STAT 824 sp 2023 Lec 10 slides

Bootstrap beyond the mean: statistical functionals

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \xrightarrow{D} N(0,1)$$

$\uparrow$

Karl B. Gregory

University of South Carolina

$\frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\hat{\sigma}_n^*}$ does a better job approximating the dist. of $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n}$ than $N(0,1)$.

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Table of Contents

Throughout, let $X_1, \dots, X_n$ be iid with distribution F

$T(F)$

Statistical functional

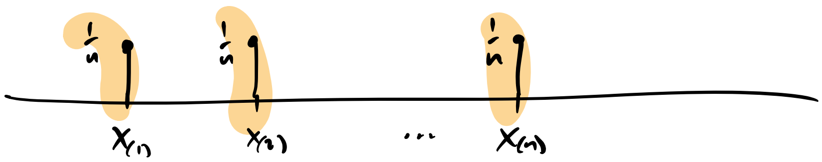
A *statistical functional* is a function $T : \mathcal{D} \to \mathbb{R}$, where $\mathcal{D}$ is the space of probability distributions.

d

- Define a quantity $\boxed{\theta_0 \in \mathbb{R}}$ of interest as $\boxed{\theta_0 = T(F)}$ has cdf $\boxed{\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i \le x)}$
- Consider *plug-in* estimator $\hat{\theta}_n = T(\hat{F}_n)$, where $\hat{F}_n = n^{-1}\sum_{i=1}^{n} \delta_{X_i}$
  Empirical distribution
- The notation δ_x represents the distribution placing unit mass on x.
- We will use F, $\hat{F}_n$ to denote distributions and their cdfs interchangeably.

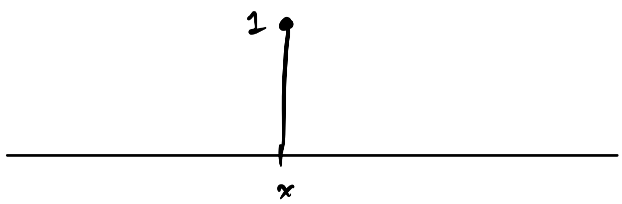
Dirac

Empirical dist $\hat{F}_n$ puts mass $\frac{1}{n}$ on each $X_1, \dots, X_n$



Defn: "Dirac delta"

δ_x is the dist. putting unit mass on the point x.



δ_x has pmf given by $p(y) = \mathbb{1}(y = x)$

Can write

$$\hat{F}_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}$$

We rely on the Glivenko-Cantelli theorem for consistency of plug-in estimators:

Theorem (Glivenko-Cantelli Theorem)

If $X_1, \ldots, X_n$ is a rs from a distribution with cdf F, then

$$P\left(\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |\hat{F}_n(x) - F(x)| = 0\right) = 1$$

Suggests $\hat{\theta}_n = T(\hat{F}_n)$ should get close to $\theta_0 = T(F)$.

$\frac{dF(x)}{dx} = f(x)$

$dF(x) = f(x)dx$

$$\begin{cases} \int x f(x)\,dx & \text{if } F \text{ has pdf } f \\ \sum_{x \in \mathcal{X}} x \cdot p(x) & \text{if } F \text{ has pmf } p \\ \vdots & \\ \mu & \end{cases}$$

Examples of statistical functionals:

1. The mean: $\mu = T(F) = \int x dF(x)$
2. The variance: $\sigma^2 = T(F) = \int (x - \int t dF(t))^2 dF(x)$
3. The τth quantile: $\xi_\tau = T(F) = \inf\{x : F(x) \geq \tau\}$
4. Shape parameter under Gamma$(\alpha, 1)$: $\alpha = T(F) =$ value of t which solves
$$\int (\log x - \Gamma'(t)/\Gamma(t))\, dF(x) = 0.$$
5. Linear functional: $\tau = \int a(x) dF(x)$
6. $\boxed{\text{The } \alpha\text{-trimmed mean: } \mu_\alpha = T(F) = (1 - 2\alpha)^{-1} \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x dF(x)}$

Exercise: Write down the plug-in estimators $T(\hat{F}_n)$ for each of the above.

① $\mu = T(F) = \int x \, dF(x)$

Plug-in estimator is

$$\hat{\mu}_n = T(\hat{F}_n) = \int x \, d\hat{F}_n(x)$$

$$= \int x \, d\left(\frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}\right)(x)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \int x \, d\delta_{X_i}(x)$$

δ_{X_i} places unit mass on X_i
has pmf $p(x) = \mathbb{1}(x = X_i)$

$$= \frac{1}{n}\sum_{i=1}^{n} \sum_{x \in \{X_i\}} x \cdot 1$$

$$= \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$= \bar{X}_n$$

② $\sigma^2 = T(F) = \int \left(x - \int t \, dF(t)\right)^2 dF(x)$

Plug-in estimator $\hat{\sigma}^2_n = T(\hat{F}_n) = \int \Big(x - \underbrace{\int t \, d\hat{F}_n(t)}_{\bar{X}_n}\Big)^2 d\hat{F}_n(x)$

$$= \int (x - \bar{X}_n)^2 \, d\left(\frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}\right)(x)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \int (x - \bar{X}_n)^2 \, d\delta_{X_i}(x)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \sum_{x \in \{X_i\}} (x - \bar{X}_n)^2 \cdot 1$$

$= \frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X}_n)^2$ = empirical variance.

(3) $$q_\tau = T(F) = \inf\left\{ x : \underbrace{\int_{(-\infty,x]} dF(t)}_{F(x) \,\leftarrow\, \text{cdf at } x} \geq \tau \right\}$$

Plug-in

$$\hat{q}_\tau = T(\hat{F}_n) = \inf\left\{ x : \int_{(-\infty,x]} d\hat{F}_n(x) \geq \tau \right\}$$

$$= \inf\left\{ x : \hat{F}_n(x) \geq \tau \right\}$$

$$= \inf\left\{ x : \frac{1}{n}\sum_{i=1}^n \mathbb{1}(X_i \leq x) \geq \tau \right\}$$

$$= \inf\left\{ x : \frac{1}{n}\sum_{i=1}^n \mathbb{1}(X_{(i)} \leq x) \geq \tau \right\}$$

$\vdots$

$$= X_{(\lceil n\tau \rceil)}$$

(4) $$\mu_\alpha = T(F) = \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} x \, dF(x)$$

plug-in: $$\hat{\mu}_\alpha = T(\hat{F}_n) = \frac{1}{1-2\alpha}\int_{\hat{F}_n^{-1}(\alpha)}^{\hat{F}_n^{-1}(1-\alpha)} x \, d\hat{F}_n(x)$$

$$= \frac{1}{1-2\alpha}\int_{X(\lceil \alpha n \rceil)}^{X(\lceil (1-\alpha)n \rceil)} x \left(\frac{1}{n}\sum_{i=1}^n \delta_{X_i}\right)(x)$$

$$= \frac{1}{1-2\alpha}\,\frac{1}{n}\sum_{i=1}^n \int_{X(\lceil \alpha n \rceil)}^{X(\lceil (1-\alpha)n \rceil)} x \, \delta_{X_i}(x)$$

$$= \frac{1}{n(1-2\alpha)}\sum_{i=\lceil \alpha n \rceil}^{\lceil (1-\alpha)n \rceil} X_{(i)}$$

Table of Contents

$$\boxed{\sqrt{n}\left(T(\hat{F}_n) - T(F)\right) \xrightarrow{D} N(0, v)}$$

$$\sqrt{n}\left(\hat{\theta}_n - \theta_0\right) \xrightarrow{D} N(0, v)$$

Bootstrap: Treat $\hat{F}_n$ as though it were the pop. dist.

Obtain $\hat{F}_n^*$ from $X_1^*, \dots, X_n^* \mid X_1, \dots, X_n \overset{iid}{\sim} \hat{F}_n$

Then use dist of $\sqrt{n}\left(T(\hat{F}_n^*) - T(\hat{F}_n)\right)$ to estimate dist. of

$$\sqrt{n}\left(T(\hat{F}_n) - T(F)\right).$$

Recall: $\hat{\theta}_n$ is MLE for θ_0, $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{D} N\left(0, \frac{1}{I_1(\theta)}\right)$

Score function $S(\theta; \underset{\sim}{X})$: $0 = S(\hat{\theta}_n; \underset{\sim}{X}) = S(\theta_0; \underset{\sim}{X}) + H(\theta_0; \underset{\sim}{X})(\hat{\theta}_n - \theta_0) + R$

Rearrange...

We want central limit theorems for $\sqrt{n}(T(\hat{F}_n) - T(F))$. $\xrightarrow{D} N(0, v)$

von Mises expansion for statistical functionals

An expansion for stat. functionals called a *von Mises expansion* lets us write

$$\sqrt{n}(T(\hat{F}_n) - T(F)) = \underline{\sqrt{n}T_F^{(1)}(\hat{F}_n - F)} + \sqrt{n}R(\hat{F}_n - F),$$

where (like a Taylor expansion but for functions $T : \mathcal{D} \to \mathbb{R}$)

- $T_F^{(1)}(\hat{F}_n - F)$ is a *von Mises derivative*.
- $R(\hat{F}_n - F)$ is a (hopefully small) remainder term.

Under some conditions (covered later) we have

$= \sqrt{n} \frac{1}{n} \sum_{i=1}^{n} \varphi_F(X_i)$

$$\sqrt{n}T^{(1)}(\hat{F}_n - F) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi_F(X_i) \xrightarrow{\text{D}} \text{Normal}(0, \text{Var}\, \varphi_F(X_1)),$$

where φ_F is called the *influence curve* of the functional T at F.

$T(F)$ $T(F+\varepsilon(\delta_x - F))$

$T((1-\varepsilon)F + \varepsilon\delta_x)$

Theorem (How to find the influence curve)

The influence curve is given by $\boxed{\varphi_F(x) = \frac{d}{d\varepsilon} T(F + \varepsilon(\delta_x - F))\Big|_{\varepsilon=0}}$.

Influence curves play an important role in *robust estimation*.

The IC expresses change in $T(F)$ due to perturbing F by adding a point mass at x.

Exercise: Find the influence curves φ_F for these functionals

1. $T(F) = \int x dF(x)$.
2. $T(F) = \int a(x) dF(x)$ for some function a.
3. $T(F) = \int (x - \int t dF(t))^2 dF(x)$.

$$\varphi_F(x) = \frac{d}{d\varepsilon} T(F + \varepsilon(\delta_x - F))\Big|_{\varepsilon=0}.$$

① $$T(F) = \int t \, dF(t)$$

$$\varphi_F(x) = \frac{d}{d\varepsilon}\Big[T(F + \varepsilon(\delta_x - F))\Big]\Big|_{\varepsilon=0}$$

$$T(F + \varepsilon(\delta_x - F)) = \int t \, d\big(F + \varepsilon(\delta_x - F)\big)(t)$$

$$= \int t \, dF(t) + \varepsilon \int t \, d\delta_x(t) - \varepsilon \int t \, dF(t)$$

$$= \mu + \varepsilon \cdot x - \varepsilon \mu$$

$$\varphi_F(x) = \frac{d}{d\varepsilon}\Big[\mu + \varepsilon x - \varepsilon \mu\Big]\Big|_{\varepsilon=0}$$

$$= (x - \mu)\Big|_{\varepsilon=0}$$

$$= x - \mu.$$

② $$\mu_a = T(F) = \int a(t) \, dF(t)$$

$$\varphi_F(x) = \frac{d}{d\varepsilon}\Big[T(F + \varepsilon(\delta_x - F))\Big]\Big|_{\varepsilon=0},$$

$$T(F + \varepsilon(\delta_x - F)) = \int a(t) \, d\big(F + \varepsilon(\delta_x - F)\big)(t)$$

$$= \int a(t)\, dF(t) + \varepsilon \underbrace{\int a(t)\, d\delta_x(t)}_{} - \varepsilon \int a(t)\, dF(t)$$

$$= \mu_a + \varepsilon\, a(x) - \varepsilon\, \mu_a$$

$$\psi_F(x) = \frac{d}{d\varepsilon}\left[\cancel{\mu_a} + \varepsilon\, a(x) - \varepsilon\, \mu_a\right]\Bigg|_{\varepsilon=0}$$

$$= a(x) - \mu_a$$

$$= a(x) - \int a(t)\, dF(t).$$

$\sqrt{n}\,\frac{1}{n}\sum_{i=1}^{n}\underbrace{\varphi_F(x_i)} = \sqrt{n}(\bar{\varphi}_n - \mu)$

von Mises

Exercise: Write down the expansion

$$\sqrt{n}(T(\hat{F}_n) - T(F)) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\varphi_F(X_i) + \sqrt{n}R(\hat{F}_n - F),$$

for the following functionals. Give the asymptotic distribution.

1. The mean: $\mu = T(F) = \int x dF(x)$.
2. The probability of a set A: $p_A = T(F) = \int_A dF(x)$.
3. A differentiable function g of the mean: $g(\mu) = T(F) = g(\int x dF(x))$.

1 $\underbrace{\sqrt{n}(\bar{x}_n - \mu)}_{\varphi_F(x) = x - \mu} = \boxed{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(x_i - \mu)} + \underbrace{R}_{=0}$

$\xrightarrow{D} N(0, \underbrace{\mathrm{Var}(x_1 - \mu)}_{\mathrm{Var}\, X_1})$.

$$\int_A f(x)\,dx\ , \quad \sum_{x\in\mathcal{X}:\, x\in A} p(x)$$

② $$p_A = T(F) = \int_A dF(x)$$

plug-in $$\hat{p}_A = T(\hat{F}_n) = \int_A d\hat{F}_n(x)$$

$$= \int_A d\left(\frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}\right)(x)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \int_A \delta_{X_i}(x)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(X_i \in A)$$

Von Mises expansion for $\sqrt{n}(\hat{p}_A - p) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \psi_F(X_i) + R$

$$\psi_F(x) = \frac{d}{d\varepsilon}\left[T\left(F + \varepsilon(\delta_x - F)\right)\right]\Big|_{\varepsilon=0}$$

$$T(F + \varepsilon(\delta_x - F)) = \int_A d\left(F + \varepsilon(\delta_x - F)\right)(t)$$

$$= \int_A dF(t) + \varepsilon \int_A d\delta_x(t) - \varepsilon \int_A dF(t)$$

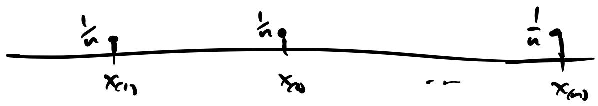
$$= p_A + \varepsilon\, \mathbb{1}(x \in A) - \varepsilon\, p_A .$$

$$\psi_F(x) = \mathbb{1}(x \in A) - p_A .$$

So $$\sqrt{n}(\hat{p}_A - p_A) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\mathbb{1}(X_i \in A) - p_A\right) + \overset{\approx 0.}{R}$$

③ $\theta_0 = T(F) = g\left(\overbrace{\int x\, dF(x)}^{\mu}\right) = g(\mu)$ $\qquad g$ known

$\hat{\theta}_n = T(\hat{F}_n) = g\left(\underbrace{\int x\, d\hat{F}_n(x)}_{\bar{x}_n}\right) = g(\bar{x}_n)$



von Mises expansion of

$$\sqrt{n}\left(T(\hat{F}_n) - T(F)\right) = \sqrt{n}\left(g(\bar{x}_n) - g(\mu)\right) \xrightarrow{D} N\left(0, [g'(\mu)]^2 \overset{\operatorname{Var} X_1}{\sigma^2}\right)$$

Delta method

$$\sqrt{n}\left(T(\hat{F}_n) - T(F)\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \varphi_F(x_i) + \sqrt{n}\, R(\hat{F}_n - F),$$

The influence function is

$$\varphi_F(x) = \frac{d}{d\varepsilon}\left[T\left(\underbrace{F + \varepsilon(\delta_x - F)}\right)\right]\Bigg|_{\varepsilon=0}$$

$$= \frac{d}{d\varepsilon}\left[g\left(\int t\, d\left(F + \varepsilon(\delta_x - F)\right)(t)\right)\right]\Bigg|_{\varepsilon=0}$$

$$= \frac{d}{d\varepsilon}\left[g\left(\underbrace{\int t\, dF(t)}_{\mu} + \varepsilon \underbrace{\int t\, d\delta_x(t)} - \varepsilon \underbrace{\int t\, dF(t)}_{\mu}\right)\right]\Bigg|_{\varepsilon=0}$$

$$= \frac{d}{d\varepsilon}\left[g\left(\mu + \varepsilon x - \varepsilon\mu\right)\right]\Bigg|_{\varepsilon=0}$$

$$= \frac{d}{d\varepsilon}\left[g\left((1-\varepsilon)\mu + \varepsilon x\right)\right]\Bigg|_{\varepsilon=0}$$

$$= \left[g'\left((1-\varepsilon)\mu + \varepsilon x \right) \left(-\mu + x \right) \right] \Big|_{\varepsilon = 0}$$

$$= g'(\mu)(x-\mu).$$

$$\varphi_F(x) = g'(\mu)(x-\mu).$$

$$\sqrt{n}\Big(\overset{T(\hat{F}_n)}{\overset{\downarrow}{g(\bar{x}_n)}} - \overset{T(F)}{\overset{\downarrow}{g(\mu)}} \Big) = \underbrace{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Big[\overset{\varphi_F(x_i)}{\overset{\downarrow}{g'(\mu)(x_i-\mu)}} \Big]}_{\xrightarrow{D} N\left(0,\, [g'(\mu)]^2 \sigma^2\right)} + \sqrt{n}\, \underset{?}{R(\hat{F}_n - F)},$$

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left[g'(\mu)(x_i-\mu) \right] = \sqrt{n}\, \frac{1}{n} \sum_{i=1}^{n} g'(\mu)(x_i-\mu)$$

$$\xrightarrow{D} N\Big(0,\, \underbrace{\mathrm{Var}\left(g'(\mu)(x_1-\mu) \right)}_{[g'(\mu)]^2 \,\mathrm{Var}\, X_1} \Big)$$

<u>Consider error term:</u>

$$\sqrt{n}\, R(\hat{F}_n - F) = \sqrt{n}\left[g(\bar{x}_n) - g(\mu) \right] - \sqrt{n}\, \frac{1}{n} \sum_{i=1}^{n} g'(\mu)(x_i-\mu)$$

$$= \sqrt{n}\Big[\underbrace{g(\bar{x}_n) - g(\mu) - g'(\mu)(\bar{x}_n-\mu)}_{} \Big]$$

$$g(\bar{x}_n) = g(\mu) + g'(\mu)(\bar{x}_n-\mu) + \cdots\cdots R$$

von Mises expansion for the τth quantile

Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} F$ with continuous density f and consider the $\tau$the quantile

$$\xi_\tau = T(F) = \inf\{x : F(x) \geq \tau\} = F^{-1}(\tau)$$

The influence function (derived in hand written notes) is

$$\boxed{\varphi_F(x) = \frac{\tau - 1(x \leq \xi_\tau)}{f(\xi_\tau)},} \qquad \text{provided } f(\xi_\tau) > 0.$$

Exercise:

1. Give the von Mises expansion of $\sqrt{n}(\hat{\xi}_\tau - \xi_\tau)$.
2. Make a conjecture about the asymptotic distribution of $\sqrt{n}(\hat{\xi}_\tau - \xi_\tau)$.

We have $\sqrt{n}R(\hat{F}_n - F) \to 0$ in probability as $n \to \infty$, by Ghosh (1971) [2].

von Mises expansion is

$$\sqrt{n}\,(\hat{q}_\tau - q_\tau) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \varphi_F(X_i) + \overbrace{\sqrt{n}\,R(\hat{F}_n - F)}^{\xrightarrow{P} 0 \text{ by Ghosh}}$$

$\hat{q}_\tau \uparrow T(\hat{F}_n)$ $\quad q_\tau \uparrow T(F)$

$$= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\underbrace{\left[\frac{\tau - \mathbb{1}(X_i \le q_\tau)}{f(q_\tau)}\right]}_{\text{has mean } 0 \text{ since}} + \sqrt{n}\,R(\hat{F}_n - F)$$

$\mathbb{E}\,\mathbb{1}(X_i \le q_\tau) = P(X_i \le q_\tau) = \tau$

$$= \sqrt{n}\,\frac{1}{n}\sum_{i=1}^{n}\left[\frac{\tau - \mathbb{1}(X_i \le q_\tau)}{f(q_\tau)}\right]$$

$$\xrightarrow{D} N\left(0, \frac{\tau(1-\tau)}{f^2(q_\tau)}\right).$$

$$\mathrm{Var}\left[\frac{\tau - \mathbb{1}(X_1 \le q_\tau)}{f(q_\tau)}\right] = \frac{1}{f^2(q_\tau)}\,\mathrm{Var}\big(\mathbb{1}(X_1 \le q_\tau)\big)$$

$$= \frac{1}{f^2(q_\tau)}\,\tau(1-\tau)$$

$\Rightarrow$

$$\sqrt{n}\,(\hat{q}_\tau - q_\tau) \xrightarrow{D} N\left(0, \frac{\tau(1-\tau)}{f^2(q_\tau)}\right)$$

Table of Contents

Consistency of bootstrap for Hadamard differentiable functionals

If T is Hadamard differentiable and $\vartheta = \operatorname{Var} \varphi_F(X_1) < \infty$, then

$$\sup_{x \in \mathbb{R}} \left| P_* \left(\sqrt{n}(T(\hat{F}_n^*) - T(\hat{F}_n)) \leq x \right) - P \left(\sqrt{n}(T(\hat{F}_n) - T(F)) \leq x \right) \right| \to 0$$

in probability as $n \to \infty$.

In the above $\hat{F}_n^* = \frac{1}{n} \sum_{i=1}^n \delta_{X_i^*}$, where $X_1^*, \dots, X_n^* | X_1, \dots, X_n \overset{\text{ind}}{\sim} \hat{F}_n$.

Many interesting statistical functionals are Hadamard differentiable (defined later).

Exercise: Given B sorted Monte Carlo reps $T^{*(1)}, \dots, T^{*(B)}$ of $T(\hat{F}_n^*)$, justify

$$\left(2 \cdot T(\hat{F}_n) - T^{*(\lceil (1-\alpha/2)B \rceil)}, 2 \cdot T(\hat{F}_n) - T^{*(\lceil (\alpha/2)B \rceil)} \right)$$

as an asymptotic $(1 - \alpha) \times 100\%$ confidence interval for $T(F)$.

Suppose $\sqrt{n}\left(T(\hat{F}_n)-T(F)\right) \sim G_n$

Then

$$P\left(G_{n,1-\alpha/2} < \sqrt{n}\left(T(\hat{F}_n)-T(F)\right) < G_{n,\alpha/2}\right) = 1-\alpha$$

$$P\left(T(\hat{F}_n) - G_{n,\alpha/2}\tfrac{1}{\sqrt{n}} < T(F) < T(\hat{F}_n) - G_{n,1-\alpha/2}\tfrac{1}{\sqrt{n}}\right)$$

So $\left(T(\hat{F}_n) - G_{n,\alpha/2}\cdot\tfrac{1}{\sqrt{n}},\ T(\hat{F}_n) - G_{n,1-\alpha/2}\tfrac{1}{\sqrt{n}}\right)$

is a $(1-\alpha)*100\%$ C.I. for $T(F)$.

Draw bootstrap samples to obtain many replicates of

$$\sqrt{n}\left(T(\hat{F}_n^*) - T(\hat{F}_n)\right) \sim \hat{G}_n$$

get this exactly by taking enough Monte Carlo draws.

Let $T^{*(1)} < \dots < T^{*(B)}$ be sorted draws of $T(\hat{F}_n^*)$

$$\hat{G}_{n,\alpha/2} = \sqrt{n}\left(T^{*(\lceil(1-\alpha/2)B\rceil)} - T(\hat{F}_n)\right)$$

$$\hat{G}_{n,1-\alpha/2} = \sqrt{n}\left(T^{*(\lceil(\alpha/2)B\rceil)} - T(\hat{F}_n)\right)$$

$$\left(T(\hat{F}_n) - \hat{G}_{n,\alpha/2} \cdot \frac{1}{\sqrt{n}},\ T(\hat{F}_n) - \hat{G}_{n,1-\alpha/2}\ \frac{1}{\sqrt{n}} \right)$$

$$= \Big(T(\hat{F}_n) - \sqrt{n} \left(T^{*(\lceil (1-\alpha/2) B \rceil)} - T(\hat{F}_n) \right) \frac{1}{\sqrt{n}},$$

$$T(\hat{F}_n) - \sqrt{n} \left(T^{*(\lceil (\alpha/2) B \rceil)} - T(\hat{F}_n) \right) \frac{1}{\sqrt{n}} \Big)$$

$$= \left(2 \cdot T(\hat{F}_n) - T^{*(\lceil (1-\alpha/2) B \rceil)},\ 2 \cdot T(\hat{F}_n) - T^{*(\lceil (\alpha/2) B \rceil)} \right)$$

$$\frac{\mu_3}{\sigma^3} = \text{skewness}$$

Exercise: Let F be the Gamma(α, β) and construct 95% bootstrap CIs for

$$T(F) = \frac{\int (x-\mu)^3 dF(x)}{\left(\int (x-\mu)^2 dF(x)\right)^{3/2}}, \text{ where } \mu = \int x dF(x).$$

Run simulations with $n = 30$ and $n = 100$ and assess coverage. Note $T(F) = \frac{2}{\sqrt{\alpha}}$.

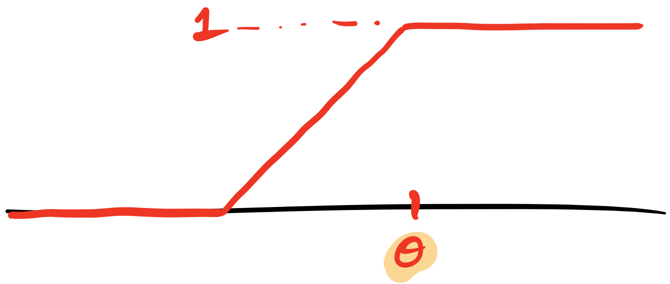
$$T(\hat{F}_n) = \frac{\frac{1}{n}\sum (x_i - \bar{x}_n)^3}{\left[\frac{1}{n}\sum (x_i - \bar{x}_n)^2\right]^{3/2}}$$

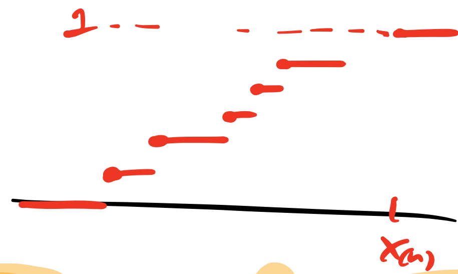
Exercise: Simulate coverage of 95% bootstrap CIs for the α-trimmed mean when

$$F = \delta \cdot \text{Gamma}(a, b) + (1 - \delta) \cdot t_2.$$

Coverage over 500 datasets under $a = 1/2,\ b = 6,\ \delta = 0.8,\ \alpha = 0.10,\ B = 500$:

n	10	20	40	80	160	320
coverage	0.74	0.90	0.87	0.92	0.94	0.92





$T(\hat{F}_n) = \inf\{x : \hat{F}_n(x) \geq 1\} = X_{(n)}$

Exercise: Let $F = \text{Uniform}(0, \theta)$. Then $\theta = T(F) = \inf\{x : F(x) \geq 1\}$.

1. Find the asymptotic distribution of $\sqrt{n}(T(\hat{F}_n) - T(F))$. $= \boxed{\sqrt{n}\,(X_{(n)} - \theta)}$
2. Find the asymptotic distribution of $\boxed{n(T(\hat{F}_n) - T(F))}$. $\to -\text{Exponential}(\theta)$
3. Consider behavior as $n \to \infty$ of the quantity

$$\sup_{x \in \mathbb{R}} \left| P_*(n(T(\hat{F}_n^*) - T(\hat{F}_n)) \leq x) - P(n(T(\hat{F}_n) - T(F)) \leq x) \right|.$$

$F_{X_{(n)}}(x) = \left(\frac{x}{\theta}\right)^n$

Does the bootstrap work in this setting?

(1) $P\left(\sqrt{n}\,(X_{(n)} - \theta) \leq x\right) = P\left(X_{(n)} \leq \theta + \frac{x}{\sqrt{n}}\right)$

$$= \begin{cases} 1 & x > 0 \\ \left(\dfrac{\theta + \frac{x}{\sqrt{n}}}{\theta}\right)^n & x \leq 0 \end{cases}$$

$$= \begin{cases} 1 & x \geq 0 \\ \underbrace{\left(1 + \frac{x/\theta}{\sqrt{n}}\right)^n} & x < 0 \end{cases}$$

$$= \left(1 + \frac{\sqrt{n}\, x/\theta}{n}\right)^n \approx e^{\sqrt{n}\, x/\theta} \to 0 \quad \text{as } n \to \infty$$

$$\to \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\boxed{\sqrt{n}\,(X_{(n)} - \theta) \xrightarrow{p} 0}$$

② $P\left(n\,(X_{(n)} - \theta) \leq x\right) = P\left(X_{(n)} \leq \theta + \frac{x}{n}\right)$

$$= F_{X_{(n)}}\left(\theta + \frac{x}{n}\right)$$

$$= \begin{cases} 1 & x \geq 0 \\ \left(\dfrac{\theta + \frac{x}{n}}{\theta}\right)^n & x < 0 \end{cases}$$

$$= \begin{cases} 1 & \\ \left(1 + \frac{x/\theta}{n}\right)^n & x < 0 \end{cases}$$

$$\to \begin{cases} 1 & x \geq 0 \\ e^{x/\theta} & x < 0 \end{cases}$$

cdf of $-$Exponential(θ).

③ $P_*\left(\underline{\sqrt{n}\left(X^*_{(n)} - X_{(n)}\right)} \le x\right) \xrightarrow{D}{}^{?}$ Exponential (θ).

$P\left(\sqrt{n}\left(X^*_{(n)} - X_{(n)}\right) = 0\right) = P(X_{(n)} \text{ drawn at least once})$

$= 1 - P(X_{(n)} \text{ never drawn})$

$= 1 - \binom{n}{0}\left(\frac{1}{n}\right)^0\left(1-\frac{1}{n}\right)^n$

$= 1 - \left(1-\frac{1}{n}\right)^n$

$\rightarrow 1 - e^{-1} \approx .63$.

Table of Contents

von Mises derivative. See the book of Luisa Fernholz [1].

The *von Mises derivative of T at F in the direction G* is defined as

$$T_F^{(1)}(G-F) = \frac{d}{d\varepsilon} T(F + \varepsilon(G-F))\Big|_{\varepsilon=0},$$

provided there exists a function φ_F, not depending on G, such that

$$T_F^{(1)}(G-F) = \int \varphi_F(x) d(G-F)(x),$$

with $\int \varphi_F(x)dF(x) = 0$; in this case $\varphi_F(x) = T_F^{(1)}(\delta_x - F)$.

The function φ_F is called the *influence curve* of the functional T at F.

Exercise: Find $T_F^{(1)}(G-F)$ and $T_F^{(1)}(\delta_x - F)$ for $T(F) = \int x dF(x)$.

von Mises expansion for M estimators

Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} F$, and for some function $\psi : \mathbb{R}^2 \to \mathbb{R}$ consider

$$\theta_0 = T(F) = \text{ value of } t \text{ which solves } \int \psi(x,t)dF(x) = 0$$
$$\hat{\theta}_n = T(F) = \text{ value of } t \text{ which solves } \int \psi(x,t)d\hat{F}_n(x) = 0$$

The von Mises derivative (derived in hand-written notes) is

$$T_F^{(1)}(G - F) = -\frac{\lambda_G(T(F))}{\lambda_F'(T(F))},$$

where $\lambda_F(t) = \int \psi(x,t)dF(x)$ and $\lambda_G(t) = \int \psi(x,t)dG(x)$, $t \in \mathbb{R}$.

Exercise:

1. Give the von Mises expansion of $\sqrt{n}(\hat{\theta}_n - \theta_0)$.
2. Make a conjecture about the asymptotic distribution of $\sqrt{n}(\hat{\theta}_n - \theta_0)$.
3. Discuss connection to maximum likelihood estimators (ψ as score function).

von Mises expansion for L estimators

Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} F$, and for some function $J : (0,1) \to \mathbb{R}$ consider

$$\theta_0 = T(F) = \int_0^1 J(u) F^{-1}(u) du$$
$$\hat{\theta}_n = T(F) = \int_0^1 J(u) \hat{F}_n^{-1}(u) du$$

The von Mises derivative (derived in hand-written notes) is

$$T_F^{(1)}(G - F) = \int_{-\infty}^{\infty} J(F(y))[F(y) - G(y)]dy.$$

Exercise:

1. Find $u_1, \dots, u_n$ such that $\hat{\theta}_n = \sum_{i=1}^n u_i X_{(i)}$, with $X_{(1)} < \dots < X_{(n)}$.
2. Identify the function J that gives the α-trimmed mean μ_α.

See handwritten notes for von Mises expansion of $\sqrt{n}(\hat{\mu}_\alpha - \mu_\alpha)$.

Central limit theorem for Hadamard differentiable functionals. See [3].

If T is a Hadamard differentiable functional then

1. $\sqrt{n}(T(\hat{F}_n) - T(F)) \to \text{Normal}(0, \vartheta)$ in distribution as $n \to \infty$, with
$$\vartheta = \int [T_F^{(1)}(\delta_x - F)]^2 dF(x).$$
2. $\sqrt{n}(T(\hat{F}_n) - T(F))/\hat{\vartheta}^{1/2} \to \text{Normal}(0, 1)$ in distribution as $n \to \infty$, with
$$\hat{\vartheta} = \int [T_{\hat{F}_n}^{(1)}(\delta_x - \hat{F}_n)]^2 d\hat{F}_n(x).$$

Result (ii) validates $T(\hat{F}_n) \pm z_{\alpha/2}\sqrt{\hat{\vartheta}/n}$ as an asymp. $(1-\alpha)100\%$ CI for $T(F)$.

Exercise: Find $\hat{\vartheta}$ for

1. $T(F) = \int x dF(x)$
2. $T(F) = g(\int x dF(x))$.

Let $\mathcal{D}$ be the space of linear combinations of probability distributions.

Hadamard differentiability

A functional $T : \mathcal{D} \to \mathbb{R}$ is *Hadamard differentiable* at $F \in \mathcal{D}$ in the direction $G \in \mathcal{D}$ if there exists a linear function $T_F^{(1)} : \mathcal{D} \to \mathbb{R}$ such that

$$\lim_{n\to\infty} \left| \frac{T(F + \varepsilon_n(G_n - F)) - T(F)}{\varepsilon_n} - T_F^{(1)}(G - F) \right| = 0,$$

for all sequences $G_n \in \mathcal{D}$ such that $\|G_n - G\|_\infty \to 0$ and $\varepsilon_n \downarrow 0$ as $n \to \infty$

Luisa Fernholz [1] gives conditions under which

1. M-estimators
2. L-estimators
3. R-estimators (rank based estimators)

satisfy Hadamard differentiability.

Quantiles do not, but asymptotic Normality of $\sqrt{n}(\hat{\xi}_\tau - \xi_\tau)$ can still be shown.

Luisa Turrin Fernholz.
Von Mises calculus for statistical functionals, volume 19.
Springer Science & Business Media, 2012.

Jayanta K Ghosh.
A new proof of the bahadur representation of quantiles and an application.
The Annals of Mathematical Statistics, pages 1957–1961, 1971.

Larry Wasserman.
All of nonparametric statistics.
Springer Science & Business Media, 2006.