

Estimateur du maximum de Vraisemblance et du maximum a posteriori

Soient n variables aléatoires $X_i, i = 1, \dots, n$ indépendantes dont la densité de probabilité est de la forme $f(x, \sigma^2) = \frac{x}{\sigma^2} \exp\left[\frac{-x^2}{2\sigma^2}\right] \mathcal{I}_{\mathcal{R}^+}(x)$ où $\sigma > 0$, $\mathcal{I}_{\mathcal{R}^+}(x)$ est la fonction indicatrice sur \mathcal{R}^+ , c'est à dire $\mathcal{I}_{\mathcal{R}^+}(x) = 1$, si $x \in \mathcal{R}^+$ et 0 sinon.

Cette loi est appelée loi de Rayleigh de paramètre σ . Typiquement, la distance D_n à laquelle une particule se trouve de son point de départ, après avoir effectué n pas d'une marche aléatoire symétrique dans le plan, suit approximativement une loi de Rayleigh.

On admettra les résultats suivants :

- $E[x_k] = \sigma \sqrt{\frac{\pi}{2}}$
- $E[x_k^2] = 2\sigma^2$
- $E[x_k^3] = (3\sqrt{\frac{\pi}{2}})\sigma^3$
- $E[x_k^4] = 8\sigma^4$

L'objectif des deux exercices suivants est de déterminer un estimateur de $\theta = \sigma^2$ en fonction des observations (z_i)

Exercice 1 Estimateur du maximum de vraisemblance

1. Écrire la vraisemblance de (z_1, \dots, z_n) en fonction de θ et montrer que cette fonction admet un maximum global unique que l'on déterminera. En déduire l'estimateur du maximum de vraisemblance $\hat{\theta}$.
2. Calculer la moyenne et la variance de $\hat{\theta}$. En déduire que l'estimateur est non-biaisé et converge vers θ .
3. À l'aide de la fonction de vraisemblance, calculer la borne de Cramer-Rao. En déduire que l'estimateur est efficace.

Exercice 2 Estimation Bayésienne

On suppose à présent que le paramètre θ est une variable aléatoire suivant une loi inverse-gamma de paramètres (α, β) de densité $f(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{\theta^{\alpha+1}} \exp\left(\frac{-\beta}{\theta}\right) \mathcal{I}_{\mathcal{R}^+}(\theta)$

1. Écrire la loi a posteriori $f(\theta|x_1, \dots, x_n)$
2. Montrer que cette loi peut s'écrire sous la forme d'une loi inverse-gamma dont on précisera les paramètres.
3. Étudier les variations en fonction de θ du logarithme de la fonction de vraisemblance.
4. En déduire l'expression de l'estimateur au sens du maximum a posteriori.
5. Exprimer cet estimateur en fonction de l'estimateur du maximum de vraisemblance obtenu à la question précédente et de α, β et n . En déduire la limite quand $n \rightarrow \infty$ de l'estimateur du maximum a posteriori. Donner une interprétation de ce résultat en terme d'influence de la loi a priori.

NB : Fonction Gamma : $\Gamma(a) = \int_0^{+\infty} x^{a-1} \exp(-x) dx$

Exercice 3 Algorithme EM et mélange de gaussiennes

Considérons le mélange de deux lois gaussiennes bidimensionnelles :

- loi 1 : $\mathcal{N}(\mu_1, \sigma_1)$, avec $\mu_1 \in \mathbb{R}$, $\sigma_1 \in \mathbb{R}_+^*$,
- loi 1 : $\mathcal{N}(\mu_2, \sigma_2)$, avec $\mu_2 \in \mathbb{R}$, $\sigma_2 \in \mathbb{R}_+^*$.

Le choix entre ces deux lois s'opère via une loi de Bernoulli $\mathcal{B}(\lambda)$.

Donner les étapes de l'algorithme EM pour l'estimation du vecteur de paramètres $\theta = (\lambda, \mu_1, \sigma_1, \mu_2, \sigma_2)$.

On cherchera donc à écrire explicitement les étapes donnant θ_{m+1} en fonction de θ_m .

Indication : suivre les étapes de ce qui a été fait dans le cours pour construire les étapes (E) et (M) du cas des lancers de pièces de monnaie.