq Z_n \leq z^*_{1-\frac{\alpha}{2}}) \approx 1 - \alpha, \end{aligned} \tag{12}$$

so that sample quantiles of the bootstrap quantities $\{Z^*_{n,b}\}_b$ should approximate the true quantiles of the distribution of Z_n.

Bootstrap Bias Estimation

The studentized pivotal interval has advantages over the other intervals discussed thus far. In particular, the studentized pivotal interval has the potential to yield a higher order approximation to $G_n(\cdot, F_0)$ compared to the asymptotic approach when the considered statistic T_n is asymptotically pivotal. Such improvements are called *asymptotic refinements*. (See Section 3.2 in Horowitz (2001).)

Another procedure through which the bootstrap can provide asymptotic refinements is bias reduction. Fortunately, this is considerably easier to illustrate. (See also Section 3.1 in Horowitz (2001)).

Suppose we are interested in a scalar parameter defined by $\theta_0 = g(E_{F_0}[X_i])$ for an unknown continuous function g (four-times differentiable) and $\mathcal{F}$ is the family of distributions with finite fourth moments. By the continuous mapping theorem, we can consistently estimate θ_0 by $\theta_n = g(\bar{X}_n)$. Note that $E[g(\bar{X}_n)] \neq g(\mu)$ in general unless g is a linear function. Hence, θ_n is a biased estimator of θ_0.

Bootstrap Bias Estimation (Contd.)

Let $G_1(\mu) := \frac{\partial g(\mu)}{\partial \mu}$ and $G_2(\mu) := \frac{\partial^2 g(\mu)}{\partial \mu^2}$. By Taylor's theorem,

$$\begin{aligned} \theta_n - \theta_0 &= G_1(\mu)^\top (\bar{X}_n - \mu) + \frac{1}{2}(\bar{X}_n - \mu)^\top G_2(\mu)(\bar{X}_n - \mu) + R_n \\ \Rightarrow E_{F_0}[\theta_n - \theta_0] &= \frac{1}{2} E_{F_0}\left[(\bar{X}_n - \mu)^\top G_2(\mu)(\bar{X}_n - \mu)\right] + O(n^{-2}). \end{aligned} \tag{13}$$

Note that the leading term on the RHS is $O(n^{-1})$.

Now consider the bootstrap analogue:

$$\begin{aligned} \theta^*_{n,b} - \theta_n &= G_1(\bar{X}_n)^\top (\bar{X}^*_{n,b} - \bar{X}_n) + \frac{1}{2}(\bar{X}^*_{n,b} - \bar{X}_n)^\top G_2(\bar{X}_n)(\bar{X}^*_{n,b} - \bar{X}_n) + R^*_n \\ \Rightarrow E^*[\theta^*_{n,b} - \theta_n] &= \frac{1}{2} E^*\left[(\bar{X}^*_{n,b} - \bar{X}_n)^\top G_2(\bar{X}_n)(\bar{X}^*_{n,b} - \bar{X}_n)\right] + O(n^{-2}), \end{aligned} \tag{14}$$

where E^* denotes the expectation with respect to the distribution function defined by the sampling procedure.

Bootstrap Bias Estimation (Contd.)

Note that we can estimate $B_n^* := E^*[\theta^*_{n,b} - \theta_n]$ to arbitrary accuracy because the distribution of the bootstrap is known.

Required input: F_n a consistent estimate of F_0, an estimate θ_n on the full data;

procedure Bootstrap Bias Estimation
  for $b \in \{1, \dots, B\}$ **do**
    $\{x_i^{(b)}\}_{i=1}^n \sim F_n$ ▷ Sample bootstrap data
    $\theta^*_{n,b} \leftarrow g(\bar{X}^*_{n,b})$ ▷ Compute the bootstrap parameter
  Return: $B^{-1} \sum_{b=1}^B \theta^*_{n,b} - \theta_n$

Note from comparison of (13) and (14), $E_{F_0}[B_n^*] = E[\theta_n - \theta_0] + O(n^{-2})$ (by another application of Taylor's theorem).

Therefore, the bias-adjusted estimator $\theta_n - B_n^*$ satisfies $E_{F_0}[\theta_n - B_n^* - \theta_0] = O(n^{-2})$, which is an improvement over $E_{F_0}[\theta_n - \theta_0] = O(n^{-1})$.

Failures of the Bootstrap

The bootstrap is appealing but one should be cautious not to treat is as a panacea for all problems. Several important econometric applications do not admit the standard bootstrap.

The following examples from DasGupta (2008) give scenarios in which the bootstrap is inconsistent:

- $T_n = \sqrt{n}(\bar{x} - \mu)$ when $Var_{F_0}(x_1) = \infty$.
- $T_n = \sqrt{n}(g(\bar{x}) - g(\mu))$ and $\nabla g(\mu) = 0$.
- $T_n = \sqrt{n}(g(\bar{x}) - g(\mu))$ and g is not differentiable at μ.
- $\mathcal{F}$ is indexed by a parameter $\theta \in \Theta$ and the support of F_θ depends on the value of θ.
- $\mathcal{F}$ is indexed by a parameter $\theta \in \Theta$ and the corresponding population value θ_0 belongs to the boundary of Θ.

Failures of the Bootstrap (Contd.)

Example 1: Let $\mathcal{F}$ be the family of distributions with unit variance. Take $g(x) = |x|$ and define $T_n = \sqrt{n}(g(\bar{X}_n) - g(\mu))$. If μ_0, then the CLT and the CMT imply $T_n \xrightarrow{d} |Z|, Z \sim \mathcal{N}(0,1)$.

Note that $T_{n,b}^* = \sqrt{n}(\bar{X}_{n,b}^* - \bar{X}_n)$, given $\bar{X}_n$, converges in distribution to $\mathcal{N}(0,1)$. Further, the joint $(T_n, T_{n,b}^*) \xrightarrow{d} (Z_1, Z_2)$ where $Z_j \overset{iid}{\sim} \mathcal{N}(0,1), j = 1, 2$. (See Example 29.7 in DasGupta, 2008.)

It then follows that when $\mu_0 = 0$, we have

$$\begin{aligned} T_{n,b}^* &= \sqrt{n}(|\bar{X}_{n,b}^*| - |\bar{X}_n|) \\ &= |\sqrt{n}(\bar{X}_{n,b}^* - \bar{X}_n) + \sqrt{n}\bar{X}_n| - |\sqrt{n}\bar{X}_n|), \end{aligned} \tag{15}$$

so that $T_{n,b}^* \xrightarrow{d} |Z_1 + Z_2| - |Z_1|$, where $Z_j \overset{iid}{\sim} \mathcal{N}(0,1)$.

Note that this is not the asymptotic distribution of T_n, so the bootstrap is inconsistent.

Failures of the Bootstrap (Contd.)

Example 2: Let $\mathcal{F}$ be the family of uniform distributions $\mathcal{U}(0, \theta)$ indexed by θ with population value $\theta_0 = 1$. Define $\theta_n = \max(\{x_i\}_{i=1}^n)$ and let $T_n = n(\theta_n - \theta_0)$. The bootstrap analogue is $T_{n,b}^* = n(\theta_{n,b}^* - \theta_n)$, where $\theta_{n,b}^*$ is the maximum of the bth bootstrap sample.

Note that

$$\begin{aligned} P_n^*(T_{n,b}^* = 0) &= P_n^*(\theta_{n,b}^* = \theta_n) \\ &= 1 - P_n^*(\theta_{n,b}^* < \theta_n) \\ &= 1 - \left(\frac{n-1}{n}\right)^n \overset{n\to\infty}{\to} 1 - e^{-1}. \end{aligned} \tag{16}$$

But $T_n \xrightarrow{d} Exp(\theta_0)$. It follows that $\lim_{n\to\infty} P_n^*(T_{n,b}^* = 0) = 1 - e^{-1}$ but $0 = \lim_{n\to\infty} P_{F_0}(T_n = 0)$ so that the bootstrap is inconsistent.

(See also Example 29.8 in DasGupta, 2008, and Example 2.5 in Horowitz, 2001.)

Failures of the Bootstrap (Contd.)

There are a number of more extensive (and more practically relevant) counter examples in the literature. Particularly notable are

- Andrews (2000), who shows that the bootstrap does not consistently estimate the distribution of a parameter estimator when the true parameter point is on the boundary of the parameter space. He considers estimation of the population mean μ subject to the constraint $\mu \geq 0$. When the population $\mu_0 = 0$, then the bootstrap is not consistent, and

- Abadie and Imbens (2008), who show that the bootstrap does not consistently estimate the distribution of matching estimates in general.

In both scenarios, alternative re-sampling schemes *do* permit consistent estimation of the respective distribution functions. See the subsampling method of Politis and Romano (1994), which samples $m < n$ observations without replacement from the empirical distribution. Another method is the m-out-of-n bootstrap (or: subsampling with replacement). See Section 2.2 of Horowitz (2001) for an overview.

Last Comments

Today's discussion has only considered the EDF as an estimator for F_n when constructing a bootstrap estimator $G_n(\cdot, F_n)$ for $G_n(\cdot, F_0)$. Needless to say, there are many more options. Some (non-exhaustive) examples are

- the parametric bootstrap, where F_n is a generative model fitted to the full data;
- the residual bootstrap, where the residuals (e.g., in a linear regression model) are shuffled;
- the wild bootstrap (Mammen, 1993), which introduces an auxiliary random variable to avoid issues with, e.g., heteroskedasisticy in the residual bootstrap;
- the block (or moving-block) bootstrap, where groups of observations are jointly sampled to account for dependence.

For further readings on bootstrap consistency, also see Fang and Santos (2019).

References

Abadie, A. and Imbens, G. W. (2008). On the failure of the bootstrap for matching estimators. *Econometrica*, 76(6):1537–1557.

Andrews, D. W. (2000). Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space. *Econometrica*, pages 399–405.

Beran, R. and Durcharme, G. R. (1991). Asymptotic theore for bootstrap methods in statistics. Les Publications CRM, Centre de recherches mathematiques, Universite de Montreal, Montreal, Canada.

DasGupta, A. (2008). *Asymptotic theory of statistics and probability*. Springer Science & Business Media.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7(1):1–26.

Efron, B. and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Fang, Z. and Santos, A. (2019). Inference on directionally differentiable functions. *The Review of Economic Studies*, 86(1):377–412.

Horowitz, J. L. (2001). The bootstrap. In *Handbook of econometrics*, volume 5, pages 3159–3228. Elsevier.

Mammen, E. (1992). *When does bootstrap work? asymptotic results and simulations*. Springer.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics*, pages 255–285.

Politis, D. N. and Romano, J. P. (1994). Large sample confidence regions based on subsamples under minimal assumptions. *Annals of Statistics*, pages 2031–2050.

van der Vaart, A. W. (2000). *Asymptotic statistics*, volume 3. Cambridge university press.

Wasserman, L. (2006). *All of nonparametric statistics*. Springer Science & Business Media.