STATISTIQUE MATHÉMATIQUE - MS ENSAI
ÉLEMENTS DE CORRECTION (chapitre 1, sept. 2025)

Échantillon gaussien et modèle exponentiel

Vous pouvez vous reporter au corrigé de l'exercice 1.3 (dans les corrigés scannés).

Considérons un n-échantillon gaussien $X = (X_1, \ldots, X_n)$ de loi $\mathcal{N}(m, \sigma^2)$. On va considérer trois cas de figures : si m est inconnu et σ connu, en ce cas notre paramètre d'intérêt est $\theta = m$; si σ est inconnu et m connu, en ce cas notre paramètre d'intérêt est $\theta = \sigma^2$; si σ et m sont inconnus, en ce cas notre paramètre d'intérêt est vectoriel, égal à $\theta = (m, \sigma^2) = (m, v)$.

• Le troisième cas $\theta = (m, \sigma^2)$ est traité dans l'exercice 1.3. Dans le corrigé, on choisit comme statistique $T = (T_1, T_2)$ où $T_1(X) = \sum_{i=1}^n X_i$ et $T_2(X) = \sum_{i=1}^n X_i^2$.

La vraisemblance s'écrit, pour $x \in \mathbb{R}^n$,

$$L(x,\theta) = \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2 \right) \right) \tag{1}$$

$$= \underbrace{(2\pi)^{-n/2}}_{h(x)} \exp\left(\frac{m}{\sigma^2} T_1(x) - \frac{1}{2\sigma^2} T_2(x) - \underbrace{\left[\frac{n}{2} \ln \sigma^2 + \frac{nm^2}{2\sigma^2} \right]}_{\beta(\theta)} \right) \tag{2}$$

On obtient la deuxième ligne en développant la première, puis en regroupant les termes qui ne dépendent pas de m et σ (et qui vont constituer $h(x)$), ensuite en regroupant les termes qui ne dépendent que de m et σ (et qui vont constituer $\beta(\theta)$). Ce qui reste doit pouvoir se mettre sous la forme d'un produit scalaire de termes en x et de termes en θ.

On aurait tout aussi bien pu choisir comme statistique $T = (T_1, T_2)$ avec $T_1(X) = \overline{X}$ et $T_2(X) = \overline{X^2}$. On aurait alors eu comme décomposition

$$L(x,\theta) = \underbrace{(2\pi)^{-n/2}}_{h(x)} \exp\left(\frac{nm}{\sigma^2} T_1(x) - \frac{n}{2\sigma^2} T_2(x) - \underbrace{\left[\frac{n}{2} \ln \sigma^2 + \frac{nm^2}{2\sigma^2} \right]}_{\beta(\theta)} \right) \tag{3}$$

On aurait pu encore choisir $T = (T_1, T_2) = (\overline{X}, S^2)$.

• Le cas $\theta = m$ peut se déduire du cas précédent. Il suffit de regrouper les termes qui ne dépendent pas de m (car σ est supposé connu). On obtient, par exemple :

$$L(x,\theta) = \underbrace{(2\pi\sigma^2)^{-n/2} \exp\left(-\frac{n}{2\sigma^2} \sum_{i=1}^n x_i^2 \right)}_{h(x)} \exp\left(\frac{nm}{\sigma^2} \sum_{i=1}^n x_i - \underbrace{\left[\frac{nm^2}{2\sigma^2} \right]}_{\beta(\theta)} \right) \tag{4}$$

La statistique naturelle est alors donnée par $T(x) = \sum_{i=1}^n x_i$.

• Le cas $\theta = \sigma^2$ peut se déduire directement de l'expression initiale de la vraisemblance. Il faut regrouper les termes qui ne dépendent pas de σ :

$$L(x,\theta) = \underbrace{(2\pi)^{-n/2}}_{h(x)} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - m)^2 - \underbrace{\frac{n}{2} \ln \sigma^2}_{\beta(\theta)} \right) \tag{5}$$

La statistique naturelle est alors $T(x) = \sum_{i=1}^n (x_i - m)^2$.

Loi multinomiale comme modèle exponentiel

On rappelle que la vraisemblance s'écrit, pour $x = (x_1, \ldots, x_p)$ et n fixé :

$$L(x,\theta) = h(x) \exp\left(\sum_{i=1}^{p-1} x_i \ln\left(\theta_i / \theta_p\right) - (-n \ln \theta_p) \right) \tag{6}$$

On a donc $\beta(\theta) = -n \ln \theta_p$.

Avant toute chose, nous devons reparamétrer pour mettre le modèle sous forme canonique. Posons

$$\lambda_i = \ln \frac{\theta_i}{\theta_p} = \ln \frac{\theta_i}{1 - \sum_{j=1}^{p-1} \theta_j} \tag{7}$$

En effet, $\sum_{i=1}^p \theta_i = 1$. On a alors

$$\widetilde{\beta}(\lambda) = \beta(\theta) \tag{8}$$

$$= -n \ln\left(1 - \sum_{i=1}^{p-1} \theta_i \right) \tag{9}$$

$$= -n \ln\left(\frac{e^{\lambda_i}}{1 + \sum_{j=1}^{p-1} e^{\lambda_j}} \right) \tag{10}$$

La vraisemblance est alors sous forme canonique :

$$L(x,\lambda) = h(x) \exp\left(\sum_{i=1}^{p-1} x_i \lambda_i - \widetilde{\beta}(\lambda) \right) \tag{11}$$

On peut alors calculer les moments de X_i en dérivant $\widetilde{\beta}$:

$$\mathbb{E}[X_i] = \frac{\partial \widetilde{\beta}}{\partial \lambda_i}(\lambda) \tag{12}$$

$$= \frac{n e^{\lambda_i}}{1 + \sum_j e^{\lambda_j}} \tag{13}$$

$$= \frac{n \theta_i / \theta_p}{1 + \sum_j \theta_j / \theta_p} \tag{14}$$

$$= n\theta_i \tag{15}$$

et de la même façon,

$$\mathbb{V}(X_i) = \frac{\partial^2 \widetilde{\beta}}{\partial \lambda_i^2}(\lambda) \tag{16}$$
$$= n\theta_i(1-\theta_i) \tag{17}$$

puis pour $i \neq j$,

$$\mathrm{cov}(X_i, X_j) = \frac{\partial^2 \widetilde{\beta}}{\partial \lambda_i \partial \lambda_j}(\lambda) \tag{18}$$
$$= -n\theta_i\theta_j \tag{19}$$

Modèle exponentiel : v.a. binomiale, échantillon exponentiel et de Bernoulli

À venir.

Échantillon gaussien dans un cadre bayésien

Le corrigé qui suit est différent de celui du polycopié et essaie de minimiser le nombre de calculs ; pour calculer la loi *a posteriori*, nous allons d'abord calculer la vraisemblance avec une unique observation et supposer que la loi *a priori* est gaussienne, centrée et réduite. Dans un second temps, nous changerons les hyperparamètres et mettrons à jour la vraisemblance *a posteriori*, puis dans un troisième temps, nous remplacerons l'unique observation par un échantillon de taille n.

Supposons donc que θ suit une loi $\mathcal{N}(0,1)$. La densité jointe de (X, θ) est alors

$$f(x,\theta) = \frac{1}{2\pi}\exp\left[-\frac{(x-\theta)^2}{2} - \frac{\theta^2}{2}\right] \tag{20}$$

La marginale en X (que l'on n'est pas obligé de calculer si l'on raisonne proportionnellement aux vraisemblance) est donnée par

$$f_X(x) = \int_{\mathbb{R}} f(x,\theta)d\theta = \frac{1}{2\pi}\int_{\mathbb{R}} \exp^{-(x-\theta)^2/2-\theta^2/2} d\theta \tag{21}$$

De

$$(x-\theta)^2 + \theta^2 = 2(\theta - x/2)^2 - x^2/2 \tag{22}$$

on en déduit que

$$f_X(x) = \frac{e^{-x^2/4}}{2\pi}\int_{-\infty}^{+\infty} \exp^{-(\theta-x/2)^2} d\theta \tag{23}$$
$$= \frac{1}{2\sqrt{\pi}} e^{-x^2/4} \tag{24}$$

On en déduit alors la loi *a posteriori* en utilisant la formule de Bayes :

$$L(\theta|x) = \frac{\frac{1}{2\pi} e^{-(\theta-x/2)^2 - x^2/4}}{\frac{1}{2\sqrt{\pi}} e^{-x^2/4}} \tag{25}$$
$$= \frac{1}{\sqrt{\pi}}\exp^{-(\theta-x/2)^2} \tag{26}$$

Il s'agit de la densité d'une loi normale de moyenne $x/2$ et de variance $1/2$.

Si maintenant on suppose que la loi *a priori* est normale de moyenne a et de variance b^2, on peut utiliser les résultats précédents en utilisant le changement de variables :

$$X = \theta + \sigma Z \tag{27}$$
$$\theta = a + bU \tag{28}$$

où $Z \sim \mathcal{N}(0,1)$ et $U \sim \mathcal{N}(0,1)$. Alors $X \sim \mathcal{N}(\theta, \sigma^2)$. On sait que la loi jointe de (X, θ) est gaussienne. Nous cherchons la loi de θ/X. Or, la loi conditionnelle d'un vecteur gaussien par un sous-vecteur gaussien est elle-même gaussienne et sa moyenne et sa variance sont données par les formules suivantes :

$$\begin{cases} \mathbb{E}[X_2|X_1] = \mathbb{E}[X_2] + \mathrm{cov}(X_2,X_1)\mathbb{V}(X_1)^{-1}(X_1 - \mathbb{E}[X_1]) \\ \mathbb{V}[X_2|X_1] = \mathbb{V}[X_2] - \mathrm{cov}(X_2,X_1)\mathbb{V}(X_1)^{-1}\mathrm{cov}(X_1,X_2) \end{cases}$$

Ici, avec $\mathbb{V}(X) = b^2 + \sigma^2$ et $\mathrm{cov}(\theta, X) = b^2$ les deux formules donnent :

$$\begin{cases} \mathbb{E}[\theta|X] = a + b^2\frac{1}{a^2+b^2}(x-a) \\ \mathbb{V}[X_2|X_1] = b^2 - b^4\frac{1}{a^2+b^2} = \frac{a^2b^2}{a^2+b^2} \end{cases} \tag{29}$$

On en déduit que la loi de $\theta|X$ est gaussienne, de moyenne et variance respective :

$$\begin{cases} m = x\frac{b^2}{a^2+b^2} + a\left(1 - \frac{b^2}{a^2+b^2}\right) \\ \tau^2 = \frac{a^2b^2}{a^2+b^2} \end{cases} \tag{30}$$

Enfin, nous remplaçons l'observation unique x par un échantillon $(x_1, .., x_n)$. Par linéarité, la moyenne de la loi *a posteriori* et la variance sont données par

$$\begin{cases} m = \lambda\overline{x} + (1-\lambda)a \\ \lambda = \frac{b^2}{\sigma^2/n + b^2} \\ \tau^2 = \left(\frac{1}{b^2} + \frac{n}{\sigma^2}\right)^{-1} \end{cases} \tag{31}$$