Clase 19: Estimación Puntual

Def: Una muestra aleatoria es una sucesión de v.a iid

$\underbrace{X_1, \dots, X_n}$ teórico o poblacional

$\underbrace{x_1, \dots, x_n}_{\text{empírico}} = X_1(\omega), \dots, X_n(\omega)$ Datos o realizaciones de la v.a.

Muestra se distribuye $f_X(x, \theta)$ θ el parámetro de la dist.

Def: un estadístico es una función de la muestra

$T = T(X_1, \dots, X_n)$. $\quad T(x_1, \dots, x_n) \in \mathbb{R}$

$T = \dfrac{\sum_{i=1}^{n} X_i}{n}$ el promedio o media muestral

Sabemos que estime a la media poblacional.

$\hat{\theta}$ es un estimador de θ.

$\hat{\theta}(x_1, \dots, x_n) = \theta_{obs} \in \mathbb{R}$

Estimamos basado en momentos

LGN: $\overline{X_n} \xrightarrow{P} E(X)$ si $E(X), Var(X) < \infty$

$Var(X) = \underbrace{E(X^2)}_{\frac{1}{n}\sum_{i=1}^{n} X_i^2} - E^2(X)$ $\qquad E(X^2) = \sum_{x \in R} x^2 p_X(x)$

$= \int_{\mathbb{R}} x^2 f_X(x)\,dx$

$\theta > 0$ en ambos

Ej: • $X_1, \dots, X_n$ iid $\mathcal{U}[0,\theta]$

$$\overline{X_n} \to \frac{\theta}{2} \Rightarrow \hat{\theta}_M = 2\overline{X_n}$$

• $X_1, \dots, X_n$ iid $\mathcal{U}[-\theta, \theta]$ $\quad Var(X_i) = \frac{(\theta+\theta)^2}{12} = \frac{\theta^2}{3}$

$\overline{X_n} \to 0$ "El 1er momento no me da info sobre θ"

Me voy al siguiente momento:

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \to E(X^2) = Var(X) + \underbrace{E^2(X)}_{=0}$$

$$\frac{1}{n}\sum X_i^2 \to \frac{\theta^2}{3}$$

$\theta > 0$

$$\hat{\theta}_M = \sqrt{3\frac{1}{n}\sum X_i^2}$$

$E(X^3) = \int x^3 f_X(x)\,dx$

Sean $X_1, \dots, X_n$ variables aleatorias i.i.d con función de probabilidad puntual dada por

k	$p_X(k)$
0	$2\theta^2 - 3\theta + 4/3$
1	$-4\theta^2 + 4\theta - 2/3$
2	$2\theta^2 - \theta + 1/3$

para $1/4 < \theta < 3/4$ hallar un estimador de momentos de θ.

$$E(X_i) = -4\theta^2 + 4\theta - \frac{2}{3} + 2\left(2\theta^2 - \theta + \frac{1}{3}\right) = 2\theta$$

$$\frac{1}{n}\sum_{i=1}^{n} X_i \to 2\theta \Rightarrow \hat{\theta}_M = \frac{\overline{X_n}}{2}$$

Estimador de Maxima Verosimilitud

Muestra $X_1, \dots, X_n \underset{iid}{\sim} X \qquad f_X(x, \theta)$

"$P(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^{n} P(X = x_i)$"

Def funcion de verosimilitud $\prod_{i=1}^{n} f_X(x_i, \theta)$

$$\hat{\theta}_{MV} = \underset{\theta \in \Theta}{\operatorname{argmáx}} \prod_{i=1}^{n} f_X(X_i, \theta)$$

Θ es el conj de posibles valores de θ.

Ej: $X_1, \dots, X_n$ iid $\mathcal{U}([0, \theta])$, $\theta > 0$.

$$\hat{\theta}_{MV} = \underset{\theta > 0}{\operatorname{argmax}} \prod_{i=1}^{n} \frac{1}{\theta} \cdot \mathbb{1}_{[0,\theta]}(X_i) =$$

$$= \underset{\theta > 0}{\operatorname{argmax}} \underbrace{\left(\frac{1}{\theta}\right)^n}_{\geq 0} \prod_{i=1}^{n} \mathbb{1}_{[0,\theta]}(X_i) = \triangle$$

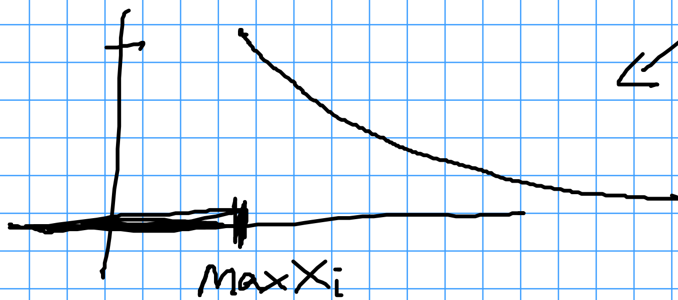
Esto se maximiza cuando $\prod \mathbb{1}_{[0,\theta]}(X_i) = 1$

$\iff \mathbb{1}_{[0,\theta]}(X_i) = 1 \quad \forall i = 1, \dots, n$

$\iff 0 \leq X_i \leq \theta \quad \forall i \iff \theta \geq \max(X_i)$

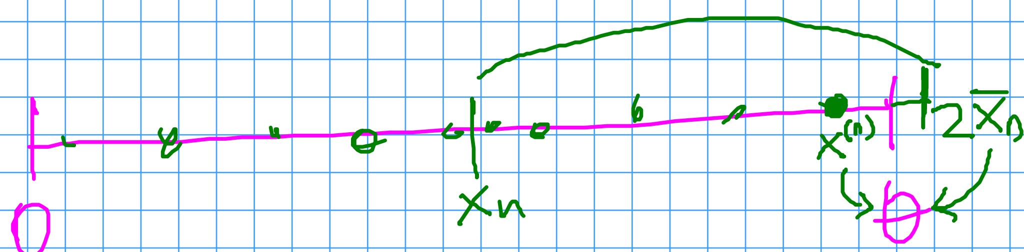
Luego $\prod \mathbb{1}_{[0,\theta]}(X_i) = \mathbb{1}_{[\max X_i, +\infty)}(\theta)$

$$\Rightarrow \quad \triangle = \underset{\theta>0}{\operatorname{argmax}} \underbrace{\left(\frac{1}{\theta}\right)^n \mathbb{1}_{[\max X_i, +\infty)}(\theta)}$$



El maximo se alcanza en $\theta = \max X_i$

$$\Longrightarrow \quad \hat{\theta}_{MV} = \max_{i=1,\dots,n} X_i$$



③ Sean $X_1, \dots, X_n$ variables aleatorias i.i.d con función de densidad dada por

$$f_X(x,\theta) = 2\theta x e^{-\theta x^2} I_{[0,+\infty)}(x) \qquad \theta > 0$$

Hallar $\hat{\theta}_{MV}$ el EMV de θ. Hallar un estimador de θ^2

$$\hat{\theta}_{MV} = \underset{\theta>0}{\operatorname{argmax}} \prod_{i=1}^{n} 2\theta x_i e^{-\theta x_i^2} \underbrace{\mathbb{1}_{[0,+\infty]}(x)}$$

si esto vale 0 $\Rightarrow$ el problema de maxi no tiene sentido

Como log es creciente podemos maximar la función de log verosimilitud.

$$\hat{\theta}_{MV} = \underset{\theta>0}{\operatorname{argmax}} \log\left(\prod_{i=1}^{n} 2\theta x_i e^{-\theta x_i^2}\right)$$

$$= \underset{\theta>0}{\operatorname{argmax}} \sum_{i=1}^{n} \log(2\theta x_i e^{-\theta x_i^2})$$

$$= \underset{\theta>0}{\operatorname{argmax}} \underbrace{\sum_{i=1}^{n} \log(2\theta)}_{n\log(2\theta)} + \underbrace{\sum_{i=1}^{n} \log X_i}_{\text{no dep de } \theta} + \underbrace{\sum_{i=1}^{n} -\theta X_i^2}_{-\theta \sum X_i^2}$$

Voy a derivar e igualar a 0. $\left(\frac{\partial}{\partial\theta}\right)$

$$\frac{2n}{2\theta} + 0 + -\sum X_i^2 = 0$$

$$\frac{n}{\theta} = \sum X_i^2 \Rightarrow \boxed{\hat{\theta}_{MV} = \frac{n}{\sum X_i^2}}$$

$\frac{\partial}{\partial\theta}$

$$-\frac{n}{\theta^2} < 0 \Rightarrow \text{el extremo es máximo.}$$

¿Cómo doy un estimador de θ^2?

Como x^2 es inyectiva en $\mathbb{R}_+$, $(\hat{\theta}_{MV})^2$ es el est. de MV para θ^2. (Como x^2 es cont si θ_{MV} era consistente, $\hat{\theta}^2_{MV}$ tmb)

$$\text{Thm} \begin{cases} X_n \xrightarrow{P} a \in \mathbb{R} \\ g \text{ cont} \quad g(X_n) \xrightarrow{P} g(a) \end{cases}$$

4. Sean $X_1, \dots, X_n$ variables aleatorias i.i.d con función de densidad dada por

$$f_X(x, \theta) = \frac{2x}{\theta^2} I_{(0,\theta)}(x) \qquad \theta > 0$$

Hallar $\hat{\theta}_{MV}$ el EMV de θ.

5. El tiempo en que el banco A demora en atender a un cliente es una variable aleatoria con la densidad dada en el item 3. El tiempo en que el banco B demora en atender a un cliente es una variable aleatoria con la densidad dada en el item 4. Se reliza un estudio en el que se recogen los tiempos de espera de 20 clientes en cada uno de los bancos obtenindose, para el banco A

2.11	2.22	2.1	1.22	3.33	2.33	3.21	3.30	2.01	3.02
3.22	3.22	2.13	3.31	1.34	1.00	1.10	1.20	3.01	1.31

y para el banco B

0.11	4.22	2.1	1.67	2.33	1.33	0.21	4.30	1.20	1.02
2.56	2.22	1.51	4.36	3.42	1.45	2.14	3.21	1.21	0.31

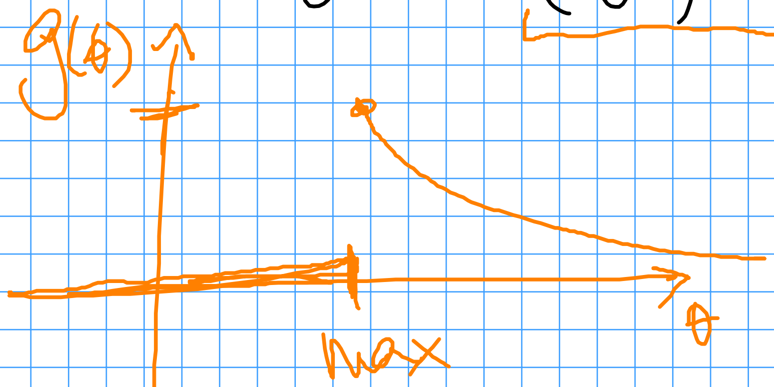
Estimar, para cada uno de los bancos, la probabilidad de que el tiempo de espera sea mayor a 2 minutos.

$$\hat{\theta}_{MV} = \underset{\theta>0}{\arg\max} \prod_{i=1}^{n} \frac{2x_i}{\theta^2} \mathbb{1}_{(0,\theta)}(x_i)$$

$$= \arg\max \left(\frac{2}{\theta^2}\right)^n \prod_{i=1}^{n} \mathbb{1}_{(0,\theta)}(x) \cdot x_i$$

$$= \arg\max \underbrace{\left(\frac{2}{\theta^2}\right)^n \mathbb{1}_{[\max X_i, +\infty)}(\theta)}_{\text{dep de } \theta} \cdot \underbrace{\prod_{i=1}^{n} x_i}_{\text{NO}}$$

$g(\theta)$



Luego $\hat{\theta}_{MV} = X_{(n)} = \max_{i=1,\dots,n} X_i$