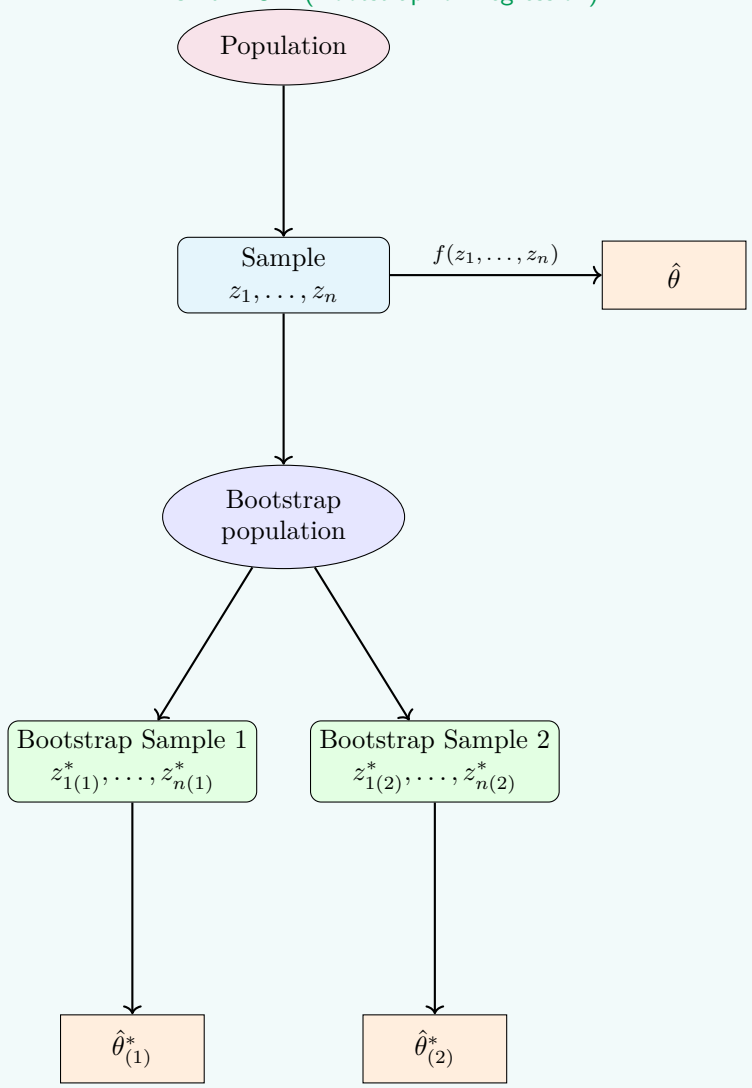
STAT 151A Lecture 20

Henry Liev

13 October 2025

Remark 0.1 (Bootstrap for Regression)



Remark 0.2

Let $\hat{\theta}$ be sample mean
$z = (6, -3, 5, 3)$
$\hat{\theta} = \bar{z} = \frac{6-3+5+3}{4} = \frac{11}{4}$
How variable is $\hat{\theta}$ across different samples from our population? i.e. what is $\text{SE}(\hat{\theta})$? 2 ways to answer

(1) Use formula for sample mean $\text{SE}(\hat{\theta}) = \frac{\sigma}{\sqrt{n}} \approx \frac{\hat{\sigma}}{\sqrt{n}} = \frac{\sqrt{\frac{1}{3}\sum(z_i - \bar{z})^2}}{\sqrt{4}} \approx 2.02$

(2) Bootstrap: Generate B bootstrap samples (or enumerate all 4^4 possibilities), $\hat{\theta}^*_{(1)}, \ldots, \hat{\theta}^*_{(B)} \to \bar{z}^*_{(1)}, \ldots, \bar{z}^*_{(B)}$
Calculate sample standard deviation of $\hat{\theta}^*$'s:
$\sqrt{\frac{1}{B-1}\sum(\hat{\theta}^*_{(i)} - \hat{\theta}^*)^2} \approx \text{SE}(\hat{\theta})$, If you use all n^n, we will get $\sqrt{\frac{n-1}{n}}\,\text{SE}(\hat{\theta})$
Can also studentized our $\hat{\theta}^*_{(i)}$

What if $\hat{\theta}$ is the median?, $\hat{\theta} = \frac{5+3}{2} = 4$ No formula for the standard error. Use the bootstrap

Remark 0.3 (Confidence Intervals (Percentile Confidence Intervals))

Line up our bootstrap statistics from largest to smallest
$\hat{\theta}^*_{(1)}, \ldots, \hat{\theta}^*_{(B)}$
Find $\left(\frac{\alpha}{2}\right) B$ element in this list and $\left(1 - \frac{\alpha}{2}\right) B$ element for the corresponding percentile
Suppose $B = 2000, \alpha = 0.05$
CI: $\left[\hat{\theta}^*_{(50)}, \hat{\theta}^*_{(1951)}\right]$