

Examen du 8 avril 2010

Durée 3h-Aucun document autorisé-Calculatrices scientifiques autorisées

1 Un test d'hypothèse simple

Soit X_1, \dots, X_n un n -échantillon de densité f par rapport à la mesure de Lebesgue sur $[0, 1]$. On veut tester

$$H_0 \quad f(x) = \frac{1}{\log 2(1+x)} \quad | \quad H_1 \quad f(x) = 1.$$

On pose pour $j = 1, \dots, n$,

$$Y_j = \frac{2 \log(1 + X_j)}{\log 2}.$$

Exprimer à partir des variables (Y_j) la procédure de test optimale. Application numérique : pour $n = 100$ et une erreur de première espèce de l'ordre de 5%, que décide t'on si la moyenne empirique observée des (Y_j) est 1,15 ?

2 Modèle linéaire

On considère le modèle de régression

$$Y_j = a^* + b^* x_j + \varepsilon_j, \quad j = 1 \dots n \geq 2.$$

Les x_j sont pour $j = 1 \dots n$ donnés et prennent au moins deux valeurs distinctes. Les variables aléatoires ε_j ($j = 1 \dots n$), sont i.i.d. de loi gaussienne centrée et de variance σ^{*2} . Les paramètres inconnus du modèle sont a^* , b^* et σ^{*2} .

- a) Calculer les estimateurs du maximum de vraisemblance des paramètres a^* , b^* et σ^{*2} .
- b) Construire un intervalle de confiance pour a^* .
- c) Tester l'hypothèse $b^* = 0$.

3 Bateau

Soit X une variable aléatoire dont la loi sur \mathbb{Z} est donnée par

$$P(X = x) = C a^{*|x|} \quad x \in \mathbb{Z}, \quad 0 < a^* < 1, \quad C > 0.$$

- a) Calculer la constante C .
- b) On observe X_1, \dots, X_n un n -échantillon de même loi que X . Montrer que l'on est ici dans le cadre d'un modèle exponentiel dont on précisera la statistique exhaustive minimale. Calculer l'estimateur du maximum de vraisemblance de $2a^*(1 - a^{*2})^{-1}$. En déduire l'estimateur du maximum de vraisemblance \hat{a}_n de a^* .
- c) Calculer l'information de Fisher du modèle et en déduire la variance asymptotique de \hat{a}_n .

4 Estimateur non paramétrique

Soit f une densité de probabilité sur \mathbb{R} . On suppose que f est deux fois continûment dérivable sur \mathbb{R} et que f'' est uniformément bornée sur \mathbb{R} . On suppose de plus que f est une fonction paire et que

$$\mu = \int_{\mathbb{R}} x^2 f(x) dx < +\infty.$$

On pose pour $\theta, x \in \mathbb{R}$, $f_{\theta}(x) = f(x - \theta)$. Soient θ^* un réel fixé, et X_1, \dots, X_n un n -échantillon de densité marginale f_{θ^*} . Le but de cet exercice est de construire et d'étudier des estimateur de $\xi^* = f(\theta^*)$.

- 1) Montrer que la moyenne empirique de l'échantillon (notée \overline{X}_n) est un estimateur sans biais de θ^* . Montrer que cet estimateur est fortement consistant étudier précisément sa vitesse de convergence.
- 2) Pour estimer ξ^* on propose l'estimateur empirique $\widehat{\xi}_n = f(\overline{X}_n)$. Quelles sont les propriétés asymptotiques de cet estimateur.
- 3) Soit (a_n) une suite de réels qui décroît vers 0, quand n tend vers l'infini. On suppose que $\lim_{n \rightarrow \infty} na_n = 0$. On pose

$$T_n = \frac{1}{2na_n} \sum_{j=1}^n \mathbf{1}_{[-a_n, a_n]}(X_j).$$

Monter que T_n est un estimateur asymptotiquement sans biais de ξ^* .

Calculer $\lim_{n \rightarrow \infty} na_n \text{Var} T_n$. en déduire que T_n converge en probabilité vers ξ^* . Quel estimateur choisir entre $\widehat{\xi}_n$ et T_n ? Commenter le résultat.

