

Ejercicios Tema 2 - Estimación. Taller 2

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Curso completo de estadística inferencial con R y Python

Contenidos

1	Estimación taller 2	1
1.1	Ejercicio 1	1
1.2	Ejercicio 2	2
1.3	Ejercicio 3	2
1.4	Ejercicio 4	2
1.5	Ejercicio 5	2
1.6	Ejercicio 6	2
1.7	Ejercicio 7	2
1.8	Ejercicio 8	2
1.9	Ejercicio 9	2
1.10	Ejercicio 10	2
1.11	Ejercicio 11	3
1.12	Ejercicio 12	3
1.13	Ejercicio 13	3
1.14	Ejercicio 14	3
1.15	Ejercicio 15	3
1.16	Ejercicio 16	4
1.17	Ejercicio 17	4
1.18	Ejercicio 18	4
1.19	Ejercicio 19	4
1.20	Ejercicio 20	4
1.21	Ejercicio 21	5
2	Soluciones	5
2.1	Solución ejercicio 1	5
2.2	Solución ejercicio 2	5
2.3	Solución ejercicio 3	6
2.4	Solución ejercicio 4	6
2.5	Solución ejercicio 5	6
2.6	Solución ejercicio 6	6
2.7	Solución ejercicio 7	7
2.8	Solución ejercicio 8	7

1 Estimación taller 2

1.1 Ejercicio 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} .

1.2 Ejercicio 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 μ y σ^2 para esta muestra.

1.3 Ejercicio 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?

1.4 Ejercicio 4

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

1.5 Ejercicio 5

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

1.6 Ejercicio 6

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

1.7 Ejercicio 7

¿Cuál es el estimador de máxima verosimilitud para el parámetro λ de una variable exponencial para una muestra de tamaño n ?

1.8 Ejercicio 8

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 el parámetro de la distribución exponencial de duración de las bombillas?

1.9 Ejercicio 9

Supongamos que X_1, X_2, \dots, X_6 es una muestra aleatoria de una variable aleatoria normal con media μ y varianza σ^2 . Hallar la constante C tal que

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

sea un estimador sin sesgo de σ^2 .

1.10 Ejercicio 10

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 λ . 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.

1.11 Ejercicio 11.

Supongamos que los años de vida de los hombres de los Estados Unidos Mexicanos están distribuidos normalmente con media μ y varianza σ^2 . Una muestra aleatoria de $n = 10000$ antecedentes de mortalidad de hombres en México se obtuvo como resultado $\bar{x} = 72.1$ hombres, $s^2 = 144$ 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.

1.12 Ejercicio 12

Supongamos que Θ_1 y Θ_2 son estimadores sin sesgo de un parámetro desconocido θ , con varianzas conocidas σ_1^2 y σ_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 $Var(\Theta)$.

1.13 Ejercicio 13

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

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

Se pide:

- a.) Demostrar que T_1 y T_2 son estimadores consistentes.
- b.) ¿Cuál estimador es más eficiente?

1.14 Ejercicio 14

Sea X_1, \dots, 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 σ^2 . ¿Cuál es el valor de C para que T sea un estimador insesgado?

1.15 Ejercicio 15

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 θ :

- a.) El método de los momentos.
- b.) El método de la máxima verosimilitud.

1.16 Ejercicio 16

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

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

- 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 χ_{2n}^2 .

1.17 Ejercicio 17

Sea X_1, \dots, 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 $Var(\tilde{b})$.

1.18 Ejercicio 18

Hallar los estimadores por método de los momentos y por máxima verosimilitud del parámetro α de la distribución de Maxwell:

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

1.19 Ejercicio 19

Sea X_1, \dots, 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 μ :

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

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: $\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$.

1.20 Ejercicio 20

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

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

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

1.21 Ejercicio 21

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 μ : $\tilde{\mu} = 2aX_1 + (1 - 2a)X_2$. Hallar el valor de a que hace que el estadístico $\tilde{\mu}$ sea el más eficiente posible.

2 Soluciones

2.1 Solución ejercicio 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 modelada por 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 \bar{X} luego $\frac{1}{b}E(X)$ de donde $b = \frac{1}{E(X)}$ y por lo tanto un estimador del parámetro b es $\hat{b} = \frac{1}{\bar{X}}$. En nuestro caso y con R

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

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

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

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

2.2 Solución ejercicio 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
```

2.3 Solución ejercicio 3

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

2.4 Solución ejercicio 4

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

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

2.5 Solución ejercicio 5

Ahora X 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 $\hat{b} = \bar{X} - \frac{19}{8}$.

2.6 Solución ejercicio 6

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

$$\begin{aligned} P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) &= P(X_1 = x_1) \cdot P(X_2 = x_2) \cdot \dots \cdot P(X_n = x_n) \\ &= \frac{\lambda^{x_1}}{x_1!} \cdot e^{-\lambda} \cdot \frac{\lambda^{x_2}}{x_2!} \cdot e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{x_n}}{x_n!} \cdot e^{-\lambda} \\ &= \frac{\lambda^{\sum_{i=1}^n x_i}}{x_1! \cdot x_2! \cdot \dots \cdot x_n!} e^{-n \cdot \lambda}. \end{aligned}$$

Así la función de verosimilitud es

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

Queremos encontrar el valor de λ que maximiza $L(\lambda | x_1, x_2, \dots, x_n)$ es decir

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

tomando logaritmos tenemos que

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

derivando respecto de λ

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

Despejando λ la ecuación

$$\frac{\sum_{i=1}^n x_i}{\lambda} - n = 0 \text{ de obtiene } \lambda = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}.$$

Luego el estimador máximo verosímil de λ es \bar{X} .

2.7 Solución ejercicio 7

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

La función de verosimilitud de la muestra para una realización de la muestra x_1, x_2, \dots, x_n es

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

Ahora $\ln(L(\lambda|x_1, x_2, \dots, x_n)) = n \cdot \ln(\lambda) - \lambda \sum_{i=1}^n x_i$ derivando e igualando a cero obtenemos $\frac{n}{\lambda} - \sum_{i=1}^n x_i = 0$ de donde $\lambda = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{x}}$.

Luego el estimador MV de λ es $\hat{\lambda} = \frac{1}{\bar{X}}$.

2.8 Solución ejercicio 8

Nos dicen que la v.a. X = duración de unas bombillas sigue una distribución $Exp(\lambda)$. Se toma una muestra de 30 bombillas, se encienden hasta que fallan y se anota el número de horas funcionando. Se obtienen x_1, x_2, \dots, x_{30} duraciones tales que $\sum_{i=1}^{30} x_i = 32000$ luego $\bar{x} = \frac{\sum_{i=1}^{30} x_i = 32000}{n} = \frac{32000}{30}$. Como la población es exponencial por el problema anterior el EMV es $\hat{\lambda} = \frac{1}{\bar{x}} = \frac{1}{\frac{32000}{30}} = \frac{3}{3200} = 0.0009375$.