Université de Paris 5 – René Descartes
UFR de Mathématiques et Informatique
45, rue des Saints-Pères 75270 Paris cedex 06

Maîtrises Maths-MASS - Statistiques générales

Partiel du 21 novembre 2003

8h30 - 11h30

Les documents ne sont pas autorisés.

Le sujet comporte deux pages.

Ex 1. Soit $(X_1, \ldots, X_n)$ un n-échantillon de de variables aléatoires de loi ayant pour densité

$$f_\theta(x) = \frac{x}{\theta} e^{-x^2/(2\theta)}$$

par rapport à la mesure de Lebesgue sur $\mathbb{R}^+$, avec $\theta > 0$.

1. Préciser le modèle statistique, et déterminer une statistique exhaustive complète et minimale. Le modèle est-il régulier ?
2. Calculer $\mathbb{E}_\theta\left[X_i^2\right]$ (on pourra faire le changement de variables $t = x^2$). En déduire l'estimateur UVMB (uniformément de variance minimale parmi les estimateurs sans biais) et efficace $\hat{\theta}_n$ de θ. Quelle est l'information de Fisher du modèle statistique ?
3. Montrer que $(\hat{\theta}_n, n \geq 1)$ est une suite d'estimateurs de θ fortement consistante, et qu'elle converge en loi vers θ à une vitesse de l'ordre de $\sqrt{n}$.
4. On définit un nouvel estimateur $T_n = c_n \hat{\theta}_n$ de θ. Montrer que son risque quadratique est proportionnel à θ^2. En déduire la valeur de c_n qui minimise ce risque, et montrer que $\hat{\theta}_n$ n'est pas un estimateur admissible de θ.

Ex 2. Soit $(X_1, \ldots, X_n)$ un n-échantillon de de variables aléatoires de loi ayant pour densité

$$f_\theta(x) = e^{-(x-\theta)} 1_{[\theta, +\infty[}(x)$$

par rapport à la mesure de Lebesgue sur $\mathbb{R}$, avec $\theta \in \mathbb{R}$.

1. Préciser le modèle statistique, et déterminer une statistique exhaustive complète et minimale, dont on décrira la loi. Le modèle est-il régulier ?
2. Calculer l'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ. Montrer que $(\hat{\theta}_n, n \geq 1)$ est une suite d'estimateurs de θ fortement consistante, et qu'elle converge en loi vers θ à une vitesse de l'ordre de n.
3. Calculer $\mathbb{E}_\theta\left[\hat{\theta}_n\right]$. En déduire l'estimateur UVMB de θ.

Ex 3. Le Loto est un jeu de hasard créé par la Française des jeux en 1976. Le parieur coche sur une grille six numéros distincts compris entre 1 et 49. À chaque tirage, six numéros distincts sont tirés au sort, avec une probabilité uniforme. Le parieur gagne le gros lot si ses numéros sont les six tirés.

Un jeune étudiant se demande comment estimer le nombre N de paris mis en jeu lors des deux tirages du 11 octobre 2003. Il suppose ce nombre constant, et les paris indépendants.

1. Il y a eu deux gagnants du gros lot lors du premier tirage, et quatre lors du second.
   (a) Déterminer, pour un pari, la probabilité p_6 de gagner le gros lot. Dans la suite, on prendra $p_6 \simeq 7,15.10^{-8}$.
   (b) Montrer qu'on peut modéliser le nombre de gagnants du gros lot lors de chaque tirage par une loi de Poisson de paramètre Np_6.
   (c) Soit X_1 (resp. X_2) le nombre de gagnants du gros lot lors du premier tirage (resp. le second). Préciser le modèle statistique associé à (X_1, X_2), déterminer une statistique exhaustive "simple". Avec la méthode des moments, construire deux estimateurs de N.
   (d) En utilisant l'inégalité de Bienaymé-Tchébycheff, construire un intervalle de confiance bilatère pour N de seuil $1 - \alpha$.
2. Pour affiner son étude, l'étudiant calcule la moyenne empirique $m = 1,789$ et la variance empirique $s^2 = 3,698$ du nombre de gagnants du gros lot au cours des deux dernières années. Il décide de modéliser N comme une variable aléatoire de loi gaussienne $\mathcal{N}(\mu, \sigma^2)$ avec $\mu = m/p_6$ et $\sigma^2 = (s^2 - m)/p_6^2$.
   (a) Justifier ce choix. Indication : $\mathbb{E}[X_1] = \mathbb{E}[\mathbb{E}[X_1|N]]$ et $\mathbb{E}[X_1^2] = \mathbb{E}\left[\mathbb{E}\left[X_1^2|N\right]\right]$.
   (b) Déterminer l'estimateur bayésien de N, qu'on exprimera en fonction des moments d'une gaussienne.