Inferencia Estadística - Guia 2

Nicolas Ferrer
e-mail: nicolas.ferrer.747@gmail.com

Junio 2020

1 Ejercicio 1

1.1 Inciso a

A partir de ahora, utilizaremos $M_k(\mathbf{X})$ para referirnos al *momento muestral* de orden k de la variable aleatoria (o secuencia de variables aleatoria) $\mathbf{X}$. Por lo tanto:

$$M_k(\mathbf{X}) = \frac{1}{n}\sum_{i=1}^{n} x_i^k \tag{1}$$

Para un vector de parámetros $\boldsymbol{\theta}$ de longitud k, el estimador del método de momentos $\tilde{\boldsymbol{\theta}}$ debe ser la solución al sistema:

$$\begin{gathered} M_1(X) = E(X) \\ M_2(X) = E(X^2) \\ \vdots \\ M_k(X) = E(X^k) \end{gathered}$$

En el caso de una variable uniforme discreta X con soporte en $\{0, 1, ..., \theta - 1, \theta\}$:

$$M_1(X) = \frac{\tilde{\theta} + 0}{2} \Rightarrow \tilde{\theta} = 2\bar{X}$$

Este estimador no resulta muy coherente, dado que $2\bar{X}$ no necesariamente será un número entero.

1.2 Inciso b

Para los datos provistos:

$$\bar{X} = 0.7 \Rightarrow \tilde{\theta} = 1.4$$

1.3 Inciso c

Si X se distribuye uniforme discreta con límites $(-\theta, \theta)$:

$$E_\theta(X) = \frac{\theta + (-\theta)}{2} = 0$$

Y por lo tanto el estimador de momentos $\tilde{\theta}$ no está definido.

2 Ejercicio 2

2.1 Inciso a

Sea $X \sim Bi(\theta, k)$, tenemos:

$$E(X) = k\theta$$
$$E(X^2) = V(X) + E(X)^2 = k\theta(1-\theta) + (k\theta)^2$$

Para obtener los estimadores de momentos $\tilde{\theta}$ y $\tilde{k}$ resolvemos el sistema:

$$M_1(x) = \tilde{\theta}\tilde{k}$$
$$M_2(x) = \tilde{\theta}\tilde{k}(1-\tilde{\theta}) + (\tilde{\theta}\tilde{k})^2$$

Entonces:

$$\tilde{\theta} = 1 - \frac{(n^{-1}\sum_{i=1}^{n} x_i^2) - \bar{x}^2}{\bar{x}}$$
$$\tilde{k} = \frac{\bar{x}}{\tilde{\theta}}$$

2.2 Inciso b

Notar que por descomposición de varianza podemos escribir:

$$(n^{-1}\sum_{i=1}^{n} x_i^2) - \bar{x}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Entonces, $\tilde{\theta}$ se transforma en:

$$\tilde{\theta} = 1 - \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}{\bar{x}}$$

Usando los datos provistos, obtenemos:

$$\tilde{\theta} = \tilde{k} = -1$$

Lo cual no resulta coherente dado que esperaríamos que ambos sean valores positivos.

3 Ejercicio 3

Para una variable Gamma(α, λ) tenemos:

$$E(X) = \frac{\alpha}{\lambda}$$
$$E(X^2) = \frac{\alpha}{\lambda^2} - \left(\frac{\alpha}{\lambda}\right)^2$$

Por lo tanto, igualando con sus contrapartes muestrales y resolviendo el sistema obtenemos:

$$\tilde{\lambda} = \frac{\bar{x}}{\bar{x}^2 + n^{-1}\sum_{i=1}^{n} x_i^2}$$
$$\tilde{\alpha} = \bar{x}\tilde{\lambda}$$

Para los datos provistos:

$$\tilde{\lambda} = \frac{5}{25 + 144/10} = \frac{50}{394} \approx 0.1269$$
$$\tilde{\alpha} = 5 * \frac{50}{394} \approx 0.6345$$

4 Ejercicio 4

A partir de ahora, utilizaremos $\hat{\theta}_{ML}$ para referirnos al estimador de máxima verosimilitud de θ, mientras que usaremos $\tilde{\theta}$ para el estimador de momentos.

4.1 Inciso a

Dado que una variable $X \sim Pois(\lambda)$ tiene $E(X) = \lambda$, el estimador de momentos de λ está dado por:

$$\tilde{\lambda} = \bar{X}$$

A partir de la función de verosimilitud para un modelo Poisson calculada en el Ejercicio 18 de la Guía 1 podemos calcular la log-verosimilitud como:

$$l(\lambda|\mathbf{x}) = -n\lambda + \sum_{i=1}^{n} x_i \ln(\lambda) - \sum_{i=1}^{n} \ln(x_i!)$$

Para obtener el estimador de máxima verosimilitud, utilizamos la condición de primer orden:

$$\frac{\partial l(\lambda|\mathbf{x})}{\partial \lambda}(\hat{\lambda}_{ML}) = 0$$

De la cual obtenemos:

$$\hat{\lambda}_{ML} = \bar{x}$$

Por lo tanto, para los datos de la muestra provista:

$$\tilde{\lambda} = \hat{\lambda}_{ML} = 2.5$$

4.2 Inciso b

Dado que una variable aleatoria $X \sim Exp(\frac{1}{\lambda})$ tiene $E(X) = \lambda$, el estimador de momentos $\tilde{\lambda}$ está dado por:

$$\tilde{\lambda} = \bar{X}$$

Una vez más, de lo visto en el ejercicio 18 de la Guía 1 podemos calcular la log-verosimilitud para este modelo:

$$l(\lambda|\mathbf{x}) = -n \ln(\lambda) - \frac{1}{\lambda} \sum_{i=1}^{n} x_i$$

De la condición de primer orden para la maximización de la log-verosimilitud obtenemos:

$$\hat{\lambda}_{ML} = \bar{X}$$

Para los datos de la muestra provista:

$$\tilde{\lambda} = \hat{\lambda}_{ML} = 2.5$$

4.3 Inciso c

Para poder obtener el estimador de momentos λ en esta distribución, primero calculamos la esperanza de X utilizando:

$$E(X) = \int_{\lambda}^{\infty} \frac{1}{\lambda} e^{1-\frac{x}{\lambda}} x \; dx$$

Al integrar esta expresión por partes[1], se obtiene $E(X) = 2\lambda$. Por lo tanto, el estimador de momentos de λ será

$$\tilde{\lambda} = \frac{\bar{X}}{2}$$

Por otro lado, la función de verosimilitud y log-verosimilitud para este modelo están dadas por:

$$L(\lambda|\mathbf{x}) = \lambda^{-n} e^{n - \frac{1}{\lambda}\sum_{i=1}^{n} x_i}$$

$$l(\lambda|\mathbf{x}) = -n\ln(\lambda) + n - \frac{1}{\lambda}\sum_{i=1}^{n} x_i$$

De la condición de primer orden para la maximización de la log-verosimilitud obtenemos:

$$\hat{\lambda}_{ML} = \bar{X}$$

Para los valores provistos para esta muestra:

$$\tilde{\lambda} = 1.25$$

$$\hat{\lambda}_{ML} = 2.5$$

4.4 Inciso d

Para una variable $X \sim \Gamma(\lambda, k)$, $E(X) = \frac{k}{\lambda}$. Por lo tanto, el estimador de momentos $\tilde{\lambda}$ con $k = 2$ será:

$$\tilde{\lambda} = \frac{2}{\bar{X}}$$

Para el modelo Gamma, la función de verosimilitud y log-verosimilitud están dadas por (con $k = 2$):

$$L(\lambda|\mathbf{x}) = \lambda^{2n} e^{-\lambda\sum_{i=1}^{n} x_i} \frac{x^n}{\Gamma(2)^n}$$

$$l(\lambda|\mathbf{x}) = 2n\ln(\lambda) - \lambda\sum_{i=1}^{n} x_i + n\left(\ln(x) - \ln(\Gamma(2))\right)$$

De la condición de primer orden para la maximización de la verosimilitud obtenemos:

$$\hat{\lambda}_{ML} = \frac{2}{\bar{X}}$$

5 Ejercicio 5

Tanto el modelo Poisson como Exponencial pertenecen a familias exponenciales, y sus estimadores máximo verosímiles son insesgados. Por lo tanto, podemos utilizar este hecho para calcular fácilmente la Cota de CR para la varianza de los EMV y analizar si esta es alcanzada. Recordar que para familias exponenciales:

$$I(\theta) = -nE\left(\frac{\partial^2}{\partial\theta^2}\log f(X;\theta)\right)$$

Para el caso del modelo **Poisson**:

$$I(\lambda) = -nE\left(\frac{x}{\lambda^2}\right) = \frac{n}{\lambda}$$

Dado que:

$$\mathrm{Var}(\hat{\lambda}) = \frac{\mathrm{Var}(x)}{n} = \frac{\lambda}{n} = I(\theta)^{-1}$$

Con lo cual se alcanza la Cota de CR y $\hat{\lambda}$ es eficiente para el modelo Poisson.

[1] También deberá usar la Regla de L'Hopital, la cual afirma que:

$$\lim_{x\to c}\frac{f(x)}{g(x)} = \lim_{x\to c}\frac{f'(x)}{g'(x)}$$

6 Ejercicio 6

6.1 Inciso a

Para obtener el estimador de momentos para θ, primero calculamos $E(X)$:

$$E(X) = \int_{-1}^{1} x \left(\frac{1 + \theta x}{2} \right) \, dx = \frac{\theta}{3}$$

Entonces, el estimador de momentos $\tilde{\theta}$ debe satisfacer $E_{\tilde{\theta}}(X) = \bar{X}$. Por lo tanto:

$$\tilde{\theta} = 3\bar{X}$$

6.2 Inciso b

Dado que $\tilde{\theta}$ es un estimador insesgado para θ (verifíquelo), tenemos $ECM(\tilde{\theta}, \theta) = \mathrm{Var}(\tilde{\theta})$. Por lo tanto, querríamos una estimación para la varianza asintótica del estimador.

Dado que las X_i se distribuyen independientemente y poseen varianza finita (verifíquelo[2]), se cumplen los supuestos del TCL, por lo tanto:

$$\sqrt{n}(\tilde{\theta} - \theta) \xrightarrow{D} N(0, 9\mathrm{Var}(X))$$

6.3 Inciso c

Dado que $\tilde{\theta}$ es insesgado y su varianza converge a cero a medida que $n \to \infty$, tenemos que:

$$\lim_{n \to \infty} ECM(\tilde{\theta}, \theta) = 0$$

Lo que equivale a afirmar que $\tilde{\theta}$ es un estimador consistente de θ.

6.4 Inciso d

Para la muestra provista, el estimador de momentos $\tilde{\theta}$ toma el valor:

$$\tilde{\theta} = 3 \left(\frac{-0.2}{4} \right) = -0.15$$

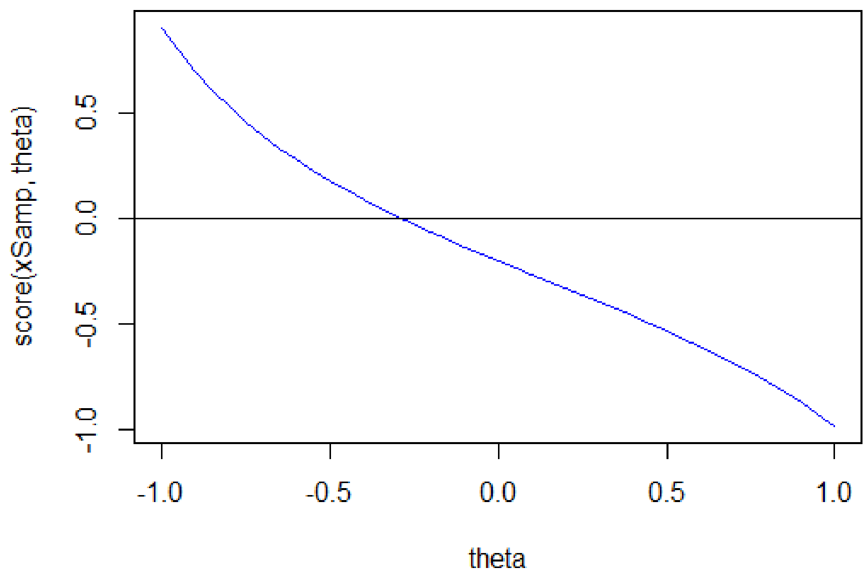
Para obtener el estimador máximo verosímil $\hat{\theta}_{ML}$, debemos igualar a cero la función score (lo que equivale a la condición de primer orden del problema de maximización de la log-verosimilitud):

$$\begin{aligned} S(\hat{\theta}_{ML}) &= \frac{\partial}{\partial \theta} \log f(\mathbf{x}; \hat{\theta}_{ML}) = 0 \\ &= \sum_{i=1}^{n} \frac{x_i}{1 + \hat{\theta}_{ML} x_i} = 0 \end{aligned}$$

A continuación se adjunta un gráfico del valor de esta función para la muestra provista en la consigna.

[2] $\mathrm{Var}(X) = \frac{1}{3} - \frac{\theta^2}{9}$

Figure 1: Función Score(θ) para muestra provista



Puede evidenciarse que la función alcanza valores cercanos a 0 para $\theta \approx 0.3$. Podemos obtener una estimación más precisa utilizando métodos numéricos. Por ejemplo, podemos implementar el método Newton-Raphson para encontrar una única raíz para la función Score utilizando la librería `uniroot` en `R`. El código utilizado estará subido al campus. Utilizando dicho método obtenemos $\hat{\theta}_{ML} \approx 0.285$.

7 Ejercicio 7

Notar que podemos escribir:

$$\psi(\lambda) = P(X = 2|\lambda) = \frac{\lambda^2 e^{-2\lambda}}{2!}$$

Por invarianza del estimador de máxima verosimilitud:

$$\hat{\psi}_{ML}(\lambda) = \psi(\hat{\lambda}_{ML})$$

Sabemos que por el ejercicio 4 que para una variable Poisson, $\hat{\lambda}_{ML} = \bar{X}$. Con los datos provistos, tenemos:

$$\bar{X} = \frac{116}{55} \Rightarrow \hat{\psi}_{ML} \approx 0.2698$$

8 Ejercicio 8

8.1 Inciso a

Sea $l(\mathbf{x}; \mu, \sigma^2)$ la función de log-verosimilitud del modelo normal, los estimadores máximo verosímiles $\hat{\mu}_{ML}, \hat{\sigma^2}_{ML}$ deben satisfacer las condiciones de primer orden:

$$\frac{\partial l(\cdot)}{\partial \mu}(\hat{\mu}_{ML}) = 0$$
$$\frac{\partial l(\cdot)}{\partial \sigma^2}(\hat{\sigma^2}_{ML}) = 0$$

A su vez, es condición necesaria que la matriz hessiana:

$$H(\hat{\theta}_{ML}) = \begin{bmatrix} \frac{\partial^2 l(\cdot)}{\partial \mu^2} & \frac{\partial^2 l(\cdot)}{\partial \mu \partial \sigma^2} \\ \frac{\partial^2 l(\cdot)}{\partial \sigma^2 \partial \mu} & \frac{\partial^2 l(\cdot)}{\partial (\sigma^2)^2} \end{bmatrix}$$

Sea semidefinida negativa. Para el modelo normal, la log-verosimilitud está dada por:

$$l(\mathbf{x}; \mu, \sigma^2) = -\frac{n}{2}\left(\ln(\sigma^2 2\pi)\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

Entonces, calculamos las derivadas a utilizar:

$$\begin{aligned}
\frac{\partial l(\cdot)}{\partial \mu} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) \\
\frac{\partial l(\cdot)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}(x_i - \mu)^2 \\
\frac{\partial^2 l(\cdot)}{\partial \mu^2} &= -\frac{n}{\sigma^2} \\
\frac{\partial^2 l(\cdot)}{\partial (\sigma^2)^2} &= \frac{n}{2\sigma^4} - \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{\sigma^6} \\
\frac{\partial l(\cdot)}{\partial \mu} &= \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) \\
\frac{\partial l(\cdot)}{\partial \mu \partial \sigma^2} &= -\frac{1}{\sigma^4}\sum_{i=1}^{n}(x_i - \mu)
\end{aligned}$$

Entonces, de las condiciones de primer orden obtenemos:

$$\begin{aligned}
\hat{\mu}_{ML} &= \bar{X} \\
\hat{\sigma^2}_{ML} &= \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n}
\end{aligned}$$

Notar que para obtener la segunda expresión utilizamos el principio de invarianza de estimadores maximo verosímiles para reemplazar $\hat{\mu}_{ML} = \bar{X}$. Para corroborar que $H(\hat{\theta}_{ML})$ es semidefinida negativa, analizamos el signo de sus primeros menores principales, D_1 y D_2, cuando los evaluamos en $\hat{\mu}_{ML}$ y $\hat{\sigma^2}_{ML}$:

$$\begin{aligned}
D_1 &= \frac{\partial^2 l(\cdot)}{\partial \mu^2} < 0 \\
D_2 &= \frac{\partial^2 l(\cdot)}{\partial \mu^2}\frac{\partial^2 l(\cdot)}{\partial (\sigma^2)^2} - 2 * \frac{\partial l(\cdot)}{\partial \mu \partial \sigma^2} > 0
\end{aligned}$$

Dado que su primer menor principal es negativo y el segundo positivo, la matriz es definida negativa y se cumplen las condiciones necesarias y suficientes para la maximización de la verosimilitud.

8.2 Inciso b

Con los valores provistos para la muestra obtenemos la estimación (recuerde utilizar la descomposición de la varianza muestral):

$$\begin{aligned}
\hat{\mu}_{ML} &= 1.7 \\
\hat{\sigma^2}_{ML} &= 5.21
\end{aligned}$$

8.3 Inciso c

Empecemos trabajando con el estimador insesgado S^2. Dado que trabajamos con una distribución normal, podemos usar de lo visto en clase que la varianza de S^2 (que al ser insesgado es igual a su ECM) es:

$$ECM(S^2, \sigma^2) = \text{Var}(S^2) = \frac{2\sigma^4}{n-1}$$

El EMV de σ^2 está dado por:

$$\hat{\sigma}^2_{ML} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})^2 = \frac{n-1}{n}S^2$$

Utilizando su relación con S^2, obtenemos:

$$E(\hat{\sigma}^2_{ML}) = \frac{n-1}{n}\sigma^2$$
$$\text{Var}(\hat{\sigma}^2_{ML}) = \frac{2(n-1)}{n^2}\sigma^4$$

Por lo tanto, utilizando la descomposición del error medio cuadrático:

$$ECM(\hat{\sigma}^2_{ML}, \sigma^2) = \frac{\sigma^4}{n^2} + \frac{2(n-1)\sigma^4}{n^2} = \frac{2n-1}{n^2}\sigma^4$$

Se puede ver que $ECM(S^2, \sigma^2) > ECM(\hat{\sigma}^2_{ML}, \sigma^2)$ se cumple si y sólo si:

$$\frac{2}{n-1} > \frac{2}{n} - \frac{1}{n^2}$$

Lo cual vale dado que:

$$\frac{2}{n-1} > \frac{2}{n} > \frac{2}{n} - \frac{1}{n^2}$$

Para cualquier $n > 1$. Entonces, el estimador insesgado posee un riesgo más alto (es decir, es menos eficiente en términos de ECM).

8.4 Inciso d

Para el caso multiparámetro en familias exponenciales, el elemento ij de la matriz de información $\mathbf{I}(\boldsymbol{\theta})$ está dado por:

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} = -E\left(\frac{\partial^2}{\partial\theta_i\theta_j}\log f(\mathbf{X}, \theta)\right)$$

A partir de los cálculos del inciso 8.a podemos ver que para este modelo:

$$\mathbf{I}(\boldsymbol{\theta}) = \begin{bmatrix} n/\sigma^2 & 0 \\ 0 & n/2\sigma^4 \end{bmatrix}$$

Dado que la matriz de información es diagonal, su inversa se obtiene de invertir los elementos en su diagonal principal. Por lo tanto, la cota de Crámer Rao será simplemente:

$$V(\boldsymbol{\theta}) \geq \mathbf{I}(\boldsymbol{\theta})^{-1} = \begin{bmatrix} \sigma^2/n & 0 \\ 0 & 2\sigma^4/n \end{bmatrix}$$

8.5 Inciso e

Sabemos que el estimador máximo verosímil de la media está dado por $\bar{X}$, y que su varianza es $\text{Var}(\bar{X}) = \sigma^2/n$. Comparando este valor con el primer elemento de la Cota de Crámer Rao, podemos evidenciar que el estimador máximo verosímil de la media es eficiente.

En contraste, para el estimador máximo verosímil de σ^2 sólo será eficiente si la media es conocida, dado que de lo contrario su varianza está dada por:

$$\text{Var}(\hat{\sigma}^2) = \frac{2(n-1)}{n^2}\sigma^4$$

Lo cual se puede ver del inciso 8.3

8.6 Inciso f

Utilizando lo visto en clase para estimadores máximo verosímiles[3], sabemos que para un modelo parametrizado por un vector $\boldsymbol{\theta}$ de dimensión $p = 2$:

$$(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\mathbf{I}(\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})' \underset{F}{\rightarrow} \chi^2_2$$

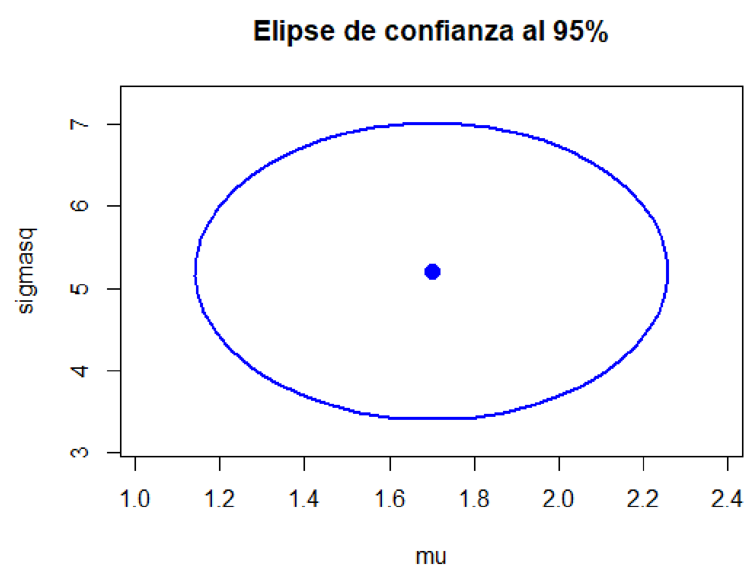
En este caso tenemos $\boldsymbol{\theta} = [\mu, \sigma^2]$. Para la matriz de información encontrada en el inciso 8.d:

$$\frac{n(\hat{\mu} - \mu)^2}{\hat{\sigma}^2} + \frac{n(\hat{\sigma}^2 - \sigma^2)^2}{2\hat{\sigma}^4} \underset{F}{\rightarrow} \chi^2_2$$

Por lo tanto, para los valores obtenidos en el inciso 8.b para $\hat{\mu}$ y $\hat{\sigma}^2$, definimos la elipse de confianza al 95% como:

$$R_{0.95} = \left\{(x, y) : \frac{100(1.7 - x)^2}{5.21} + \frac{100(5.21 - y)^2}{2(5.21)^2} \leq c\right\}$$

Donde c es el cuantil $1 - \alpha$ de una variable chi-cuadrado con dos grados de libertad ($c = 5.99$). Si graficamos en `R`:



[3] Ver por ejemplo página 22 de Slide 5.

9 Ejercicio 9

Sea $\{X_i\}_{i=1}^n$ una muestra aleatoria de una población Uniforme (continua) con soporte en $[0, \theta]$, nos interesa encontrar $\hat{\theta}_{ML}$. Sabemos que para cualquier X_i:

$$f(x;\theta) = \begin{cases} 1/\theta & \text{si } 0 \leq x \leq \theta \\ 0 & \text{en otro caso.} \end{cases}$$

Por lo cual la función de verosimilitud será:

$$f(\mathbf{x}, \theta) = \prod_{i=1}^{n} \frac{1}{\theta} \text{ para } x_i \in [0, \theta]$$

Al derivar esta expresión[4] en búsqueda de una condición de primer orden obtenemos:

$$\frac{\partial f(\mathbf{x}, \theta)}{\partial \theta} = -\frac{n}{\theta^{n-1}} < 0 \; \forall \; n, \theta > 0$$

Por lo tanto, la verosimilitud es siempre decreciente como función de θ para un dado n. Por lo tanto, querríamos elegir el valor de θ más bajo posible como su estimador máximo verosímil. No obstante, la función de verosimilitud sólo está definida para $x_i : x_i < \theta$. Por lo tanto, el estimador máximo verosímil será $\hat{\theta}_{ML} = X_{(n)} = \max\{X_i\}_{i=1}^n$.

10 Ejercicio 10

Dado un estimador W para θ, definimos el error cuadrático medio asociado a W como:

$$\begin{aligned}
ECM(W, \theta) &= E[(W - \theta)^2] \\
&= E[(W - E(W) + E(W) - \theta)^2] \\
&= E[(W - E(W))^2 + 2(W - E(W))(E(W) - \theta) + (E(W) - \theta)^2] \\
&= E[(W - E(W))^2] + 2E[(W - \underbrace{E(W)}_{\text{constante}})\underbrace{(E(W) - \theta)}_{\text{constante}}] + E[(E(W) - \theta)^2] \\
&= E[(W - E(W))^2 + 2(E(W) - E(W))(E(W) - \theta) + (E(W) - \theta)^2] \\
&= \underbrace{E[(W - E(W))^2]}_{=\mathrm{Var}(W)} + \underbrace{E[(E(W) - \theta)^2]}_{Sesgo_\theta^2(W)}
\end{aligned}$$

Que era lo que queríamos demostrar.

11 Ejercicio 11

Para la ver que $a + bW$ es insesgado para $a + b\theta$:

$$E[a + bW] = a + bE[W] = a + b\theta$$

En el caso de W^2:

$$E[W^2] = \mathrm{Var}(W) + E[W]^2 = \mathrm{Var}(W) + \theta^2$$

Entonces,

$$\mathrm{Sesgo}(E[W^2], \theta^2) = E[W^2] - \theta^2 = \mathrm{Var}(W)$$

[4] Generalmente trabajamos con la *log-verosimilitud* porque facilita las cuentas, en este caso, no es necesario.

12 Ejercicio 12

Para cualquier estimador W de θ podemos descomponer su error cuadrático medio $R(W,\theta)$ como:

$$R(W,\theta) = \text{Sesgo}_\theta^2(W) + \text{Var}(W)$$

En este ejercicio trabajaremos con una secuencia de variables $\{X_i\} \overset{iid}{\sim} Bi(p)$ con un tamaño muestral dado. La suma de variables aleatorias binomiales es en sí una variable aleatoria binomial con un tamaño de experimento igual a la suma del tamaño de las muestras asociadas a cada variable. Por lo tanto, en este ejercicio, n se refiere a la suma del tamaño de cada uno de los experimentos binomiales X_i, y $\bar{X}$ a la proporción de éxitos en la suma de experimentos.

Para el estimador máximo verosímil $\hat{p}_1 = \bar{X}$:

$$\begin{aligned}
\text{Sesgo}_p(\hat{p}_1) &= E(\bar{X}) - p = 0 \\
\text{Var}(\hat{p}_1) &= \frac{p(1-p)}{n} \\
R(\hat{p}_1, p) &= \frac{p(1-p)}{n}
\end{aligned}$$

Para el estimador bayesiano $\hat{p}_1 = (\sqrt{n/4} + n\bar{X})/(n+\sqrt{n})$:

$$\begin{aligned}
\text{Sesgo}_p(\hat{p}_2) &= \frac{\sqrt{n/4} + nE(\bar{X})}{n+\sqrt{n}} - p \\
&= \frac{\sqrt{n}(0.5-p)}{n+\sqrt{n}} \\
\text{Var}(\hat{p}_2) &= \left(\frac{n}{n+\sqrt{n}}\right)^2 \text{Var}(\bar{X}) \\
&= \frac{np(1-p)}{(n+\sqrt{n})^2} \\
R(\hat{p}_2, p) &= \frac{n(0.5-p)^2}{(n+\sqrt{n})^2} + \frac{np(1-p)}{(n+\sqrt{n})^2} \\
&= \frac{n}{(n+\sqrt{n})^2}\left((0.5-p)^2 + p(1-p)\right)
\end{aligned}$$

12.1 Inciso b

Notar que para $p = 0.5$:

$$\begin{aligned}
R(\hat{p}_1, 0.5) - R(\hat{p}_2, 0.5) &= 0.25\left(\frac{1}{n} - \frac{n}{(n+\sqrt{n})^2}\right) \\
&= 0.25\frac{2\sqrt{n}+1}{(n+\sqrt{n})^2} > 0 \text{ para todo } n > 0
\end{aligned}$$

Es decir, para un n fijo, siempre podemos encontrar un valor de p arbitrariamente cerca de 0.5 tal que $R(\hat{p}_1, p) > R(\hat{p}_2, p)$. Dicho límite queda caracterizado por la expresión $R(\hat{p}_1, p) = R(\hat{p}_2, p)$

12.2 Inciso c

Dado que el sesgo del estimador $\hat{p}_2$ tiende a cero a medida que n tiende a infinito y su varianza es siempre menor, lo preferiríamos en caso de que n sea lo suficientemente grande. Si n es pequeño, el sesgo puede ser significativo (particularmente si p se encuentra muy lejos de 0.5), por lo cual preferiríamos el estimador $\hat{p}_1$.

13 Ejercicio 13

13.1 Inciso a

Sabemos que S^2 es un estimador insesgado de la varianza[5]. A su vez, sabemos que para muestras provenientes de una población normal se cumple[6]:

$$(n-1)\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$$

Por lo tanto:

$$\text{Var}(S^2) = \left(\frac{\sigma^2}{n-1}\right)^2 \underbrace{\text{Var}(\chi^2_{n-1})}_{=2(n-1)} = \frac{2\sigma^4}{n-1}$$

Entonces:

$$\begin{aligned}
\text{Sesgo}_{\sigma^2}(bS^2) &= (b-1)\sigma^2 \\
\text{Var}(bS^2) &= b^2\frac{2\sigma^4}{n-1} \\
R(bS^2, \sigma^2) &= (b-1)^2\sigma^4 + b^2\frac{2\sigma^4}{n-1} \\
&= \sigma^4\left((b-1)^2 + \frac{2b^2}{n-1}\right)
\end{aligned}$$

13.2 Inciso b

Quiero elegir el b que minimice $R(bS^2, \sigma^2)$. Usando la condición de primer orden:

$$\begin{aligned}
\frac{\partial R(bS^2, \sigma^2)}{\partial b} &= 0 \\
\sigma^4\left(2(b-1) + \frac{4b}{n-1}\right) &= 0 \Rightarrow b^* = \frac{n-1}{n+1}
\end{aligned}$$

14 Ejercicio 14

14.1 Inciso a

Al buscar la condición de primer orden del problema de maximización de verosimilitud, encontramos que no podemos obtener una solución analítica para el estimador máximo verosímil de θ:

$$\begin{aligned}
S(\hat{\theta}_{ML}) &= \frac{\partial}{\partial\theta}\log f(\mathbf{x}; \hat{\theta}_{ML}) = 0 \\
&= 2\sum_{i=1}^{n}\frac{(x_i - \hat{\theta}_{ML})}{1 + (x_i - \hat{\theta}_{ML})^2} = 0
\end{aligned}$$

[5] Ejercicio 8.c de Guía 1.

[6] Ver página 32 de Slide 1.

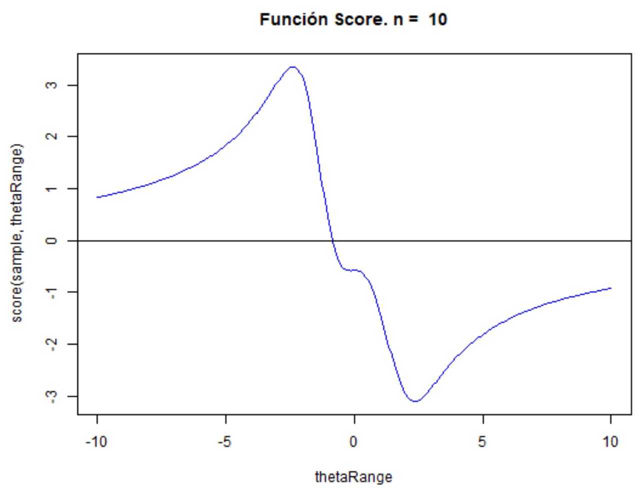
14.2 Inciso b

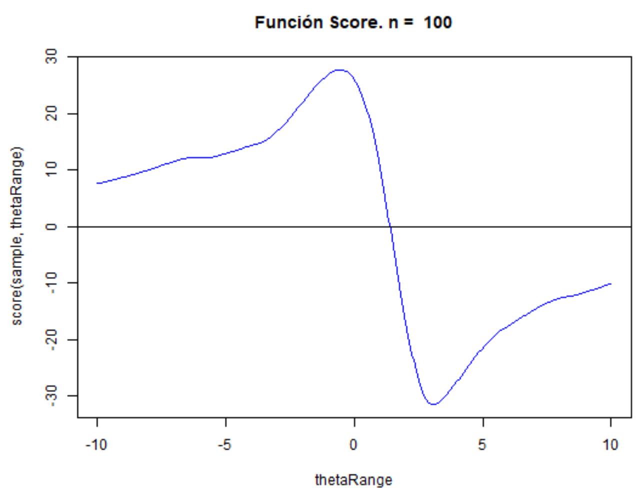
Ya encontramos la función $S(\theta)$ en el inciso anterior. Para obtener la función hessiano $H(\theta)$ simplemente derivamos una vez más.

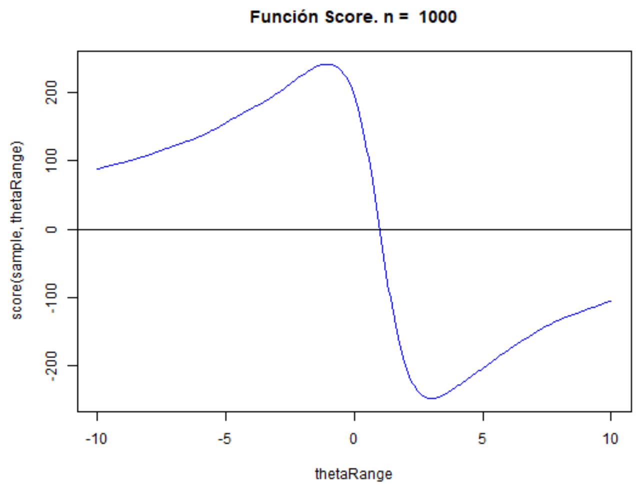
$$
\begin{aligned}
H(\theta) &= \frac{\partial^2}{\partial \theta^2} \log f(\mathbf{x}; \hat{\theta}_{ML}) = 0 \\
&= 2 \sum_{i=1}^{n} \frac{(x_i - \theta)^2 - 1}{\left(1 + (x_i - \theta)^2\right)^2}
\end{aligned}
$$

14.3 Inciso c

A continuación se presentan gráficos de la función Score para cada una de las poblaciones Cauchy creadas utilizando `R`.







Puede evidenciarse que la función muestra un comportamiento mas "suave" a medida que la población crece. Al igual que en el ejercicio anterior, utilizamos la función `uniroot` para obtener los estimadores MVE por métodos numéricos. Los códigos usados serán subidos al campus. Los valores obtenidos fueron:

$$\hat{\theta}_{n=10} = -0.847$$
$$\hat{\theta}_{n=100} = 1.381$$
$$\hat{\theta}_{n=1000} = 0.98$$

Notar que el valor estimado se acerca al valor verdadero a medida que n crece.

15 Ejercicio 15

15.1 Inciso a

Estimador de momentos

Dado $\{X\}_{i=1}^{n} \sim U(0, \theta)$, sabemos por el ejercicio 1.a de esta guía que el estimador de momentos de θ está dado por $\tilde{\theta} = 2\bar{X}$

Para este estimador:

$$\text{Sesgo}^2_\theta(\tilde{\theta}) = \left[E(2\bar{X}) - \theta\right]^2 = 0$$
$$\text{Var}(\tilde{\theta}) = \frac{4Var(X)}{n} = \frac{\theta^2}{3n}$$

Por lo tanto:

$$ECM(\tilde{\theta}, \theta) = \frac{\theta^2}{3n}$$

Estimador alternativo

Sea $X_{(n)} = \max\{X_i\}_{i=1}^{n}$, podemos aplicar los pasos del inciso 11.b de la Guía 1, para obtener su función de densidad, la cual necesitaremos para calcular la esperanza y varianzas de este estimador. Entonces tenemos:

$$\begin{aligned}
f_{X_{(n)}} &= n\frac{x^{n-1}}{\theta^n} \\
E(X_{(n)}) &= \frac{n}{n+1}\theta \\
\mathrm{Var}(X_{(n)}) &= \frac{2n\theta^2}{(n+2)(n+1)^2}
\end{aligned}$$

Llamemos $\hat{\theta} = \frac{n+1}{n}X_{(n)}$, entonces:

$$\begin{aligned}
\mathrm{Sesgo}^2_\theta(\hat{\theta}) &= \left[E(\hat{\theta}) - \theta\right]^2 = 0 \\
\mathrm{Var}(\hat{\theta}) &= \left(\frac{n+1}{n}\right)^2 \mathrm{Var}(X_{(n)}) = \frac{2\theta^2}{n(n+2)}
\end{aligned}$$

Por lo tanto:

$$ECM(\hat{\theta}, \theta) = \frac{2\theta^2}{n(n+2)}$$

15.2 Inciso b

Ambos estimadores son insesgados y su varianza tiende a cero a medida que n tiende a infinito, con lo cual son consistentes.

15.3 Inciso c

Dado que ambos estimadores son insesgados, su eficiencia relativa está dada por:

$$e(\tilde{\theta}, \hat{\theta}) = \frac{\mathrm{Var}(\tilde{\theta})}{\mathrm{Var}(\hat{\theta})} = \frac{\theta^2}{3n}\frac{n(n+2)}{2\theta^2} = \frac{n+2}{6}$$

15.4 Inciso d

Dado que ambos estimadores son insesgados, de acuerdo al criterio de eficiencia relativa nos quedaríamos con el estimador $\hat{\theta}$ siempre que $n > 4$ dado que posee menor varianza.

16 Ejercicio 16

16.1 Inciso a

Dado que $E(X) = \theta$, estamos trabajando con una variable X tal que:

$$\begin{aligned}
f_X(x;\theta) &= \frac{e^{-\frac{x}{\theta}}}{\theta} \\
F_X(x) &= 1 - e^{-\frac{x}{\theta}}
\end{aligned}$$

Para el EMV $\hat{\theta}$ tenemos:

$$\begin{aligned}
\mathrm{Sesgo}^2_\theta(\hat{\theta}) &= \left[E(\bar{X}) - \theta\right]^2 = 0 \\
\mathrm{Var}(\tilde{\theta}) &= \frac{\mathrm{Var}(\bar{X})}{n} = \frac{\theta^2}{n}
\end{aligned}$$

Por lo tanto:

$$ECM(\hat{\theta}, \theta) = \frac{\theta^2}{n}$$

Siguiendo los pasos del ejercicio 11 de la Guía 1 para el mínimo de variables aleatorias, podemos llegar a la siguiente función de densidad para $X_{(1)}$:

$$f_{X_{(1)}}(x; \theta) = \frac{n}{\theta} e^{-\frac{n}{\theta} x}$$

Comparando esta función de densidad con la correspondiente a X, podemos ver que $X_{(1)} \sim \text{Exp}(\theta/n)$ con $E(X_{(1)}) = \frac{\theta}{n}$ y $\text{Var}(X_{(1)}) = \left(\frac{\theta}{n}\right)^2$. Por lo tanto, para el estimador $W = nX_{(1)}$:

$$Sesgo_\theta^2(W) = \left[nE(X_{(1)}) - \theta\right]^2 = 0$$
$$\text{Var}(W) = n^2 \text{Var}(X_{(1)}) = \theta^2$$

Por lo tanto:

$$ECM(W, \theta) = \theta^2$$

Ambos estimadores son insesgados, pero la varianza de $\hat{\theta}$ es menor para cualquier $n > 1$, por lo cual es preferible utilizar el primer estimador.

16.2 Inciso b

En este caso, la varianza de W no converge a cero a medida que crece la población, con lo cual sólo $\hat{\theta}$ es consistente.

17 Ejercicio 17

17.1 Inciso a

Utilizando que el modelo Beta pertenece a la familia exponencial (ver ejercicio 4.g de Guía 1), podemos utilizar la siguiente definición para encontrar la Información de Fisher:

$$\begin{aligned} I(\theta) &= ni(\theta) \\ &= -nE\left(\frac{\partial^2}{\partial \theta^2} \log f(X; \theta)\right) \\ &= -nE\left(-\frac{1}{\theta^2}\right) \\ &= \frac{n}{\theta^2} \end{aligned}$$

17.2 Inciso b

Para este modelo, la verosimilitud y log-verosimilitud están dadas por:

$$L(\theta, \mathbf{x}) = \theta^n \left(\prod_{i=1}^n x_i\right)^{(\theta-1)}$$

$$l(\theta, \mathbf{x}) = n \log(\theta) + (\theta - 1) \sum_{i=1}^n \log x_i$$

Por lo tanto, la función score será:

$$S(\theta|\mathbf{x}) = \frac{\partial}{\partial \theta} l(\theta|\mathbf{x}) = \frac{n}{\theta} + \sum_{i=1}^n \log x_i$$

Igualando $S(\theta|\mathbf{x}) = 0$, encontramos que el estimador máximo verosimil $\hat{\theta}$ es:

$$\hat{\theta} = -\frac{n}{\sum_{i=1}^{n} \log x_i}$$

Si este estimador fuese insesgado, podríamos utilizar el hecho de que es función de un estadístico suficiente para afirmar que es UMVUE (por Lehman-Scheffé). No obstante, se puede demostrar que dicha propiedad no se cumple[7].

17.3 Inciso c

Por normalidad asintótica del estimador máximo verosímil:

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} N\left(0, i(\theta)^{-1}\right)$$

Entonces, para cualquier estimador máximo verosímil podemos construir un intervalo de confianza aproximado para θ de significancia α como:

$$P\left(\hat{\theta} - \frac{Z_{1-\alpha/2}}{\sqrt{ni(\hat{\theta})}} \leq \theta \leq \hat{\theta} + \frac{Z_{1-\alpha/2}}{\sqrt{ni(\hat{\theta})}}\right) \approx 1 - \alpha$$

Notar que podemos reemplazar $ni(\theta) = I(\theta)$. Por lo tanto, los límites del intervalo de confianza aproximado para θ de significancia α son:

$$IC_{\alpha}(\theta) \approx \hat{\theta} \pm Z_{1-\alpha/2}\frac{\hat{\theta}}{\sqrt{n}} = \hat{\theta}\left(1 \pm \frac{Z_{1-\alpha/2}}{\sqrt{n}}\right)$$

Donde $Z_{1-\alpha/2}$ es tal que:

$$\Phi(Z_{1-\alpha/2}) = 1 - \alpha/2$$

Siendo Φ la función de distribución de una variable normal estándar.

17.4 Inciso d

Para un nivel de significancia $\alpha = 5\%$, tenemos $Z_{1-\alpha/2} = 1.96$. Por lo tanto, para $n = 100$ y $\hat{\theta} = 7$ tenemos:

$$IC_{\alpha=0.05}(\theta) \approx 7\left(1 \pm \frac{1.96}{\sqrt{100}}\right) = [5.628, 8.372]$$

18 Ejercicio 18

18.1 Inciso a

Dado que las x están fijas, es fácil ver que Y es un modelo de locación respecto a ε con $\mu = \beta x$. Por lo tanto,

$$Y \sim N(\beta x, \sigma^2)$$

Dado que las x están dadas, el modelo está parametrizado por β y σ^2.

[7]Para ver una demostración de su sesgo, puede ver el ejemplo 6.2.4 de Hogg & Craig. *Introduction to Mathematical Statistics*

18.2 Inciso b

Dada la distribución de Y, la verosimilitud y log-verosimilitud para $\theta = (\beta, \sigma^2)$ están dada por:

$$L(\theta|\mathbf{y},\mathbf{x}) = \left(\frac{1}{2\sqrt{\sigma^2\pi}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2\right)$$
$$l(\theta|\mathbf{y},\mathbf{x}) = -\frac{n}{2}\left[\log 2\pi + \log\sigma^2\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \beta x_i)^2$$

Entonces, de las condiciones de primer orden respecto a β, σ^2:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} y_i x_i}{\sum_{i=1}^{n} x_i^2}$$
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\beta}x_i)^2$$

Utilizando que $\sum_{i=1}^{n} x_i^2/n = 1$, podemos simplificar estas expresiones a:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} y_i x_i}{n}$$
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i^2) - \hat{\beta}^2$$

18.3 Inciso c

Utilizando que el modelo normal es una familia exponencial, calculamos $I(\theta)$ a partir de la segunda derivada de la log-verosimilitud:

$$\begin{aligned} I(\beta) &= -E\left[\frac{\partial^2}{\partial\beta^2}\log f(\mathbf{x};\beta,\sigma^2)|\mathbf{x}\right] \\ &= -E\left[-\frac{\sum_{i=1}^{n} x_i^2}{\sigma^2}|\mathbf{x}\right] \\ &= E\left[\frac{n}{\sigma^2}|\mathbf{x}\right] = \frac{n}{\sigma^2} \end{aligned}$$

Donde el último paso proviene de $\sum_{i=1}^{n} x_i^2/n = 1$. Notar que $\mathbf{x}$ está fijo, con lo cual condicionamos sobre su valor al tomar esperanza y calcular varianza.

Para saber si $\hat{\beta}$ es eficiente, calculamos su varianza y la comparamos con la Cota de Crámer Rao.

$$\begin{aligned}
\text{Var}(\hat{\beta}) &= \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i|\mathbf{x}\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} x_i^2 \text{Var}(y_i|\mathbf{x}) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} x_i^2 \text{Var}(\beta x_i + \varepsilon|\mathbf{x}) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} x_i^2 \left(\text{Var}(\beta x_i|\mathbf{x}) + \text{Var}(\varepsilon_i|\mathbf{x})\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} x_i^2 \sigma^2 \\
&= \frac{\sigma^2}{n} = I(\beta)^{-1}
\end{aligned}$$

Con lo cual $\text{Var}(\hat{\beta})$ alcanza la Cota de CR, por lo cual es eficiente.

18.4 Inciso c

Para encontrar la distribución de $\hat{\beta}$ notar que:

$$\begin{aligned}
\hat{\beta} &= \frac{1}{n}\sum_{i=1}^{n} y_i x_i \\
&= \frac{1}{n}\sum_{i=1}^{n} \beta x_i^2 + x_i \varepsilon_i \\
&= \beta + \sum_{i=1}^{n} \frac{x_i}{n}\varepsilon_i
\end{aligned}$$

Entonces, condicional en $\mathbf{x}$, $\hat{\beta}$ es una combinación lineal de ε_i provenientes de una población $N(0,1)$. Por lo tanto, $\hat{\beta}$ se distribuye normal con media y varianza:

$$\begin{aligned}
E(\hat{\beta}|\mathbf{x}) &= \beta + \sum_{i=1}^{n} \frac{x_i}{n} E(\varepsilon_i|\mathbf{x}) = \beta \\
\text{Var}(\hat{\beta}|\mathbf{x}) &= \frac{\sigma^2}{n} \text{ (por el inciso anterior)}
\end{aligned}$$

18.5 Inciso e

Dado que $\hat{\beta}$ es insesgado, su ECM es igual a:

$$ECM(\hat{\beta}, \beta) = \text{Var}(\hat{\beta}) = \frac{\sigma^2}{n}$$

Dado que su varianza tiende a cero cuando $n \to \infty$, $\hat{\beta}$ es consistente.

19 Ejercicio 19

19.1 Inciso a

Para este modelo, la log-verosimilitud de θ está dada por:

$$l(\theta|\mathbf{x}) = -n\log 2 + 3n\log\theta + 2\sum_{i=1}^{n}\log x_i - \theta\sum_{i=1}^{n} x_i$$

Entonces, la condición de primer orden de la maximización de la log-verosimilitud está dada por:

$$\begin{aligned}\frac{\partial}{\partial\theta} l(\theta|\mathbf{x}) &= 0\\ \frac{3n}{\theta} - \sum_{i=1}^{n} x_i &= 0\\ \Rightarrow \hat{\theta} &= \frac{3}{\bar{x}}\end{aligned}$$

19.2 Inciso b

Sabemos por normalidad asintótica del estimador máximo verosímil que a medida que $n \to \infty$ su varianza converge a la inversa de la matriz de información. Es fácil mostrar que este modelo corresponde a una familia exponencial. Por tanto:

$$\begin{aligned}I(\theta) &= -E\left[\frac{\partial^2}{\partial\theta^2}\log f(\mathbf{x};\theta)|\mathbf{x}\right]\\ &= -E\left[-\frac{3n}{\theta^2}\right]\\ &= \frac{3n}{\theta^2}\end{aligned}$$

Entonces, la varianza asintótica de $\hat{\theta}$ es igual a $I(\theta)^{-1} = \theta^2/3n$.

19.3 Inciso c

Utilizando la normalidad asintótica del estimador máximo verosímil:

$$\sqrt{I(\hat{\theta})}(\hat{\theta}-\theta) \xrightarrow{D} N(0,1)$$

Por lo tanto, podemos construir un intervalo de confianza para θ con un nivel de significancia de $\alpha = 5\%$ como:

$$\hat{\theta} \pm Z_{1-\alpha/2}\frac{\hat{\theta}}{\sqrt{3n}}$$

Reemplazando $Z_{0.975} = 1.96$ y operando, los límites del intervalo pueden escribirse como:

$$\hat{\theta}\left(1 \pm \frac{1.96}{\sqrt{3n}}\right)$$

19.4 Inciso d

Para los datos provistos, $\bar{x} = 1.728 \Rightarrow \hat{\theta} \approx 1.736$. Por lo tanto, el intervalo de confianza estimado se vuelve:

$$1.736\left(1 \pm \frac{1.96}{\sqrt{3*250}}\right) \approx [1.61, 1.86]$$

20 Ejercicio 20

Para este ejercicio, podemos proponer el modelo $\{X_i\}_{i=1}^n \overset{iid}{\sim} Ber(p)$, donde cada variable aleatoria representa la intención de voto de un habitante de la CABA y el éxito se define como intención de votar por A.

20.1 Inciso a

Dado que para el modelo propuesto $n_{si} = \sum_{i=1}^n x_i$, podemos escribir al estimador $\hat{p}_{JP}$ como:

$$\hat{p}_{JP} = \frac{1}{n+20}\left(10 + \sum_{i=1}^n x_i\right)$$

Por lo tanto:

$$\begin{aligned}
E(\hat{p}_{JP}) &= \frac{1}{n+20}\left(10 + \sum_{i=1}^n E(x_i)\right) \\
&= \frac{1}{n+20}\left(10 + \sum_{i=1}^n p\right) \\
&= \frac{np+10}{n+20}
\end{aligned}$$

El sesgo de $\hat{p}_{JP}$ está dado por:

$$\text{Sesgo}_p(\hat{p}_{JP}) = \frac{np+10-np-20p}{n+20} = \frac{10-20p}{n+20} \neq 0 \text{ para todo } p \neq 0.5$$

La varianza de $\hat{p}_{JP}$ está dada por:

$$\begin{aligned}
\text{Var}(\hat{p}_{JP}) &= \left(\frac{1}{n+20}\right)^2 \sum_{i=1}^n \text{Var}(x_i) \\
&= \frac{np(1-p)}{(n+20)^2}
\end{aligned}$$

Por lo tanto, a partir de la descomposición del error medio cuadrático:

$$\begin{aligned}
ECM(\hat{p}_{JP}, p) &= \text{Sesgo}_p^2(\hat{p}_{JP}) + \text{Var}(\hat{p}_{JP}) \\
&= \left(\frac{10-20p}{n+20}\right)^2 + \frac{np(1-p)}{(n+20)^2}
\end{aligned}$$

Finalmente, queremos comparar el error medio cuadrático de ambos estimadores. En el caso del estimador insesgado $\hat{p}$, su error medio cuadrático es igual a su varianza:

$$ECM(\hat{p}, p) = \text{Var}(\hat{p}) = \frac{p(1-p)}{n}$$

La diferencia de ECM es igual a:

$$\begin{aligned}
ECM(\hat{p}_{JP}, p) - ECM(\hat{p}, p) &= \left(\frac{10-20p}{n+20}\right)^2 + \frac{np(1-p)}{(n+20)^2} - \frac{p(1-p)}{n} \\
&= \left(\frac{10-20p}{n+20}\right)^2 + p(1-p)\left(\frac{n^2-(n+20)^2}{(n+20)^2 n}\right) \\
&= \left(\frac{10-20p}{n+20}\right)^2 + p(1-p)\left(\frac{n^2-(n^2+20n+400)}{(n+20)^2 n}\right) \\
&= \frac{1}{(n+20)^2}\left((10-20p)^2 - p(1-p)\frac{400+20n}{n}\right)
\end{aligned}$$

Reemplazando $n = 1000$ de acuerdo a la información provista en el ejercicio, la diferencia se vuelve:

$$ECM(\hat{p}_{JP}, p) - ECM(\hat{p}, p) = \frac{1}{(1020)^2}\left((10 - 20p)^2 - p(1-p)\frac{20400}{1000}\right)$$

Si igualamos a cero y buscamos las raíces de la ecuación, encontramos que los estimadores son equivalentes en términos de ECM para valores aproximados de p de 0.38 y 0.61. Valores por fuera de este intervalo favorecen al estimador insesgado, mientras que valores mas cercanos a 0.5 favorecen a $\hat{p}_{JP}$.

21 Ejercicio 21

21.1 Inciso a

Para este modelo, la log-verosimilitud está dada por:

$$l(\alpha|\mathbf{t}) = n \log \alpha + \sum_{i=1}^{n} \log t_i - \frac{\alpha}{2}\sum_{i=1}^{n} t_i^2$$

Entonces, de la condición de primer orden del problema de maximización, se obtiene:

$$\hat{\alpha} = \frac{2n}{\sum_{i=1}^{n} \log t_i^2}$$

21.2 Inciso b

Para los valores provistos en el ejercicio, tenemos:

$$\hat{\alpha} = \frac{500}{510.58} \approx 0.979$$