

# BOOTSTRAP

STAT 432

Spring 2020

DALPIAZ

POPULATION DISTRIBUTION



RANDOM VARIABLES

$X_1, X_2, \dots, X_n \sim p(x|0)$

$X_1, X_2, \dots, X_n$

(POTENTIAL) REALIZED VALUES

2.1, 1.3, ..., 3.4

REALIZED VALUES

$\theta$

PARAMETER

$\hat{\theta}$

ESTIMATOR

$\hat{\theta}$

DISTRIBUTION?



USE FOR INTERVAL  
ESTIMATE

WHAT IF  $p(x|0)$  IS UNKNOWN?

↳ ESTIMATE IT ! THEN SIMULATE !

## Bootstrap Resample

- A SAMPLE OF SIZE  $n$  WITH REPLACEMENT DRAWN FROM THE ORIGINAL SAMPLE.
- SAME AS A RANDOM SAMPLE OF SIZE  $n$  FROM EDF

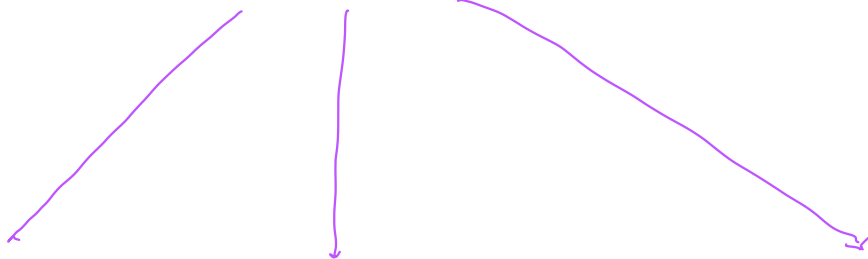
## Bootstrap Replicate

- A STATISTIC CALCULATED ON A

BOOTSTRAP RESAMPLE

# Simulation

$$p(x|y)$$



$$x^{(n1)} = (x_1^{(n1)}, x_2^{(n1)}, \dots, x_n^{(n1)})$$

$$x^{(n2)} = (x_1^{(n2)}, x_2^{(n2)}, \dots, x_n^{(n2)})$$

...

$$x^{(n3)} = (x_1^{(n3)}, x_2^{(n3)}, \dots, x_n^{(n3)})$$



$$S(x^{(n1)})$$



$$S(x^{(n2)})$$

...



$$S(x^{(n3)})$$

# BOOTSTRAP

SAMPLE WITH REPLACEMENT  
OF SIZE  $n$

ORIGINAL RANDOM SAMPLE  
FROM POPULATION

$$X = (x_1, x_2, \dots, x_n)$$

REPEAT  $B$  TIMES

$$X^{*1} = (x_1, x_2, \dots, x_n) \quad X^{*2} = (x_1, x_2, \dots, x_n) \quad \dots \quad X^{*B} = (x_1, x_2, \dots, x_n)$$

BOOT  
RESAMPLES

$$S(X^{*1})$$

$$S(X^{*2})$$

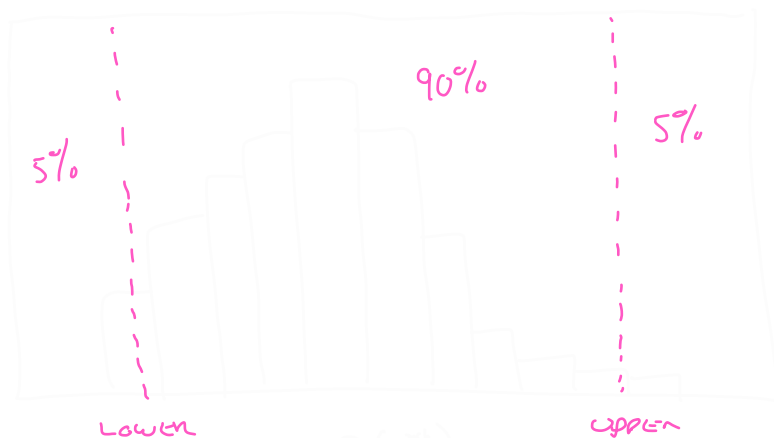
$$S(X^{*B})$$

BOOT  
REPLICATES

# BOOTSTRAP CONFIDENCE INTERVAL

(PERCENTILE METHOD)

HISTOGRAM OF BOOT REPS



90% CI