

Problema 3: Màxima Versemblança 1

Gerard Barrachina Josep de Cid Albert Ribes
Kerstin Winter

28 de septiembre de 2017

Considerem un experiment aleatori en què mesurem una determinada variable aleatòria X , que segueix una distribució gaussiana univariada, cosa que escrivim $X \sim N(\mu, \sigma^2)$. Prenem N mesures independents de X i obtenim una mostra aleatòria simple $\{x_1, \dots, x_N\}$, on cada x_n es una realització de X , per $n = 1, \dots, N$. Es demana:

1. **Escriviu la funció de densitat de probabilitat per un x_n qualsevol i construïu la funció log-versemblança (negativa) de la mostra.**

$$f(x_j|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_j-\mu)^2}{2\sigma^2}}$$

$$lh = - \sum_{n=1}^N \left[\ln(f(x_n|\mu, \sigma^2)) \right]$$

$$lh = - \sum_{n=1}^N \left[\ln \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x_j-\mu)^2}{2\sigma^2}} \right) \right]$$

$$lh = - \sum_{n=1}^N \left[\ln \left(\frac{1}{\sigma\sqrt{2\pi}} \right) + \ln \left(e^{-\frac{(x_j-\mu)^2}{2\sigma^2}} \right) \right]$$

$$lh = - \sum_{n=1}^N \left[-\ln(\sigma\sqrt{2\pi}) - \frac{(x_j-\mu)^2}{2\sigma^2} \right]$$

$$lh = \sum_{n=1}^N \left[\ln(\sigma\sqrt{2\pi}) + \frac{(x_j-\mu)^2}{2\sigma^2} \right]$$

2. **Trobeu els estimadors de màxima versemblança $\hat{\mu}$ i $\hat{\sigma}^2$ per μ i per σ^2 , a partir de la mostra**

Debemos encontrar los valores de σ^2 y μ que hacen que lh sea mínimo (puesto que hemos cambiado el signo a la función de densidad). Para ello su derivada debe ser 0 y su segunda derivada en ese punto debe dar un resultado positivo.

La derivada de la función de máxima verosimilitud (likelihood) respecto de σ^2 es:

$$\begin{aligned}
lh &= \sum_{n=1}^N \left[\ln(\sigma\sqrt{2\pi}) + \frac{(x_j - \mu)^2}{2\sigma^2} \right] \\
\frac{\partial lh}{\partial \sigma^2} &= \sum_{n=1}^N \left[\left(\frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{\sqrt{2\pi}}{2\sigma} \right) + \left(\frac{(x_j - \mu)^2}{2} \left(-\frac{1}{\sigma^4} \right) \right) \right] \\
\frac{\partial lh}{\partial \sigma^2} &= \sum_{n=1}^N \left[\frac{1}{2\sigma^2} - \frac{(x_j - \mu)^2}{2\sigma^4} \right] \\
\frac{\partial lh}{\partial \sigma^2} &= \frac{N}{2\sigma^2} - \frac{1}{2\sigma^4} \sum_{n=1}^N (x_j - \mu)^2
\end{aligned}$$

Entonces para que $\frac{\partial lh}{\partial \sigma^2} = 0$ se tiene que cumplir que:

$$\begin{aligned}
\frac{N}{2\sigma^2} &= \frac{1}{2\sigma^4} \sum_{n=1}^N (x_j - \mu)^2 \\
\frac{N2\sigma^4}{2\sigma^2} &= \sum_{n=1}^N (x_j - \mu)^2 \\
N\sigma^2 &= \sum_{n=1}^N (x_j - \mu)^2 \\
\sigma^2 &= \frac{1}{N} \sum_{n=1}^N (x_j - \mu)^2
\end{aligned}$$

La derivada de la función de máxima verosimilitud (likelihood) respecto de μ es:

$$\begin{aligned}
lh &= \sum_{n=1}^N \left[\ln(\sigma\sqrt{2\pi}) + \frac{(x_j - \mu)^2}{2\sigma^2} \right] \\
\frac{\partial lh}{\partial \mu} &= \sum_{n=1}^N \left[0 + \frac{1}{2\sigma^2} (2(x_j - \mu)(-1)) \right] \\
\frac{\partial lh}{\partial \mu} &= \sum_{n=1}^N \left[-\frac{2(x_j - \mu)}{2\sigma^2} \right] \\
\frac{\partial lh}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{n=1}^N \left[-(x_j - \mu) \right] \\
\frac{\partial lh}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{n=1}^N (\mu - x_j) \\
\frac{\partial lh}{\partial \mu} &= \frac{N\mu}{\sigma^2} - \frac{1}{\sigma^2} \sum_{n=1}^N x_j
\end{aligned}$$

Entonces para que $\frac{\partial lh}{\partial \mu} = 0$ se tiene que cumplir que

$$\frac{N\mu}{\sigma^2} = \frac{1}{\sigma^2} \sum_{n=1}^N x_j$$

$$\frac{N\mu\sigma^2}{\sigma^2} = \sum_{n=1}^N x_j$$

$$N\mu = \sum_{n=1}^N x_j$$

$$\mu = \frac{1}{N} \sum_{n=1}^N x_j$$

Para tener un extremo deberían cumplirse estas condiciones:

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^N (x_j - \mu)^2$$

$$\mu = \frac{1}{N} \sum_{n=1}^N x_j$$

Que son las fórmulas clásicas para medir la media muestral y la varianza.

3. Demostreu que realment són mínims (i no extrems qualsevol)

Para demostrar que realmente son mínimos y no máximos hay que fijarse en la segunda derivada. Como se ha explicado anteriormente, se tratará de mínimos si la imagen de las dos segundas derivadas en esos puntos es positiva, y se tratará de máximos si son negativas.

Las derivadas son:

$$lh = \sum_{n=1}^N \left[\ln(\sigma\sqrt{2\pi}) + \frac{(x_j - \mu)^2}{2\sigma^2} \right]$$

$$\frac{\partial lh}{\partial \sigma^2} = \frac{N}{2\sigma^2} - \frac{1}{2\sigma^4} \sum_{n=1}^N (x_j - \mu)^2$$

$$\frac{\partial^2 lh}{\partial \sigma^4} = \left[\frac{N}{2} \left(-\frac{1}{\sigma^4} \right) \right] - \left[\frac{1}{2} \sum_{n=1}^N (x_j - \mu)^2 \left(-\frac{2}{\sigma^6} \right) \right]$$

$$\frac{\partial^2 lh}{\partial \sigma^4} = -\frac{N}{2\sigma^4} + \frac{1}{\sigma^6} \sum_{n=1}^N (x_j - \mu)^2$$

Evaluamos en qué casos es mayor que 0:

$$-\frac{N}{2\sigma^4} + \frac{1}{\sigma^6} \sum_{n=1}^N (x_j - \mu)^2 > 0$$

$$\frac{1}{\sigma^6} \sum_{n=1}^N (x_j - \mu)^2 > \frac{N}{2\sigma^4}$$

Puesto que σ^6 y N son positivos podemos hacer:

$$\frac{1}{N} \sum_{n=1}^N (x_j - \mu)^2 > \frac{\sigma^6}{\sigma^4}$$

$$\frac{2}{N} \sum_{n=1}^N (x_j - \mu)^2 > \sigma^2$$

Y puesto que lo estamos evaluando en ese punto podemos hacer la sustitución:

$$2\sigma^2 > \sigma^2$$

$$2 > 1$$

Da valores positivos en todos los casos, por lo tanto realmente es un mínimo. En el caso de μ :

$$lh = - \sum_{n=1}^N \left[\ln(f(x_n | \mu, \sigma^2)) \right]$$

$$\frac{\partial lh}{\partial \mu} = \frac{N\mu}{\sigma^2} - \frac{1}{\sigma^2} \sum_{n=1}^N x_j$$

$$\frac{\partial^2 lh}{\partial \mu^2} = \frac{N}{\sigma^2} + 0$$

$$\frac{\partial^2 lh}{\partial \mu^2} = \frac{N}{\sigma^2}$$

Puesto que N y σ^2 son positivos, la derivada segunda también lo es, y por lo tanto se trata de un mínimo.

4. **Calculeu els biaixos dels dos estimadors. Determineu si l'estimador per μ és consistent**
5. **Calculeu la variança de l'estimador per μ de 2 maneres (que han de coincidir), a triar de les 3 següents. Pista: useu que si $X_n, X_m \sim N(\mu, \sigma^2)$, llavors:**

$$\mathbb{E}[X_n \cdot X_m] = \begin{cases} \mu^2 & \text{si } n \neq m \\ \mu^2 + \sigma^2 & \text{si } n = m \end{cases}$$

- (a) Insertant directament el valor de l'estimador i utilitzant les propietats de la variança;
- (b) Usant la coneguda fórmula $Var[\hat{\theta}] = \mathbb{E}[\hat{\theta}] - (\mathbb{E}[\hat{\theta}])^2$;
- (c) Utilitzant la definició directa de la variança $Var[\hat{\theta}] = \mathbb{E}[(\mathbb{E}[\hat{\theta}] - \hat{\theta})^2]$

6. Per determinar la velocitat màxima d'un prototip d'avió es van fer 15 proves independents (i caríssimes!) amb els resultats
422.2, 418.7, 425.6, 420.3, 425.8, 423.1, 431.5, 428.2, 438.3, 434.0, 411.3, 417.2, 413.5, 441.3, 423.0
on els valors venen expressats en m/s. Suposant que aquesta velocitat màxima és gaussiana, quins valors estimaríeu per μ i per σ ?
7. Fixeu unes μ i σ^2 al vostre gust i genereu 1000 mostres aleatòries simples Gaussianes de mida $N = 50$ (noteu que no cal emmagatzemar-les); utilitzeu les mostres (que han de ser i.i.d. i independents entre si) per estimar els biaixos i variàncies dels valors teòrics μ i σ^2 . Compareu els resultats respecte els valors teòrics dels estimadors.