Université Paris Dauphine — Année 2002-2003
U.F.R. Mathématiques de la Décision
MD3 - Statistique Inférentielle

Examen partiel 20/03/2003

Les 3 parties sont indépendantes. Seules les réponses soigneusement justifiées seront prises en compte.

Partie 1 : Question de cours

On considère le modèle statistique paramétrique $(\mathcal{X}, \mathcal{A}, P_\theta : \theta \in \Theta)$, que l'on suppose dominé par une mesure positive et σ-finie μ, ainsi qu'une variable aléatoire X à valeurs dans $\mathcal{X}$, un n-échantillon $(X_1, \ldots, X_n)$ de X et une statistique $T = \varphi(X_1, \ldots, X_n)$.

1 Définir l'exhaustivité de T (ou de φ), puis énoncer sans démonstration le théorème de factorisation. Donner la définition d'une statistique exhaustive minimale.

2 Application : on considère la loi P_X admettant pour densité relativement à la mesure de Lebesgue la fonction $f(., \theta)$ donnée par

$$f(x, \theta) = \frac{\alpha {x_0}^\alpha}{x^{\alpha - 1}} \mathbf{1}_{[x_0, +\infty[} \qquad x \in \mathbf{R}$$

où x_0 et α sont strictement positifs. En justifiant la réponse, indiquer si cette loi appartient à la famille exponentielle, dans les deux cas suivants :

(i) x_0 est donné et le paramètre du modèle est $\theta = \alpha$,
(ii) le paramètre du modèle est le vecteur $\theta = (x_0, \alpha)$.

Dans chaque cas, mettre en évidence une statistique exhaustive minimale pour le paramètre θ.

Partie 2 : Exercice 1

Soit X une variable aléatoire réelle admettant pour densité de probabilité :

$$f_X(x;\lambda,\alpha,\theta) = \begin{cases} \left(\dfrac{1}{c}\right)\exp\left(-\dfrac{(x-\lambda)}{\theta}\right)x^{\alpha-1} & \text{si} \quad x \geq \lambda \\ 0 & \text{sinon} \end{cases}.$$

On considére un n-échantillon $(X_1,\ldots,X_n)$ iid comme X.

1 On suppose que $\alpha = 1$.

On définit les statistiques $L_n = \inf(X_1,\ldots,X_n)$ et $T_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i - L_n$.

a) Calculer c, $\mathbb{E}(X)$ et $\mathbb{V}(X)$.

b) Montrer que L_n converge en probabilité vers λ
(Indication : calculer $\mathbb{P}(|L_n - \lambda| > \delta)$).

En déduire que T_n converge en probabilité vers θ.

c) Montrer que, si $\left(\dfrac{k_n}{n}\right)$ tend vers 0 lorsque n tend vers l'infini, alors $Z_n = k_n\left(\dfrac{L_n-\lambda}{\theta}\right)$ converge en moyenne quadratique vers 0 (Indication : calculer $\mathbb{E}(Z_n)$ et $\mathbb{V}(Z_n)$).

d) Montrer que $\sqrt{n}\left(\dfrac{T_n-\theta}{\theta}\right)$ converge en loi vers une $\mathcal{N}(0,1)$.

2 On suppose maintenant que $\lambda = 0$ et $\alpha = 2$.

On définit la statistique $H_n = \dfrac{n}{\displaystyle\sum_{i=1}^{n}\dfrac{1}{X_i}}$.

a) Calculer c, $\mathbb{E}(X)$ et $\mathbb{V}(X)$.

b) Montrer que H_n converge presque sûrement vers θ.

Partie 3 : Exercice 2

Soit X une variable aléatoire réelle de fonction de répartion continue F. On considère un n-échantillon $(X_1, \ldots, X_n)$ iid comme X. $\left(X_{(1)}, \ldots, X_{(n)}\right)$ désigne l'échantillon ordonné dans le sens croissant. On associe à chaque variable aléatoire X_i son rang (indice) i d'observation et son rang R_i dans l'échantillon ordonné.

1 Montrer que, pour tout $k \in \{1, \ldots, n\}$, $\mathbb{P}(R_i = k) = \dfrac{1}{n}$, quel que soit $i \in \{1, \ldots, n\}$. Calculer $\mathbb{E}(R_i)$ et $\mathbb{V}(R_i)$.
$(\sum_{k=1}^{n} k = \dfrac{n(n+1)}{2}$ et $\sum_{k=1}^{n} k^2 = \dfrac{n(n+1)(2n+1)}{6})$.

2 On définit par r_s le coefficient de corrélation linéaire empirique entre l'échantillon des indices (i) et celui des rangs (R_i) :

$$r_s = \frac{\sum_{i=1}^{n} \left(R_i - \overline{R}\right)\left(i - \overline{i}\right)}{\sqrt{\sum_{i=1}^{n} \left(R_i - \overline{R}\right)^2 \sum_{i=1}^{n} \left(i - \overline{i}\right)^2}}$$

où $\overline{i} = \dfrac{1}{n} \sum_{i=1}^{n} i$ et $\overline{R} = \dfrac{1}{n} \sum_{i=1}^{n} R_i$.

r_s porte le nom de coefficient de Spearman.

Montrer que $\sum_{i=1}^{n} \left(R_i - \overline{R}\right)^2 = \dfrac{n\left(n^2-1\right)}{12}$.

En déduire que $r_s = 1 - \dfrac{6}{n(n-1)(n+1)} \sum_{i=1}^{n} (R_i - i)^2$.

3 Calculer $\mathbb{E}(r_s)$ et $\mathbb{V}(r_s)$.