Résumé

1. On suppose $X_1, \ldots, X_n$ des observations i.i.d suivant une loi paramétrée $\mathbb{P}_{\theta}$.

3. Le modèle Bayésien met à jour la distribution sur θ avec les données et obtient la loi a posteriori $\mathbb{P}(\theta | X_1, \ldots, X_n)$ avec le théorème de Bayes: $\mathbb{P}(\theta | X_1, \ldots, X_n) \propto \mathbb{P}(X_1, \ldots, X_n | \theta) \mathbb{P}(\theta)$

5. On peut très bien considérer des lois a priori impropres (ex. uniforme sur $\mathbb{R}$: $\pi(\theta) \propto 1$ sur $\mathbb{R}$) ...

4. Une loi de densité proportionnelle à g est dite impropre si $\int g = +\infty$.

6. ... si la loi a posteriori est propre c-à-d que **la constante de normalisation** $\int \mathbb{P}(X_1, \ldots, X_n|\theta)\mathbb{P}(\theta)\mathrm{d}\theta$ est finie.

7. On dit que la loi a priori est conjuguée pour le modèle $X|\theta$ si elle appartient à la même famille de la loi a posteriori.

8. L'estimateur de Bayes correspond à la moyenne de cette distribution appelée moyenne a posteriori.

9. BvM: La loi a posteriori est asymptotiquement normale de moyenne $\widehat{\theta}_{MV}$ et de variance $I(\theta)^{-1}$.

10. Ainsi, si $n \to +\infty$, la loi a priori devient négligeable: on retrouve l'approche fréquentiste.

11. Calculer cet estimateur nécessite de connaître la constante de normalisation pour avoir une densité.

12. Si le modèle est conjugué, alors ce calcul est déjà connu.

14. Avec des échantillons $\theta_1, \ldots, \theta_m \sim \theta | X$, on peut estimer une moyenne a posteriori empirique $\frac{1}{m} \sum_{i=1}^{m} \theta_i$.

13. Si ce n'est pas le cas, la constante de normalisation est très souvent intractable (difficile à calculer).

15. Pour générer de tels échantillons, on utilise des méthodes dites de Monte Carlo.

Abus de notation: $\mathbb{P}$ d'une variable aléatoire désigne sa densité (continue ou discrète)

2. Le modèle Bayésien considère que θ est une **variable aléatoire** à modéliser par une loi a priori $\mathbb{P}(\theta) = \pi$, la distribution des données (vraisemblance) devient alors conditionnelle: $\mathbb{P}(X|\theta)$.

Résumé

1. On suppose $X_1, \ldots, X_n$ des observations i.i.d suivant une loi paramétrée $\mathbb{P}_\theta$.
2. Le modèle Bayésien considère que θ est une **variable aléatoire** à modéliser par une loi a priori $\mathbb{P}(\theta) = \pi$, la distribution des données (vraisemblance) devient alors conditionnelle: $\mathbb{P}(X|\theta)$.
3. Le modèle Bayésien met à jour la distribution sur θ avec les données et obtient la loi a posteriori $\mathbb{P}(\theta|X_1, \ldots, X_n)$ avec le théorème de Bayes: $\mathbb{P}(\theta|X_1, \ldots, X_n) \propto \mathbb{P}(X_1, \ldots, X_n|\theta)\mathbb{P}(\theta)$
4. Une loi de densité proportionnelle à g est dite impropre si $\int g = +\infty$.

> Abus de notation: $\mathbb{P}$ d'une variable aléatoire désigne sa densité (continue ou discrète)

5. On peut très bien considérer des lois a priori impropres (ex. uniforme sur $\mathbb{R}$: $\pi(\theta) \propto 1$ sur $\mathbb{R}$) ...
6. ... si la loi a posteriori est propre c-à-d que **la constante de normalisation** $\int \mathbb{P}(X_1, \ldots, X_n|\theta)\mathbb{P}(\theta)\mathrm{d}\theta$ est finie.
7. On dit que la loi a priori est conjuguée pour le modèle $X|\theta$ si elle appartient à la même famille de la loi a posteriori.
8. L'estimateur de Bayes correspond à la moyenne de cette distribution appelée moyenne a posteriori.
9. BvM: La loi a posteriori est asymptotiquement normale de moyenne $\widehat{\theta}_{MV}$ et de variance $I(\theta)^{-1}$.
10. Ainsi, si $n \to +\infty$, la loi a priori devient négligeable: on retrouve l'approche fréquentiste.
11. Calculer cet estimateur nécessite de connaître la constante de normalisation pour avoir une densité.
12. Si le modèle est conjugué, alors ce calcul est déjà connu.
13. Si ce n'est pas le cas, la constante de normalisation est très souvent intractable (difficile à calculer).
14. Avec des échantillons $\theta_1, \ldots, \theta_m \sim \theta|X$, on peut estimer une moyenne a posteriori empirique $\frac{1}{m}\sum_{i=1}^m \theta_i$.
15. Pour générer de tels échantillons, on utilise des méthodes dites de Monte Carlo.

Statistiques Bayésiennes

Hicham Janati

hjanati@insea.ac.ma

INSEA