Ejercicios Tema 2 - Estimación

Ricardo Alberich, Juan Gabriel Gomila y Arnau Mir

Curso completo de estadística inferencial con R y Python

Estimación

1. El fabricante SMART_LED fabrica bombillas led inteligentes y de alta gama. Supongamos que la vida de de estas bombillas sigue una distribución exponencial de parámetro λ. Si tomamos una muestra aleatoria de tamaño n de estas bombillas y representamos por X_i la duración de la $i-$ésima bombilla para $i = 1, \ldots, n$, ¿cuál es la función de densidad conjunta de la muestra?

2. Sean $X_1, X_2, \ldots, X_{10}$ variables aleatorias que son una muestra aleatoria simple de una v.a. X.

   a. Dividimos la muestra en dos partes: de forma que la primera son los 5 primeros valores y la segunda los restantes. ¿Son independientes las dos partes?
   b. Volvemos a dividir la muestra en dos partes: la primera está formada por los 5 valores más pequeños y la segunda por el resto. ¿Son independientes las dos partes?

3. Un fabricante de motores pone a prueba 6 motores sobre el mismo prototipo de coche de competición. Para probar que los motores tienes las mismas prestaciones se someten a distintas pruebas en un circuito. Las velocidades máximas en 10 vueltas al circuito de cada motor tras la prueba son $190, 195, 193, 177, 201$ y 187 en Km/h. Estos valores forman una muestra aleatoria simple de la variable $X =$ velocidad máxima de un motor en 10 vueltas. Se pide calcular los valores observados de los siguientes estadísticos de la muestra:

   a. $\overline{X}$.
   b. $\tilde{S}^2$.
   c. Mediana.
   d. $X_{(4)}$ (valor que ocupa el cuarto lugar ordenados los valores de menor a mayor).

4. ¿Cuál es la probabilidad de que el máximo de de una muestra de tamaño $n = 10$ de una v.a. uniforme en el intervalo $(0, 1)$ sea mayor que 0.9? ¿Cuál es la probabilidad de sea menor que $\frac{1}{2}$?

5. Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria simple de una variable aleatoria normal de parámetros μ y σ. Denotemos por $X_{(1)} \leq X_{(2)} \leq, \ldots, \leq X_{(n)}$ la muestra ordenada de menor a mayor.

   a. Calcular la funciones de densidad del mínimo $X_{(1)}$ y del máximo $X_{(n)}$
   b. ¿Alguna de estas variables sigue una distribución normal?

6. Consideremos la muestra aleatoria simple $X_1, X_2, \ldots, X_n$ de una v.a X de media μ y varianza σ^2 desconocidas. Definimos
$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ y } T = \frac{\sqrt{n}\cdot(\overline{X} - \mu)}{\sigma}.$$

   a. ¿Cuál es la distribución de T?
   b. ¿Es T un estadístico?

7. Consideremos la muestra aleatoria simple $X_1, X_2, \ldots, X_n$ de tamaño $n = 10$ de una v.a X normal estándar. Calculad $P\left(2.56 < \sum\limits_{i=1}^{10} X_i^2 < 18.31\right)$.

8. Consideremos la muestra aleatoria simple $X_1, X_2, \ldots, X_n$ de tamaño $n = 10$ de una v.a X normal $N(\mu = 2, \sigma = 4)$. Definimos la siguiente variable aleatoria $Y = \frac{\sum\limits_{i=1}^{10}(X_i-2)^2}{16}$. Calculad $P(Y \leq 2.6)$

Soluciones

1. Cada X_i sigue una ley $Exp(\lambda)$ la densidad es

$$f_{X_i}(x_i) = \left\{ \begin{array}{ll} \lambda \cdot \mathrm{e}^{-\lambda \cdot x_i} & \text{si } x_i > 0 \\ 0 & \text{si } x_i \leq 0 \end{array} \right.$$

Así la densidad conjunta de la muestra es

$$f(x_1, x_2, \ldots x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \ldots \cdot f_{X_n}(x_n) = f_{X_i}(x_i) = \left\{ \begin{array}{ll} \lambda^n \cdot \mathrm{e}^{-\lambda \cdot \sum_{i=1}^n x_i} & \text{si } x_i > 0 \text{ para todo } i = 1, 2, \ldots, n \\ 0 & \text{si } x_i \leq 0 \text{ para algún } i = 1, 2, \ldots, n \end{array} \right. .$$

2. En el primer caso las muestras son independientes, saber los resultados del los 5 primeros no aporta información sobre los 5 últimos. En el segundo caso sí que aporta información pues los valores de la segunda parte deben ser mayores que todos los de la primera parte, luego no son independientes.

3. Lo calcularemos con R

```
x=c(190,195,193,177,201,187)
x
```

```
## [1] 190 195 193 177 201 187
```

```
n=length(x)
n # tamaño de la muestra
```

```
## [1] 6
```

```
mean(x)# media
```

```
## [1] 190.5
```

```
var(x)# variana muestral con la función var
```

```
## [1] 66.3
```

```
sum((x-mean(x))^2)/(n-1) # variana muestral calculada directamente con R
```

```
## [1] 66.3
```

```
median(x)
```

```
## [1] 191.5
```

```
sort(x) # muestra ordenada
```

```
## [1] 177 187 190 193 195 201
```

```
sort(x)[4] # M_(4) el cuarto valor de la muestra ordenada
```

```
## [1] 193
```

4. La primera probabilidad es $P(\max\{X_1, \ldots, X_n\} \geq 0.9) = 1 - P(\max\{X_1, \ldots, X_n\} \leq 0.9) = 1 - P(X_1 \leq 0.9, X_2 \leq 0.9, \ldots, X_n \leq 0.9) = 1 - P(X_1 \leq 0.9) \cdot P(X_2 \leq 0.9) \cdot \ldots \cdot P(X_n \leq 0.9) = 1 - 0.9^10 = 0.6513.$

La segunda es $P(\max\{X_1,\ldots,X_n\}\leq 0.1)=P(X_1\leq 0.1, X_2\leq 0.1,\ldots,X_n\leq 0.1)=P(X_1\leq 0.1)\cdot P(X_2\leq 0.1)\cdot\ldots\cdot P(X_n\leq 0.1)=0.1^{10}=10^{-10}$

Hemos utilizado que la distribución uniforme

$$P(X_i\leq x)=\begin{cases}0 & \text{si } x\leq 0\\ x & \text{si } 0<x<1\\ 1 & \text{si } x\geq 1\end{cases}$$

5. Sea F_X la distribución de la variable que se muestrea entonces $F_{X_i}=F_X$ para $i=1,2,\ldots,n$.

$P(\max\{X_1,\ldots,X_n\}\leq x)=P(X_1\leq x, X_2\leq x,\ldots,X_n\leq x)=P(X_1\leq x)\cdot P(X_2\leq x)\cdot\ldots\cdot P(X_n\leq x)=F_{X_1}(x)\cdot F_{X_1}(x)\ldots\cdot F_{X_n}(x)=F_X(x)^n$.

$P(\min\{X_1,\ldots,X_n\}\leq x)=1-P(\min\{X_1,\ldots,X_n\}\geq x)=1-P(X_1\geq x, X_2\leq x,\ldots,X_n\geq x)=1-(P(X_1\geq x)\cdot P(X_2\geq x)\cdot\ldots\cdot P(X_n\geq x))=1-P(X_1\geq x)\cdot P(X_2\geq x)\cdot\ldots\cdot P(X_n\geq x)=1-(1-P(X_1\leq x))\cdot(1-P(X_1\leq x))\cdot\ldots\cdot(1-P(X_1\leq x))=1-(1-F_X(x))\cdot(1-F_X(x))\ldots\cdot(1-F_X(x))=1-(1-F_X(x))^n$.

No son normales. Se deja como ejercicio derivar la función de distribución del máximo (o del mínimo) para el caso $n=2$ y comprobar que no es una gaussiana.

6. Ahora tenemos una muestra aleatoria simple de una distribución de media μ y desviación típica sigma y como siempre tenemos el estadśitico $\overline{X}$.

    a. Nos piden la distribución de $T=\frac{\sqrt{n}\cdot(\overline{X}-\mu)}{\sigma}$ operando

$$T=\frac{\sqrt{n}\cdot(\overline{X}-\mu)}{\sigma}=\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

Ahora sabemos que la distribución de T por el Teorema Central de Límite converge en distribución a una normal estándar cuando $n\to\infty$.

Además si las variables fueran normales T seguirá distribución normal estándar.

b. Claro que T es un estadístico, ya que estadístico es cualquier función de una muestra. Además si nos fijamos bienes simplemente la tipificación del estadístico $\overline{X}$.

7. Como se una muestra de una normal estándar tenemos que $\mu=0$ y $sigma=1$

Así que si denotamos por $Y=\sum_{i=1}^n$, resulta que Y es la suma de normales estándar $N(\mu=0,\sigma=1)$, idénticamente distribuidas y por lo tanto sabemos que Y sigue una ley $N(n\cdot 0=0, n\cdot\sqrt{\cdot}\sigma=\sqrt{10})$. Ahora podemos operar

$P\left(1.56<\sum_{i=1}^n<18.31\right)=P(2.56<Y<18.31)=P(Y<18.31)-P(Y<2.56)=0.9664497-0.6010246=0.3654251$.

```r
pnorm(18.31,mean=0,sd=sqrt(10))
```

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

```r
pnorm(2.56,mean=0,sd=sqrt(10))
```

```
## [1] 0.7908986
```

```r
pnorm(18.31,mean=0,sd=sqrt(10))-pnorm(2.56,mean=0,sd=sqrt(10))
```

```
## [1] 0.2091014
```

o también, tipificando $Z=\frac{Y}{\sqrt{10}}$ es una $N(0,1)$

```r
pnorm(18.31/sqrt(10),mean=0,sd=1)
```

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

```r
pnorm(2.56/sqrt(10),mean=0,sd=1)
```

```
## [1] 0.7908986
```

```r
pnorm(18.31/sqrt(10),mean=0,sd=1)-pnorm(2.56/sqrt(10),mean=0,sd=1)
```

```
## [1] 0.2091014
```

obtenemos el mismo resultado.

8. Consideremos la muestra aleatoria simple $X_1, X_2, \ldots, X_n$de tamaño $n = 10$ de una v.a X normal $N(\mu = 2, \sigma = 4)$. Definimos la siguiente variable aleatoria $Y = \frac{\sum_{i=1}^{10} (X_i - 2)^2}{16}$. Calculad $P(Y \leq 2.6)$

Notemos que $Z_i = \frac{X_i - 2}{4}$ son variables $N(0,1)$ para $i = 1.2, \ldots, 10$

Ahora

$$Y = \frac{\sum_{i=1}^{10} (X_i - 2)^2}{16} = \sum_{i=1}^{10} \left(\frac{X_i - 2}{4}\right)^2 = \sum_{i=1}^{10} Z_i^2 = \chi_{10}^2.$$

Luego $Y = \chi_{10}^2$ es una v.a. χ^2 con 10 grados de libertad. Ya podemos calcular la probabilidad pedida $P(Y \leq 2.6) = P(\chi_{10}^2 \leq 2.6) = 0.010663$.

El cálculo lo hemos hecho con

```r
pchisq(2.6,df=10)
```

```
## [1] 0.01066303
```