# Ejercicios Tema 2 - Estimación (continuación)

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Curso completo de estadística inferencial con R y Python

## Estimación (continuación)

1. Supongamos que la cantidad de lluvia registrada en una cierta estación meteorológica en un día determinado está distribuida uniformemente en el intervalo $(0, b)$. Nos dan la siguiente muestra de los registros de los últimos 10 años en ese día:

   $$0, 0, 0.7, 1, 0.1, 0, 0.2, 0.5, 0, 0.6$$

   Estimar el parámetro $b$ a partir de su estimador $\tilde{b}$.

2. Supongamos que el grado de crecimiento de unos pinos jóvenes en metros de altura en un año es una variable aleatoria normal con media y varianza desconocidas. Se registran los crecimientos de 5 árboles y los resultados son: $0.9144, 1.524, 0.6096, 0.4572$ y $1.0668$ metros. Calcular los valores estimados de $\mu$ y $\sigma^2$ para esta muestra.

3. $X$ es una variable geométrica con parámetro $p$. Dada una muestra aleatoria de $n$ observaciones de $X$, cuál es el estimador de $p$ por método de los momentos?

4. Se supone que el número de horas que funciona una bombilla LED es una variable exponencial con parámetro $\lambda$. Dada una muestra de $n$ duraciones, calcular el estimador por método de los momentos para $\lambda$.

5. Si se supone que $X$ esta distribuida uniformemente en el intervalo $(b - \frac{1}{4}, b + 5)$, ¿cuál es el estimador por método de los momentos para a $b$ en base a una muestra aleatoria de $n$ observaciones?

6. Supongamos que $X$ es una variable aleatoria de Poisson con parámetro $\lambda$. Dada una muestra aleatoria de $n$ observaciones de $X$, cuál es el estimador de máxima verosimilitud para a $\lambda$?

7. Supongamos que $X$ está distribuida uniformemente en el intervalo $\left(b - \frac{1}{2}, b + \frac{1}{2}\right)$. ¿Cuál es el estimador de máxima verosimilitud para a $b$ Dada una muestra aleatoria de tamaño $n$ para $X$?

8. ¿Cuál es el estimador de máxima verosimilitud para al parámetro $\lambda$ de una variable exponencial para a una muestra de tamaño $n$?

9. Se registran los tiempos de duración de 30 bombillas. Supongamos que el tiempo de duración de estas bombillas es una variable exponencial. Si la suma de los tiempos $\sum x_i = 32916$ horas, ¿cuál es el estimador de máxima verosimilitud para al parámetro de la distribución exponencial de duración de les bombillas?

10. Supongamos que $X_1, X_2, \ldots, X_6$ es una muestra aleatoria de una variable aleatoria normal con media $\mu$ y varianza $\sigma^2$. Hallar la constante $C$ tal que

    $$C \cdot \left((X_1 - X_2)^2 + (X_3 - X_4)^2 + (X_5 - X_6)^2\right),$$

    sea un estimador sin sesgo de $\sigma^2$.

11. Un vendedor de coches piensa que el número de ventas de coches nuevos que hace en un día no festivo es una variable aleatoria de Poisson con parámetro $\lambda$. Examinando los registros del año anterior (que

tuvo 310 días no festivos), observa que vendió un total de 279 coches. Calcular por método de la máxima verosimilitud la probabilidad que no venda ningún coche el próximo día de trabajo.

12. Supongamos que los años de vida de los hombres de los Estados Unidos Mexicanos están distribuidos normalmente con media $\mu$ y varianza $\sigma^2$. Una muestra aleatoria de $n = 10000$ antecedentes de mortalidad de hombres en México se obtuvo como resultado $\overline{x} = 72.1$ hombres , \hbox{$s^2 = 144$ años. Estimar por método de máxima verosimilitud la probabilidad que un hombre mexicano viva hasta los 50 años de edad y la probabilidad que un hombre no llegue a los 90 años.

13. Supongamos que $\Theta_1$ y $\Theta_2$ son estimadores sin sesgo de un parámetro desconocido $\theta$, con varianzas conocidas $\sigma_1^2$ y $\sigma_2^2$, respectivamente. Se pide:

   a.) Demostrar que $\Theta = (1-a)\cdot\Theta_1 + a\cdot\Theta_2$ también es insesgado para cualquier valor de $a$.

   b.) Hallar el valor de $a$ que minimiza $Var(\Theta)$.

14. Sea $X$ una variable aleatoria $N(\mu,\sigma)$ con $\sigma$ conocida y $X_1,\ldots,X_n$ una muestra aleatoria simple de $X$. Consideremos los siguientes estimadores del parámetro $\lambda = \mu^2$:

$$\begin{aligned} T_1 &= \overline{X}^2 = \left(\frac{\sum\limits_{i=1}^{n} X_i}{n}\right)^2, \\ T_2 &= \frac{\sum\limits_{i=1}^{n} X_i^2}{n} - \sigma^2. \end{aligned}$$

   Se pide:

   a.) Demostrar que $T_1$ y $T_2$ son estimadores consistentes.

   b.) ¿Cuál estimador es más eficiente?

15. Sea $X_1,\ldots,X_{2n}$ una muestra aleatoria simple de una variable aleatoria $N(\mu,\sigma)$. Sea:

$$T = C\left(\left(\sum_{i=1}^{2n} X_i\right)^2 - 4n\sum_{i=1}^{n} X_{2i}X_{2i-1}\right)$$

   un estimador del parámetro $\sigma^2$. ¿Cuál es el valor de $C$ para que $T$ sea un estimador insesgado?

16. Una variable aleatoria $X$ sigue la distribución de Rayleigh con parámetro $\theta > 0$ si es una variable aleatoria con valores $x > 0$ y función de densidad:

$$f(x) = \frac{x}{\theta}e^{-\frac{x^2}{2\theta}}.$$

   Hallar estimadores del parámetro $\theta$:

   a.) El método de los momentos.

   b.) El método de la máxima verosimilitud.

17. Consideremos una variable aleatoria $X$ que es $Exp(\lambda)$. Sea $X_1,\ldots,X_n$ una muestra aleatoria simple de $X$. Consideremos los siguientes estimadores del parámetro $\frac{1}{\lambda^2}$:

$$\begin{aligned} T_1 &= \overline{X}^2, \\ T_2 &= \frac{1}{2n}\sum_{i=1}^{n} X_i^2. \end{aligned}$$

a.) ¿Son estimadores insesgados?

b.) ¿Cuál de los dos estimadores es más eficiente? Indicación: La variable aleatoria $2\lambda\sum_{i=1}^{n} X_i$ es una variable $\chi^2_{2n}$.

18. Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria $X$ uniforme en el intervalo $[0, b]$. Consideremos el siguiente estimador del parámetro $b$:

$$\tilde{b} = a\sum_{i=1}^{n} X_i,$$

donde $a$ es una constante. Calcular $a$ para que $\tilde{b}$ sea un estimador insesgado de $b$, en este caso calcular $Var(\tilde{b})$.

19. Hallar los estimadores por método de los momentos y por máxima verosimilitud del parámetro $\alpha$ de la distribución de Maxwell:

$$f(x) = \frac{4}{\alpha^3\sqrt{\pi}}x^2e^{-\frac{x^2}{\alpha^2}},\ x \geq 0,\ \alpha > 0.$$

20. Sea $X_1, \cdots, X_n$ una muestra aleatoria simple de una variable aleatoria $X$ con $E(X) = \mu$ y $Var(X) = \sigma^2$. Consideremos los siguientes estimadores del parámetro $\mu$:

$$T_1 = \overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}, \quad T_2 = k\sum_{i=1}^{n} yX_i.$$

Se pide:

a.) El valor de la constante $k$ para que $T_2$ sea insesgado.

b.) Demostrar que $T_1$ y $T_2$ son consistentes.

c.) ¿Cuál estimador es más eficiente? Ayuda: $\sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}$.

21. Sea $X_1, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria $X$ tal que $F_X$ depende de un parámetro desconocido $\lambda$ con $E(X) = \lambda$ y $Var(X) = \lambda^2$. Consideremos el siguiente estimador de $\lambda$:

$$\tilde{\lambda} = \frac{1}{2}\left(\frac{1}{m}(X_1 + \cdots X_m) + \frac{1}{n-m}(X_{m+1} + \cdots X_n)\right),$$

con $m = \frac{n}{3}$, en el que suponemos que $n$ es múltiplo de 3. Hallar la varianza de $\tilde{\lambda}$.

22. Sea $X$ una variable aleatoria tal que $E(X) = \mu$ y $Var(X) = \sigma^2$. Sea $X_1, X_2$ una muestra aleatoria simple de $X$ de tamaño 2. Consideremos el siguiente estimador del parámetro $\mu$: $\tilde{\mu} = 2aX_1 + (1-2a)X_2$. Hallar el valor de $a$ que hace que el estadístico
$\tilde{\mu}$ sea el más eficiente posible.

# Soluciones

1. Cargemos los datos en R

```
muestra_lluvia=c(0,0,0.7,1,0.1,0,0.2,0.5,0,0.6)
```

Nos dicen que los datos de la muestra provienen de una población modela da pir una v.a. $X$ con distribución $U(0, b)$. Entonces $E(X) = \frac{1}{b-0} = \frac{1}{b}$. por el método de los momentos estimamos $E(X)$ por $\overline{X}$ luego $\frac{1b}{=}E(X)$ de donde $b = \frac{1}{E(X)}$ y por lo tanto un estimador del parámetro $b$ es $\hat{b} = \frac{1}{X}$. En nuestro caso y con R

```
media_lluvia=mean(muestra_lluvia)
media_lluvia
```

```
## [1] 0.31
```

```
bhat=1/media_lluvia
bhat
```

```
## [1] 3.225806
```

2. Tenemos que $X$ = crecimiento en metros de un pino joven en un año sigue una ley $N(\mu, \sigma)$ de parámetros desconocidos. Tenemos un muestra que cargamos con R

```
muestra_pinos=c(0.9144, 1.524, 0.6096, 0.4572 ,1.0668)
mean(muestra_pinos)
```

```
## [1] 0.9144
```

```
var(muestra_pinos)
```

```
## [1] 0.1741932
```

```
n=length(muestra_pinos)
n
```

```
## [1] 5
```

```
media_pinos=sum(muestra_pinos)/n
media_pinos
```

```
## [1] 0.9144
```

```
varianza_pinos=sum(muestra_pinos^2)/n-media_pinos^2
varianza_pinos
```

```
## [1] 0.1393546
```

3. Si $X$ una v.a. discreta con distribución $Ge(p)$ con $D_X = \{0, 1, 2, \ldots\}$ en este caso sabemos que $E(X) = \frac{1}{p}$, como $\overline{X}$ es un estimador de $E(X)$ podemos operar y $\hat{p} = \frac{1}{\overline{X}}$ es un estimador por el método de los momentos del parámetro $p$.

4. Ahora $X$ una $Exp(\lambda)$. La solución es similar que el caso anterior (no en vano la exponencial es la versión contniua de la v.a. geométrica).

Sabemos que $E(X) = \frac{1}{\lambda}$ luego un estimador del parámetro $\lambda$ de una población exponencial es $\hat{\lambda} = \frac{1}{\overline{X}}$.

5. Ahora $X$ es sigue una ley $U(b - \frac{1}{4}, b + 5)$ entonces $E(X) = \frac{b - \frac{1}{4} + b + 5}{2} = \frac{2 \cdot b + \frac{19}{4}}{2} = b + \frac{19}{8}$. Así $\hat{b} = \overline{X} - \frac{19}{8}$.

6. Si $X$ es una variable $Po(\lambda)$ y tenemos una m.a.s $X_1, X_2, \ldots, X_n$ de esa v.a; así su función de probabilidad es $P(X_i = x_i) = \frac{\lambda^{x_i}}{x_i!} \cdot \mathrm{e}^{-\lambda}$ si $x_i=0,1,2,. . . . La distribución de la muestra es

$$
\begin{array}{rcl}
P\left(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right) & = & P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \ldots \cdot P(X_n = x_n) \\
& = & \frac{\lambda^{x_1}}{x_1!} \cdot \mathrm{e}^{-\lambda} \cdot \frac{\lambda^{x_2}}{x_2!} \cdot \mathrm{e}^{-\lambda} \cdot \ldots \frac{\lambda^{x_n}}{x_n!} \cdot \mathrm{e}^{-\lambda} \\
& = & \frac{\lambda^{\sum_{i=1}^{n} x_i}}{x_1! \cdot x_2! \cdot \ldots \cdot x_n!} \mathrm{e}^{-n \cdot \lambda}.
\end{array}
$$

Así la función de verosimilitud es

$$L(\lambda|x_1,x_2\ldots,x_n)=\frac{\lambda^{\sum_{i=1}^{n}x_i}}{x_1!\cdot x_2!\cdot\ldots\cdot x_n!}\cdot e^{-n\cdot\lambda}$$

Queremos encontrar el valor de $\lambda$ que máximiza $L(\lambda|x_1,x_2,\ldots,x_n)$ es decir

$$\arg\max_{\lambda} L(\lambda|x_1,x_2,\ldots,x_n)$$

tomando logaritmos tenemos que

$$\begin{aligned}
\ln\left(L(\lambda|x_1,x_2\ldots,x_n)\right) &= \ln\left(\frac{\lambda^{\sum_{i=1}^{n}x_i}}{x_1!\cdot x_2!\cdot\ldots\cdot x_n!}+e^{-n\cdot\lambda}\right)\\
&= \ln\left(\frac{\lambda^{\sum_{i=1}^{n}x_i}}{x_1!\cdot x_2!\cdot\ldots\cdot x_n!}\right)-n\cdot\lambda\\
&= \ln\left(\lambda^{\sum_{i=1}^{n}x_i}\right)-\ln\left(x_1!\cdot x_2!\cdot\ldots\cdot x_n!\right)-n\cdot\lambda\\
&= \left(\sum_{i=1}^{n}x_i\right)\cdot\ln(\lambda)-\ln\left(x_1!\cdot x_2!\cdot\ldots\cdot x_n!\right)-n\cdot\lambda
\end{aligned}$$

derivando respecto de $\lambda$

$$\frac{\partial}{\partial\lambda}\ln\left(L(\lambda|x_1,x_2\ldots,x_n)\right)=\frac{\partial}{\partial\lambda}\left(\left(\sum_{i=1}^{n}x_i\right)\cdot\ln(\lambda)-\ln\left(x_1!\cdot x_2!\cdot\ldots\cdot x_n!\right)-n\cdot\lambda\right)=\frac{\sum_{i=1}^{n}x_i}{\lambda}-n.$$

el máximo se alacanza para el $\lambda$ que cumpla la ecuación
$\frac{\sum_{i=1}^{n}x_i}{\lambda}-n=0$ de donde $\lambda=\frac{\sum_{i=1}^{n}x_i}{n}=\overline{x}$.

Luego el estimador máximo verosimil de $\lambda$ es $\overline{X}$.

7. Si $X$ es $U\left(b-\frac{1}{2},b+\frac{1}{2}\right)$ su densidad es $f_X(x)=\begin{cases}\frac{1}{6} & \text{si } b-\frac{1}{2}<x<b+\frac{1}{2}\\ 0 & \text{en otro caso}\end{cases}$ su función de verosimilitud es constante NO TIENE EMV.

8. Las v.a. de la muestra son $X_i$y tienen distribución $Exp(\lambda)$. Sus densidades son $f_{X_i}(x_i)=\lambda\cdot\mathrm{e}^{-\lambda x_i}$ si $x_i>0$ y cero en el resto de casos.

La función de verosimilitu de la muestra para una realización de la muestra $x_i,x_2,\ldots,x_n$ es

$$\begin{aligned}
L(\lambda|x_1,x_2,\ldots,x_n) &= f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots x_n)=f_{X_1}(x_1)\cdot f_{X_1}(x_1)\cdot\ldots\cdot f_{X_n}(x_n)\\
&= \lambda\mathrm{e}^{-\lambda x_1}\cdot\lambda\mathrm{e}^{-\lambda x_2}\cdot\ldots\cdot\lambda\mathrm{e}^{-\lambda x_n}\\
&= \lambda^n\cdot\mathrm{e}^{-\lambda\cdot(x_1+x_2+\cdots+x_n)}=\lambda^n\cdot\mathrm{e}^{-\lambda\sum_{i=1}^{n}x_i}
\end{aligned}$$

Ahora $\ln\left(L(\lambda|x_1,x_2,\ldots,x_n)\right)=n\cdot\ln(\lambda)-\lambda\sum_{i=1}^{n}x_i$ derivando e igualando a cero obtenemos $\frac{n}{\lambda}-\sum_{i=1}^{n}x_i=$ de donde $\lambda=\frac{n}{\sum_{i=1}^{n}x_i}=\frac{1}{\overline{x}}$.

Luego el estimador MV de $\lambda$ es $la\hat{m}bda = \frac{1}{\overline{X}}$.