Université René Descartes

UFR de Mathématiques et Informatique

# EXAMEN DE STATISTIQUE

MAITRISE de MATHEMATIQUES et MAITRISE MASS

1ère Session - Année 2002-2003

30 Janvier 2003- 9h-12h

(Le sujet comporte 3 pages. Aucun document autre que le polycopié du cours n'est autorisé)

## Exercice 1

Soit $X_1, X_2, \cdots, X_n$ des variables aléatoires réelles indépendantes telles que, pour $k = 1, 2, \cdots, n$, la loi de $X_k$ est la loi normale $\mathcal{N}(k\theta, 1)$ o $\theta$ est un paramètre réel inconnu.

On pose $a_n = \sum_{k=1}^n k^2$ et on rappelle que $a_n = \dfrac{n(n+1)(2n+1)}{6}$.

1. Montrer que le modèle statistique associé à $X = (X_1, X_2, \cdots, X_n)$ est exponentiel et régulier.
   Justifier que $T_n = \dfrac{1}{a_n} \displaystyle\sum_{k=1}^{n} kX_k$ est une statistique exhaustive complète et minimale.
2. On veut estimer le paramètre $\theta$.
   (a) Déterminer l'estimateur du maximum de vraisemblance $\widehat{\theta}_n$ de $\theta$; donner sa loi et sa fonction de risque.
   Justifier que $\widehat{\theta}_n$ est une statistique régulière et un estimateur UVMB, efficace et faiblement consistant de $\theta$.

(b) Soit $m \in \mathbb{R}$ donné. Quelle est la loi a posteriori du paramètre $\theta$ sachant que $X_1 = x_1, X_2 = x_2, \cdots, X_n = x_n$ si la loi a priori est la loi normale $\mathcal{N}(m, 1)$ ? En déduire l'estimateur bayésien $\theta_n^\star$ associé ; donner sa loi et sa fonction de risque.

(c) Calculer les valeurs des fonctions de risque de $\widehat{\theta}_n$ et $\theta_n^\star$ en $\theta = m$. L'estimateur $\widehat{\theta}_n$ est-il optimal ?

3. On veut tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$ o $\theta_0 \in \mathbb{R}$ est donné. Exprimer, à l'aide de la fonction de répartition de la loi $\mathcal{N}(0, 1)$, la fonction puissance du test déterministe $\varphi$ de région critique $[|\widehat{\theta}_n - \theta_0| > c]$ , $c \in \mathbb{R}_+^\star$. Montrer que ce test est UPPB à son niveau.

   Donner la valeur $c_\alpha$ de $c$ pour que ce test soit de niveau $\alpha$ ( on exprimera $c_\alpha$ à l'aide de la fonction réciproque de la fonction de répartition de la loi $\mathcal{N}(0, 1)$ ).

   Justifier que $]\widehat{\theta}_n - c_\alpha, \widehat{\theta}_n + c_\alpha[$ est un intervalle de confiance au niveau $1 - \alpha$ pour $\theta$.

   Dans le cas où $n = 10$, donner la valeur de $c_\alpha$ pour $\alpha = 0,05$ ( on donne $\sqrt{385} = 19,62$).

**Exercice 2**

Soit $Y = (Y_1, Y_2, \cdots, Y_n)$ un vecteur aléatoire gaussien de matrice de covariance égale à $\sigma^2 I_n$, $\sigma^2 > 0$ étant inconnu.

On suppose que, pour $k = 1, 2, \cdots, n$, $Y_k = a + bx_k + cx_k^2 + \epsilon_k$ , o $a$, $b$ et $c$ sont des paramètres réels inconnus, $x_1, x_2, \cdots, x_n$ des réels donnés tous distincts tels que $\sum_{k=1}^n x_k = 0$ et $\sum_{k=1}^n x_k^3 = 0$ et $(\epsilon_1, \epsilon_2, \cdots, \epsilon_n)$ un n-échantillon de la loi $\mathcal{N}(0, \sigma^2)$.

On pose $t = \dfrac{1}{n}\displaystyle\sum_{k=1}^n x_k^2$ et $v = \dfrac{1}{n}\displaystyle\sum_{k=1}^n (x_k^2 - t)^2$.

1. On pose $\lambda = a + c\, t$ et $\beta = (\lambda, b, c)$.

   Montrer que l'on peut écrire, matriciellement, $E(Y) = X\beta$ où $X$ est une matrice que l'on déterminera.

   Calculer ${}^tXX$ et en déduire les estimateurs des moindres carrés de $\lambda$, $a$, $b$ et $c$.

   Donner l'estimateur de $\sigma^2$, noté $\widehat{\sigma}^2$, optimal dans la classe des estimateurs sans biais.

   Donner les lois des estimateurs précédents ( on donnera des expressions où les valeurs de $x_1, x_2, \cdots, x_n$ n'interviennent que par celles de $t$ et $v$).

2. On suppose, dans cette question, $c = 0$. Déterminer les estimateurs des moindres carrés de $a$ et $b$.

3. Justifier que, pour le test de l'hypothèse $H_0 : c = 0$ contre l'hypothèse $H_1 : c \neq 0$, la statistique du test de Fisher est égale à $\dfrac{\left(\sum_{k=1}^{n} (x_k^2 - t)\ Y_k\right)^2}{n\ v\ \widehat{\sigma}^2}$.
   Sous l'hypothèse $H_0$, donner la loi de cette statistique.