# Bootstrapping

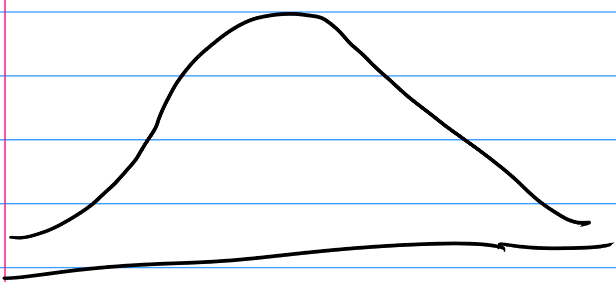
Construct unbiased standard errors and confidence intervals.

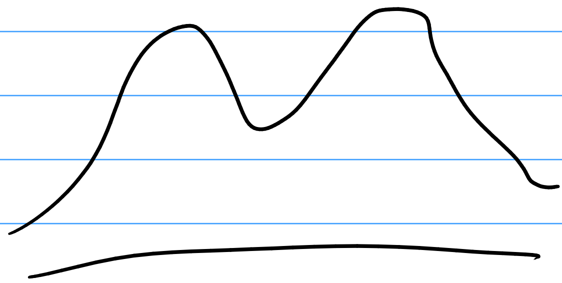
$$X = \{X_1, \ldots, X_n\} \overset{iid}{\sim} F(\theta)$$

$$\hat{\theta} \to \theta \qquad \hat{\theta} \overset{\cdot}{\sim} N(0,1)$$

95% CI

$$\hat{\theta} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}}$$





$X$ contains all the information to estimate $F$, so long as $n \to \infty$

$$\hat{F}_n \longrightarrow F$$

estimator of the distribution function

What is a good estimator for $F$?

$$\hat{F}_n(x) = \begin{cases} 0, & x < X_{(1)} \\ \frac{i}{n} & X_{(i)} < x < X_{(i+1)}, \; i=1,\dots,n-1 \\ 1, & X_{(n)} \leq x \end{cases}$$

Glivenko-Cantelli Theorem

$$\hat{F}_n(x) \longrightarrow F(x)$$

will converge uniformly as $n \to \infty$

Truth

$$F \to X \longrightarrow \hat{F}_n$$

$X \to \hat{\theta}$ What is the sampling distribution of $\hat{\theta}$?

$\hat{F}_n$ is extremely close to $F$

therefore, if we sample with replacement from $X$, that is equivalent to sampling from $\hat{F}_n$.

The Bootstrap Method.

(1). Draw a new sample $X_b^*$ (* Boot sample, b indexed sample) from $X$ of size $n$., with replacement.

(2). Compute $\hat{\theta}_b^* = T(X_b^*)$, and store it.

(3). Repeat steps 1 and 2, B times, to obtain replicates $T^* = (\hat{\theta}_1^*, \hat{\theta}_2^*, \dots, \hat{\theta}_B^*)$

$B = n$

$T^* \sim G_{\hat{\theta}}$

$T^*$ follows the theoretical distribution of $\hat{\theta}$

$$SE(\hat{\theta}) = \sqrt{\frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\theta}_b^* - \bar{T}^*\right)^2}$$

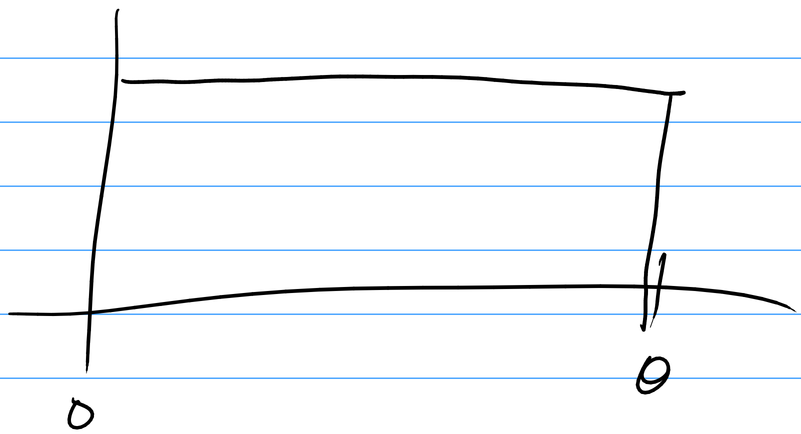
$$\bar{T}^* = \frac{1}{B} \sum_{b=1}^{B} \hat{\theta}_b^*$$

$(1-\alpha)100\%$

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

$$P\left(X \le T^*_{\alpha/2}\right) = \alpha/2$$

$$X \sim U(0, \theta)$$



$X_{(n)} < \theta$