Module ST21
Corrigé du sujet de l'examen du 6 novembre 2006

Exercice 1

1. Montrer *que pour tout* $\theta \in [0,2]$, f_θ *est bien une densité de probabilité sur* $[0,1]$:

$f_\theta(x) = \theta x + 1 - \frac{\theta}{2}$ est positive ou nulle sur $[0,1]$ et on a bien $\int_0^1 (\theta x + 1 - \frac{\theta}{2})dx = \frac{\theta}{2} + 1 - \frac{\theta}{2} = 1$.

2. *Montrer que l'espérance de chaque* X_i *est égales à* $\frac{1}{2} + \frac{\theta}{12}$:

$E[X_i] = \int_0^1 x(\theta x + 1 - \frac{\theta}{2})dx = \frac{\theta}{3} + \frac{1}{2}(1 - \frac{\theta}{2}) = \frac{1}{2} + \frac{\theta}{12}$.

3. *Montrer que la variance de chaque* X_i *est égale à* $\frac{1}{12} - \frac{\theta^2}{144}$:

$Var[X_i] = E[(X_i)^2] - E[X_i]^2 = \int_0^1 x^2(\theta x + 1 - \frac{\theta}{2})dx - (\frac{1}{2} + \frac{\theta}{12})^2 = \frac{\theta}{4} + \frac{1}{3} - \frac{\theta}{6} - \frac{1}{4} - \frac{\theta}{12} - \frac{\theta^2}{144} =$

$= \frac{1}{12} - \frac{\theta^2}{144}$

4. *Montrer que l'estimateur*

$$\hat{\theta}_1(X_1,...,X_n) = 12(\frac{X_1 + ... + X_n}{n} - \frac{1}{2})$$

est sans biais et qu'il converge en moyenne quadratique :

On a $E[\hat{\theta}_1(X_1,...,X_n)] = 12(\frac{1}{2} + \frac{\theta}{12} - \frac{1}{2}) = \theta$, donc $\hat{\theta}_1$ est sans biais. Son erreur quadratique est alors égale à sa variance, qui est $Var[\hat{\theta}_1(X_1,...,X_n)] =$ $= 144 Var[\frac{X_1 + ... + X_n}{n}] = \frac{144n}{n^2} Var[X_1] = \frac{12 - \theta^2}{n}$, l'avant dernière égalité venant de l'indépendance des variables X_1, …, X_n. On a bien $\frac{12 - \theta^2}{n} \xrightarrow[n \to +\infty]{} 0$, d'où la convergence demandée.

5. *Montrer que l'estimateur bayésien* $\hat{\theta}_2$ *associé à* L *est donné par*

$$\hat{\theta}_2(X_1,...,X_n) = \frac{\int_0^2 \theta^2[\prod_{i=1}^n(\theta X_i + 1-\theta/2)]d\theta}{\int_0^2 \theta[\prod_{i=1}^n(\theta X_i + 1-\theta/2)]d\theta} :$$

$\hat{\theta}_2(X_1,...,X_n)$ minimise, par rapport à θ', l'application $\theta' \to \int_0^2 (\theta-\theta')^2 f(\theta) f_\theta(X_1,...,X_n)d\theta = \int_0^2 (\theta-\theta')^2 \frac{\theta}{2}[\prod_{i=1}^n(\theta X_i + 1-\theta/2)]d\theta$. En annulant sa dérivée par rapport à θ' on obtient s'égalité demandée. On retrouve le même résultat en se souvenant que l'estimateur bayésien associé à la fonction de perte considérée est l'espérance conditionnelle, qui est l'espérance par rapport à la loi conditionnelle donnée par la densité

$$f_x(\theta) = \frac{f(\theta) f_\theta(x)}{\int_0^2 f(\theta) f_\theta(x)d\theta} = \frac{\theta[\prod_{i=1}^n(\theta X_i + 1-\theta/2)}{\int_0^2 \theta[\prod_{i=1}^n(\theta X_i + 1-\theta/2)]d\theta}.$$

6. *On considère maintenant que sur-estimer le paramètre est deux fois plus grave que de le sous-estimer, ce que l'on modélise par la fonction de perte* $L^*(\theta',\theta) = \begin{cases} 2(\theta'-\theta) & \text{si } \theta' > \theta \\ \theta-\theta') & \text{si } \theta' < \theta \end{cases}$
Ecrire la fonction dont la minimisation donne l'estimateur bayésien associé à L^* *:*

Comme dans la question précédente, l'estimateur bayésien associé à L^* minimise, par rapport à θ', l'application $\theta' \to \int_0^2 L^*(\theta',\theta) f(\theta) f_\theta(X_1,...,X_n)d\theta$, qui est ici (à θ' donné, on considère $[0,\theta']$ et $[\theta',0]$ séparément) :

$$\int_0^{\theta'} 2(\theta'-\theta)\frac{\theta}{2}[\prod_{i=1}^n(\theta X_i + 1-\theta/2)d\theta + \int_{\theta'}^2 (\theta-\theta')\frac{\theta}{2}[\prod_{i=1}^n(\theta X_i + 1-\theta/2)d\theta$$

Notons que malgré l'aspect peu engageant de cette fonction la minimisation se fait rapidement sur ordinateur.

Exercice 2

1. *Afin d'estimer les paramètres* π_a, π_b, π_c, m_a, m_b, m_c, σ_a^2, σ_b^2, *et* σ_c^2 *on prélève 1000 spécimens dont on note* x_1, …, x_{1000} *l'espèce (chaque* x_i *est dans* $\Omega = \{a,b,c\}$*) et* y_1, …, y_{1000} *le poids. Proposer des estimateurs pour* π_a, m_a, *et* σ_a^2 :

π_a peut être estimé par la proportion des poissons a dans l'échantillon prélevé, ce qui s'écrit : $\hat{\pi}_a = \frac{\sum_{i=1}^{1000} 1_{[x_i=a]}}{1000}$. En considérant uniquement les poissons de l'espèce a dans l'échantillon prélevé, la moyenne peut alors être estimée par l'estimateur classique du

maximum de vraisemblance, qui s'écrit : $\hat{m}_a = \dfrac{\sum_{i=1}^{1000} y_i 1_{[x_i=a]}}{\sum_{i=1}^{1000} 1_{[x_i=a]}}$ (la forme du numérateur montre que l'on somme uniquement les poids des poissons a, et le dénominateur est simplement le nombre des poissons a dans l'échantillon prélevé). De même, la variance peut être estimée par l'estimateur classique $\hat{\sigma}_i^2 = \dfrac{\sum_{i=1}^{1000} (y_i - \hat{m})^2 {}_a 1_{[x_i=a]}}{\sum_{i=1}^{1000} 1_{[x_i=a]}}$, avec les mêmes commentaires concernant le numérateur et le dénominateur.

2. *Donner la méthode bayésienne de classification des 9000 poissons restant minimisant la probabilité d'erreur :*

Minimiser la probabilité d'erreur revient à rechercher la stratégie bayésienne associée à la fonction de perte « 0-1 ». En appliquant la formule (2.10), page 78 du polycopié (avec $\lambda_{ij} = 1$ pour $i \neq j$ et $\lambda_{ij} = 0$ pour $i = j$), on a

$$\hat{s}(y_i) = \arg\max_{\omega \in \{a,b,c\}} \pi_\omega f_\omega(y_i) = \begin{cases} a \ si \ \pi_a f_a(y_i) > \pi_b f_b(y_i) \ et \ \pi_a f_a(y_i) > \pi_c f_c(y_i) \\ b \ si \ \pi_b f_b(y_i) > \pi_a f_a(y_i) \ et \ \pi_b f_b(y_i) > \pi_c f_c(y_i) \\ c \ si \ \pi_c f_c(y_i) > \pi_a f_a(y_i) \ et \ \pi_c f_c(y_i) > \pi_b f_b(y_i) \end{cases} \qquad (1)$$

3. *Ecrire la formule, sous forme d'une intégrale, donnant la probabilité de mal classer un poisson de l'espèce a :*

Notons A l'ensembles des nombres réels y tels que $\hat{s}(y) = a$, B l'ensembles des nombres réels y tels que $\hat{s}(y) = b$, et C l'ensembles des nombres réels y tels que $\hat{s}(y) = c$. Un poisson de l'espèce a sera mal classé lorsque son poids y se trouve dans $B \cup C$. Sachant que la loi de son poids suit la densité f_a, la probabilité pour que cela arrive est $\int_{B \cup C} f_a(y) dy$.

4. *On suppose que l'espèce a est précieuse et l'on souhaite limiter la probabilité de l'erreur consistant à classer un poisson de cette espèce dans les espèces b ou c à ε. Décrire rapidement comment doit-on modifier la méthode précédente pour atteindre cet objectif :*

Pour diminuer la probabilité de mal classer l'espèce a, on remplace (1) ci-dessus par (on arrive à ces formules en considérant (2.10), page 78 du polycopié avec $\lambda_{ba} = \lambda_{ca} = \alpha$, $\lambda_{ab} = \lambda_{cb} = \lambda_{ac} = \lambda_{bc} = 1$, et $\lambda_{aa} = \lambda_{bb} = \lambda_{cc} = 0$) :

$$\hat{s}_\alpha(y_i) = \begin{cases} a \ si \ \alpha\pi_a f_a(y_i) > \pi_b f_b(y_i) \ et \ \alpha\pi_a f_a(y_i) > \pi_c f_c(y_i) \\ b \ si \ \pi_b f_b(y_i) > \alpha\pi_a f_a(y_i) \ et \ \pi_b f_b(y_i) > \pi_c f_c(y_i) \\ c \ si \ \pi_c f_c(y_i) > \alpha\pi_a f_a(y_i) \ et \ \pi_c f_c(y_i) > \pi_b f_b(y_i) \end{cases}, \qquad (2)$$

avec $\alpha > 1$. Comme ci-dessus, considérons A^α, B^α, et C^α les ensembles correspondants à $\hat{s}_\alpha$ donnée par (2). Pour tout y_i, on peut dire que $\hat{s}_\alpha(y_i) = a$ pour α suffisamment grand. Cela signifie que la fonction indicatrice de A^α tend vers 1 « simplement » (pour tout y_i, $1_{A^\alpha}(y_i) \xrightarrow[\alpha \to +\infty]{} 1$) ; de plus, cette convergence est « monotone » (pour tout y_i, $\alpha_1 \geq \alpha_2$ implique $1_{A^{\alpha_1}}(y_i) \geq 1_{A^{\alpha_2}}(y_i)$). Il en résulte (théorème de la convergence monotone de Lebesgue) que $E_a[1_{A^\alpha}] = \int_{A^{Ga}} f_a(y)dy \xrightarrow[\alpha \to +\infty]{} 1$, et donc l'erreur $\int_{B^\alpha \cup C^\alpha} f_a(y)dy = 1 - \int_{A^\alpha} f_a(y)dy \xrightarrow[\alpha \to +\infty]{} 0$. Pour tout ε, on peut donc choisir $\alpha > 1$ tel que cette erreur est inférieure à ε.

Notons que la diminution de la proportion des poissons mal classés de l'espèce a est accompagnée de l'augmentation des proportions des poissons mal classés des espèces b et c ; de plus, la proportion totale des poissons mal classés (indépendamment de l'espèce) sera ici plus grande que dans 3., où cette proportion est minimale (la « proportion » est assimilée à la « probabilité » - via la loi des grands nombres - pour une pêche suffisamment riche).