# STAT 824 sp 2023 Lec 10 slides

# Bootstrap beyond the mean: Statistical functionals

Karl B. Gregory

University of South Carolina

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 $F$

## Statistical functional

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

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

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

### Theorem (Glivenko-Cantelli Theorem)

*If $X_1, \dots, X_n$ 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 $\theta_0 = T(F)$.

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 $\tau$th quantile: $\xi_\tau = T(F) = \inf\{x : F(x) \geq \tau\}$
4. Shape parameter under $\text{Gamma}(\alpha, 1)$: $\alpha = T(F) =$ 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. The $\alpha$-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.

# Table of Contents

We want central limit theorems for $\sqrt{n}(T(\hat{F}_n) - T(F))$.

## von Mises expansion for statistical functionals

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

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

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

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

where $\varphi_F$ is called the *influence curve* of the functional $T$ at $F$.

## Theorem (How to find the influence curve)

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

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 $\varphi_F$ for these functionals

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

**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: $g(\mu) = T(F) = g(\int x dF(x))$.

## von Mises expansion for the $\tau$th 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

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

**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].

# 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 $T(F)$.

**Exercise:** Let $F$ be the Gamma$(\alpha, \beta)$ and construct 95% bootstrap CIs for

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

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

**Exercise:** Simulate coverage of 95% bootstrap CIs for the $\alpha$-trimmed mean when

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

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 |

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

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

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

Does the bootstrap work in this setting?

# 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 $\varphi_F$, 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 $\varphi_F$ 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 ($\psi$ 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 $\alpha$-trimmed mean $\mu_\alpha$.

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.