Corrigé de l'axamen 1 de ST21, 2004-2005.

Exercice 1.

On considère cinq variables aléatoires réelles indépendantes X_1, …, X_5 suivant chacune la loi Gaussienne de moyenne m et de variance σ^2. On se propose d'estimer m par les estimateurs suivants :

$$\hat{m}_1(X_1,...,X_5) = \frac{1}{5}(X_1 + ... + X_5)$$

$$\hat{m}_2(X_1,...,X_5) = \frac{1}{5}(X_1 + X_2) + \frac{1}{4}(X_3 + X_4 + X_5)$$

$$\hat{m}_3(X_1,...,X_5) = \frac{1}{10}(2X_1 + 3X_2) + \frac{1}{8}(X_3 + 2X_4 + X_5)$$

1. Lesquels parmi ces estimateurs sont sans biais ?

Solution : $\hat{m}_1$ et $\hat{m}_3$ sont sans biais (calcul trivial)

2. Parmi les estimateurs sans biais, lequel est le plus intéressant au sens de l'erreur quadratique moyenne ?

Solution : Pour un estimateur sans biais son e.q.m. est égale à sa variance. Les observations étant indépendantes on a

$$Var[\hat{m}_1] = \frac{1}{25}(Var[X_1] + ... + Var[X_5]) = \frac{\sigma^2}{5} \text{ et}$$

$$Var[\hat{m}_3] = \frac{4\sigma^2 + 9\sigma^2}{100} + \frac{1}{64}(\sigma^2 + 4\sigma^2 + \sigma^2) = (\frac{13}{100} + \frac{6}{64})\sigma^2$$

On trouve $Var[\hat{m}_1] < Var[\hat{m}_3]$; $\hat{m}_1$ est donc plus intéressant

Exercice 2

On considère l'ensemble de trois classes $\Omega = \{\theta_1, \theta_2, \theta_3\}$ muni de la probabilité a priori $\pi_1 = 0.2$, $\pi_2 = 0.3$, et $\pi_3 = 0.5$. Les trois densités de la loi de l'observation conditionnelles aux classes θ_1, θ_2, θ_3 sont définies sur [0,1] par $f_1 = 2(1-x)$, $f_2 = 1$, et $f_3 = 2x$.

1. On considère la fonction de perte « 0-1 ». Donner la stratégie bayésienne de classification et calculer la probabilité de l'erreur ;

Solution : En appliquant soigneusement la formule 2.10 du polycopié (page 78) au cas de trois classes on trouve

$$\hat{s}_B(x) = \begin{cases} \theta_1 \ \ si \ \ \pi_1 f_1(x) \geq \pi_2 f_2(x) \ \ et \ \ \pi_1 f_1(x) \geq \pi_3 f_3(x) \\ \theta_2 \ \ si \ \ \pi_2 f_2(x) \geq \pi_1 f_1(x) \ \ et \ \ \pi_2 f_2(x) \geq \pi_3 f_3(x) \\ \theta_3 \ \ si \ \ \pi_3 f_3(x) \geq \pi_1 f_1(x) \ \ et \ \ \pi_3 f_3(x) \geq \pi_2 f_2(x) \end{cases}$$

ce qui donne

$$\hat{s}_B(x) = \begin{cases} \theta_1 \ \ si \ \ x \in [0, 0.25] \\ \theta_2 \ \ si \ \ x \in [0.25, 0.3] \\ \theta_3 \ \ si \ \ x \in [0.3, 1] \end{cases}.$$

La probabilité d'erreur est

$$P[Erreur] = \pi_1 \int_{0.25}^{1} f_1(x)dx + \pi_2 \left(\int_{0}^{0.25} f_2(x)dx + \int_{0.3}^{1} f_2(x)dx \right) + \pi_3 \int_{0}^{0.3} f_3(x)dx \approx 0.4425$$

2. On considère que ne pas détecter θ_1 coûte trois fois plus cher que les autres erreurs. Donner la fonction de perte correspondante, calculer la stratégie bayésienne, calculer la perte moyenne ;

Solution : en suivant la définition de la fonction de perte, on écrit directement :

$$L(\theta_i, \theta_j) = \lambda_{ij} = \begin{cases} 0 \ si \ \ \theta_i = \theta_j \\ \lambda \ \ si \ \ \theta_i \neq \theta_j \ \ et \ \ j \neq 1 \\ 3\lambda \ \ si \ \ \theta_i \neq \theta_j \ \ et \ \ j = 1 \end{cases}$$

En appliquant à nouveau la formule 2.10 du polycopié (page 78) à cette nouvelle fonction de perte on trouve

$$\hat{s}_B(x) = \begin{cases} \theta_1 \ \ si \ \ 3\pi_1 f_1(x) \geq \pi_2 f_2(x) \ \ et \ \ 3\pi_1 f_1(x) \geq \pi_3 f_3(x) \\ \theta_2 \ \ si \ \ \pi_2 f_2(x) \geq 3\pi_1 f_1(x) \ \ et \ \ \pi_2 f_2(x) \geq \pi_3 f_3(x) \\ \theta_3 \ \ si \ \ \pi_3 f_3(x) \geq 3\pi_1 f_1(x) \ \ et \ \ \pi_3 f_3(x) \geq \pi_2 f_2(x) \end{cases}$$

ce qui donne

$$\hat{s}_B(x) = \begin{cases} \theta_1 \ \ si \ \ x \in [0, \frac{6}{11}] \\ \\ \theta_3 \ \ si \ \ x \in [\frac{6}{11}, 1] \end{cases}$$

La perte moyenne est $3\pi_1 \int_{\frac{6}{11}}^{1} f_1(x)dx + \pi_2 \int_{0}^{1} f_2(x)dx + \pi_3 \int_{0}^{\frac{6}{11}} f_3(x)dx \approx 0.5727$. Attention, ce n'est pas la probabilité d'erreur, cette dernière s'écrivant $\pi_1 \int_{\frac{6}{11}}^{1} f_1(x)dx + \pi_2 \int_{0}^{1} f_2(x)dx + \pi_3 \int_{0}^{\frac{6}{11}} f_3(x)dx$.

3. Pourquoi peut-on affirmer sans calcul que la probabilité d'erreur de la stratégie du point 2. est supérieure ou égale à la probabilité d'erreur de la stratégie de celle du point 1. ? On répondra en une phrase sans aucune formule.

Solution : La stratégie Bayésienne correspondant à la fonction de perte « 0-1 » du point 1. minimise la probabilité d'erreur, donc la probabilité d'erreur du point 1. est nécessairement plus petite (ou égale à) que celle du point 2.

Exercice 3

On considère l' ensemble des paramètres $\Theta = [0,1]$ et les densités de probabilités f_θ définies sur l' ensemble des nombres réels par $f_\theta(x) = 1_{[\theta-0.5,\,\theta+0.5]}(x)$ (fonctions indicatrices des ensembles $[\theta-0.5, \theta+0.5]$).

1. Pour une observation x, on considère la stratégie $\hat{s}(x) = x$. Calculer le Risque associé à la fonction de perte quadratique $L(\theta', \theta) = (\theta' - \theta)^2$.

Solution : $R(\hat{s},\theta) = \int_0^1 (\theta - x)^2 \, f_\theta(x)dx = \int_{\theta-05}^{\theta+0.5} (\theta - x)^2 \, dx = \int_{-05}^{0.5} x^2 dx = \frac{1}{12}$

2. On dispose d' une information selon laquelle "le paramètre est plutôt vers 1", ce que l' on modélise par la loi de probabilité donnée sur $\Theta = [0,1]$ par la densité $f(\theta) = 3\theta^2$
Montrer que la stratégie Bayésienne $\hat{s}_B$ vérifie :

$$\hat{s}_B(x) = \frac{\int_0^1 \theta^3 1_{[x-0.5,\,x+0.5]}(\theta)d\theta}{\int_0^1 \theta^2 1_{[x-0.5,\,x+0.5]}(\theta)\,d\theta}$$

Solution : On peut procéder de deux façons :

(1) Selon (1.17) page 73, $\hat{s}_B(x)$ est le paramètre θ' qui minimise $\varphi(\theta') = \int_0^1 (\theta - \theta')^2 \, g(\theta, x)d\theta$, avec $g(\theta, x) = f(\theta) f_\theta(x)$. Sachant que φ est une parabole « inversée », il suffit de chercher la racine de la dérivée, ce qui donne $\hat{s}_B(x)$ ci dessus.

(2) Sachant que la solution Bayésienne correspondant à la fonction de perte quadratique est l'espérance conditionnelle $\hat{s}_B(x) = \int_R \theta f_x(\theta)d\theta$, on calcule f_x :

$$f_x(\theta) = \frac{g(\theta, x)}{\int_R g(\theta, x)} = \frac{f(\theta) f_\theta(x)}{\int_R f(\theta) f_\theta(x)}$$

Calculer $\hat{s}_B$. **Solution** : L'observation x est dans $[-0.5,1.5]$. Pour $x \le 0.5$ on a $\hat{s}_B(x) = \dfrac{[\theta^4/4]_0^{x+0.5}}{[\theta^3/3]_0^{x+0.5}} = \dfrac{3}{4}(x+0.5)$, et pour $0.5 \le x$ on a $\hat{s}_B(x) = \dfrac{[\theta^4/4]_{x-0.5}^{1}}{[\theta^3/3]_{x-0.5}^{1}} = \dfrac{3[1-(x-0.5)^4]}{4[1-(x-0.5)^3]}$.