Sorbonne Université, Master 1
Cours : A. Ben-Hamou

MU4MA015, Statistique, 2023-2024
TD : A. Godichon, A. Guyader et M. Sangnier

TD 7 : Approche bayésienne

Exercice 1 (Lois conditionnelles I)
Soient X et Y deux variables aléatoires réelles de loi jointe de densité sur $\mathbb{R}^2$:

$$(x, y) \mapsto \frac{1}{2\pi} \exp\left(-x^2 + xy - \frac{y^2}{2}\right).$$

1. Déterminer :
   (a) les lois marginales de X et Y ;
   (b) la loi conditionnelle de Y sachant $X = x$ et celle de X sachant $Y = y$.
2. Vérifier que la loi de (X, Y) est celle d'un vecteur gaussien sur $\mathbb{R}^2$, dont on précisera la moyenne et la matrice de variance-covariance.

Exercice 2 (Lois conditionnelles II)
Soient X et Y deux variables aléatoires indépendantes et $S = X + Y$. Dans chacun des trois cas suivants, déterminer la loi de S, la loi de X sachant S, ainsi que l'espérance conditionnelle $\mathbb{E}[X \mid S]$.

1. $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$, avec $\lambda, \mu > 0$.
2. $X \sim \Gamma(r, \lambda)$ et $Y \sim \Gamma(t, \lambda)$, avec $r, t, \lambda > 0$.
3. $X \sim \mathcal{B}(n, p)$ et $Y \sim \mathcal{B}(m, p)$, avec $n, m \in \mathbb{N}^*$ et $p \in]0, 1[$.

Indication :
— Si X suit une loi de Poisson de paramètre λ (respectivement une loi Gamma de paramètres (r, λ)), alors sa fonction caractéristique vaut $\exp(\lambda(e^{it} - 1))$ (respectivement $(1 - it/\lambda)^{-r}$).
— Si $p \in [0, 1]$ et $0 \leq n \leq N$ sont deux entiers naturels avec pN entier, on dit que X suit une loi hypergéométrique de paramètres (N, p, n) si $\mathbb{P}(X = k) = \binom{pN}{k}\binom{(1-p)N}{n-k}/\binom{N}{n}$ si $\max(0, n - (1 - p)N) \leq k \leq \min(pN, n)$, auquel cas son espérance vaut np.

Exercice 3 (Tirages à pile ou face)
On considère une expérience de tirage de pile–ou–face. Soit P_θ la loi de Bernoulli de paramètre $\theta \in]0, 1[$. Le modèle considéré est $\mathcal{P} = \{P_\theta,\ \theta \in \Theta\}$, avec $\Theta \subset]0, 1[$. On dispose par ailleurs d'observations $X_1, \ldots, X_n$ à valeurs dans $\{0, 1\}$.

1. Dans cette question, $\Theta = \{a, b\}$, avec $0 < a < b < 1$. On note Π la loi sur l'ensemble Θ définie par

$$\Pi[\{a\}] = \Pi[\{b\}] = \frac{1}{2}.$$

   On pose le modèle suivant pour le cas d'une observation ($n = 1$) :

$$\boldsymbol{\theta} \sim \Pi \qquad \text{et} \qquad X \mid \boldsymbol{\theta} \sim P_{\boldsymbol{\theta}}\,.$$

   (a) Calculer la loi a posteriori de $\boldsymbol{\theta} \mid X$.
   (b) Donner la loi jointe de $(X, \boldsymbol{\theta})$ et la loi marginale de X.
   (c) On pose le modèle suivant pour les observations $\mathbf{X} = (X_1, \ldots, X_n)$ et pour $\boldsymbol{\theta}$:

$$\boldsymbol{\theta} \sim \Pi \qquad \text{et} \qquad \mathbf{X} \mid \boldsymbol{\theta} \sim P_{\boldsymbol{\theta}}^{\otimes n}\,.$$

   Déterminer la loi a posteriori $\boldsymbol{\theta} \mid \mathbf{X}$ ainsi que l'espérance a posteriori $\mathbb{E}[\boldsymbol{\theta} \mid \mathbf{X}]$.

2. Thomas Bayes (1763) considère le problème suivant. Une boule de billard roule sur une ligne de longueur 1, avec une probabilité uniforme de s'arrêter en un point. Supposons qu'elle s'arrête en θ. Une deuxième boule roule n fois dans les mêmes conditions, et on note $X_i = 1$ si le i-ème lancer s'est arrêté avant la première boule (et $X_i = 0$ sinon). Bayes se demande : connaissant $\mathbf{X}$, quelle inférence peut-on mener sur θ ?
   (a) Ecrire cette expérience dans un formalisme bayésien.
   (b) Répondre à la question de Bayes en calculant la densité a posteriori.

Exercice 4 (Lois conjuguées)
Montrer que les familles de lois a priori suivantes sont conjuguées, pour $n \geq 1$. De plus, on donnera dans chaque cas l'expression de la moyenne a posteriori.

1. La famille des lois Gamma(a, b), $a > 0$, $b > 0$, pour $\mathcal{P} = \{\mathcal{E}(\lambda)^{\otimes n},\ \lambda > 0\}$.
2. La famille des lois Beta(a, b), $a > 0$, $b > 0$, pour $\mathcal{P} = \{\mathcal{B}(p)^{\otimes n},\ p \in]0, 1[\}$.
3. La famille des lois gaussiennes $\mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$, pour $\mathcal{P} = \{\mathcal{N}(\theta, 1)^{\otimes n},\ \theta \in \mathbb{R}\}$.

Exercice 5 (Empirical Bayes et lois exponentielles) Soit $\mathcal{P} = \{P_\theta^{\otimes n},\ \theta > 0\}$, avec $P_\theta = \mathcal{E}(\theta)$ (loi exponentielle). On dispose d'observations $\mathbf{X} = (X_1, \ldots, X_n)$ de loi $P_{\boldsymbol{\theta}}^{\otimes n}$ sachant $\boldsymbol{\theta}$, et l'on cherche à choisir une loi a priori sur $\boldsymbol{\theta}$ parmi l'ensemble des lois $\Pi_\lambda = \mathcal{E}(\lambda)$ pour $\lambda > 0$. On se propose de déterminer λ par une méthode bayésienne empirique.

1. En quoi consiste cette méthode ?
2. On va construire $\widehat{\lambda}$ par la méthode du maximum de vraisemblance marginal. Rappeler le principe de cette méthode.
3. Soit $\lambda > 0$. Montrer que, lorsque $\boldsymbol{\theta} \sim \Pi_\lambda$, la loi marginale de X_1 a pour densité $f_\lambda(x) = \frac{\lambda}{(\lambda+x)^2}\mathbf{1}_{x \geq 0}$.
4. Soit $\lambda > 0$ et $\boldsymbol{\theta} \sim \Pi_\lambda$. Calculer la densité marginale de $\mathbf{X} = (X_1, \ldots, X_n)$.
5. En déduire que l'estimateur du maximum de vraisemblance marginal est $\widehat{\lambda} = \overline{X}_n$. Quelle est la pseudo-loi a posteriori suggérée par la méthode bayésienne empirique ?

Exercice 6 (Empirical Bayes et lois normales)
On se place dans le cadre du modèle gaussien $\mathcal{P} = \{P_\theta^{\otimes n},\ \theta \in \mathbb{R}\}$ avec $P_\theta = \mathcal{N}(\theta, 1)$. On dispose d'observations $\mathbf{X} = (X_1, \ldots, X_n)$ de loi $P_{\boldsymbol{\theta}}^{\otimes n}$ sachant $\boldsymbol{\theta}$, et l'on cherche à choisir une loi a priori sur $\boldsymbol{\theta}$ parmi l'ensemble des lois $\Pi_\mu = \mathcal{N}(\mu, 1)$ pour $\mu \in \mathbb{R}$. On se propose de déterminer μ par une méthode bayésienne empirique.

1. Montrer que la loi marginale de $\mathbf{X}$ est celle d'un vecteur gaussien et déterminer $\widehat{\mu}$, l'estimateur du maximum de vraisemblance marginal.
2. Quelle est la pseudo-loi a posteriori suggérée par la méthode bayésienne empirique ?

Exercice 7 (Laplace, mélange et bayésien hiérarchique)
On considère l'expérience botanique suivante : une plante située à l'origine de l'axe des abscisses émet une graine qui est ensuite soumise à un mouvement aléatoire dû à l'effet d'un vent de force T et atterrit à l'abscisse $X \,\big|\, T \sim \mathcal{N}(0, 2T)$. La force du vent est inconnue et modélisée par une loi exponentielle $T \sim \mathcal{E}(\lambda)$, $\lambda > 0$.

1. Ecrire la densité (marginale) f de X comme une intégrale que l'on ne cherchera pas à calculer.
2. Calculer la fonction génératrice $M_X(s) = \mathbb{E}[e^{sX}]$ pour tout s tel que $|s| < \sqrt{\lambda}$.
3. Montrer que si Y suit une loi de Laplace de paramètre $\mu > 0$, i.e. $f_Y(y) = \frac{\mu}{2}e^{-\mu|y|}$, alors $M_Y(s) = \mu^2/(\mu^2 - s^2)$, pour tout $|s| < \mu$.
4. Identifier la loi marginale de X et interpréter le résultat dans un contexte bayésien hiérarchique : quelle loi a priori correspond à la hiérarchie $\theta \,\big|\, T \sim \mathcal{N}(0, 2T)$ et $T \sim \mathcal{E}(\alpha)$ pour α une constante positive fixée ?

Exercice 8 ($\star$) (Famille de Dirichlet et modèle multinomial)
On rappelle que pour $a = (a_1, \dots, a_K) \in {\mathbb{R}_+^*}^K$, la loi de Dirichlet $\mathrm{Dir}(a)$ est la loi sur le simplexe $\mathcal{S}_K = \{x = (x_1, \dots, x_K) \in [0,1]^K,\ \sum_{k=1}^K x_k = 1\}$ de densité (par rapport à la mesure de Lebesgue sur $\mathcal{S}_K$)

$$x \mapsto \frac{1}{B(a)} \prod_{k=1}^K x_k^{a_k - 1}, \qquad B(a) = \frac{\prod_{k=1}^K \Gamma(a_k)}{\Gamma(\sum_{k=1}^K a_k)}.$$

1. (a) Montrer que si $Z = (Z_1, \dots, Z_K) \sim \mathrm{Dir}(a)$, alors Z_1 suit une loi Beta que l'on déterminera (on pourra considérer le changement de variables $x_i = (1 - x_1) y_i$, pour $2 \le i \le K - 1$).
   (b) Montrer que la famille des lois de Dirichlet $\mathrm{Dir}(a_1, \dots, a_K)$ est conjuguée pour le modèle

$$\mathcal{P} = \left\{ \left(\sum_{k=1}^K p_k \delta_k \right)^{\otimes n},\ p = (p_1, \dots, p_K) \in \mathcal{S}_K \right\}.$$

2. Soient $X_1, \dots, X_n$ des variables i.i.d. à valeurs dans $\{1, 2, \dots, K\}$, de loi $p = (p_1, \dots, p_K)$ (i.e. $\mathbb{P}(X_1 = j) = p_j$). Notons, pour $j = 1, \dots, K$,

$$N_j = \sum_{i=1}^n \mathbb{1}_{X_i = j}.$$

   (a) Quelle est la loi de N_j pour $j \in \{1, \dots, K\}$?
   (b) Montrer que la loi jointe des N_j est donnée par

$$\mathbb{P}\left(N_j = n_j,\ 1 \le j \le K\right) = \binom{n}{n_1, \dots, n_K} \prod_{j=1}^K p_j^{n_j} \mathbb{1}_{\{\sum_{j=1}^K n_j = n\}},$$

   où $\binom{n}{n_1, \dots, n_K} = n!/(n_1! \dots n_K!)$ désigne le nombre de façons de regrouper n éléments en K groupes de tailles respectives $n_1, \dots, n_K$.
   On dit que le vecteur $(N_1, \dots, N_K)$ suit la loi multinomiale de paramètres $(n, p_1, \dots, p_K)$.
   (c) Montrer que la famille des lois de Dirichlet $\mathrm{Dir}(a)$, $a \in (\mathbb{R}_+^*)^K$, est conjuguée pour le modèle des lois multinomiales $\mathcal{P} = \{\mathrm{Mult}(n, p_1, \dots, p_K),\ p = (p_1, \dots, p_K) \in \mathcal{S}_K\}$.

Exercice 9 ($\star$) (A posteriori séquentiel et information)
On se place dans le cadre bayésien suivant :

$$\begin{aligned} \boldsymbol{\theta} &\sim \Pi, & \mathrm{d}\Pi &= \pi \, \mathrm{d}\nu, \\ X_1, \dots, X_n \,\big|\, \boldsymbol{\theta} &\sim P_{\boldsymbol{\theta}}^{\otimes n}, & \mathrm{d}P_\theta &= p_\theta \, \mathrm{d}\mu. \end{aligned}$$

1. Montrer que la loi a posteriori de $\boldsymbol{\theta}$ sachant (X_1, X_2), dans le modèle d'origine où $\boldsymbol{\theta}$ suit la loi a priori Π, coïncide avec la loi a posteriori de $\boldsymbol{\theta}$ sachant X_2 dans le modèle où l'on prend $\Pi[\cdot \,|\, X_1]$ comme loi a priori.
2. Généraliser le résultat précédent à $\boldsymbol{\theta} \,|\, X_1, \dots, X_n$. Est-ce différent de prendre l'a posteriori directement par rapport à toutes les observations ou séquentiellement observation par observation ?
3. L'ordre des X_i dans le conditionnement $\boldsymbol{\theta} \,|\, X_1, \dots, X_n$ a-t-il de l'importance ? Qu'en serait-il si $X_1, \dots, X_n \,|\, \boldsymbol{\theta}$ ne suivait pas une loi produit ?

Exercice 10 (Estimateur de Laplace-Bayes)
Une urne contient N boules, dont r boules rouges et $N - r$ boules noires, avec r inconnu. On tire n boules sans remise dans l'urne, $n \le N$, et l'on note X le nombre de boules rouges tirées. On considère le cadre bayésien suivant :

$$\begin{aligned} R &\sim \mathrm{Unif}(\{0, 1, \dots, N\}) \\ X \,\big|\, R &\sim P_R. \end{aligned}$$

1. Sachant R, déterminer la loi P_R qui correspond au modèle décrit.
2. Déterminer la loi a posteriori de R sachant X. Indication : avec les conventions usuelles sur les coefficients binomiaux, on a la relation
$$\sum_{k=0}^{N} \binom{k}{m}\binom{N-k}{n-m} = \binom{N+1}{n+1}.$$
3. Quelle est la probabilité, sachant X, que le $(n+1)^{\text{ième}}$ tirage donne une boule rouge ?

Exercice 11 (Identifiabilité)
On s'intéresse à une colonie d'insectes pondeurs d'œufs. On suppose que pour le site observé, le nombre X d'œufs suit une loi de Poisson de paramètre $\lambda > 0$ et que chaque œuf pondu donne, indépendamment des autres et de X, un insecte avec une probabilité $p \in]0,1[$. On note Y le nombre de naissances, de sorte qu'on peut écrire
$$Y = \sum_{i=1}^{X} \varepsilon_i$$
où les ε_i sont i.i.d. de loi de Bernoulli de paramètre p.

1. On suppose qu'on observe à la fois X et Y.
   (a) Déterminer la loi conditionnelle de Y sachant $X = x$.
   (b) Écrire le modèle statistique associé. Est-il identifiable ?
   (c) Supposons que l'on observe des couples indépendants $(X_1, Y_1), \ldots, (X_n, Y_n)$, de même loi que (X, Y). Proposer un estimateur de λ et de p.
2. On n'observe plus que Y.
   (a) Quelle est la loi de Y ?
   (b) Le modèle associé est-il identifiable si le paramètre d'intérêt est toujours le couple (λ, p) ?

Exercice 12 (A priori de Jeffreys)
Soit $\mathcal{P} = \{P_\theta,\, \theta \in \Theta\}$ un modèle régulier.

1. Rappeler la définition de l'a priori de Jeffreys.
2. Déterminer l'a priori de Jeffreys dans chacun des cas suivants, en vérifiant qu'il s'agit bien de modèles réguliers :
   (a) $\Theta = \mathbb{R}$ et $P_\theta = \mathcal{N}(\theta, 1)$ (loi normale) ;
   (b) $\Theta =]0,1[$ et $P_\theta = \mathcal{B}(\theta)$ (loi de Bernoulli) ;
   (c) $\Theta = \mathbb{R}_+^*$ et $P_\theta = \mathcal{P}(\theta)$ (loi de Poisson).