

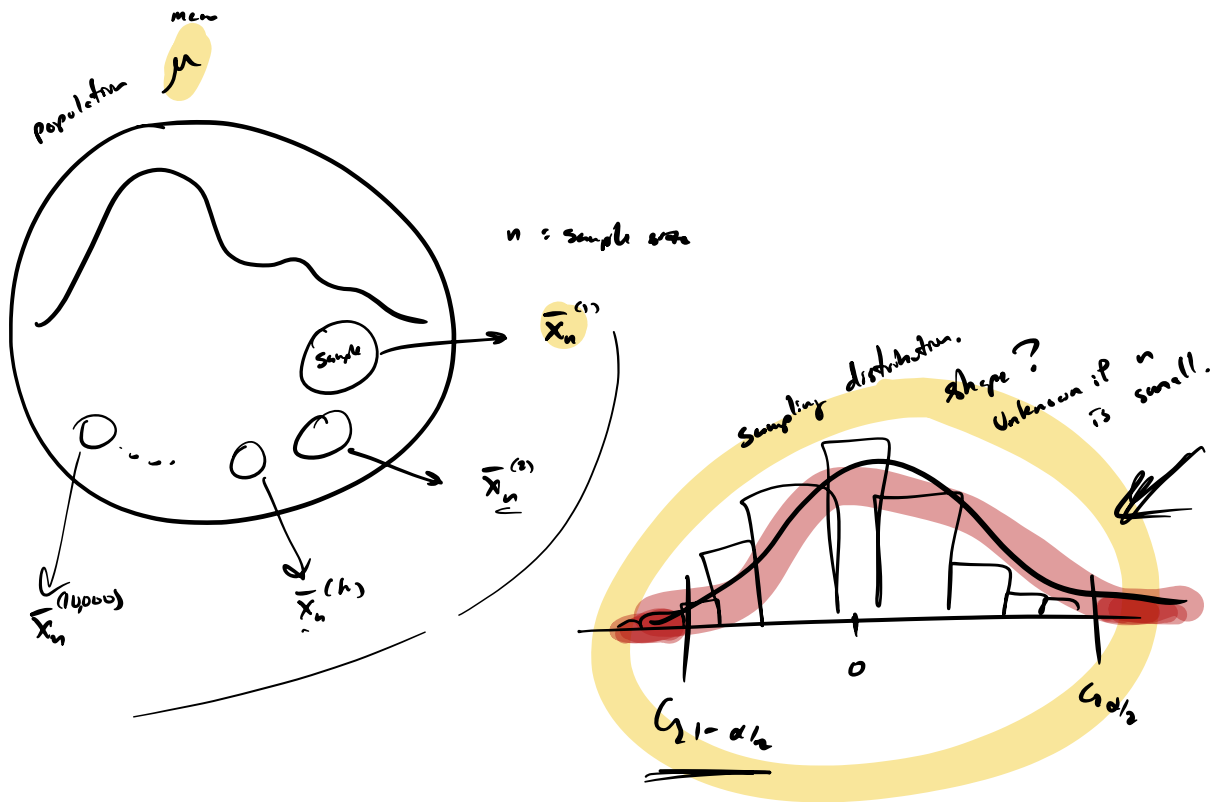
STAT 824 sp 2023 Lec 08 slides

Bootstrap for the mean

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.



If $\sqrt{n}(\bar{X}_n - \mu) \sim G$, G is known, then

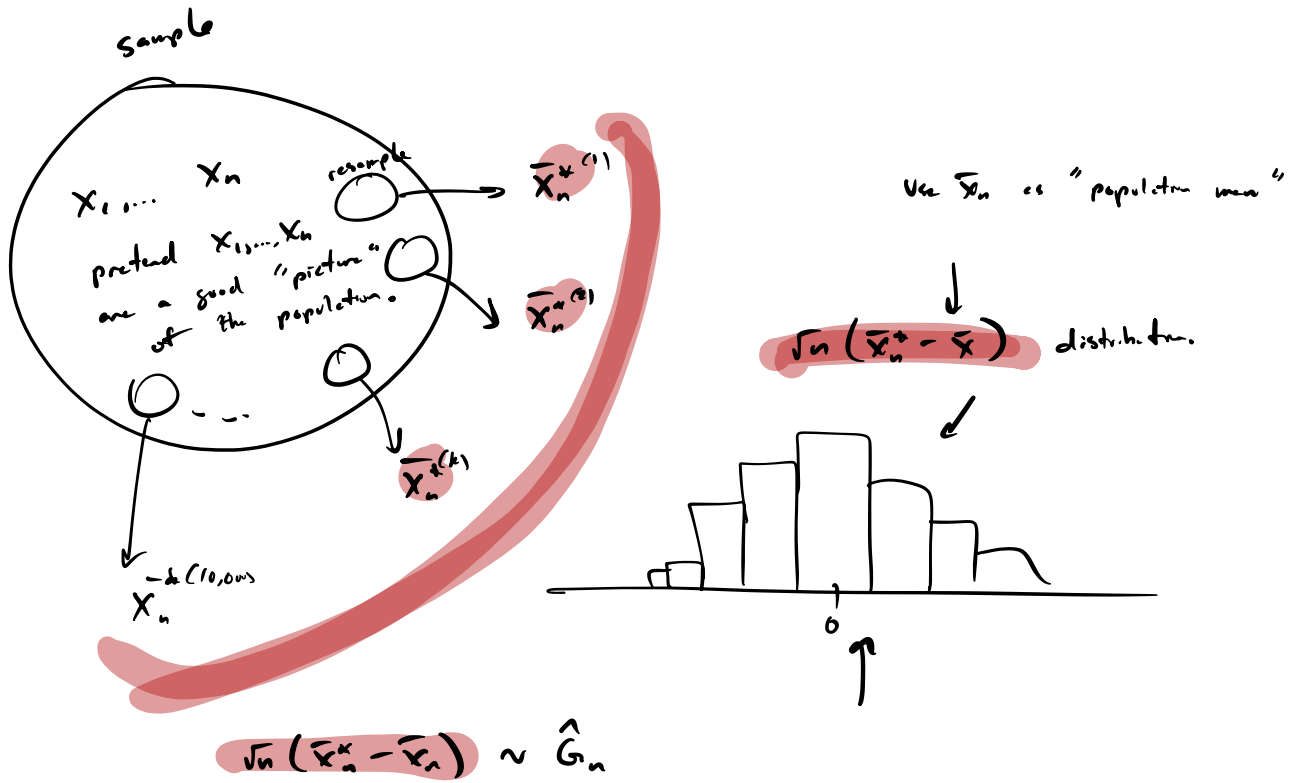
$$P(G_{1-d/2} < \sqrt{n}(\bar{X}_n - \mu) < G_{d/2}) = 1-d$$

$$\Rightarrow P\left(\frac{1}{\sqrt{n}}G_{1-d/2} - \bar{X}_n < -\mu < \frac{1}{\sqrt{n}}G_{d/2} - \bar{X}_n\right) = 1-d$$

$$P\left(\bar{X}_n - \frac{1}{\sqrt{n}}G_{d/2} < \mu < \bar{X}_n - \frac{1}{\sqrt{n}}G_{1-d/2}\right)$$

So $(1-d) \cdot 100\%$ C.S. for μ is

$$\left(\bar{X}_n - \frac{1}{\sqrt{n}}G_{d/2}, \bar{X}_n - \frac{1}{\sqrt{n}}G_{1-d/2}\right)$$



$$\left(\bar{X}_n - \frac{1}{\sqrt{n}} \hat{G}_{2n, d/2}, \bar{X}_n - \frac{1}{\sqrt{n}} \hat{G}_{2n, 1-d/2} \right)$$

Table of Contents

- 1 Bootstrap
- 2 Consistency of the bootstrap for the mean

The *bootstrap* is a method for estimating sampling distributions.

It is useful in many contexts. For now we focus on the mean of iid data:

Pivot quantities for the mean

We wish to estimate the sampling distributions of the pivots

$$T_{n,U} = \sqrt{n}(\bar{X}_n - \mu) \quad \text{or} \quad T_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \quad \text{or} \quad T_{n,S} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n},$$

$\overset{\text{unstandardized}}{\uparrow}$
 $\overset{\text{standardized}}{\uparrow}$
 $\overset{\text{studentized}}{\uparrow}$

where X_1, \dots, X_n are iid with $\mathbb{E}X_1 = \mu$, $\text{Var} X_1 = \sigma^2$ and

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{and} \quad \hat{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

We will call these the *unstandardized*, *standardized*, and *studentized* pivots for μ .

Application: build confidence intervals for μ based on these pivot quantities.

cdfs of pivot quantities for the mean

Define the cdfs of the pivots as

$$G_{n,U}(x) = P(\sqrt{n}(\bar{X}_n - \mu) \leq x)$$

$$G_n(x) = P(\sqrt{n}(\bar{X}_n - \mu)/\sigma \leq x)$$

$$G_{n,S}(x) = P(\sqrt{n}(\bar{X}_n - \mu)/\hat{\sigma}_n \leq x)$$

for all $x \in \mathbb{R}$.

The bootstrap can be used to get estimators $\hat{G}_{n,U}$, \hat{G}_n , and $\hat{G}_{n,S}$ of these cdfs.

Exercise: Give CIs for μ when $G_{n,U}$, G_n , and $G_{n,S}$ known.

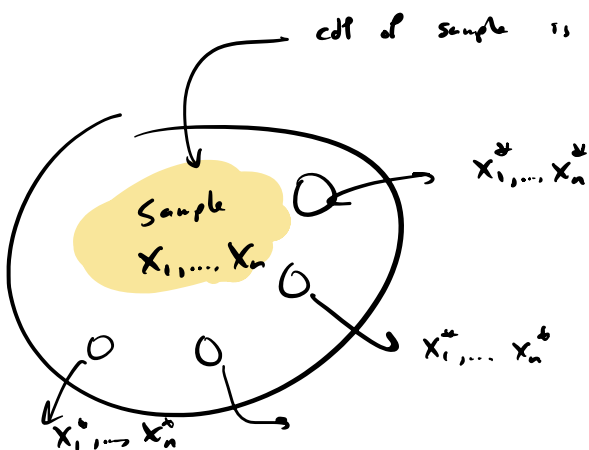
$$\left(\bar{x}_n - \frac{1}{\sqrt{n}} G_{2n,U, d/2}, \bar{x}_n + \frac{1}{\sqrt{n}} G_{2n,U, 1-d/2} \right)$$

$$P \left(G_{2n, 1-d/2} < \sqrt{n} \frac{(\bar{x}_n - \mu)}{\sigma} < G_{2n, d/2} \right) = 1-d$$

$$P \left(\bar{x}_n - \frac{\sigma}{\sqrt{n}} G_{2n, d/2} < \mu < \bar{x}_n + \frac{\sigma}{\sqrt{n}} G_{2n, 1-d/2} \right) = 1-d$$

$$P \left(G_{2n,S, 1-d/2} < \sqrt{n} \frac{(\bar{x}_n - \mu)}{\hat{\sigma}_n} < G_{2n,S, d/2} \right) = 1-d$$

$$P \left(\bar{x}_n - \frac{\hat{\sigma}_n}{\sqrt{n}} G_{2n,S, d/2} < \mu < \bar{x}_n + \frac{\hat{\sigma}_n}{\sqrt{n}} G_{2n,S, 1-d/2} \right) = 1-d$$



$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$$

↑ puts prob. mass $\frac{1}{n}$ on each of X_1, \dots, X_n .

$$G_{2n,U}(x) = P \left(\sqrt{n} (\bar{x}_n - \mu) \leq x \right)$$

$$G_{2n,U}(x) = P \left(\sqrt{n} (\bar{x}_n - \mu) / \sigma \leq x \right)$$

$$G_{2n,S}(x) = P \left(\sqrt{n} (\bar{x}_n - \mu) / \hat{\sigma}_n \leq x \right)$$

IID bootstrap for the mean

Introduce iid rvs $X_1^*, \dots, X_n^* | X_1, \dots, X_n$ with cdf $F_n(x) = n^{-1} \sum_{i=1}^n \mathbf{1}(X_i \leq x)$.

The iid bootstrap estimators $\hat{G}_{n,U}$, \hat{G}_n , and $\hat{G}_{n,S}$ of $G_{n,U}$, G_n , and $G_{n,S}$ are given by

$$\hat{G}_{n,U}(x) = P(\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x | X_1, \dots, X_n)$$

$$\hat{G}_n(x) = P(\sqrt{n}(\bar{X}_n^* - \bar{X}_n) / \hat{\sigma}_n \leq x | X_1, \dots, X_n)$$

$$\hat{G}_{n,S}(x) = P(\sqrt{n}(\bar{X}_n^* - \bar{X}_n) / \hat{\sigma}_n^* \leq x | X_1, \dots, X_n)$$



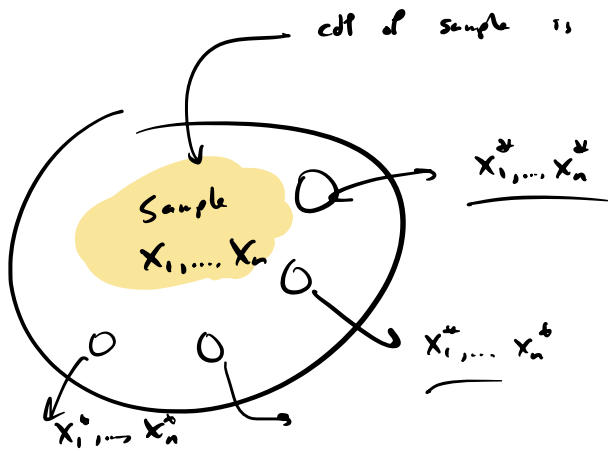
for all $x \in \mathbb{R}$, with $\bar{X}_n^* = n^{-1} \sum_{i=1}^n X_i^*$, and $(\hat{\sigma}_n^*)^2 = n^{-1} \sum_{i=1}^n (X_i^* - \bar{X}_n^*)^2$.

Idea is to ask how the pivot behaves when F_n is the population cdf.

Exercise: Show that

- 1 $\mathbb{E}[X_1^* | X_1, \dots, X_n] = \bar{X}_n$
- 2 $\text{Var}[X_1^* | X_1, \dots, X_n] = \hat{\sigma}_n^2$.

Discuss: How to build CIs after obtaining $\hat{G}_{n,U}$, \hat{G}_n , and $\hat{G}_{n,S}$.



$$\begin{aligned}
 \textcircled{1} \quad \mathbb{E}[\bar{X}_n^* | X_1, \dots, X_n] &= \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i^* \mid X_1, \dots, X_n\right] \\
 &\uparrow \\
 \mathbb{E}_x \bar{X}_n^* = \bar{X}_n &= \mathbb{E}[X_1^* | X_1, \dots, X_n] \\
 &= \sum_{x \in \{X_1, \dots, X_n\}} x \cdot p(x) \\
 &= \sum_{x \in \{X_1, \dots, X_n\}} x \cdot \frac{1}{n} \\
 &= \sum_{i=1}^n X_i \cdot \frac{1}{n} \\
 &= \bar{X}_n
 \end{aligned}$$

empirical variance of X_1, \dots, X_n .

$X_1^* | X_1, \dots, X_n$ has pmf given by $p(x) = \frac{1}{n} \mathbb{1}(x \in \{X_1, \dots, X_n\})$

$$\begin{aligned}
 \textcircled{2} \quad \text{Var}[X_1^* | X_1, \dots, X_n] &= \mathbb{E}[X_1^{*2} | X_1, \dots, X_n] - \left(\mathbb{E}[X_1^* | X_1, \dots, X_n]\right)^2 \\
 &\uparrow \\
 \text{Var}_{X_1^*} X_1^* &= \sum_{x \in \{X_1, \dots, X_n\}} x^2 \cdot \frac{1}{n} - \bar{X}_n^2 \\
 &= \frac{1}{n} \sum_{i=1}^n X_i^2 - (\bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2
 \end{aligned}$$

Bootstrap notation

Let P_* , \mathbb{E}_* and Var_* be operators such that

$$P_*(\cdot) = P(\cdot | X_1, \dots, X_n)$$

$$\mathbb{E}_*(\cdot) = \mathbb{E}(\cdot | X_1, \dots, X_n)$$

$$\text{Var}_*(\cdot) = \text{Var}(\cdot | X_1, \dots, X_n),$$

representing conditional probability, expectation, and variance, given X_1, \dots, X_n .

So P_* , \mathbb{E}_* and Var_* treat X_1^*, \dots, X_n^* as random and X_1, \dots, X_n as fixed.

$$G_{n,U}(x) = P(\sqrt{n}(\bar{X}_n - \mu) \leq x).$$

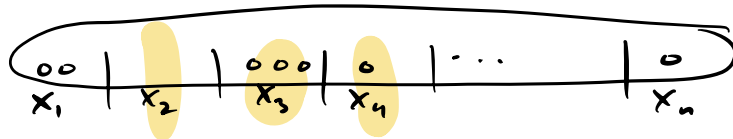
Discuss: The bootstrap estimator $\hat{G}_{n,U}$ of $G_{n,U}$ is given by

$$\hat{G}_{n,U}(x) = P_*(\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x)$$

for all $x \in \mathbb{R}$. Can we compute this? $= P(\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \leq x | X_1, \dots, X_n)$

\bar{X}_n^* has a discrete dist.

What is the total number of unique bootstrap samples?



$n-1$ "wells"

n "balls"

$$\frac{(n+n-1)!}{(n-1)! n!} = \binom{2n-1}{n-1}$$

total unique bootstrap samples.

Instead of computing $\hat{G}_{n,U}$, \hat{G}_n , and $\hat{G}_{n,S}$ exactly, we use MC approximation.

Monte Carlo approximation to bootstrap estimators $\hat{G}_{n,U}$, \hat{G}_n , and $\hat{G}_{n,S}$

For $b = 1, \dots, B$ for large B (≥ 500 , say):

1 Draw $X_1^{*(b)}, \dots, X_n^{*(b)}$ with replacement from X_1, \dots, X_n .

2 Compute

$$T_{n,U}^{*(b)} = \sqrt{n}(\bar{X}_n^{*(b)} - \bar{X}_n)$$

or $T_n^{*(b)} = \sqrt{n}(\bar{X}_n^{*(b)} - \bar{X}_n) / \hat{\sigma}_n$

or $T_{n,S}^{*(b)} = \sqrt{n}(\bar{X}_n^{*(b)} - \bar{X}_n) / \hat{\sigma}_n^{*(b)}$

better



where $\bar{X}_n^{*(b)} = n^{-1} \sum_{i=1}^n X_i^{*(b)}$ and $(\hat{\sigma}_n^{*(b)})^2 = n^{-1} \sum_{i=1}^n (X_i^{*(b)} - \bar{X}_n^{*(b)})^2$.

We now retrieve (MC-approximated) quantiles of $\hat{G}_{n,U}$, \hat{G}_n , and $\hat{G}_{n,S}$ as

$$\hat{G}_{n,U}^{-1}(u) = T_{n,U}^{*(\lceil uB \rceil)} \quad \text{or} \quad \hat{G}_n^{-1}(u) = T_n^{*(\lceil uB \rceil)} \quad \text{or} \quad \hat{G}_{n,S}^{-1}(u) = T_{n,S}^{*(\lceil uB \rceil)}$$

after sorting the realizations of each pivot in ascending order.

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{D} N(0,1) \quad \text{C.L.T.}$$

$$P(-z_{\alpha/2} < \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} < z_{\alpha/2}) \rightarrow 1 - \alpha \quad \text{c.s. } n \rightarrow \infty.$$

$$\Rightarrow \bar{X}_n \pm z_{\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}} \quad \leftarrow \text{replace } \sigma \text{ with } \hat{\sigma}_n.$$

$$T_{n,U} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \sim G_{n,U}$$

$$T_{n,U}^{(A)} \leq \dots \leq T_{n,U}^{(B)}$$

$$P(G_{n,U,1-\alpha/2} \leq \sqrt{n}(\bar{X}_n - \mu) \leq G_{n,U,\alpha/2}) = 1 - \alpha$$

$$T_{n,U}^{(B)} = \sqrt{n}(\bar{X}_n^{(B)} - \bar{X}_n)$$

$$P\left(\bar{X}_n - G_{n,U,\alpha/2} \frac{1}{\sqrt{n}} < \mu < \bar{X}_n - G_{n,U,1-\alpha/2} \frac{1}{\sqrt{n}}\right) = 1 - \alpha$$

$(1-\alpha) \times 100\%$ c.s. for μ is

$$\left(\bar{X}_n - G_{n,U,\alpha/2} \frac{1}{\sqrt{n}}, \bar{X}_n - G_{n,U,1-\alpha/2} \frac{1}{\sqrt{n}}\right)$$

$$= \left(\bar{X}_n - T_{n,U}^{(1-\alpha/2)} \frac{1}{\sqrt{n}}, \bar{X}_n - T_{n,U}^{(\alpha/2)} \frac{1}{\sqrt{n}}\right)$$

$$= \left(\bar{X}_n - \left(\sqrt{n}(\bar{X}_n^{(1-\alpha/2)} - \bar{X}_n)\right) \frac{1}{\sqrt{n}}, \bar{X}_n - \left(\sqrt{n}(\bar{X}_n^{(\alpha/2)} - \bar{X}_n)\right) \frac{1}{\sqrt{n}}\right)$$

$$= \left(2\bar{X}_n - \bar{X}_n^{(1-\alpha/2)}, 2\bar{X}_n - \bar{X}_n^{(\alpha/2)}\right)$$

$$T_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \sim G_n \rightarrow \left(\bar{X}_n - G_{n,\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}}, \bar{X}_n - G_{n,1-\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}}\right)$$

$$T_{n,s}^{+(b)} = \frac{\sqrt{n}(\bar{X}_n^{+(b)} - \bar{X}_n)}{\hat{\sigma}_n^{+(b)}}$$

$$T_{n,s} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \sim U_{n,s}$$

$$T_{n,s}^{+(1)} \leq \dots \leq T_{n,s}^{+(B)}$$

$$\rightarrow \left(\bar{X}_n - \left[G_{n,s, \alpha/2} \right] \frac{\hat{\sigma}_n}{\sqrt{n}}, \bar{X}_n - G_{n,s, 1-\alpha/2} \frac{\hat{\sigma}_n}{\sqrt{n}} \right)$$

$$\left(\bar{X}_n - T_{n,s}^{+((1-\alpha/2)B)} \frac{\hat{\sigma}_n}{\sqrt{n}}, \bar{X}_n - T_{n,s}^{+(\alpha/2)B} \frac{\hat{\sigma}_n}{\sqrt{n}} \right)$$

Discuss: Why not just use

μ

$\sqrt{n}(\bar{X}_n - \mu)/\sigma \rightarrow \text{Normal}(0, 1)$ in distribution as $n \rightarrow \infty$

and the corresponding asymptotically correct interval $\bar{X}_n \pm z_{\alpha/2} \hat{\sigma}_n/\sqrt{n}$?

Exercise: Compare via simulation the performance of these intervals:

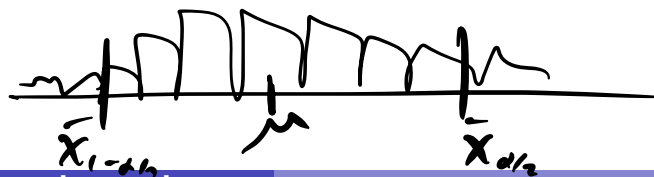
1 $(\bar{X}_n - \overbrace{\Phi^{-1}(1 - \alpha/2)}^{z_{\alpha/2}} \hat{\sigma}_n/\sqrt{n}, \bar{X}_n - \overbrace{\Phi^{-1}(\alpha/2)}^{-z_{\alpha/2}} \hat{\sigma}_n/\sqrt{n})$. Asymptotic.

2 $(2\bar{X}_n - \bar{X}_n^{*(\lceil(1-\alpha/2)B\rceil)}, 2\bar{X}_n - \bar{X}_n^{*(\lceil(\alpha/2)B\rceil)})$. The $G_{n,U}$ -based interval.

3 $(\bar{X}_n - \hat{G}_{n,S}^{-1}(1 - \alpha/2) \hat{\sigma}_n/\sqrt{n}, \bar{X}_n - \hat{G}_{n,S}^{-1}(\alpha/2) \hat{\sigma}_n/\sqrt{n})$. The $G_{n,S}$ -based.

4 $(\bar{X}_n^{*(\lceil(\alpha/2)B\rceil)}, \bar{X}_n^{*(\lceil(1-\alpha/2)B\rceil)})$. Called the *percentile interval*.

for small n when the population distribution is non-Normal.



Sampling from a Gamma(1.5, 4)

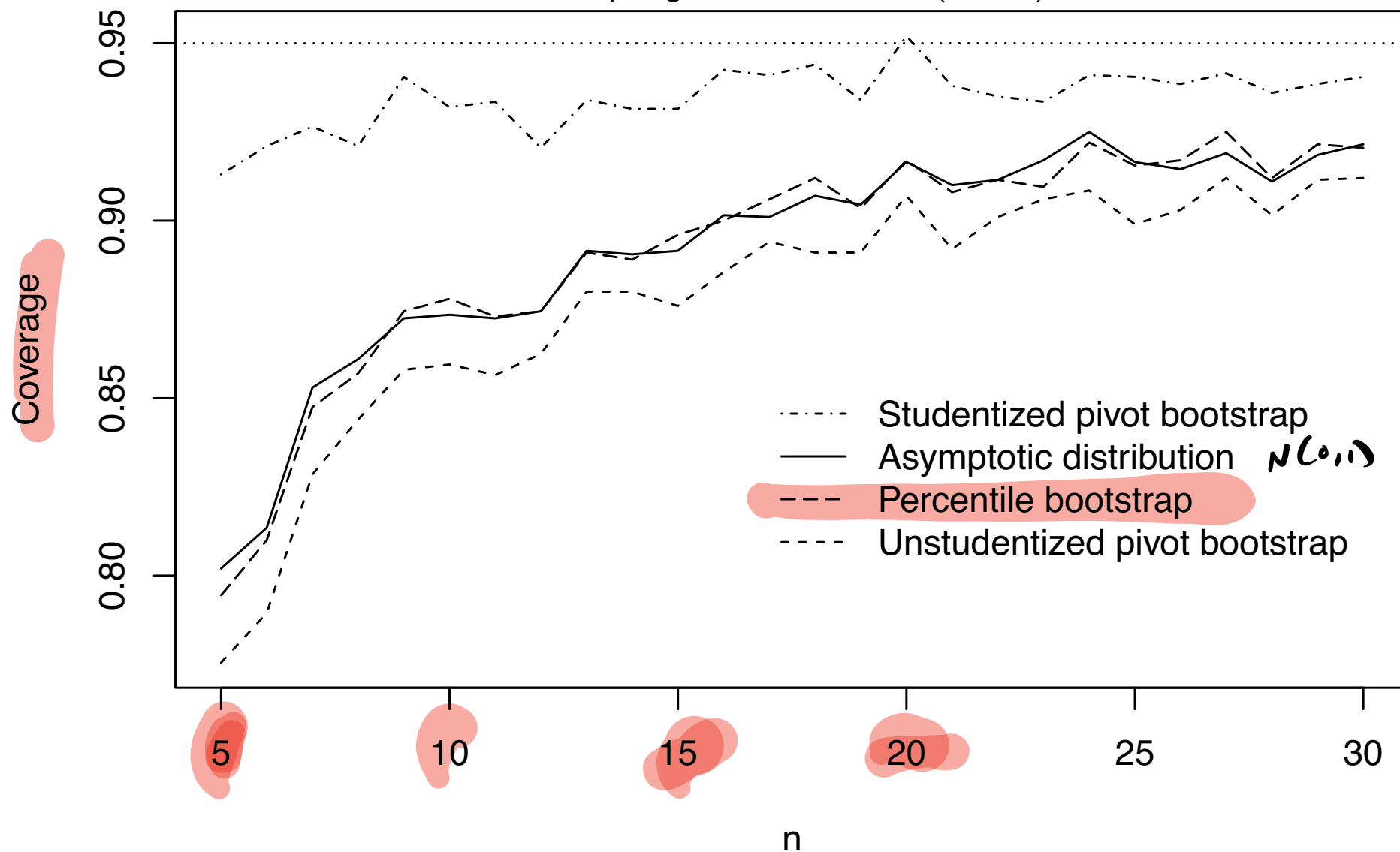


Table of Contents

- 1 Bootstrap
- 2 Consistency of the bootstrap for the mean
- 3 Other bootstrap applications

Want to show that the bootstrap dist of $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}}$ approaches the true dist. as $n \rightarrow \infty$.

$$T_n = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \quad \text{"standardized"}$$

We now present a result on the consistency of the bootstrap.

Draw X_1^*, \dots, X_n^* , compute \bar{X}_n^* . $\mathbb{E}_{X^*} X_i^* = \bar{X}_n$ $\text{Var}_{X^*} X_i^* = \hat{\sigma}_n^2$.

Bootstrap "works" for the mean

Let X_1, \dots, X_n be iid with $\mathbb{E}X_1 = \mu$, $\text{Var} X_1 = \sigma^2 < \infty$. Then

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

$$\sup_{x \in \mathbb{R}} \left| P_* \left(\frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\hat{\sigma}_n} \leq x \right) - P \left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x \right) \right| \rightarrow 0 \text{ w.p. 1 as } n \rightarrow \infty.$$

with prob 1

cdf of $\frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\hat{\sigma}_n} \mid X_1, \dots, X_n$

cdf of $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$

We could also write this as $\|\hat{G}_n - G_n\|_\infty \rightarrow 0$ w.p. 1 as $n \rightarrow \infty$.

$$\sup_{x \in \mathbb{R}} |\hat{G}_n(x) - G_n(x)| \rightarrow 0 \text{ w.p. 1.}$$

Consequently the coverage probability of the interval

$$(\bar{X}_n - \hat{G}_n^{-1}(1 - \alpha/2)\sigma/\sqrt{n}, \bar{X}_n - \hat{G}_n^{-1}(\alpha/2)\sigma/\sqrt{n})$$

converges to $(1 - \alpha)$ w.p. 1 as $n \rightarrow \infty$.

Discuss: Sufficient to show $\|\hat{G}_n - \Phi\|_\infty \rightarrow 0$ w.p. 1 as $n \rightarrow \infty$.

$$\sup_{x \in \mathbb{R}} |U_n(x) - a_n(x)| \rightarrow 0 \quad \text{u.p.t.}$$

How to show:

① We know:

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x \right) - \Phi(x) \right| \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

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

② Sufficient to show

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\hat{\sigma}_n} \leq x \right) - \Phi(x) \right| \rightarrow 0 \quad \text{u.p.t. as } n \rightarrow \infty.$$

Use B-E theorem.

$X_1^*, \dots, X_n^* \mid X_1, \dots, X_n$ iid mean \bar{X}_n , variance $\hat{\sigma}_n^2$. B-E gives

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P} \left(\frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\hat{\sigma}_n} \leq x \mid X_1, \dots, X_n \right) - \Phi(x) \right| \leq \frac{C \mathbb{E} \left[|X_1^* - \bar{X}_n|^3 \mid X_1, \dots, X_n \right]}{\hat{\sigma}_n^3 \sqrt{n}}$$

$$\sup_{x \in \mathbb{R}} \left| P_* \left(\frac{\sqrt{n}(\bar{X}_n^* - \bar{X}_n)}{\hat{\sigma}_n} \leq x \right) - \Phi(x) \right| \leq C \frac{\mathbb{E}_* |X_1^* - \bar{X}_n|^3}{\hat{\sigma}_n^3 \sqrt{n}}$$

$$(\mathbb{E}|X+Y|^p)^{\frac{1}{p}} \leq (\mathbb{E}|X|^p)^{\frac{1}{p}} + (\mathbb{E}|Y|^p)^{\frac{1}{p}}.$$

$$\begin{aligned} \left(\mathbb{E}_* |X_1^* - \bar{X}_n|^3 \right)^{\frac{1}{3}} &\leq \left(\mathbb{E}_* |X_1^*|^3 \right)^{\frac{1}{3}} + \left(\mathbb{E}_* |\bar{X}_n|^3 \right)^{\frac{1}{3}} \\ &= 2 \left(\mathbb{E}_* |X_1^*|^3 \right)^{\frac{1}{3}} \\ &= \left(|\bar{X}_n|^3 \right)^{\frac{1}{3}} \\ &= \left(|\mathbb{E}_* X_1^*|^3 \right)^{\frac{1}{3}} \\ &\stackrel{\text{Jensen's}}{\leq} \left(\mathbb{E}_* |X_1^*|^3 \right)^{\frac{1}{3}} \end{aligned}$$

so

$$\begin{aligned} \mathbb{E}_* |X_1^* - \bar{X}_n|^3 &\leq 2^3 \mathbb{E}_* |X_1^*|^3 \\ &= 2^3 \sum_{x \in \{X_1, \dots, X_n\}} |x|^3 \cdot \frac{1}{n} \\ &= 2^3 \frac{1}{n} \sum_{i=1}^n |X_i|^3. \end{aligned}$$

$$\text{Then } C \frac{\mathbb{E}_* |X_1^* - \bar{X}_n|^3}{\hat{\sigma}_n^3 \sqrt{n}} \leq C \frac{2^3 \frac{1}{n} \sum_{i=1}^n |X_i|^3}{\hat{\sigma}_n^3 \sqrt{n}}$$

set $Y_i = |X_i|^3$.
 $p = 2/3$, $\mathbb{E} |Y_i|^{2/3} = \mathbb{E} (|X_i|^3)^{2/3} = \mathbb{E} |X_i|^2 < \infty$

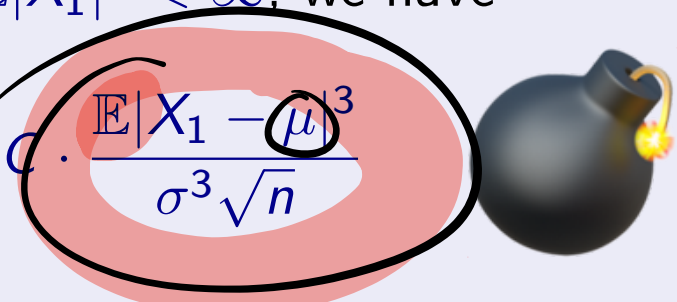
Then $\frac{\sum_{i=1}^n Y_i}{n^{3/2}} \xrightarrow{\text{w.p. 1.}} 0$

Then if $\mathbb{E} |Y_1|^p < \infty$, $n^{-1/\ominus} \sum_{i=1}^n Y_i \rightarrow 0$ w.p. 1.

We will prove that the bootstrap works with the following amazing theorem:

Berry–Esseen theorem

For X_1, \dots, X_n iid with $\mathbb{E}X_1 = \mu$, $\text{Var} X_1 = \sigma^2$, and $\mathbb{E}|X_1|^3 < \infty$, we have

$$\sup_{x \in \mathbb{R}} \left| P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x\right) - \Phi(x) \right| \leq C \cdot \frac{\mathbb{E}|X_1 - \mu|^3}{\sigma^3 \sqrt{n}}$$


for each $n \geq 1$ and $0 \leq C \leq \sqrt{\frac{2}{\pi}} \left(\frac{5}{2} + \frac{12}{\pi}\right) < 5.05$. See pg 361 of [1].

Exercise: Prove the bootstrap works with B–E and results on next slide.

M-Z

Special case of Marcinkiewz–Zygmund SLLN

Let Y_1, \dots, Y_n be iid, $p \in (0, 1)$. Then if $\mathbb{E}|Y_1|^p < \infty$, $n^{-1/p} \sum_{i=1}^n Y_i \rightarrow 0$ w.p. 1.

Minkowski's inequality

$$\left(\mathbb{E}|a+b|^p\right)^{\frac{1}{p}} \leq \left(\mathbb{E}|a|^p\right)^{\frac{1}{p}} + \left(\mathbb{E}|b|^p\right)^{\frac{1}{p}}$$

For any rvs $X, Y \in \mathbb{R}$, $p \in (1, \infty)$, we have $\left(\mathbb{E}|X+Y|^p\right)^{\frac{1}{p}} \leq \left(\mathbb{E}|X|^p\right)^{\frac{1}{p}} + \left(\mathbb{E}|Y|^p\right)^{\frac{1}{p}}$.

Jensen's inequality

If $g : \mathbb{R} \rightarrow \mathbb{R}$ is a convex function, then for any rv X we have

$$g(\mathbb{E}X) \leq \mathbb{E}g(X),$$

provided $\mathbb{E}|X| < \infty$ and $\mathbb{E}|g(X)| < \infty$.

$$\frac{\mathbb{E}|X_i^* - \bar{X}_n|^3}{\hat{\sigma}_n^3 \sqrt{n}}$$

Our bootstrap result concerns only the standardized pivot $T_n = \sqrt{n}(\bar{X}_n - \mu)/\sigma$.

But these pivots are the ones we can use (since we don't know σ):

$$T_{n,U} = \sqrt{n}(\bar{X}_n - \mu) \quad \text{and} \quad T_{n,S} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} \leftarrow \text{better.}$$

We can also use B-E to prove that the bootstrap “works” for these pivots.

Discuss: Is one better than the other??

Table of Contents

- 1 Bootstrap
- 2 Consistency of the bootstrap for the mean
- 3 Other bootstrap applications**

Fisher's z transformation

Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be iid realizations of (X, Y) . Let

$$\rho = \text{corr}(X, Y) \quad \text{and} \quad \hat{\rho} = \frac{\sum_{i=1}^n (X_i - \bar{X}_n)(Y_i - \bar{Y}_n)}{\sqrt{\sum_{i=1}^n (X_i - \bar{X}_n)^2 \sum_{i=1}^n (Y_i - \bar{Y}_n)^2}}.$$

If (X, Y) are bivariate Normal then $\sqrt{n}(\zeta(\hat{\rho}) - \zeta(\rho)) \xrightarrow{D} \text{Normal}(0, 1)$, where

$$\zeta(\rho) = \frac{1}{2} \log \left(\frac{1 + \rho}{1 - \rho} \right). \quad (\text{can show by delta method})$$

Then an asymptotic $(1 - \alpha) \times 100\%$ CI for ρ is given by

$$\left[\zeta^{-1} \left(\zeta(\hat{\rho}) - \frac{z_{\alpha/2}}{\sqrt{n}} \right), \zeta^{-1} \left(\zeta(\hat{\rho}) + \frac{z_{\alpha/2}}{\sqrt{n}} \right) \right], \quad \zeta^{-1}(z) = \frac{e^{2z} - 1}{e^{2z} + 1}.$$

Exercise: Compare asymp. interval with percentile boot interval in simulation:

- 1 When (X, Y) bivariate Normal with correlation $\rho = 1/2$.
- 2 When $Y|X \sim \text{Normal}(X, \sigma^2)$, $X \sim \text{Expo}(\lambda)$, $\text{corr}(X, Y) = \sqrt{\lambda^2 / (\lambda^2 + \sigma^2)}$.

 Krishna B Athreya and Soumendra N Lahiri.
Measure theory and probability theory.
Springer Science & Business Media, 2006.