

MA34B-Estadística - Pauta Control 1
Profesora: Nancy Lacourly
Auxiliares: Felix Carrasco - Claudio Pareja

P1

- (a) Es natural que para $i = 1, \dots, n$ definamos $X_i = \begin{cases} 1 & \text{si } i \text{ quiere entrar} \\ 0 & \sim \end{cases}$.

Con esto obtenemos una MAS $\{X_i\}_{i=1}^n$ con $X_i \sim \text{Bernoulli}(\theta)$, ie, se tiene la densidad $\theta^{X_i}(1-\theta)^{1-X_i}$. Con θ la proporción a estimar. Además los X_i son i.i.d. pues es un muestreo sin reemplazo. Luego la función de verosimilitud queda:

$$\begin{aligned} f_n(\underline{x}/\theta) &= \prod_{i=1}^n f_{X_i}(X_i/\theta) \\ &= \prod_{i=1}^n \theta^{X_i}(1-\theta)^{1-X_i} \\ &= \theta^{\sum X_i} (1-\theta)^{n-\sum X_i} \end{aligned}$$

Tomando logaritmo

$$L = \ln(f_n) = \ln(\theta) \sum_{i=1}^n X_i + \ln(1-\theta) \left(n - \sum_{i=1}^n X_i \right)$$

y luego derivando se tiene:

$$\begin{aligned} \frac{\partial L}{\partial \theta} &= \frac{1}{\theta} \sum_{i=1}^n X_i + \left(n - \sum_{i=1}^n X_i \right) \frac{-1}{1-\theta} \\ &= \sum_{i=1}^n X_i \left(\frac{1}{\theta} + \frac{-1}{1-\theta} \right) - \frac{n}{1-\theta} \end{aligned}$$

Igualamos a 0 y despejamos:

$$\begin{aligned} \sum_{i=1}^n X_i \left(\frac{1}{\hat{\theta}} + \frac{-1}{1-\hat{\theta}} \right) - \frac{n}{1-\hat{\theta}} &= 0 \\ \sum_{i=1}^n X_i \left(\frac{1}{\hat{\theta}} + \frac{-1}{1-\hat{\theta}} \right) &= \frac{n}{1-\hat{\theta}} \\ \hat{\theta} &= \frac{1}{n} \sum_{i=1}^n X_i \end{aligned}$$

Notamos que la cantidad de veces que X_i es distinta de 0 es exactamente el número de interesados, luego $\hat{\theta}_{MV} = \frac{z}{n}$.

- (b) Notemos que $\mathbb{E}(X_i) = \theta$ (ver tabla de distribuciones), luego

$$\mathbb{E}(\hat{\theta}_{MV}) = \mathbb{E}\left(\frac{z}{n}\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} n\theta = \theta$$

Además $\mathbb{V}(X_i) = \theta(1 - \theta)$ con lo que tenemos:

$$\mathbb{V}(\hat{\theta}_{MV}) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) = \frac{1}{n^2} n\theta(1 - \theta) = \frac{\theta(1 - \theta)}{n}$$

Y así $\lim_{n \rightarrow \infty} \mathbb{V}(\hat{\theta}_{MV}) = 0$.

Con las 2 propiedades anteriores tenemos que el estimador es consistente.

- (c) Calculamos la cota como $I_n(\theta) = nI(\theta)$ pues tenemos las hipótesis necesarias (independencia de la muestra y del dominio).

$$\begin{aligned} I(\theta) &= -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} \left(\ln(f(x/\theta)) \right) \right] = -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} \left(x \ln(\theta) + (1-x) \ln(1-\theta) \right) \right] \\ &= -\mathbb{E} \left[\frac{\partial}{\partial \theta} \left(\frac{x}{\theta} - \frac{1-x}{1-\theta} \right) \right] \\ &= -\mathbb{E} \left[\frac{-x}{\theta^2} - \frac{-1(1-x)}{(1-\theta)^2} \right] \quad (\mathbb{E}(X) = \theta) \\ &= +\frac{1}{\theta} + \frac{1}{1-\theta} = \frac{1}{\theta(1-\theta)} \end{aligned}$$

Y así por la cota de Cramer-Rao tenemos:

$$\mathbb{V}(\hat{\theta}_{MV}) \geq \frac{1}{nI(\theta)} = \frac{\theta(1-\theta)}{n}$$

Como el estimador alcanza la cota, es de mínima varianza.

- (d) Por ser $\Pi(\theta)$ una distribución debe integrar 1 sobre su dominio, como es función de θ y esta es una proporción su dominio es $[0, 1]$, luego:

$$\begin{aligned} \int_0^1 \Pi(\theta) &= 1 \\ \int_0^1 a\theta(1-\theta) &= 1 \\ a \left(\frac{\theta^2}{2} - \frac{\theta^3}{3} \right)_0^1 &= 1 \\ \frac{1}{2} - \frac{1}{3} &= \frac{1}{a} \\ \frac{1}{6} &= \frac{1}{a} \end{aligned}$$

Con lo que $a = 6$.

- (e)

$$\begin{aligned} \xi(\theta/x) &= \frac{\theta^{n\bar{X}}(1-\theta)^{n-n\bar{X}} 6\theta(1-\theta)}{\int_0^1 \theta^{n\bar{X}}(1-\theta)^{n-n\bar{X}} 6\theta(1-\theta) d\theta} \\ &= \frac{\theta^{1+n\bar{X}}(1-\theta)^{n+1-n\bar{X}}}{\int_0^1 \theta^{1+n\bar{X}}(1-\theta)^{n+1-n\bar{X}} d\theta} \end{aligned}$$

Reconocemos dentro de la integral una distribución Beta salvo por constantes (constantes respecto a θ) que las agregamos:

$$\xi(\theta/x) = \frac{\theta^{1+n\bar{X}}(1-\theta)^{n+1-n\bar{X}}}{\int_0^1 \theta^{1+n\bar{X}}(1-\theta)^{n+1-n\bar{X}} d\theta} \frac{\frac{\Gamma(2+n\bar{X}+n+2-n\bar{X}+1)}{\Gamma(2+n\bar{X}+1)\Gamma(n+2-n\bar{X}+1)}}{\frac{\Gamma(2+n\bar{X}+n+2-n\bar{X}+1)}{\Gamma(2+n\bar{X}+1)\Gamma(n+2-n\bar{X}+1)}}$$

Así abajo queda justo la función densidad $Beta(2+n\bar{X}, n+2-n\bar{X})$ integrada sobre todo su dominio por lo que el denominador vale 1, con lo que $\xi(\theta/x) \sim Beta(2+n\bar{X}, n+2-n\bar{X})$.

(f) Como la función de perdida es cuadrática:

$$\hat{\theta}_B = \mathbb{E}(\theta/X) = \frac{n\bar{X} + 2}{n + 4}$$

Que lo obtenemos de la tabla de distribuciones.

(g) Tenemos $\mathbb{E}(\hat{\theta}_B) = \frac{1}{n+4} (\sum \mathbb{E}(X_i) + 2) = \frac{n\theta+2}{n+4}$. Luego $Sesgo(\hat{\theta}_B) = \left| \frac{n\theta+2}{n+4} - \theta \right|$.
Además $\mathbb{V}(\hat{\theta}_B) = \frac{1}{(n+4)^2} \mathbb{V}(\sum X_i + 2) = \frac{1}{(n+4)^2} \mathbb{V}(\sum X_i) = \frac{1}{(n+4)^2} n \mathbb{V}(X_i) = \frac{n\theta(1-\theta)}{(n+4)^2}$.
Con todo lo anterior podemos calcular el ECM:

$$\begin{aligned} ECM(\hat{\theta}_B) &= \mathbb{V}(\hat{\theta}_B) + (sesgo)^2 \\ &= \frac{n\theta(1-\theta)}{(n+4)^2} + \frac{(2-4\theta)^2}{(n+4)^2} \end{aligned}$$

Como $\hat{\theta}_{MV}$ es insesgado se tiene que:

$$ECM(\hat{\theta}_{MV}) = \mathbb{V}(\hat{\theta}_{MV}) = \frac{\theta(1-\theta)}{n}$$

(h) Consideremos el siguiente cociente:

P2

(a) Por definición:

$$\mathbb{E}(x) = \int_{x_0}^{\infty} x \theta x_0^{\theta} x^{-\theta-1} dx = \frac{\theta x_0^{\theta} x^{-\theta+1}}{-\theta+1} \Big|_{x_0}^{\infty} = \frac{\theta x_0}{\theta-1}$$

Además:

$$\mathbb{E}(x^2) = \int_{x_0}^{\infty} x^2 \theta x_0^{\theta} x^{-\theta-1} dx = \frac{\theta x_0^{\theta} x^{-\theta+2}}{-\theta+2} \Big|_{x_0}^{\infty} = \frac{\theta x_0}{\theta-2}$$

Luego:

$$\mathbb{V}(x) = \mathbb{E}(x^2) - (\mathbb{E}(x))^2 = \frac{\theta x_0}{\theta-2} + \frac{\theta^2 x_0^2}{(\theta-1)^2}$$

(b)

$$f(x) = \theta x_0^{\theta} x^{-\theta-1} = \theta x_0^{\theta} \exp(-\ln(x)(\theta+1))$$

Luego por el teorema de Darmois-Koopman podemos considerar $\hat{\theta}_{suf} = \sum_{i=1}^n \ln(x_i)$.

- (c) La función de verosimilitud queda definida como $f = \prod_{i=1}^n \theta x_0^\theta x_i^{-\theta-1} = \theta^n x_0^{n\theta} \prod_{i=1}^n x_i^{-\theta-1}$.

Tomamos logaritmo y derivamos:

$$\begin{aligned} \ln(f) &= \ln(\theta^n x_0^{n\theta} \prod_{i=1}^n x_i^{-\theta-1}) = n\ln(\theta) + n\theta\ln(x_0) - \sum_{i=1}^n (\theta+1)\ln(x_i) \\ \Rightarrow \frac{\partial \ln(f)}{\partial \theta} &= \frac{n}{\theta} + n\ln(x_0) - \sum_{i=1}^n \ln(x_i) \end{aligned}$$

Igualando a 0 tenemos que:

$$\begin{aligned} \frac{n}{\hat{\theta}} + n\ln(x_0) - \sum_{i=1}^n \ln(x_i) &= 0 \\ \frac{n}{\hat{\theta}} &= -n\ln(x_0) + \sum_{i=1}^n \ln(x_i) \\ \hat{\theta} &= \frac{n}{\sum_{i=1}^n \ln(x_i) - n\ln(x_0)} \end{aligned}$$

Y así es claro que es función del estadístico suficiente. Los estimadores de máxima verosimilitud son asintoticamente normales y eficientes, luego su varianza es la cota de Cramer-Rao. Primero veamos la información de Fisher:

$$\begin{aligned} I_n(\theta) &= nI(\theta) = -n\mathbb{E}\left(\frac{\partial^2}{\partial \theta^2} \ln(f)\right) \\ &= -n\mathbb{E}\left(\frac{\partial^2}{\partial \theta^2} \ln(\theta) + \theta\ln(x_0) - (\theta+1)\ln(x)\right) \\ &= -n\mathbb{E}\left(\frac{\partial}{\partial \theta} \theta^{-1} + \ln(x_0) - \ln(x)\right) = -\mathbb{E}(-\theta^{-2}) = n\theta^2 \end{aligned}$$

Luego la varianza asintótica es

$$\frac{\theta^2}{n}.$$

P3

- (a) Utilizamos la definición de esperanza y varianza

$$\mathbb{E}(x) = \int_{\theta}^{\theta+a} \frac{x}{a} dx = \frac{1}{2a} ((\theta+a)^2 - \theta^2) = \frac{\theta^2 + 2\theta a + a^2 - \theta^2}{2a} = \frac{2\theta + a}{2}$$

$$\begin{aligned} \mathbb{E}(x^2) &= \int_{\theta}^{\theta+a} x^2 \frac{1}{a} dx = \frac{1}{3a} ((\theta+a)^3 - \theta^3) \\ &= \frac{1}{3a} (\theta^3 + 3\theta^2 a + 3\theta a + a^3 - \theta^3) \\ &= \frac{3\theta^2 + 3\theta a + a^2}{3} \end{aligned}$$

$$\text{Luego } \mathbb{V}(x) = \mathbb{E}(x^2) - \mathbb{E}(x)^2 = a^2/12.$$

(b) Utilizamos el primer momento:

$$\begin{aligned}\mathbb{E}(x) = \bar{x} &\Leftrightarrow \frac{2\theta + a}{2} = \bar{x} \\ 2\theta &= 2\bar{x} - a \\ \hat{\theta}_{MM} &= \bar{x} - \frac{a}{2}\end{aligned}$$

Ahora veamos la esperanza y varianza:

$$\mathbb{E}(\hat{\theta}_{MM}) = \mathbb{E}(\bar{x}) - \frac{a}{2} = \frac{2\theta + a}{a} - \frac{a}{2} = \theta$$

$$\mathbb{V}(\hat{\theta}_{MM}) = \mathbb{V}(\bar{x}) = \frac{na^2}{12n^2} = \frac{a^2}{12n} \xrightarrow{n \rightarrow \infty} 0$$

Con estas dos tenemos que $\hat{\theta}_{MM}$ es consistente. Al ser consistente tenemos que converge en media cuadrática y luego en probabilidad.

(c) La fn de verosimilitud es $f_n(\underline{x}/\theta) = \left(\frac{1}{a}\right)^n$ que es constante y distinta de 0 si

$$\forall i = 1, \dots, n \quad x_i \in [\theta, \theta + a].$$

De esto obtenemos que $\hat{\theta} \leq \min\{x_i\}$ y $\max\{x_i\} \leq \hat{\theta} + a$. Luego existirá solución si $\min\{x_i\} \leq \max\{x_i\} - a$ y será único cuando $\min\{x_i\} = \max\{x_i\} - a$, en este caso se tendrá $\hat{\theta}_{MV} = \min\{x_i\}$.