Sea $X_1, \ldots, X_n$ una MAS del modelo de Poisson con parámetro $\theta > 0$. Es decir, las v.a. X_i son iid con ley de Poisson dada por

$$\mathbb{P}_\theta(X_i = k) = \frac{e^{-\theta}\theta^k}{k!} \qquad k = 0, 1, 2, \ldots$$

Interesa estimar la función
$g(\theta) = e^{-\theta} = \mathbb{P}_\theta(X_i = 0)$
para lo cual se proponen los siguientes estimadores

$$\hat{g}_1(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{x_i = 0\}}$$

$$\hat{g}_2(x) = a^{\sum_{i=1}^{n} x_i}$$

donde a es una constante positiva.

a) Determine el espacio de parámetros

En este caso, como se trata de una poisson, $\Theta = (0, \infty)$

Recordar que el espacio de parámetros es el conjunto de todos los valores que puede tomar θ.

Buscamos el sesgo de ambos estimadores

Sesgo : $b_\theta(\hat{g}) = \mathbb{E}_\theta(\hat{g}(x)) - g(\theta)$.

Para el $\hat{g}_1$ tenemos que

$$b_\theta(\hat{g}_1) = \mathbb{E}_\theta(\hat{g}_1(x)) - g(\theta)$$

$$= \mathbb{E}_\theta\left(\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{x_i=0\}}\right) - g(\theta)$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_\theta(\mathbb{1}_{\{x_i=0\}}) - \underbrace{g(\theta)}_{*}$$

↓

linealidad de la esperanza
propiedad de probabilidad

* Consideremos $Y_n = \mathbb{1}_{\{x_i=0\}}$ para simplificar notamos que es una v.a. Bernoulli (por ser una indicatriz) y tenemos que

$$\mathbb{P}_\theta(Y_i=1) = \mathbb{P}_\theta(X_i=0) = g(\theta)$$

↳ por enunciado, esta es la función que buscamos estimar

$$\mathbb{P}_\theta(Y_i=0) = \mathbb{P}_\theta(X_i>0) = 1 - g(\theta)$$

↳ "en otro caso"

Continuamos calculando esa esperanza:
$\mathbb{E}$ de una Bernulli

$$\mathbb{E}_\theta(\mathbb{1}_{\{x_i=0\}}) = \mathbb{E}_\theta(Y_i) = \cancel{0\,\mathbb{P}(Y_i=0)} + 1\,\mathbb{P}(Y_i=1)$$
$$= g(\theta)$$

Volviendo al sesgo

$$= \underbrace{\frac{1}{n}\sum_{i=1}^{n} g(\theta)}_{\text{como esto es una constante}} - g(\theta)$$

$$= \cancel{\frac{1}{n}}\,\cancel{n}\cdot g(\theta) - g(\theta) = 0$$

$\Rightarrow$ por lo tanto, $\hat{g}_1$ es insesgado.

Propiedad vista en clase:
El ecm de un estimador insesgado es la varianza. Por lo tanto

$$ecm_\theta(\hat{g}) = \mathbb{E}_\theta((\hat{g}(x) - g(\theta))^2)$$
$$= \mathbb{V}ar_\theta(\hat{g}(\vec{X}))$$

entonces tenemos que

$$\mathbb{V}ar(\hat{g}_1(X)) = \mathbb{V}ar\left(\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\{X_i=0\}}\right)$$

$$= \frac{1}{n^2}\mathbb{V}ar\left(\sum_{i=1}^{n} \mathbb{1}_{\{X_i=0\}}\right)$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \underbrace{\mathbb{V}ar(\mathbb{1}_{\{X_i=0\}})}_{☾} + \underbrace{\sum_{i\neq j} cov(\mathbb{1}_{X_i}, \mathbb{1}_{X_j})}_{\star}\right)$$

★ como las v.a. son iid, esta covarianza es igual a 0

☾ $\mathbb{V}ar(\mathbb{1}_{\{X_i=0\}}) = e^{-\theta}(1-e^{-\theta}) = g(\theta)(1-g(\theta))$ → Varianza de una Bernoulli

Volviendo a la varianza original

$$\mathbb{V}ar(\hat{g}_1(x)) = \frac{1}{n^2}\, n \cdot g(\theta)(1-g(\theta)) =$$

$$= \frac{g(\theta)(1-g(\theta))}{n} = \frac{e^{-\theta}(1-e^{-\theta})}{n}$$

Notar que este ecm decrece (tiende a 0) cuando n tiende a infinito

"Consistencia en media cuadrática"

Repetir para $\hat{g}_2$. Calculamos su sesgo

$$b_\theta(\hat{g}_2) = \mathbb{E}_\theta(\hat{g}_2(x)) - g(\theta)$$

$$= \mathbb{E}_\theta\left(a^{\sum_{i=1}^n x_i}) - g(\theta)\right)$$

Recordar que la suma n v.a. Poisson iid con parámetro θ se distribuye como Poisson de parámetro $n\theta$. Entonces si $S = \sum_{i=1}^n x_i$ tenemos

$$\mathbb{P}_\theta(S=k) = \frac{e^{-n\theta}(n\theta)^k}{k!} \qquad k = 0, 1, 2, \ldots$$

Entonces

esperanza de una poisson

$$\mathbb{E}_\theta(\hat{g}_2(x)) = \mathbb{E}_\theta(a^S) = \sum_{k=0}^{\infty} a^k \frac{e^{-n\theta}(n\theta)^k}{k!}$$

$$= e^{-n\theta} \sum_{k=0}^{\infty} \frac{(an\theta)^k}{k!}$$

serie de la exponencial

$$= e^{-n\theta + an\theta}$$

$$= e^{-\theta n(1-a)}$$

simplemente factorizar

Entonces para que sea insesgado tiene que pasar lo siguiente:

$$e^{-\theta n(1-a)} - e^{-\theta} = 0$$

$$\Rightarrow e^{-\theta n(1-a)} = e^{-\theta}$$

$$\Rightarrow n(1-a) = 1$$

$$\Rightarrow a = 1 - \frac{1}{n}$$

Calculamos ahora el error cuadrático medio de $\hat{g}_2$

$$ecm_\theta(\hat{g}_2(X)) = \mathbb{E}_\theta((a^S - e^{-\theta})^2)$$

$$= \mathbb{E}_\theta(a^{2S} - 2a^S \underbrace{e^{-\theta}}_{cte} + \underbrace{e^{-2\theta}}_{cte})$$

$$= \mathbb{E}_\theta(a^{2S}) - 2e^{-\theta}\mathbb{E}_\theta(a^S) + e^{-2\theta}$$

por linealidad de la esperanza

$$= \underbrace{\mathbb{E}_\theta(a^{2S})} - 2e^{-\theta n(1-a)}e^{-\theta}$$

y ahora calculamos ★ esto de acá

$$\mathbb{E}_\theta(a^{2S}) = \sum_{k=0}^{\infty} a^{2k}\frac{e^{-n\theta}(n\theta)^k}{k!} = e^{-n\theta}\sum_{k=0}^{\infty}\frac{(a^2 n\theta)^k}{k!}$$

$$= e^{-n\theta + a^2 n\theta} = e^{-\theta n(1-a^2)}$$

Luego

$$ecm_\theta(g_2(x)) = e^{-\theta n(1-a^2)} - 2e^{-\theta n(1-a)-\theta} + e^{-2\theta}$$

En el caso $a = 1 - \frac{1}{n}$ obtenemos

$$ecm_\theta(g_2(x)) = e^{-\theta n(1-(1-\frac{1}{n})^2)} - 2e^{-\theta n(\cancel{1}-\cancel{1}+\frac{1}{n})-\theta} + e^{-2\theta}$$

$$= e^{-\theta n(\frac{2}{n}-\frac{1}{n^2})} \underbrace{- 2e^{-\theta n(\frac{1}{n})-\theta}} + e^{-2\theta}$$

$$= e^{-\theta n(\frac{2}{n}-\frac{1}{n^2})} - 2e^{-2\theta} + e^{-2\theta}$$

$$= e^{-\theta \cancel{n}\cdot\frac{2}{\cancel{n}}} e^{-\frac{\theta}{n}} - e^{-2\theta},$$

$$= e^{-2\theta}\left(e^{-\frac{\theta}{n}} - 1\right)$$

Luego, comparamos ambos estimadores

$$\left.\begin{aligned} ecm_\theta(\hat{g}_1) &= \frac{e^{-\theta}(1-e^{-\theta})}{n} \\ ecm_\theta(\hat{g}_2) &= e^{-2\theta}(e^{-\frac{\theta}{n}} - 1) \end{aligned}\right\}$$

Con esto, concluimos que

$$ecm_\theta(g_2) < ecm_\theta(g_1) \quad \forall \theta \in \Theta \text{ si } n \geq 2$$

(puede quedar propuesto)

1) Explicar porqué $g_1(X) = 2\bar{X}_n$ es inadmisible:

Recordando el significado de admisibilidad visto en clase: Un estimador es inadmisible cuando hay alguno que tiene un error cuadrático medio menor para cualquier valor que tome el estimador.

¿Entonces por qué es inadmisible? Simplemente porque podemos construir un estimador con un error cuadrático medio menor, en algún punto de $\theta \in \Theta$ o incluso para cualquier punto

Tenemos que $\mathbb{E}_\theta(\bar{X}_n) = \frac{\theta}{2}$ y $\mathbb{V}ar_\theta(\bar{X}_n) = \frac{\theta^2}{12n}$

Luego $\mathbb{E}(2\bar{X}_n) = \theta$ y $\mathbb{V}ar_\theta(2\bar{X}_n) = \frac{\theta^2}{3n}$

con esto se puede calcular el sesgo con $g(\theta) = \theta$

$$b_\theta(\hat{g}_1) = \mathbb{E}_\theta(\hat{g}(\theta)) - \theta = 0$$

por ser insesgado, el $ecm_\theta(2\bar{X}_n) = \mathbb{V}ar_\theta(2\bar{X}_n) = \frac{\theta^2}{3n}$

ahora definamos $Y_n = máx(X_1, \dots, X_n)$
y el estimador $\hat{g}_2(X_1, \dots, X_n) = Y_n$

determinamos su error cuadrático medio

$$ecm_\theta(\hat{g}_2) = \mathbb{E}_\theta[(g_2(x) - \theta)^2]$$

$$= \int_0^\theta (y-\theta)^2 \frac{ny^{n-1}}{\theta^n} dy$$

→ ahora
por
com

$$= \frac{2\theta^2}{(n+1)(n+2)}$$

por qué

$$\mathbb{P}(Y_n = y) = \begin{cases} \dfrac{ny^{n-1}}{\theta^n} & \text{para } y \in [0,\theta] \\ 0 & \text{"} \end{cases}$$

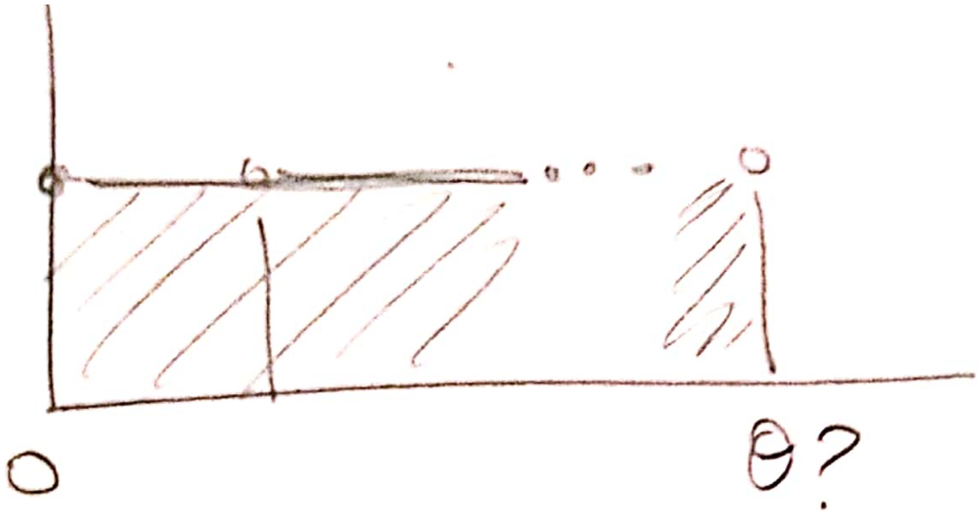
y con esto, tenemos el valor de $ecm(\hat{g}_2)$

con esto, para $n=2$

$$ecm_\theta(\hat{g}_2) = \frac{2\theta^2}{3\cdot 4} = \frac{\theta^2}{6}$$

→ son iguales $\forall \theta \in (0,\infty)$

$$ecm_\theta(\hat{g}_1) = \frac{\theta^2}{3\cdot 2} = \frac{\theta^2}{6}$$



¿Qué pasa para $n \geq 3$

$$ecm(\hat{g}_1) = \frac{\theta^2}{3 \cdot 3} = \frac{\theta^2}{9}$$

$$ecm(\hat{g}_2) = \frac{\cancel{2}^{\,1}\theta^2}{\underbrace{\cancel{(3+1)}}_{\cancel{4}\;2}\underbrace{(3+2)}_{5}} = \frac{\theta^2}{10}$$

$\Rightarrow ecm(\hat{g}_2) < ecm(\hat{g}_1)$
$\forall \theta \in (0, \infty$

$\therefore 2\bar{X}_n$ es inadmisible