Sea $X_1, X_2, \dots, X_n$ una m.a. de $X \sim F(x)$ y nos interesa estimar

$$\theta = T(F)$$

$$\Rightarrow \quad \hat{\theta}_n = T(\hat{F}_n) \leftarrow \text{estimador plug-in}$$

pero quiero un estimador por intervalo del $(1-\alpha) \times 100\,\%$ de confianza para θ.

Asumiendo normalidad asintótica

$$\frac{\hat{\theta}_n - \theta}{se(\hat{\theta}_n)} \xrightarrow{D} Z \sim N(0,1)$$

$$L_n = \hat{\theta}_n - z_{\alpha/2}\, se(\hat{\theta}_n)$$
$$U_n = \hat{\theta}_n + z_{\alpha/2}\, se(\hat{\theta}_n), \quad \text{con } se(\hat{\theta}_n) = \sqrt{Var(\hat{\theta}_n)}$$

El problema es calcular $Var(\hat{\theta}_n)$

pero $Var_F(\hat{\theta}_n) \overset{1}{\approx} \hat{Var}_{\hat{F}_n}(\hat{\theta}_n) \overset{2}{\approx} \hat{Var}_{\hat{F}_n}(\hat{\theta}_n)$

similar al plug-in (1) — método Bootstrap (2)

(A) Necesitamos $\hat{F}_n(x) \approx F(x) \quad \forall x \in \mathbb{R}$, en general se cumple si n es suf grande

(B) En general es una muy buena aproximación, solo necesitamos generar muchas simulaciones

Método Bootstrap

Tenemos $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$ la m.a. observada de $X \sim F(x)$

$$1 = X_1^{*1}, X_2^{*1}, \ldots, X_n^{*1} \overset{m.a.}{\sim} \hat{F}_n(x) \longrightarrow \hat{\theta}_n^{*1}$$

$$2 = X_1^{*2}, X_2^{*2}, \ldots, X_n^{*2} \overset{m.a.}{\sim} \hat{F}_n(x) \longrightarrow \hat{\theta}_n^{*2}$$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

$$B = X_1^{*B}, X_2^{*B}, \ldots, X_n^{*B} \overset{m.a.}{\sim} \hat{F}_n(x) \longrightarrow \hat{\theta}_n^{*B}$$

$$\underline{Var_{Boot}(\hat{\theta}_n)} = \frac{1}{B} \sum_{t=1}^{B} \left(\hat{\theta}_n^{*t} - \bar{\hat{\theta}}_n^{*} \right)^2$$

$$\text{con} \quad \bar{\hat{\theta}}_n^{*} = \frac{1}{B} \sum_{t=1}^{B} \hat{\theta}_n^{*t}$$

Existen al menos 3 maneras de obtener intervalos del $(1-\alpha) \times 100\%$ de confianza vía Bootstrap

1) Asumiendo normalidad asintótica

$$\hat{\theta}_n \mp z_{\alpha/2} \, se_{Boot}(\hat{\theta}_n)$$

2) Percentil $\rightarrow \hat{\theta}_n^{*1}, \hat{\theta}_n^{*2}, \dots, \hat{\theta}_n^{*B}$ muestra Bootstrap

calculamos los percentiles

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

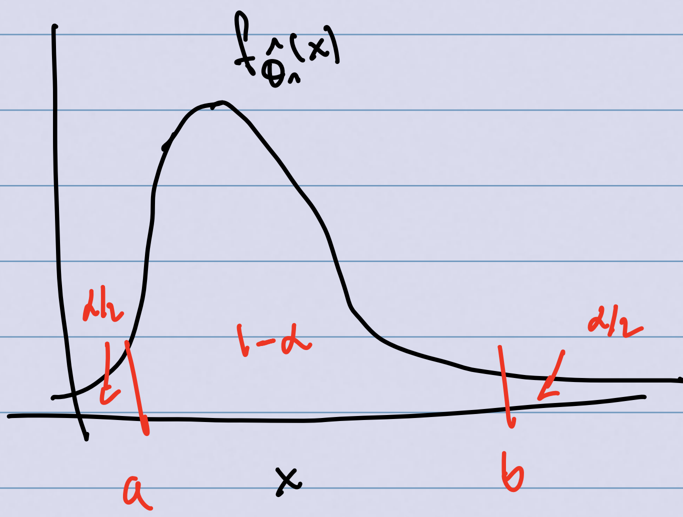
3) Pivotal

$$\left(2\hat{\theta}_n - \hat{\theta}^*_{n,1-\alpha/2}, \ 2\hat{\theta}_n - \hat{\theta}^*_{n,\alpha/2}\right)$$

Justificación

- Normal, es necesario al menos realizar simulaciones para verificar que $\hat{\theta}_n$ se distribuye normal

- Percentil



¿Cómo tomo a y b?

$$a = \hat{\theta}_{n,\alpha/2} \approx \hat{\theta}^*_{n,\alpha/2}$$
$$b = \hat{\theta}_{n,1-\alpha/2} \approx \hat{\theta}^*_{n,1-\alpha/2}$$

En realidad lo que sucede es que la distribución de $\hat{\theta}_n^{*b}$, $b=1,\dots,B$ generados por el Bootstrap debe aproximar

$\hat{\theta}_n^{b}$, $b=1,\dots,B$ generados por el esquema de muestreo (mágico $X_i \sim F(x)$)

- Método pivotal

Vamos a hacer mundo irreal

$$R_n = \hat{\theta}_n - \theta$$

v.a. desconocido pero fijo

$$H(r) = P(R_n \leq r)$$

Un intervalo del $(1-\alpha)\times 100\%$ de confianza para θ está dado por

$$L_n = \hat{\theta}_n - H^{-1}(1-\alpha/2) = \hat{\theta}_n - (\hat{\theta}_{n,1-\alpha/2} - \theta)$$
$$U_n = \hat{\theta}_n - H^{-1}(\alpha/2) = \hat{\theta}_n - (\hat{\theta}_{n,\alpha/2} - \theta)$$

$$P(L_n \leq \theta \leq U_n) = P(L_n - \theta_n \leq \theta - \theta_n \leq U_n - \theta_n)$$
$$= P(\hat{\theta}_n - U_n \leq \hat{\theta}_n - \theta \leq \hat{\theta}_n - L_n)$$

$$= P(\hat{\theta}_n - U_n \leq R_n \leq \hat{\theta}_n - L_n)$$

$$= P(R_n \leq \hat{\theta}_n - L_n) - P(R_n \leq \hat{\theta}_n - U_n)$$

$$= P(R_n \leq \hat{\theta}_n - (\hat{\theta}_n - H^{-1}(1-\alpha/2)))$$

$$- P(R_n \leq \hat{\theta}_n - (\hat{\theta}_n - H^{-1}(\alpha/2)))$$

$$= P(R_n \leq H^{-1}(1-\alpha/2)) - P(R_n \leq H^{-1}(\alpha/2))$$

$$H(H^{-1}(1-\alpha/2)) - H(H^{-1}(\alpha/2))$$

$$= 1-\alpha/2 - \alpha/2 = 1-\alpha$$

Mundo real $\quad R_n = \hat{\theta}_n - \theta$

¿Qué hace Bootstrap?

$$\hat{R}_n^* = \hat{\theta}_n^* - \hat{\theta}_n$$

$$\Rightarrow L_n^* = \hat{\theta}_n - (\hat{\theta}^*_{n,1-\alpha/2} - \hat{\theta}_n) = 2\hat{\theta}_n - \hat{\theta}^*_{n,1-\alpha/2}$$

$$U_n^* = \hat{\theta}_n - (\hat{\theta}^*_{n,\alpha/2} - \hat{\theta}_n) = 2\hat{\theta}_n - \hat{\theta}^*_{n,\alpha/2}$$