Ejercicio 5

Sean $X_1, X_2, \dots X_n$ *iid* $\sim F$ Encontrar los estimadores de máxima verosimilitud para los siguientes parámetros y compararlos con los respectivos estimadores por el método de los momentos:

1. p si la distribución es Ber(p)
2. λ si la distribución es $\mathcal{P}(\lambda)$
3. p si la distribución es Geo(p)
4. μ y σ^2 si la distribución es $N(\mu, \sigma^2)$
5. a y b si la distribución es $\mathcal{U}[a, b]$.

5. $X_1, \dots, X_n$ iid $\sim \mathcal{U}[a,b]$ $(x_1, \dots, x_n)$ valores observadores

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in [a,b] \\ 0 & \text{si no} \end{cases}$$

estimadores por máxima verosimilitud para a, b

* función de verosimilitud

$$L(a,b) = \prod_{i=1}^{n} f_X(x_i) = \begin{cases} \dfrac{1}{(b-a)^n} & \text{si todos los } x_i \in [a,b] \\ 0 & \text{si no} \end{cases}$$

* tomamos logaritmo de $L(a,b)$

$$\ell(a,b) = \log(L(a,b)) \underset{x_i \in [a,b]}{=} \log\left(\frac{1}{(b-a)^n}\right) = \log\left((b-a)^{-n}\right)$$
$$= -n \log(b-a)$$

$$\frac{\partial \ell}{\partial a}(a,b) = -n \frac{-1}{b-a} = \frac{n}{b-a} > 0 \quad \text{no tiene puntos críticos}$$

$$\frac{\partial \ell}{\partial b}(a,b) = -n \frac{1}{b-a} = -\frac{n}{b-a} < 0 \quad \text{no tiene puntos críticos}$$

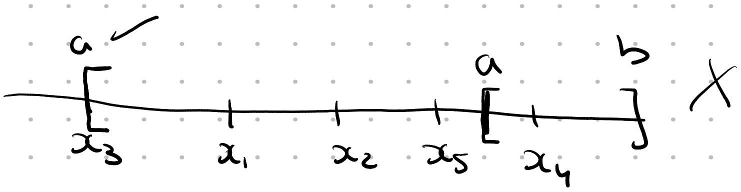
$$L(a,b) = \begin{cases} \dfrac{1}{(b-a)^n} & \text{si } x_i \in [a,b] \\ 0 & \text{si no} \end{cases}$$

$\dfrac{1}{(b-a)^n}$ es más grande mientras más chico sea $b-a$

$b > a$

→ queremos que b sea lo más chico posible y que a sea lo más grande posible

$\mathcal{U}[a,b]$ $(x_1, \dots, x_n)$ valores observados



los valores observados tienen que pertenecer al intervalo $[a,b]$

→ queremos que a sea lo más grande posible y además $a \leq x_i$ para todo $i \in \{1, \dots, n\}$

$$\hat{a} = \min\{x_1, \dots, x_n\}$$

→ queremos que b sea lo más chico posible y además $x_i \leq b$ para todo $i \in \{1, \dots, n\}$

$$\hat{b} = \max\{x_1, \dots, x_n\}$$

Sesgo de un estimador

$\hat{\Theta}_n$ estimador de Θ

$$\begin{aligned} \text{sesgo}(\hat{\Theta}_n) &= E(\hat{\Theta}_n - \Theta) \\ &= E(\hat{\Theta}_n) - E(\Theta) \\ &= E(\hat{\Theta}_n) - \Theta \end{aligned}$$

Decimos que $\hat{\Theta}_n$ es insesgado si $\text{sesgo}(\hat{\Theta}_n) = 0$ o equivalentemente si $E(\hat{\Theta}_n) = \Theta$.

ejemplo: $\sigma_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ y $s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2$ son estimadores de la varianza

σ_n^2 no es insesgado

s_n^2 es insesgado

Ejercicio 9

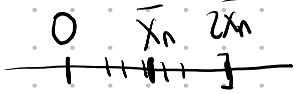
Sean $X_1, \dots, X_n$ $iid \sim \mathcal{U}[0, \theta]$. Interesa estimar el valor de θ.

1. Hallar el estimador de θ por el método de los momentos.
2. Estudiar su sesgo, varianza y error cuadrático medio.
3. Demostrar que el estimador de máxima verosimilitud de θ es X_n^*, el máximo de los valores muestrales.

$X_1, \dots, X_n$ iid $\sim \mathcal{U}[0, \theta]$

① estimador de θ por el método de los momentos

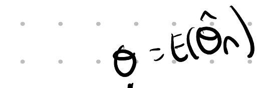
$$\overline{X}_n \xrightarrow{cs} \underbrace{E(X_1)}_{= \frac{\theta}{2}}$$



$$\overline{X}_n \xrightarrow{cs} \frac{\theta}{2}$$

estimador de θ: $\hat{\theta}_n = 2\overline{X}_n$

② * sesgo de $\hat{\theta}_n$



$$\text{sesgo}(\hat{\theta}_n) = E(\hat{\theta}_n) - \theta$$

$$E(\hat{\theta}_n) = E(2\overline{X}_n) = 2E\left(\frac{X_1 + \dots + X_n}{n}\right)$$

$$= \frac{2}{n} E(X_1 + \dots + X_n)$$

$$= \frac{2}{n}\left(\underbrace{E(X_1)}_{\theta/2} + \dots + \underbrace{E(X_n)}_{\theta/2}\right)$$

$$= \frac{\cancel{2}}{\cancel{n}} \cancel{n} \frac{\theta}{\cancel{2}}$$

$$= \theta$$

$$\text{sesgo}(\hat{\theta}_n) = E(\hat{\theta}_n) - \theta = \theta - \theta = 0$$

$\Rightarrow \hat{\theta}_n$ es un estimador insesgado de θ

* $\text{Var}(\hat{\theta}_n)$

$$\text{Var}(\hat{\theta}_n) = \text{Var}(2\overline{X}_n) = 4\text{Var}(\overline{X}_n)$$

$$= 4\,\text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right)$$

$$= \frac{4}{n^2} \text{Var}(X_1 + \dots + X_n)$$

$X \sim U[a,b]$

$Var(X) = \frac{(b-a)^2}{12}$

$$= \frac{4}{n^2}\left(\underbrace{Var(X_1)}_{=\frac{\theta^2}{12}} + \dots + \underbrace{Var(X_n)}_{\frac{\theta^2}{12}}\right)$$ porque las X_i son independientes

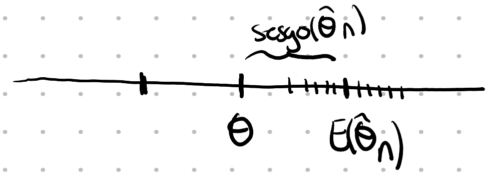
$$= \frac{4}{n^2} n \frac{\theta^2}{12}$$

$$= \frac{\theta^2}{3n}$$

Error cuadrático medio

$\hat{\theta}_n$ estimador de θ

$$ECM(\hat{\theta}_n) = E((\hat{\theta}_n - \theta)^2) = Var(\hat{\theta}_n) + sesgo(\hat{\theta}_n)^2$$



$$ECM(\hat{\theta}_n) = E((\hat{\theta}_n - \theta)^2)$$ $$Var(\hat{\theta}_n) = E(\hat{\theta}_n^2) - E(\hat{\theta}_n)^2$$

$$= E(\hat{\theta}_n^2 - 2\hat{\theta}_n\theta + \theta^2)$$

$$= E(\hat{\theta}_n^2) - 2\theta E(\hat{\theta}_n) + \theta^2$$

$$= E(\hat{\theta}_n^2) - E(\hat{\theta}_n)^2 + E(\hat{\theta}_n)^2 - 2\theta E(\hat{\theta}_n) + \theta^2$$

$$= Var(\hat{\theta}_n) + (\underbrace{E(\hat{\theta}_n) - \theta}_{sesgo(\hat{\theta}_n)})^2$$

$$= Var(\hat{\theta}_n) + sesgo(\hat{\theta}_n)^2$$

* $ECM(\hat{\theta}_n)$

$$ECM(\hat{\theta}_n) = Var(\hat{\theta}_n) + \underbrace{sesgo(\hat{\theta}_n)^2}_{=0}$$

$$= \frac{\theta^2}{3n}$$

Teorema central del límite

$X_1, \dots, X_n$ MAS de X

$E(X) = \mu$, $Var(X) = \sigma^2$

Consideramos $\overline{X_n}$

* $E(\overline{X_n}) = \mu$

$$E(\overline{X_n}) = E\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n} E(X_1 + \dots + X_n)$$

$$= \frac{1}{n}(\underbrace{E(X_1)}_{E(X)=\mu} + \dots + \underbrace{E(X_n)}_{E(X)=\mu})$$

$$= \frac{1}{\cancel{n}} \cancel{n} \mu$$

* $Var(\overline{X_n}) = \frac{\sigma^2}{n}$

$$Var(\overline{X_n}) = Var\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2} Var(X_1 + \dots + X_n)$$

porque las X_i son independientes

$$= \frac{1}{n^2}(\underbrace{Var(X_1)}_{''Var(X)=\sigma^2} + \dots + \underbrace{Var(X_n)}_{''Var(X)=\sigma^2})$$

$$= \frac{1}{n^2} n \sigma^2$$

$$= \frac{1}{n} \sigma^2$$

TCL: para n grande $\overline{X_n}$ tiene distribución normal

$$\overline{X_n} \sim_{aprox} \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$$

$$\frac{(\overline{X_n} - \mu)}{\sqrt{\sigma^2/n}} = \frac{(\overline{X_n} - \mu)}{\sigma} \sqrt{n} \sim_{aprox} \mathcal{N}(0,1)$$

Ejercicio 2

Los resistores de cierto tipo de tienen resistencias que en promedio son de $\mu = 200$ ohms, con una desviación estándar de $\sigma = 10$ ohms. Se toman (al azar) 25 de estos resistores y se conectan (en forma independiente) en un circuito.

1. Calcular la probabilidad (aproximada) de que la resistencia promedio de los 25 resistores este entre 199 y 202 ohms.
2. Calcular la probabilidad (aproximada) de que la resistencia total de los 25 resistores no sea mayor que 5100 ohms.

$X =$ valor de la resistencia de un resistor

$E(X) = 200$

$Var(X) = \sigma^2 = 100$

$X_1, \dots, X_{25}$ MAS de X

① $P(199 \leq \overline{X_{25}} \leq 202)$

$E(\overline{X_{25}}) = E(X) = 200$

$$Var(\overline{X_{25}}) = \frac{Var(X)}{25} = \frac{100}{25} = 4 = 2^2$$

por el TCL: $\overline{X_{25}} \sim_{aprox} \mathcal{N}(200, 2^2)$

$$P(199 \leq \overline{X_{25}} \leq 202) = P\left(\frac{199-200}{2} \leq \frac{\overline{X_{25}}-200}{2} \leq \frac{202-200}{2}\right)$$

$$= P\Big(-\frac{1}{2} \leq \underbrace{\frac{\overline{X_{25}}-200}{2}}_{\sim_{aprox} \mathcal{N}(0,1)} \leq 1\Big)$$

$$\approx \phi(1) - \phi(-\tfrac{1}{2})$$

$$= \phi(1) - (1 - \phi(\tfrac{1}{2}))$$

$$= \phi(1) + \phi(\tfrac{1}{2}) - 1$$

$$= 0{,}8413 + 0{,}6915 - 1$$

$$= 0{,}5328$$

z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944

② $P(X_1 + X_2 + \dots + X_{25} \leq 5100)$

Forma 1:

$$P(X_1 + \dots + X_{25} \leq 5100) = P\left(\frac{X_1 + \dots + X_{25}}{25} \leq \frac{5100}{25} \right)$$

$\overline{X}_{25} \sim_{aprox} \mathcal{N}(200, 2^2)$

$$= P(\overline{X}_{25} \leq 204)$$

$$= P\left(\frac{\overline{X}_{25} - 200}{2} \leq \frac{204 - 200}{2} \right)$$

$$= P\left(\frac{\overline{X}_{25} - 200}{2} \leq 2 \right)$$

$\sim_{aprox} \mathcal{N}(0,1)$

$$\approx \phi(2)$$

$$= 0{,}9772$$

z	0.00	0.01	0.02	0.03	0.04
0.0	0.5000	0.5040	0.5080	0.5120	0.5160
0.1	0.5398	0.5438	0.5478	0.5517	0.5557
0.2	0.5793	0.5832	0.5871	0.5910	0.5948
0.3	0.6179	0.6217	0.6255	0.6293	0.6331
0.4	0.6554	0.6591	0.6628	0.6664	0.6700
0.5	0.6915	0.6950	0.6985	0.7019	0.7054
0.6	0.7257	0.7291	0.7324	0.7357	0.7389
0.7	0.7580	0.7611	0.7642	0.7673	0.7703
0.8	0.7881	0.7910	0.7939	0.7967	0.7995
0.9	0.8159	0.8186	0.8212	0.8238	0.8264
1.0	0.8413	0.8438	0.8461	0.8485	0.8508
1.1	0.8643	0.8665	0.8686	0.8708	0.8729
1.2	0.8849	0.8869	0.8888	0.8907	0.8925
1.3	0.9032	0.9049	0.9066	0.9082	0.9099
1.4	0.9192	0.9207	0.9222	0.9236	0.9251
1.5	0.9332	0.9345	0.9357	0.9370	0.9382
1.6	0.9452	0.9463	0.9474	0.9484	0.9495
1.7	0.9554	0.9564	0.9573	0.9582	0.9591
1.8	0.9641	0.9649	0.9656	0.9664	0.9671
1.9	0.9713	0.9719	0.9726	0.9732	0.9738
2.0	0.9772	0.9778	0.9783	0.9788	0.9793
2.1	0.9821	0.9826	0.9830	0.9834	0.9838
2.2	0.9861	0.9864	0.9868	0.9871	0.9875
2.3	0.9893	0.9896	0.9898	0.9901	0.9904

Forma 2:

$\Gamma\ X_1, \dots, X_n$ MAS X

$E(X) = \mu, \ Var(X) = \sigma^2$

$$E(X_1 + \dots + X_n) = \underbrace{E(X_1)}_{=E(X)} + \dots + \underbrace{E(X_n)}_{=E(X)} = n\mu$$

$$\left(E(X_1 + \dots + X_n) = E(n \overline{X}_n) = n E(\overline{X}_n) = n\mu \right)$$

$$Var(X_1 + \dots + X_n) = Var(X_1) + \dots + Var(X_n) = n\sigma^2$$

$$\left(Var(X_1 + \dots + X_n) = Var(n\bar{X}_n) = n^2 Var(\bar{X}_n) = n^2 \frac{\sigma^2}{n} = n\sigma^2 \right)$$

<u>TCL</u>: si n es grande $X_1 + \dots + X_n$ tiene distribución normal

$$X_1 + \dots + X_n \sim_{aprox} \mathcal{N}(n\mu, n\sigma^2)$$

$$\frac{(X_1 + \dots + X_n) - n\mu}{\sigma} \sqrt{n} \sim_{aprox} \mathcal{N}(0,1)$$

L

$X_1, \dots, X_{25}$ MAS de X

$E(X) = 200$

$Var(X) = 100$

$E(X_1 + \dots + X_{25}) = 5000$

$Var(X_1 + \dots + X_{25}) = 2500 = 50^2$

TCL: $X_1 + \dots + X_{25} \sim_{aprox} \mathcal{N}(5000, 50^2)$

$$P(X_1 + \dots + X_{25} \leq 5100) = P\left(\frac{X_1 + \dots + X_{25} - 5000}{50} \leq \frac{5100 - 5000}{50} \right)$$

$$= P\Big(\underbrace{\frac{X_1 + \dots + X_{25} - 5000}{50}}_{\sim \mathcal{N}(0,1)} \leq 2 \Big)$$

$$\approx \phi(2)$$

$$= 0{,}9772$$