Ex. Data $X_1, \dots, X_n \sim F$

$$T_n = \frac{1}{n}(X_1 + \cdots + X_n)$$

Bootstrap data: $X_1^*, \dots, X_n^* \sim \hat{F}_n$

$$\hat{T}_n = \frac{1}{n}(X_1^* + \cdots + X_n^*)$$

$$\mathbb{V}[T_n^* \mid X_1, \dots, X_n] = \frac{1}{n^2} \sum_{i=1}^{n} \mathbb{V}[X_i^* \mid X_1, \dots, X_n]$$

$$= \frac{1}{n} \mathbb{V}[X_1^* \mid X_1, \dots, X_n]$$

$$\mathbb{E}[X_1^* \mid X_1, \dots, X_n] = \int x \, d\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} X_i = \bar{X}_n$$

$$\mathbb{E}[X_1^{*2} \mid X_1, \dots, X_n] = \int x^2 \, d\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} X_i^2$$

$$\mathbb{V}[T_n^* \mid X_1, \dots, X_n] = \frac{1}{n}\left(\left(\frac{1}{n} \sum_{i=1}^{n} X_i^2 \right) - \left(\frac{1}{n} \sum_{i=1}^{n} X_i \right)^2 \right)$$

$$= \hat{\sigma}_n^2 / n$$

so we dont even need to approximate this quantity.

Bootstrap Pivotal Confidence Interval

$$\theta = T(F) \qquad \hat{\theta}_n = T(\hat{F}_n)$$

$$R_n = \hat{\theta}_n - \theta \qquad H(x) = \mathbb{P}[R_n \le x]$$

$$(a_n, b_n) = \left(\hat{\theta}_n - H^{-1}\left(1-\frac{\alpha}{2}\right), \hat{\theta}_n - H^{-1}\left(\frac{\alpha}{2}\right)\right)$$

$$\mathbb{P}[\theta \in (a_n, b_n)] = \mathbb{P}[a_n - \hat{\theta}_n \le \theta - \hat{\theta}_n \le b_n - \hat{\theta}_n]$$

$$= \mathbb{P}[\hat{\theta}_n - b_n \le \hat{\theta}_n - \theta \le \hat{\theta}_n - a_n]$$

$$= \mathbb{P}\left[H^{-1}\left(\frac{\alpha}{2}\right) \le R_n \le H^{-1}\left(1-\frac{\alpha}{2}\right)\right]$$

$$= H\left(H^{-1}\left(1-\frac{\alpha}{2}\right)\right) - H\left(H^{-1}\left(\frac{\alpha}{2}\right)\right)$$

$$= 1 - \frac{\alpha}{2} - \frac{\alpha}{2} = 1 - \alpha$$

(a_n, b_n) is an exact $1-\alpha$ CI for θ.

But, we do not know $H(x)$

Assume $\hat{\theta}_n^* - \hat{\theta}_n$ is a good approximation for $\hat{\theta}_n - \theta$, where $\hat{\theta}_n^* = g(X_1^*, \ldots, X_n^*)$ is our bootstrap approx of $\hat{\theta}_n$

we can define

$$R_n^* = \hat{\theta}_n^* - \hat{\theta}_n \qquad H^*(x) = \mathbb{P}[R_n^* \le x \mid X_1, \ldots, X_n]$$

Can further approximate

$$\hat{H}_B^*(x) = \frac{1}{B} \sum_{b=1}^{n} \mathbb{1}(R_{n,b}^* \le x)$$

where $R_{n,1}^*, \ldots, R_{n,B}^*$ are iid copies of R_n^*

$$C_n = \left(\hat{\theta}_n - (\hat{H}_B^*)^{-1}(1 - \tfrac{\alpha}{2}),\ \hat{\theta}_n - (\hat{H}_B^*)^{-1}(\tfrac{\alpha}{2}) \right)$$

<u>Theorem</u> Under weak conditions of $T(F)$,

$$\lim_{n \to \infty} \mathbb{P}(T(F) \in C_n) = 1 - \alpha .$$