# Feuille de travaux dirigés 4 : Modélisation bayésienne

**Exercice 1** (Modèle gaussien)**:**
On considère le modèle bayésien suivant sur $\mathcal{X} = \mathbb{R}$ :

$$\begin{cases} X|\theta \sim \mathcal{N}(\theta, \sigma^2) \\ \boldsymbol{\theta} \sim \mathcal{N}(\mu_0, \tau_0^2) \end{cases}$$

où $\sigma^2, \mu_0$ et $\tau_0^2$ sont des constantes supposées connues.

1. Quel est l'espace des paramètres ? Donnez la loi a posteriori $\pi(\theta|x)$.
2. On considère maintenant un échantillon i.i.d. $Y = (X_1, \dots, X_n)$, où $X_i \sim X$. Quelle est le modèle pour $Y$ ? Donnez la loi a posteriori $\pi(\theta|y)$. Que se passe-t-il lorsque $n \to \infty$ ?

**Exercice 2** (Mélange d'opinions)**:**
Une expérience aléatoire a deux résultats possibles (succès ou échec). On note $X$ la variable aléatoire valant 1 en cas de succès, 0 en cas d'échec. $X$ est supposée suivre une loi de Bernoulli de paramètre $\theta$ inconnu. On considère seulement deux valeurs possibles pour $\theta$ : $\theta \in \{\theta_1 = 0.2, \theta_2 = 0.6\}$.

1. Écrire le modèle statistique, détailler l'espace des paramètres.
2. Le premier expert accorde une confiance égale en les deux possibilités pour $\theta$. Autrement dit, son prior est $\pi_1(\theta_1) = \pi_1(\theta_2) = 0.5$. Donnez la loi a posteriori $(\pi_1(\theta_i|x = 1))_{i=1,2}$ et $\pi_1(\theta_i|x = 0)_{i=1,2}$.
3. Même question pour un deuxième expert qui croit a priori plus à la seconde alternative : son prior est $\pi_2(\theta_1) = 1/4$, $\pi_2(\theta_2) = 3/4$.
4. La loi *predictive a posteriori* (sachant l'observation $X = x$) est par définition, la loi sur $\mathcal{X}$ dont la densité par rapport à la mesure de référence est donnée par

$$p(y) = \int_\Theta p(y|\theta)\pi(\mathrm{d}\theta|x),$$

   où $\pi(\,\cdot\,|x)$ est la loi a posteriori. Quelle est cette prédictive a posteriori pour le prior $\pi_2$, lorsque $x = 1$ ?

On observe maintenant un échantillon *i.i.d.* $X = (X_1, \dots, X_n)$.

5. Montrer que les lois a posteriori $\pi(\theta|x)$ ne dépendent que de $s = \sum_{j=1}^n x_j$.
6. On suppose que $s = n/2$. Écrire la loi a posteriori $\pi_2(\theta|x)$ $(x \in \{0,1\}^n, \sum_i x_i = s, \theta \in \{\theta_1, \theta_2\})$ pour l'a priori $\pi_2$. Que se passe-t-il lorsque $n \to \infty$ ? Même question pour l'a priori $\pi_1$. Plus généralement, le comportement lorsque $n$ tend vers l'infini dépend-il de l'a priori ?

.

**Exercice 3:**
Une modélisation classique du trafic téléphonique repose sur l'hypothèse suivante, dite Poissonienne : le nombre d'appels des abonnés pendant un intervalle de temps $\Delta t$ est supposé

suivre une loi de Poisson, et les nombres d'appels sur des intervalles de temps disjoints sont indépendants. Soit $\mathbf{X} = (X_1, \ldots, X_n)$ les nombres d'appels observés sur $n \in \mathbb{N}^*$ intervalles de temps successifs. L'hypothèse est alors que les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes et identiquement distribuées sur un espace probabilisé $\Omega$ selon une loi de Poisson notée $\mathcal{P}oiss(\theta)$, avec $\theta \in \Theta := ]0, \infty[$ inconnu. n dispose maintenant d'une information a priori $\pi$ sur le paramètre $\theta$. On note $\pi$ la loi a priori et $\boldsymbol{\theta} \sim \pi$ la variable aléatoire à valeurs dans $\Theta = ]0, \infty[$ associée. On choisit pour $\pi$ une loi Gamma, $\pi = \mathcal{G}amma(a, \lambda)$ avec $a > 0$, $\lambda > 0$ des quantités fixées par l'utilisateur en fonction de sa connaissance a priori sur $\boldsymbol{\theta}$. Soit $\pi(\,\cdot\,|\mathbf{x})$ la loi a posteriori sachant l'observation $\mathbf{X}(\omega) = \mathbf{x} \in \mathcal{X}^n$, pour un certain $\omega \in \Omega$. Dans la suite, on notera également $\pi(\theta)$ la densité de la loi a priori évaluée en $\theta \in \Theta$ et $\pi(\theta|\mathbf{x})$ la densité de la loi a posteriori.

1. Montrer que pour tout $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{X}^n$, la loi a posteriori $\pi(\,\cdot\,|\mathbf{x})$ est une loi Gamma dont on précisera les paramètres en fonction de $\mathbf{x}, a, \lambda$.
2. Quelle est l'espérance a posteriori $\mathbb{E}(\boldsymbol{\theta}|\mathbf{x})$ ?
3. On appelle $\theta_0$ le « vrai » paramètre (inconnu) régissant les observations. En utilisant la loi des grands nombres, montrer que $\mathbb{E}(\boldsymbol{\theta}|\mathbf{X})$ converge presque sûrement vers une limite que l'on précisera, lorsque $n \to \infty$.
4. La variance a posteriori est définie par $\mathbb{V}\mathrm{ar}(\boldsymbol{\theta}|\mathbf{x}) = \mathbb{V}\mathrm{ar}(Y)$ où $Y$ suit la loi a posteriori $\pi(\,\cdot\,|\mathbf{x})$. Que vaut $\mathbb{V}\mathrm{ar}(\theta|\mathbf{X})$ ? Montrer que $\mathbb{V}\mathrm{ar}(\boldsymbol{\theta}|\mathbf{X})$ converge presque sûrement vers une limite que l'on précisera.