MAÎTRISE DE MATHÉMATIQUES 2003-2004
STATISTIQUES

5. Statistiques exhaustives complètes, théorème de Lehmann-Scheffé, modèles exponentiels

5.1. Soit $(X_1,\ldots,X_n)$ un n-échantillon de loi P_θ, $\theta \in \Theta$. Indiquer des statistiques exhaustives pour le paramètre θ. S'agit-il des statistiques complètes? Trouver ensuite des estimateurs efficaces pour θ.

a) $\Theta = [0,1]$ et $\mathrm{P}_\theta = \mathcal{B}(1,\theta)$ (loi de Bernoulli);
b) $\Theta = \mathbb{R}_+$ et $\mathrm{P}_\theta = \mathcal{P}(\theta)$ (loi de Poisson);
c) $\Theta = \mathbb{R}$ et $\mathrm{P}_\theta = \mathcal{N}(\theta,1)$ (loi normale avec variance connue);
d) $\Theta = \mathbb{R}_+^*$ et $\mathrm{P}_\theta = \mathcal{N}(0,\theta^2)$ (loi normale avec espérance connue);
e) $\Theta = \mathbb{R} \times \mathbb{R}_+^*$ et $\mathrm{P}_\theta = \mathcal{N}(m,\sigma^2)$ (loi normale);
f) $\Theta = \mathbb{R}_+$ et $\mathrm{P}_\theta = \mathcal{E}(\theta)$ (loi exponentielle).

5.2. Soit $(X_1,\ldots,X_n)$ un n-échantillon de loi uniforme sur $[\mu,\theta]$, $-\infty < \mu < \theta < \infty$.

a) Indiquer une statistique exhaustive pour le paramètre (μ,θ).
b) Supposons que μ est connu et que le seul paramètre est θ. Trouver une statistique exhaustive complète pour θ. Peut-on trouver l'estimateur efficace pour θ? (On pourra faire $\mu = 0$.)
c)* Supposons cette fois que $\theta = \mu + 1$, est que le seul paramètre est μ. Montrer que la statistique trouvée au premier point est exhaustive mais n'est pas complète.

5.3. Soit $(X_1,\ldots,X_n)$ un n-échantillon de loi géomètrique de paramètre $p \in]0,1[$. Trouver une statistique suffisante pour p. Est-elle complète?

5.4. Soit $(X_1,\ldots,X_n)$ un n-échantillon de loi $\mathcal{B}(1,1/(1+\theta))$, $\theta > 0$. Montrer que $X_1+\ldots+X_n$ statistique exhaustive. Est-elle complète?

5.5. Soit $(X_1,\ldots,X_n)$ un n-échantillon de loi $\mathcal{N}(\theta,\theta)$, $\theta > 0$. Montrer que $(S,T) := (X_1 + \ldots + X_n, X_1^2 + \ldots + X_n^2)$ est une statistique suffisante pour θ. Calculer l'espérance de $S^2 - T$. Que peut-on conclure?

5.6. Soit $(X_1,\ldots,X_n)$ un n-échantillon de loi $\gamma(p,\lambda)$.

a) Montrer que $(X_1 \ldots X_n, X_1 + \ldots + X_n)$ est une statistique suffisante pour le paramètre (p,λ).
b) Supposons ensuite que λ est connu. Donner une statistique suffisante pour p.
c) Montrer que $\overline{X}$ est une statistique exhaustive pour le paramètre $\theta > 0$, lorsque la densité de l'échantillon est $f_\theta(x) = \frac{1}{6\theta^4} x^3 e^{-x/\theta} \mathbb{1}_{]0,\infty[}(x)$.

5.7*. Soit $((X_1,Y_1)\ldots,(X_n,Y_n))$ un n-échantillon de loi $\mathcal{N}_2\left(\begin{pmatrix} m \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho \\ \rho & \tau^2 \end{pmatrix}\right)$. Montrer que $(\sum_{i=1}^n X_i, \sum_{i=1}^n Y_i, \sum_{i=1}^n X_i^2, \sum_{i=1}^n Y_i^2, \sum_{i=1}^n X_i Y_i)$ est une statistique exhaustive pour $(m,\mu,\sigma^2,\tau^2,\rho)$.

5.8. Soit $(X_1,\dots,X_n)$ un n-échantillon de loi $\mathcal{B}(1,p)$. On prend comme paramètre la variance $\theta = p(1-p)$. Utiliser le théorème de Lehmann-Scheffé pour trouver l'estimateur efficace de θ.

5.9. Soit $(X_1,\dots,X_n)$ un n-échantillon de loi $\mathcal{P}(\lambda)$. On fixe $r \in \mathbb{N}$ et soit $\theta_r = e^{-\lambda}\lambda^r/(r!)$.

a) Soit $r = 0$ et on note $S = X_1 + \ldots + X_n$ et $N_0 = \sum_{i=0}^{n} \mathbb{1}_{\{X_i=0\}}$.

- i) On pose $T = \mathbb{E}(\frac{1}{n}N_0 \mid S)$. Montrer que $T = \mathbb{E}(\mathbb{1}_{\{X_1=0\}} \mid S)$.
- ii) Montrer que $\mathbb{E}(\mathbb{1}_{\{X_1=0\}} \mid S = s) = \mathbb{P}(X_1 = 0 \mid S = s)$, $s \in \mathbb{N}$.
- iii) En déduire que $\mathbb{E}(\mathbb{1}_{\{X_1=0\}} \mid S = s) = \left(\frac{n-1}{n}\right)^s$, $s \in \mathbb{N}$ et que $T = \left(1 - \dfrac{1}{n}\right)^{n\overline{X}}$.
- iv) Montrer que T est un estimateur sans biais et efficace de θ_0.

b)* Reprendre un raisonnement identique pour trouver un estimateur sans biais efficace de θ_r, pour $r \in \mathbb{N}^*$.

5.10. Soit $(X_1,\dots,X_n)$ un n-échantillon de loi gaussienne $\mathcal{N}(m,\sigma^2)$. Trouver l'estimateur efficace de $p := \Phi(-\mu/\sigma)$, où Φ est la fonction de répartition d'une variable gaussienne standard (la fonction de Laplace). On pourra considérer d'abord le cas où σ^2 est connu.

5.11. Soient $(T_1,\dots,T_r)$, r estimateurs sans biais d'un paramètre θ et on note $\mathrm{cov}(T_i,T_j) = \sigma_{ij}$, $i,j \in \{1,\dots,r\}$.

a) Parmi les combinaisons linéaires des T_i, trouver l'estimateur sans biais de variance minimale.

b) Si $\sigma_{ij} = 0$ pour $i \neq j$, que vaut cette variance minimale?

c) On dispose de r échantillons indépendants de tailles n_i, $i = 1,\dots,r$, de lois $\mathcal{N}(m_i,\sigma^2)$. Proposer un estimateur sans biais de σ^2. Est-il efficace?

5.12. Indiquer lesquels des modèles suivants sont des modèles exponentiels:

$\{\mathcal{B}(1,\theta) : \theta \in [0,1])\}$; $\{\mathcal{P}(\theta) : \theta > 0)\}$; $\{\mathcal{G}(\theta) : \theta \in [0,1])\}$; $\{\mathcal{E}(\theta) : \theta > 0)\}$;

$\{\gamma(p,\lambda) : \theta = (p,\lambda) \in \mathbb{R}_+^* \times \mathbb{R}_+^*\}$; $\{\mathcal{N}(m,\sigma^2) : \theta = (m,\sigma^2) \in)\mathbb{R} \times \mathbb{R}_+^*\}$;

$$\left\{\mathcal{N}_2\left(\begin{pmatrix} m \\ \mu \end{pmatrix}, \begin{pmatrix} \sigma^2 & \rho \\ \rho & \tau^2 \end{pmatrix}\right) : (m,\mu,\sigma^2,\tau^2,\rho) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^* \times \mathbb{R}_+^* \times \mathbb{R}\right\}.$$