Statistique inférentielle - TD 3

Exercice 1.

Soient $(X_1, ..., X_n)$ i.i.d. de loi uniforme sur $[0, \theta]$. On se place dans un cadre bayésien, où l'on suppose que le paramètre θ suit la distribution a priori de densité

$$g_{a,b}(\theta) = ba^b \theta^{-(b+1)} \mathbb{1}_{\theta > a},$$

où a et b sont strictement positifs. Déterminer l'estimateur bayésien de θ dans chacun des cas suivants :

1. on considère la perte quadratique $l(\theta, \theta') = (\theta - \theta')^2$
2. on considère la perte L^1 $l(\theta, \theta') = |\theta - \theta'|$

Exercice 2.

On dispose de n observations i.i.d. $(X_1, ..., X_n)$ de loi $\mathcal{N}(\theta, 1)$ et on considère la loi a priori $\mathcal{N}(\tau, \sigma^2)$.

1. Calculer l'estimateur bayésien $\hat{\theta}$ de θ pour la perte quadratique.
2. Comparer le risque de Bayes de $\bar{X} = n^{-1} \sum_{i=1}^n X_i$ et $\hat{\theta}_n$.

Exercice 3.

On observe n variables aléatoires i.i.d. $(X_1, ..., X_n)$ de loi de Poisson de paramètre $\mu > 0$ inconnu.

1. On choisit comme loi a priori pour μ la loi $\Gamma(t, \lambda)$ où $t > 1$, et $\lambda > 0$. Quelle est la densité jointe de $X_1, ..., X_n$ et μ par rapport à la loi $(d\mathbb{N}^{\otimes n}) \otimes dx$, où $d\mathbb{N}$ désigne la mesure de comptage sur $\mathbb{N}$ et dx la mesure de Lebesgue sur $\mathbb{R}^+$? En déduire la loi a posteriori de μ.
2. On choisit la fonction de perte $l(\mu, z) = \mu^{-1}(\mu - z)^2$ (perte lorsque le vrai paramètre est μ et que l'on décide z). Quel est l'estimateur de Bayes $\tilde{\mu}_n$ de μ ?
3. Quelle est la fonction de risque de $\tilde{\mu}_n$? En déduire le risque de Bayes.

Exercice 4.

On considère $(X_1, ..., X_n)$ i.i.d. de densité inconnue f. On considère un estimateur à noyau de f par

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{X_i - x}{h}\right),$$

où K est une fonction positive, paire (i.e. $K(u) = K(-u)$), $\int K(u)du = 1$ et $\int u^2 K(u)du < \infty$. On suppose que f est deux fois dérivable avec ses dérivées jusqu'à l'ordre 2 bornées.

1. Montrer que $E[\hat{f}(x)] = f(x) + O(h^2)$.
2. Montrer que $Var(\hat{f}(x)) = O(1/(nh))$, et donc que $E[(\hat{f}(x) - f(x))^2]^{1/2} = O(h^2 + 1/(nh))$. Quelle est la vitesse optimale pour h ?
3. On considère à présent que les X_i sont dans $\mathbb{R}^k$, f étant donc une fonction à k variables C^2. On considère alors
$$\hat{f}(x) = \frac{1}{nh^k} \sum_{i=1}^n K_k\left(\frac{X_i - x}{h}\right),$$
où $K_k(a_1, ..., a_k) = K(a_1)K(a_2)...K(a_k)$. Que valent $E[\hat{f}(x)]$ et $Var(\hat{f}(x))$? Commentaires ?