DSXS-315: Statistique bayésienne

DM

A rendre pour le 05/02/2026
Les parties 2 et 3 sont des parties préliminaires au BE du 05/02/2026.

The use of any IA tool is forbidden for this homework. You can use human intelligence and discuss it among you. You are asked to explicitly acknowledging this fact by handwriting the following sentence at the beginning of your homework :

I hereby certify that I did not use any IA tool for this homework.

and then sign.

Première partie : Modèle binomial

On reprend l'exemple du premier cours. Une boule de billard W roule sur une ligne de longueur un, avec une probabilité uniforme de s'arrêter n'importe où. Supposons qu'elle s'arrête en θ. Une deuxième boule O roule alors n fois dans les mêmes conditions, et on note $X_i = 1$ si la i-ème boule s'arrête à gauche de W, et $X_i = 0$ sinon. On notera $\mathbf{X} = (X_1, \dots, X_n)$ et $S = X_1 + \cdots + X_n$.

I.1. Montrez que le modèle paramétrique est donné par $f(x \mid \theta) = \theta^s(1-\theta)^{n-s}$ avec $s = \sum x_i$. Quelle est la loi a priori ?

I.2. Calculez l'estimateur du maximum de vraisemblance du modèle, que l'on notera dans la suite $\hat{\theta}$.

I.3. Montrez que S est une statistique exhaustive, i.e., que la loi des données conditionnellement à S ne dépend pas de θ.

I.4. Calculez la loi a priori de Jeffreys.

On rappelle que la loi beta de paramètre $\alpha, \beta > 0$, notée $\mathcal{B}e(\alpha, \beta)$ est la loi de densité

$$f_{\alpha,\beta}(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \mathbb{1}\left(0 \le x \le 1\right).$$

Sa moyenne vaut $\frac{\alpha}{\alpha+\beta}$ et son mode $\frac{\alpha-1}{\alpha+\beta-2}$.

I.5. Calculez et identifiez la loi a posteriori pour une loi a priori $\mathcal{B}e(\alpha, \beta)$. La famille des lois beta est-elle conjuguée pour le modèle paramétrique considéré ?

I.6. On considère maintenant une loi a priori $\mathcal{B}e(\alpha, \beta)$. Calculez la moyenne a posteriori $\mathbb{E}(\theta \mid \mathbf{X})$ et le maximum a posteriori (MAP) $\arg\max_\theta \pi(\theta \mid \mathbf{X})$.

I.7. Le MAP est-il consistent ? asymptotiquement normal ? Mêmes questions pour la moyenne a posteriori. Vous justifierez vos réponses.
Dans la suite de l'exercice, on considère une loi a priori uniforme et

$$Z = \frac{\sqrt{n}}{\omega}(\eta - \hat{\theta}) \quad \text{avec} \quad \omega = \sqrt{\hat{\theta}(1-\hat{\theta})},$$

où η une variable aléatoire qui, conditionnellement aux observations $\mathbf{X}$, est distribuée selon la loi a posteriori $\pi(\cdot \mid \mathbf{X})$.

Version du 27 janvier 2026.

I.8. Soit $\pi'(\cdot \mid \mathbf{X})$ la loi de Z conditionnellement aux observations $\mathbf{X}$: montrez que

$$\pi'(z \mid \mathbf{X}) = \frac{\omega}{\sqrt{n}}\pi\left(\hat{\theta} + \frac{\omega z}{\sqrt{n}} \mid \mathbf{X}\right)$$

et déduisez-en à l'aide de la question 5 que

$$\pi'(z \mid \mathbf{X}) \propto \left(\hat{\theta} + \frac{\omega z}{\sqrt{n}}\right)^{n\hat{\theta}} \left(1 - \hat{\theta} - \frac{\omega z}{\sqrt{n}}\right)^{n(1-\hat{\theta})}.$$

I.9. Déduisez de la question précédente que conditionnellement aux observations $\mathbf{X}$, Z converge en loi vers une loi normale standard lorsque $n \to \infty$.

I.10. Quel résultat du cours aurait pu vous permettre de retrouver directement le résultat de la question précédente ? Interprétez le facteur de normalisation ω à l'aune de ce résultat.

Deuxième partie : Dimension 1, modèle gaussien à moyenne et variance inconnues

La loi Gamma de paramètre $\alpha, \beta > 0$, notée $\mathcal{G}(\alpha, \beta)$, est la loi sur $\mathbb{R}$ de densité

$$g(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \mathbb{1}\left(x > 0\right).$$

Pour une loi gaussienne $N(\mu, \sigma^2)$ de moyenne μ et de variance σ^2, on définit la précision τ comme l'inverse de la variance : $\tau = 1/\sigma^2$. Dans cet exercice on considère le modèle Gaussien à moyenne et précision inconnues, i.e., $f(\,\cdot \mid \mu, \theta)$ la densité de la loi normale $N(\mu, 1/\tau)$.

II.1. Explicitez la vraisemblance $f(x \mid \mu, \tau)$ du modèle pour un échantillon $x = (x_1, \ldots, x_n)$ de taille $n \geq 1$.

II.2. Rappelez une famille de lois a priori conjuguées pour le modèle gaussien à variance connue.

II.3. Montrez que les lois Gamma forment une famille de lois conjuguées pour le modèle paramétrique gaussien à moyenne connue et paramétrée par la précision τ.

II.4. On retourne maintenant au modèle paramétrique $\{f(\cdot \mid \mu, \tau) : \mu, \tau)\}$ complet. Au vu des deux questions précédentes, justifiez succinctement (quelques phrases maximum, et pas de calcul) le choix de la loi a priori

$$\pi(\mu, \tau) \propto \exp\left(-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right) \times \tau^{\alpha+1} \exp\left(-\beta\tau\right).$$

Donnez une interprétation simple de cette loi.

II.5. Cette famille de lois a priori est-elle conjuguée pour le modèle gaussien considéré ?

II.6. Montrez que si $\tau \sim \mathcal{G}(\alpha, \beta)$ et $\mu \mid \tau \sim \mathcal{N}(\mu_0, 1/(n_0\tau))$, alors

$$\mu \mid \tau, x \sim \mathcal{N}\left(\frac{n}{n+n_0}\overline{x} + \frac{n_0}{n+n_0}\mu_0, \frac{1}{(n+n_0)\tau}\right)$$

et

$$\tau \mid x \sim \mathcal{G}\left(\alpha + \frac{n}{2}, \beta + \frac{1}{2}\sum_{i=1}^n (x_i - \bar{x})^2 + \frac{nn_0}{2(n+n_0)}\left(\bar{x} - \mu_0\right)^2\right).$$

avec $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$ la moyenne empirique. Cette famille de lois a priori est-elle conjuguée pour le modèle gaussien considéré ?

Troisième partie : Modèle gaussien à variance connue en dimension générale

Dans cette partie on fixe $d, n \geq 1$ et une matrice $X \in \mathbb{R}^n \times \mathbb{R}^d$ et on considère le modèle paramétrique

$$f(y \mid \beta, \sigma) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y - X\beta\|^2\right), \ y \in \mathbb{R}^n,$$

pour $\beta \in \mathbb{R}^d, \sigma > 0$ et où $y \in \mathbb{R}^n$ est l'échantillon complet des observations.

<u>III.1.</u> Quel est le nom de la densité $f(\cdot \mid \beta, \sigma)$?

<u>III.2.</u> Montrez que pour toute matrice positive M et toute matrice définie positive M_0, $M + M_0$ est inversible.

<u>III.3.</u> On considère un a priori gaussien de moyenne $\beta_0 \in \mathbb{R}^d$ et de matrice de variance-covariance $\sigma^2 M_0^{-1}$, avec $M_0 \in \mathbb{R}^d \times \mathbb{R}^d$ une matrice symétrique définie positive. Montrez que la densité a posteriori est la loi normale de moyenne $(X^\top X + M_0)^{-1}(X^\top y + M_0\beta_0)$ et de matrice de variance-covariance $\sigma^2(X^\top X + M_0)^{-1}$.
<u>Indication :</u> vous pourrez utiliser l'identité

$$\beta^\top A\beta - 2\beta^\top v = (\beta - A^{-1}v)^\top A(\beta - A^{-1}v) - v^\top A^{-1}v$$

où A est une matrice symétrique définie positive et v un vecteur.

<u>III.4.</u> Calculez le gradient de J où $J(\beta) = \|y - X\beta\|^2$.