$X_1, \dots, X_n$ i.i.d from distribution with cdf $F$ $(F(x) = P(X_i \le x))$

- GOAL: CI FOR $F^{-1}(0.5)$
  ↑ POPULATION MEDIAN
- POINT ESTIM: SAMPLE MEDIAN $\tilde{m}_n$
- put $R_n = \tilde{m}_n - F^{-1}(0.5)$

  if distrib of $R_n$ known, then

$$P\left( \mu_n(\alpha/2) \le \tilde{m}_n - F^{-1}(0.5) \le \mu_n(1-\alpha/2) \right) = 1-\alpha$$

  ↑ $\alpha/2$-quantile $R_n$ ; ↑ $(1-\alpha/2)$-quantile $R_n$

$$\rightsquigarrow P\left( \tilde{m}_n - \mu_n(1-\alpha/2) \le F^{-1}(0.5) \le \tilde{m}_n - \mu_n(\alpha/2) \right)$$

---

Here it seems better to consider $\tilde{R}_n = \dfrac{\tilde{m}_n}{F^{-1}(0.5)}$

$$\rightsquigarrow P\left( \tilde{\mu}_n(\alpha/2) \le \frac{\tilde{m}_n}{F^{-1}(0.5)} \le \tilde{\mu}_n(1-\alpha/2) \right) = 1-\alpha$$

$$\downarrow$$

$$P\left( \frac{\tilde{m}_n}{\tilde{\mu}_n(1-\alpha/2)} \le F^{-1}(0.5) \le \frac{\tilde{m}_n}{\tilde{\mu}_n(\alpha/2)} \right)$$

A) monte-Carlo principle

- if $X_i \sim \mathrm{Exp}\left(\frac{1}{90}\right)$, $f_X(x) = \lambda \exp\{-\lambda x\} \mathbb{1}\{x>0\}$, $\lambda = \frac{1}{90}$

$EX_i = 90$

(0) FOR $b = 1, \dots, B$
(1) GENERATE $X_1^*, \dots, X_n^*$ FROM $\mathrm{Exp}\left(\frac{1}{90}\right)$
(2) CALCULATE $R_n^* = \tilde{m}_n^* - F^{-1}(0.5)$

$\rightsquigarrow R_1^*, \dots, R_B^*$ $\rightsquigarrow \mu_n^*(\alpha/2)$, $\mu_n^*(1-\alpha/2)$

↑ SAMPLE QUANTILES

## B) PARAMETRIC BOOTSTRAP

- assume that $X_i \sim Exp(\lambda)$, $\lambda$ unknown

$\rightsquigarrow \hat{\lambda}_n = \frac{1}{\bar{X}_n}$ (MLE)

- (1) CHANGE LINE (2) TO

$X_1^*, \dots, X_n^*$ FROM $Exp(\hat{\lambda}_n)$

## C) NONPARAMETRIC BOOTSTRAP

- (2): $X_1^*, \dots, X_n^*$ FROM $\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{X_i \le x\}$

↑ edf

$\Leftrightarrow$ SIMPLE RANDOM SAMPLE OF SIZE $n$ FROM $X_1, \dots, X_n$ WITH REPLAC

# NONPARAMETRIC BOOTSTRAP (THEORY)

$X_1, \dots, X_n$ i.i.d. $\sim F$

- interest is in estimating some characteristic of distribution $F$, say $T(F)$ (e.g. $T(F) = \mathbb{E}X_i$, $T(F) = F^{-1}(0.5)$, ...)

**Remark** Sole i.i.d. STRUCTURE $\mathcal{D}_0$
$\rightsquigarrow$ it cannot be used for time-series $\mathcal{D}_0$.

- put $R_n = \sqrt{n}(T(F_n) - T(F))$
  ($F_n$ ↑ $\bar{x}_n$, $\tilde{m}_n$)

- NB: estimates $H_n(x) = P(R_n \le x)$ with RHS $\frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\{R^*_{n,b} \le x\} = : H^*_{n,B}(x)$,
  where $R^*_{n,b} = T(F^*_{n,b}) - T(F_n)$

Note that $H^*_{n,B}(x) \xrightarrow[B\to\infty]{a.s.} H^*_n(x) = P_{F_n}(R^*_n \le x)$
↑ cdf of $R^*_n = T(F^*_{n,b}) - T(F_n)$ from $x_1, \dots, x_n$ fixed

- ~~NB cont~~
- ~~unknown $H_n(x)$ is estimated by $H^*_n(x)$~~
- PLUG-IN PRINCIPLE: $H_n(x) = P_F(R_n \le x)$ is estimated by $H^*_n(x) = P_{F_n}(R^*_n \le x)$ is estimated by
- MONTE CARLO: $H^*_{n,B}(x) = \frac{1}{B}\sum_{b=1}^{B} \mathbb{1}\{R^*_{n,b} \le x\}$.

WE AIM AT SHOWING THAT: DETERMINISTIC (UNKNOWN)

$$\sup_{x \in \mathbb{R}} \left| H_n^*(x) - H_n(x) \right| \xrightarrow[n \to \infty]{a.s.} 0 \qquad (4)$$

$H_n^*$: RANDOM - DEPENDS

---

**Theorem** (4) HOLDS FOR $T(F) = \mathbb{E}X_1$ provided that $\mathrm{var}(X_1) < \infty$.

---

In more detail

$$\sup_{x \in \mathbb{R}} \left| P_{F_n}\left(\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \le x\right) - P_F\left(\sqrt{n}(\bar{X}_n - \mathbb{E}X_1) \le x\right) \right| \xrightarrow[n \to \infty]{a.s.} 0$$

CLT: $\sqrt{n}(\bar{X}_n - \mathbb{E}X_1) \xrightarrow[n \to \infty]{d} N(0, \mathrm{var}(X_1))$

$\Rightarrow$ Thm implies that also $\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \xrightarrow[n \to \infty]{d} N(0, \mathrm{var}(X_1))$ a.s.

---

**Theorem** (4) HOLDS FOR $T(F) = g(\mathbb{E}X_1)$, provided that $\mathrm{var}(X_1) < \infty$ and $g \in C^1(\mathcal{U}(\mathbb{E}X_1))$

continuous part domain of a neighb. of $\mathbb{E}X_1$

Example:

(i) $T(F_n) = \log(\bar{X}_n)$

(ii) $\binom{X_1}{Y_1}, \dots, \binom{X_n}{Y_n}$ i.i.d., then

$$\underset{n \times p}{\mathbb{X}} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \quad \underset{n \times 1}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$$

$$\hat{\beta}_n = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T Y = \left(\sum_{i=1}^n X_i X_i^T\right)^{-1} \sum_{i=1}^n X_i Y_i$$

$$= \left(\frac{1}{n}\sum_{i=1}^n X_i X_i^T\right)^{-1} \frac{1}{n}\sum_{i=1}^n X_i Y_i = g(\bar{W}_n, \bar{Z}_n)$$

$\bar{W}_n = \frac{1}{n}\sum X_i$, $\frac{1}{n}\sum X_i Y_i$

**Theorem** (a) HOLDS PROVIDED $T(F)$ is Hadamard differentiable (VIEWED AS A FCTIONAL $\cdot = \longrightarrow T(\cdot)$) at $F$.

**Remark:**

* Had. diff. implies that

$$\sqrt{n}\left(T(F_n) - T(F)\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} L(X_i) + \underbrace{o_p(1)}_{\text{REMAINDER TERM}} \xrightarrow{p} 0$$

* HD generalizes $\Delta$-theorem:

e.g. $T(F) = g(\mathbb{E}X_1)$, $g: \mathbb{R} \to \mathbb{R}$

$$\rightsquigarrow \sqrt{n}\left(T(F_n) - T(F)\right) = g'(\mathbb{E}X_1) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mathbb{E}X_i) + o_p(1)$$

## Bootstrap vs. normal approximation

e.g. $X_1, \dots, X_n$ i.i.d $\mathrm{Exp}(\lambda)$

$X_1, \dots, X_n$ i.i.d

- GOAL: CI for $\mathbb{E}X_i$

CLT: $\sqrt{n}\,(\bar{X}_n - \mathbb{E}X_i) \xrightarrow[n\to\infty]{d} N(0, \mathrm{var}X_i)$

$$\frac{\sqrt{n}\,(\bar{X}_n - \mathbb{E}X_i)}{S_n} \xrightarrow[n\to\infty]{d} N(0,1) \rightsquigarrow$$

$$\rightsquigarrow P\left(-u_{1-\alpha/2} \le \frac{\sqrt{n}\,(\bar{X}_n - \mathbb{E}X_i)}{S_n} \le u_{1-\alpha/2}\right) \xrightarrow[n\to\infty]{} 1-\alpha$$

$\rightsquigarrow$ Asympt. CI for $\mathbb{E}X_i$:

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

Bootstrap: $R_n^* = \sqrt{n}\left(\bar{X}_n^* - \bar{X}_n\right) \rightsquigarrow$

Studentized bootstrap: $\tilde{R}_n^* = \dfrac{\sqrt{n}\,(\bar{X}_n^* - \bar{X}_n)}{S_n} \rightsquigarrow$

now we compare approx. of $P\left(\frac{\sqrt{n}(\bar{X}_n - \mathbb{E}X_i)}{S_n} \le x\right)$

↗ NORMAL APPROX. $\Phi(x)$

↑ STUDENT. BOOT APPROX. $\widehat{H}_{n,B}(x)$

↑ def of $\widehat{R}^*_{n,1}, \dots \widehat{R}^*_{n,B}$

## LIMITS OF NONP BOOTSTRAP

- $X_1, \dots, X_n$ ARE NOT i.i.d
- does not work for $T(F_n) = \max_{1 \le i \le n} \{X_i\}$
- it might be sensitive to outlying values

## BOOTSTRAP CONF. INTERVALS

$T_n = T(F_n)$

I. NORMAL CI - one recalculates $T_n \rightsquigarrow T^*_{n,1}, \dots T^*_{n,B}$

$\left(T_n(\hat{F}_n) \mp u_{1-\alpha/2}\ S.E(T_n)\right)$

$\sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(T^*_{n,b} - \bar{T}^*_{n,\cdot}\right)^2}$

II. Bootstrap basic CI

- one resamples $R^*_n = \sqrt{n}\left(T^*_n - T_n\right) \rightsquigarrow R^*_{n,1}, \dots R^*_{n,B}$

sample quantiles $\rightsquigarrow$ unknown quantiles

$u^*_{n,B}(\alpha/2)$, $u^*_{n,B}(1-\alpha/2)$ are estimates of quantiles of the distrib of $R_n = \sqrt{n}\left(T_n - T(F)\right)$

$\rightsquigarrow$ if NB works then:

$$P\left(u^*_{n,B}(\alpha/2) \le \sqrt{n}\left(T^*_n - T(F)\right) \le u^*_{n,B}(1-\alpha/2)\right) \simeq 1-\alpha$$

$$\Vert$$

$$P\left(T^*_n - \frac{u^*_{n,B}(1-\alpha/2)}{\sqrt{n}} \le T(F) \le T_n - \frac{u^*_{n,B}(\alpha/2)}{\sqrt{n}}\right)$$

$\rightsquigarrow$ CI for $T(F)$: $\left( T_n - \frac{\hat{\mu}^*_{n,B}(1-\alpha/2)}{\sqrt{n}} \,,\, T_n - \frac{\hat{\mu}^*_{n,B}(\alpha/2)}{\sqrt{n}} \right)$

$$\| $$

$\left( 2T_n - q^*_{n,B}(1-\alpha/2) \,,\, 2T_n - q^*_{n,B}(\alpha/2) \right)$

where $q^*_{n,B}(\alpha)$ = quantile of $T^*_{n,1}, \ldots, T^*_{n,B}$

III. Percentile CI

$\left( q^*_{n,B}(\alpha/2) \,,\, q^*_{n,B}(1-\alpha/2) \right)$

IV. Studentized boot. CI

- one resamples $\tilde{R}^*_n = \frac{\sqrt{n}(T^*_n - T_n)}{S.E.(T^*_n)}$ $\rightsquigarrow \tilde{R}^*_{n,1}, \ldots \tilde{R}^*_{n,B}$

to mimic $R_n = \frac{\sqrt{n}(T_n - T(F))}{S.E.(T_n)}$

$\rightsquigarrow$ CI for $T(F)$:

$\left( T_n - \tilde{\mu}_{n,B}(1-\alpha/2)\, S.E.(T_n) \,,\, T_n - \tilde{\mu}_{n,B}(\alpha/2)\, S.E.(T_n) \right)$

# PARAMETRIC BOOTSTRAP

Let $\underset{\sim}{X}_1, \dots, \underset{\sim}{X}_n$ have the joint cdf $F(\cdot; \underset{\sim}{\theta}_X)$.

Then $\underset{\sim}{X}^*_{1,b}, \dots, \underset{\sim}{X}^*_{n,b}$ are generated from $F(\cdot; \hat{\underset{\sim}{\theta}}_n)$, where $\hat{\underset{\sim}{\theta}}_n$ is an estimate of $\underset{\sim}{\theta}_X$.

Ex. $X_1, \dots, X_n$ i.i.d $U(0,\theta)$ i.e. $f(x,\theta) = \frac{1}{\theta} \mathbb{1}\{x \in [0,\theta]\}$.

$\Rightarrow$ MLE $\hat{\theta}_n = \max_{1 \le i \le n} \{X_i\}$

$R_n = \hat{\theta}_n - \theta_X$ (or better $R_n = \frac{\hat{\theta}_n}{\theta_X}$)

$X^*_{1,b}, \dots, X^*_{n,b}$ i.i.d from $U(0, \hat{\theta}_n)$

## Goodness-of-fit testing

Let $\underset{\sim}{X}_1, \dots, \underset{\sim}{X}_n$ be r.v.d from $F$

$H_0 : F \in \mathcal{F} = \{F(x;\theta), \theta \in \Theta\}$ | Eg. $\mathcal{F} = \{\text{Exp}(\lambda), \lambda > 0\}$

$KS_n = \sup_x |F_n(x) - F(x; \hat{\theta}_n)|$ | Ex. $F(x;\lambda) = 1 - e^{-\lambda x}$

$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}(X_i \le x)$ | $F(x; \hat{\lambda}_n) = 1 - e^{-\hat{\lambda}_n}$

$\hat{\lambda}_n = \frac{1}{\bar{X}_n}$

(1) FOR $b = 1, \dots, B \leftarrow 999$

(2) $X^*_{1,b}, \dots, X^*_{n,b}$ i.i.d from $F(x; \hat{\lambda}_n)$

(3) $KS^*_{n,b} = \sup_{x \in \mathbb{R}} |\hat{F}_{n,b}(x) - F(x; \hat{\theta}^*_{n,b})|$

(4) Estim of p-value: $\dfrac{1 + \sum_{b=1}^{B} \mathbb{1}\{KS^*_{n,b} \ge KS_n\}}{1 + B}$

# Testing hypotheses and bootstrap

- one can use duality of CI and testing
- alternatively let $T_n$ be a test statistic whose large values speak against $H_0$

Ex $X_1, \dots, X_n$ i.i.d $F_X$ } indep $H_0 : F_X^{-1}(0.5) = F_Y^{-1}(0.5)$
$Y_1, \dots, Y_m$ i.i.d $F_Y$

$$T_{n,m} = \left| \log\left( \frac{\text{SAMPLE MEDIAN OF } X_1, \dots, X_n}{\text{SAMPLE MEDIAN OF } Y_1, \dots, Y_m} \right) \right|$$

- Let $T_n$ should be resampled under $H_0$

indep { $X_{1,b}^*, \dots, X_{n,b}^*$ — non-parametric bootstrap from $X_1 - \hat{\mu}_n^X, \dots, X_n - \hat{\mu}_n^X$
$Y_{1,b}^*, \dots, Y_{m,b}^*$ — " — $Y_1 - \hat{\mu}_m^Y, \dots, Y_m - \hat{\mu}_m^Y$

- recalculate $T_{n,b}^*$

- p-value $\dfrac{1 + \sum_{b=1}^{B} \mathbb{1}\{ |T_{n,b}^*| \geq T_n \}}{1 + B}$

Ex Comparing expectation $T_n = |\bar{X}_n - \bar{Y}_n|$, $T_n = \dfrac{|\bar{X}_n - \bar{Y}_n|}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}}$, $p = 2(1 - \Phi_t(T_n))$

## Pairwise comparisons and bootstrap

$Y_{11}, \dots, Y_{1n_1}$ i.i.d $F_1$ }
$\vdots$ indep
$Y_{K1}, \dots, Y_{Kn_K}$ i.i.d $F_K$ }

$\mu_1 = F_1^{-1}(0.5)$ or
$\mu_K = F^{-1}(0.5)$

$H_0^{(r,s)}$ : $\mu_r = \mu_s$ $H_1^{(r,s)}$ : $\mu_r \neq \mu_s$

$$T_{rs} = \left| \log\left( \frac{\hat{\mu}_r}{\hat{\mu}_s} \right) \right|$$ ← sample median of the $Y_{r1}, \dots, Y_{rn_r}$

$T_{rs}^{(b)}$ $T_b^{(r,s)*} = \left| \log\left( \frac{\hat{\mu}_{r,b}^*}{\hat{\mu}_{s,b}^*} \right) \right|$

NONP BOOT SAMPLE
RD $Y_{r1}, \dots, Y_{rn_r}$

$T_b^{(max)} = \max_{j,k} |\dots|$

- p-value

ADJUSTED p-value: $p_{adj}^{(j,k)} = \dfrac{1 + \sum_{b=1}^{B} \mathbb{1}\{T_b^{(max)} \geq T_b^{(j,k)}\}}{1+B}$

## Computing expectations

$$T^{(j,k)} = \frac{|\bar{Y}_{j\cdot} - \bar{Y}_{k\cdot}|}{\sqrt{\frac{S_j^2}{n_j} + \frac{S_k^2}{n_k}}}$$

$$\bar{Y}_{j\cdot} = \frac{1}{n_j}\sum_{i=1}^{n_j} Y_{ji}, \quad S_j^2 = \frac{1}{n_j - 1}\sum_{i=1}^{n_j}(Y_{ji} - \bar{Y}_{j\cdot})^2$$

$$\bar{Y}_{k\cdot} = \frac{1}{n_k}\sum_{i=1}^{n_k} Y_{ki}$$

cdf of t-distribution with $\nu$ degrees of freedom

$$p^{(j,k)} = 2\left(1 - G_\nu(|t_0^{(j,k)}|)\right)$$

$$p_{min,b}^{*} = \min_{j,k} p_b^{*(j,k)}$$

$$p_{adj}^{(j,k)} = \frac{1 + \sum_{b=1}^{B} \mathbb{1}\{p_{min,b}^{*} \leq p^{(j,k)}\}}{1+B}$$

## Examples

Quantile regression

$\begin{pmatrix} Y_1 \\ X_1 \end{pmatrix}, \dots, \begin{pmatrix} Y_n \\ X_n \end{pmatrix}$ i.i.d

- Let $F^{-1}_{Y|X}(\tau)$ be the conditional $\tau$-quantile of $Y$ given $X$

$F^{-1}_{Y|X}(\tau) = \alpha_\tau + \beta_\tau X$

- model: $\begin{pmatrix} Y_1 \\ X_1 \end{pmatrix}, \dots, \begin{pmatrix} Y_n \\ X_n \end{pmatrix}$ i.i.d ("cases")

$\rightsquigarrow$ nonp. bootstrap $\begin{pmatrix} Y_1^* \\ X_1^* \end{pmatrix}, \dots, \begin{pmatrix} Y_n^* \\ X_n^* \end{pmatrix}$ i.i.d from $F_n$

# TIME SERIES

AR(1): $X_{n+1} = \mu + \rho(X_n - \mu) + \varepsilon_{n+1}$ , $n = 1, \dots, T$

- GOAL: CI for $\mu$

## model-based bootstrap

- calculate $\hat{\varepsilon}_{n+1} = X_{n+1} - \hat{\mu} - \hat{\rho}(X_n - \hat{\mu})$ , $n = 1, \dots, T-1$
- center: $\hat{\varepsilon}_n^{(c)} = \hat{\varepsilon}_n - \frac{1}{T-1}\sum_{n=1}^{T-1} \hat{\varepsilon}_n$ , $n = 1, \dots, T$
- FOR $b = 1, \dots, B$:
  $\hat{\varepsilon}^*_{1,b}, \dots, \hat{\varepsilon}^*_{T,b}$ as i.i.d from $\hat{\varepsilon}_2, \dots, \hat{\varepsilon}_T$
  $X^*_{1,b} = \hat{\varepsilon}^*_{1,b}$ , $X^*_{2,b} = \hat{\rho} X^*_{1,b} + \hat{\varepsilon}^*_{2,b}$ , $\dots$ $X^*_{n,b} = \hat{\rho} X^*_{n-1,b} + \hat{\varepsilon}^*_{n,b}$
  $\hat{\mu}^*_b = \frac{1}{T}\sum_{n=1}^{T} X^*_{n,b}$ ← st. dev. of $\hat{\mu}^*_1, \dots, \hat{\mu}^*_B$

CI: $\left(\hat{\mu}_T \mp \kappa_{1-\alpha/2} \cdot S.E.(\hat{\mu}_T)\right)$ , $S.E. = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\mu}^*_b - \hat{\mu}^*_{\cdot})^2}$

## JACK-KNIFE (leave-one out) - ancestor of bootstrap

- $X_1, \dots, X_n$ data
- let $\hat{\theta}_n = T_n(X_1, \dots, X_n)$
  $\hat{\theta}^{(-i)}_n = T(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)$
- jack-knife pseudo-values:
  $\tilde{\theta}^*_{n,i} = n\hat{\theta}_n - (n-1)\hat{\theta}^{(-i)}_n$

→ $S.E.(\hat{\theta}_n) = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\tilde{\theta}^*_{n,i} - \bar{\theta}^*_n)^2} \cdot \frac{1}{n}$ , $\bar{\theta}^*_n = \frac{1}{n}\sum_{i=1}^{n}\tilde{\theta}^*_{n,i}$

$= \frac{1}{n} \cdot$ st. dev. of $\tilde{\theta}^*_{n,1}, \dots, \tilde{\theta}^*_{n,n}$

- works for $\hat{\theta}_n = g(\bar{X}_n)$ (nice; $\bar{X}_n$ ← mean of i.i.d)
- doesn't work for quantiles