STATISTIQUE MATHÉMATIQUE - MS ENSAI
ÉLEMENTS DE CORRECTION (chapitre 2, sept. 2024)

Estimateur du maximum de vraisemblance pour un échantillon gaussien

Partant d'un n-échantillon gaussien $X = (X_1,\dots,X_n)$ de loi $N(m,\sigma^2)$, on souhaite estimer $\theta = (m,\sigma^2)$. Dans toute la suite, on posera $v = \sigma^2$. La log-vraisemblance de cet échantillon est

$$l(x,\theta) = -\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2 - \frac{n}{2}\ln\sigma^2 - \frac{n}{2}\ln(2\pi) \quad (1)$$

$$= -\frac{1}{2v}\sum_{i=1}^{n}(x_i - m)^2 - \frac{n}{2}\ln v - \frac{n}{2}\ln(2\pi) \quad (2)$$

Le gradient de cette fonction de deux variables en $\theta = (m, v)$ est le vecteur colonne

$$\nabla_l(\theta) = \left(\frac{1}{v}\sum_{i=1}^{n}(x_i - m); \frac{1}{2v^2}\sum_{i=1}^{n}(x_i - m)^2 - \frac{n}{2v}\right)' \quad (3)$$

Le gradient s'annule en un unique point critique $\widehat{\theta}$ dont les coordonnées sont données par

$$m = \overline{x} \text{ et } v = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = s^2 \quad (4)$$

La matrice hessienne, en θ, de la fonction l est donnée par

$$H_l(\theta) = \begin{pmatrix} \frac{\partial^2}{\partial m^2} l(x,\theta) & \frac{\partial^2}{\partial m \partial v} l(x,\theta) \\ \frac{\partial^2}{\partial v \partial m} l(x,\theta) & \frac{\partial^2}{\partial v^2} l(x,\theta) \end{pmatrix} \quad (5)$$

avec

$$\frac{\partial^2}{\partial m^2} l(x,\theta) = -\frac{n}{v} \quad (6)$$

$$\frac{\partial^2}{\partial v^2} l(x,\theta) = -\frac{1}{v^3}\sum_{i=1}^{n}(x_i - m)^2 + \frac{n}{2v^2} \quad (7)$$

$$\frac{\partial^2}{\partial v \partial m} l(x,\theta) = \frac{\partial^2}{\partial m \partial v} l(x,\theta) \quad (8)$$

$$= -\frac{1}{v^2}\sum_{i=1}^{n}(x_i - m) \quad (9)$$

Ces expressions s'entendent à x fixé, x étant une observation du n-échantillon.

Si nous évaluons maintenant cette matrice hessienne au point critique $\widehat{\theta} = (\overline{x}, s^2)$ où $\overline{x}$ et s^2 sont respectivement la moyenne et la variance empirique de l'observation, alors

$$\frac{\partial^2}{\partial m^2} l(x,\widehat{\theta}) = -\frac{n}{s^2} \quad (10)$$

$$\frac{\partial^2}{\partial v^2} l(x,\widehat{\theta}) = -\frac{1}{s^6}\sum_{i=1}^{n}(x_i - \overline{x})^2 + \frac{n}{2s^4} \quad (11)$$

$$= -\frac{ns^2}{s^6} + \frac{n}{2s^4} = -\frac{n}{2s^4} \quad (12)$$

$$\frac{\partial^2}{\partial v \partial m} l(x,\theta) = -\frac{1}{v^2}\sum_{i=1}^{n}(x_i - m) \quad (13)$$

$$= -\frac{1}{s^2}\sum_{i=1}^{n}(x_i - \overline{x}) = 0 \quad (14)$$

Finalement,

$$H_l(\widehat{\theta}) = \begin{pmatrix} -\frac{n}{s^2} & 0 \\ 0 & -\frac{n}{2s^4} \end{pmatrix} \quad (15)$$

Cette matrice est clairement définie négative et par conséquent $\widehat{\theta}$ est l'unique maximum de vraisemblance. L'estimateur du maximum de vraisemblance est donc

$$\widehat{\theta}_{MV} = \left(\overline{X}, S^2\right) \quad (16)$$

Estimateur du maximum de vraisemblance pour la loi uniforme sur $[\theta, \theta+1]$

La vraisemblance d'un échantillon de cette loi s'écrit

$$L(x,\theta) = 1_{[\theta \le x_{(1)} \le x_{(n)} \le \theta + 1]}(x) \quad (17)$$

Cette fonction (de θ) vaut 1 ou 0. Pour une observation fixée x de l'échantillon, elle est nulle si $\theta > x_{(1)}$ ou bien si $\theta < x_{(n)} - 1$. Elle est constante et vaut 1 dans l'intervalle $[x_{(n)} - 1, x_{(1)}]$. Ainsi, toute valeur de θ dans cet intervalle réalise le maximum de la vraisemblance:

Tout $\widehat{\theta}$ choixit dans l'intervalle (aléatoire) $[X_{(n)} - 1, X_{(1)}]$ est un estimtateur du maximum de vraisemblance.

Estimateur gaussien de la moyenne *a posteriori*

À la fin du chapitre 1, nous avons trouvé l'expression de la loi *a posteriori* d'un échantillon gaussien dont la moyenne inconnue suit une loi *a priori* gaussienne, d'hyperparamètres a (pour sa moyenne) et b^2 (pour sa variance). La loi *a posteriori* est gaussienne, de moyenne m et de variance τ^2 avec :

$$\begin{cases} m &= \overline{x}\frac{b^2}{b^2+\sigma^2/n} + a\frac{\sigma^2/n}{b^2+\sigma^2/n} \\ \tau^2 &= \frac{b^2\sigma^2/n}{b^2+\sigma^2/n} \end{cases} \quad (18)$$

m est une combinaison linéaire convexe de la moyenne de l'échantillon et de la moyenne *a priori*. Les coefficients de cette combinaison linéaire sont proportionnels aux variances de l'échantillon et variance *a priori*.

L'estimateur de la moyenne *a posteriori* est donc $\widehat{\theta}$ défini par :

$$\widehat{\theta} = \overline{X}\frac{b^2}{b^2+\frac{\sigma^2}{n}} + a\left(1 - \frac{b^2}{b^2+\frac{\sigma^2}{n}}\right) \quad (19)$$

Estimateur du coefficient de corrélation linéaire empirique

$(X_1,Y_1),\dots,(X_n,Y_n)$ est un n-échantillon d'un vecteur gaussien (X,Y) dont le coefficient de corrélation linéaire est noté ρ. Le coefficient de corrélation empirique est par définition égal à

$$\widehat{\rho}_n = \frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\left[\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 \times \sum_{i=1}^{n}\left(Y_i - \overline{Y}\right)^2\right]^{1/2}} \quad (20)$$

Montrons que $\sqrt{n}(\widehat{\rho}_n - \rho)$ est asymptotiquement normale et que

$$\sqrt{n}(\widehat{\rho}_n - \rho) \rightsquigarrow \mathrm{N}\left(0, (1-\rho^2)^2\right) \tag{21}$$

La méthode de stabilisation de la variance amène à chercher une fonction différentiable ψ dont la dérivée vaut $(1-\rho^2)^{-1}$. On pose donc naturellement $\psi(\rho) = \operatorname{arctanh} \rho$ et une nouvelle application de la méthode delta montre que

$$\sqrt{n}(\psi(\widehat{\rho}_n) - \psi(\rho)) \rightsquigarrow \mathrm{N}(0,1) \tag{22}$$