Recordando... supongamos que

$X_1, X_2, \ldots, X_n$ es una m.a. de $X \sim F(x)$ y nos interesa hacer inferencia sobre

$$\kappa = \mathbb{E}\left((X-\mu)^3\right) / \sigma^3$$

$\Rightarrow$ sabemos que podemos estimarlo puntualmente vía

$$\hat{\kappa}_n = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu}_n)^3 / \hat{\sigma}_n^3$$

que sería el estimador plug-in (en donde veo $F(x)$ inserto $\hat{F}_n(x)$, pero en este caso lo tengo que hacer en el numerador y denominador)

Pero y si quiero un intervalo de confianza... una opción es asumir normalidad asintótica, i.e.

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

con $se(\hat{\kappa}_n) = \sqrt{Var(\hat{\kappa}_n)}$ $\Rightarrow$ el reto es calcular $\underline{\underline{Var(\hat{\kappa}_n)}}$

además también tendríamos que verificar que $\hat{K}_n$ sigue una distribución normal asintótica

Pero vamos a saltar este "$\underline{\underline{\text{detalle}}}$" y concentrarnos en

$Var(\hat{K}_n)$

De manera general siempre asumimos que los datos que tenemos son una m.a. $X_1, X_2, \dots, X_n$ de $F(x)$ y queremos estimar

$$\theta = T(F)$$

$\Rightarrow \quad \hat{\theta}_n = T(\hat{F}_n)$ pero buscamos una estimación puntual, como ya vimos necesitamos

$$\underline{\underline{Var(\hat{\theta}_n)}}$$

es importante notar que $\underset{\underline{\underline{F}}}{Var}(\hat{\theta}_n)$ depende de $\underline{\underline{F}}$

¿por qué?

$$\hat{\theta}_n = T(\hat{F}_n) = g(X_1, X_2, \dots, X_n)$$

es una función de la m.a. ≡ v.a.i.i.d

$X \sim F(x)$

pero estamos aproximando $F(x)$ usando $\hat{F}_n(x)$

① Podemos aproximar

$$\mathrm{Var}_F(\hat{\theta}_n) \approx \mathrm{Var}_{\hat{F}_n}(\hat{\theta}_n)$$

es a nivel analítico

que puede ser muy difícil

② Aproximo $\mathrm{Var}_{\hat{F}_n}(\hat{\theta}_n) \approx \widehat{\mathrm{Var}}_{\hat{F}_n}(\hat{\theta}_n)$

vía simulación

Entonces tengo

$$\underbrace{\mathrm{Var}_F(\hat{\theta}_n)}_{\text{analítica}} \approx \underbrace{\mathrm{Var}_{\hat{F}_n}(\hat{\theta}_n)}_{\text{analítica}} \approx \underbrace{\widehat{\mathrm{Var}}_{\hat{F}_n}(\hat{\theta}_n)}_{\text{por simulación}}$$

necesitamos

$F(x) \approx \hat{F}_n(x)$, en general $\underline{n}$ suf grande

esta aproximación generalmente es buena, pero necesitamos muchas simulaciones

¿Cómo aproximo?

Ley Débil de los grandes números (LDGN)

Sea $X_1, X_2, \dots, X_n$ una m.a de $X \sim F(x)$, $\mu = \mathbb{E}(X)$

$$\Rightarrow \bar{X}_n \xrightarrow{P} \mu = \mathbb{E}(X) = \int x \, dF(x) = T(F)$$

¿que dice este resultado?

① Un buen estimador de μ es $\bar{X}_n$ ok ok
de hecho $\bar{X}_n$ es consistente para estimar a μ

② Si quisieramos aproximar la integral

$$\mu = \mathbb{E}(X) = \int x \, dF(x)$$

$\Rightarrow$ genero m.a. $X_1, X_2, \dots, X_n$ de $F(x)$
y luego simplemente

$\underline{\underline{\bar{X}_n}}$ ¡para $\underline{\underline{n}}$ muy grande!

Obviamente necesito ser capaz de generar m.a. de $F(x)$

El resultado anterior se puede generalizar para

$$\theta = T(F) = \int h(x) \, dF(x)$$

Si $X_1, X_2, \dots, X_n$ es una m.a. de $X \sim F(x)$

$$\Rightarrow \hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} h(X_i) \xrightarrow{p} \theta$$

$\Rightarrow$ de nuevo para aproximar la integral

$$\theta = T(F) = \int h(x)\,dF(x)$$

puedo generar una m.a. $X_1, X_2, \dots, X_n$ de $X \sim F(x)$ y luego simplemente

Esto de hecho recibe el nombre de <u>integración de Monte Carlo</u> y forma parte de los métodos Monte Carlo.

Vamos a pensar....

Sea $X_1, X_2, \dots X_n$ una m.a. de $X \sim F(x)$ y quiero aproximar $\sigma^2 = Var(X)$

$$\sigma^2 = T(F) = \int (x-\mu)^2\,dF(x)$$

$$\hat{\sigma}_n = \frac{1}{n}\sum_{i=1}^{n} (X_i - \hat{\mu}_n)$$

$\Rightarrow \quad \hat{\sigma}_n \xrightarrow{P} \sigma^2$ $\quad$ n suf grande

Pero ahora ¿cuál es el problema?

$$\theta = T(F)$$

$$\hat{\theta}_n = T(\hat{F}_n) = h(X_1, X_2, \dots, X_n)$$

¿Cómo estimar $Var(\hat{\theta}_n)$?

Luego vamos a pensar en el mundo <u>irreal</u> en el que podemos generar m.a. de $F(x)$

$$1 \quad X_1^1, X_2^1, \dots, X_n^1 \overset{m.a.}{\sim} F(x) \Rightarrow \hat{\theta}_n^1 = h(X_1^1, \dots, X_n^1)$$

$$2 \quad X_1^2, X_2^2, \dots, X_n^2 \overset{m.a.}{\sim} F(x) \Rightarrow \hat{\theta}_n^2 = h(X_1^2, \dots, X_n^2)$$

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

$$B \quad X_1^B, X_2^B, \dots, X_n^B \overset{m.a.}{\sim} F(x) \Rightarrow \hat{\theta}_n^B = h(X_1^B, \dots, X_n^B)$$

$$Var(\hat{\theta}_n) \approx \frac{1}{B} \sum_{i=1}^{B} \left(\hat{\theta}_n^i - \bar{\hat{\theta}}_n \right)^2$$

$$\hat{\Theta}_n = \frac{1}{B}\sum_{i=1}^{B} \hat{\Theta}_n^{i}$$

Esto sería una excelente aproximación a $V_F(\hat{\Theta}_n)$

problema $\rightarrow$ no conocemos $F(x)$, por lo que

pero conocemos $\hat{F}_n(x)$, podemos replicar el proceso anterior simulando de $\hat{F}_n(x)$

Mundo Bootstrap

$$\begin{array}{llll}
1 & X_1^{*1}, X_2^{*1}, \dots, X_n^{*1} \overset{m.a.}{\sim} \hat{F}_n(x) & \Rightarrow & \hat{\Theta}_n^{*1} = h(X_1^{*1}, \dots, X_n^{*1}) \\
2 & X_1^{*2}, X_2^{*2}, \dots, X_n^{*2} \overset{m.a.}{\sim} \hat{F}_n(x) & \Rightarrow & \hat{\Theta}_n^{*2} = h(X_1^{*2}, \dots, X_n^{*2}) \\
\vdots & \vdots & & \vdots \\
B & X_1^{*B}, X_2^{*B}, \dots, X_n^{*B} \overset{m.a.}{\sim} \hat{F}_n(x) & \Rightarrow & \hat{\Theta}_n^{*B} = h(X_1^{*B}, \dots, X_n^{*B})
\end{array}$$

¿Cómo simulo de $\hat{F}_n(x)$?

$\hat{F}_n(x)$ pone masa $1/n$ a cada punto $X_1 \rightarrow X_n$ de la m.a. observada

¿Cómo genero m.a. de $\hat{F}_n(x)$?

$$\Rightarrow \widehat{Var}_{\hat{F}_n}(\hat{\theta}_n) = \frac{1}{B}\sum_{i=1}^{n}\left(\hat{\theta}_n^{*i} - \hat{\theta}_n^{*}\right)^2$$