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 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 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?

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 a λ?

1.7 Ejercicio 7

¿Cuál es el estimador de máxima verosimilitud para al parámetro λ de una variable exponencial para a 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 al parámetro de la distribución exponencial de duración de les bombillas?

2 Soluciones

2.1 Solución ejercicio 1

Carguemos 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 $\overline{X}$ luego $\stackrel{1b}{=} E(X)$ de donde $b = \frac{1}{E(X)}$ y por lo tanto un estimador del parámetro b es $\hat{b} = \frac{1}{\overline{X}}$. 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
```

```r
var(muestra_pinos)
```

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

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

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

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

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

```r
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 $D_X = \{0, 1, 2, \ldots\}$ 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.

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 $\hat{\lambda} = \frac{1}{\overline{X}}$.

2.5 Solución ejercicio 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 $\hat{b} = \overline{X} - \frac{19}{8}$.

2.6 Solución ejercicio 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

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

Así la función de verosimilitud es

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

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

$$\text{arg max}_\lambda L(\lambda | x_1, x_2, \ldots, x_n)$$

tomando logaritmos tenemos que

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

derivando respecto de λ

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

Despejando λ de la ecuación

$$\frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0$$

se obtiene que $\lambda = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}$.

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

2.7 Solución ejercicio 7

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

La función de verosimilitud de la muestra para una realización de la muestra $x_i, x_2, \ldots, x_n$ 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 $\frac{n}{\lambda} - \sum_{i=1}^{n} x_i = 0$ de donde $\lambda = \frac{n}{\sum_{i=1}^{n} x_i} = \frac{1}{\overline{x}}$.

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

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 $x_1, x_2, \ldots, x_{30}$ duraciones tales que $\sum_{i=1}^{30} x_i = 32000$ luego $\overline{x} = \frac{\sum_{i=1}^{30} x_i = 32000}{n} = \frac{32000}{30}$. Como la población es exponencial por el problema anterior el EMV es $\hat{\lambda} = \frac{1}{\overline{x}} = \frac{1}{\frac{32000}{30}} = \frac{3}{3200} = 0.0009375$.