TD 5 : tests

Les questions marquées d'un astérisque **(*)** *sont facultatives.*

Exercice 1. (Propriétés des tests) Soit $(X_i)_{i\geq 1}$ une suite de v.a. i.i.d. de variance σ^2 connue et d'espérance μ. Soit $\mu_0 \in \mathbb{R}$. On cherche à tester l'hypothèse H0 : $\mu = \mu_0$ contre l'hypothèse H1 : $\mu \neq \mu_0$.

1. Pour chaque $n \geq 1$, étant donné le n-échantillon $(X_1, \dots, X_n)$ et un réel $\alpha \in (0, 1)$:
   (a) construire un test de niveau α utilisant l'inégalité de Bienaymé-Tchebychev,
   (b) construire un test de niveau asymptotique α, fondé sur le théorème central limite.

Dans la suite, on suppose les X_i gaussiens.

2. Ces tests sont-ils biaisés ? Ces suites de tests sont-elles consistantes ?
3. **(*)** Même question si on considère l'hypothèse alternative H1 : $\mu = \mu_1$ pour $\mu_1 \neq \mu_0$.
4. Montrer que les tests fondés sur le TCL sont uniformément plus puissants que ceux fondés sur l'inégalité de Bienaymé-Tchebychev.

Exercice 2. Soient $X_1, \dots, X_n$ des variables aléatoires telles que

$$\forall i \in \{1, \dots, n\}, \quad X_i = \theta + \epsilon_i,$$

où $\epsilon_1, \dots, \epsilon_n$ sont des variables aléatoires indépendantes, centrées et sous-Gausienne, i.e., $\mathbb{E}[e^{\lambda X}] \leq e^{\frac{\lambda^2}{2}}$, $\forall \lambda \in \mathbb{R}$. On veut tester $H_0 : \theta \leq 0$ contre $H_1 > 0$. On estime θ par $\hat{\theta} = n^{-1} \sum_{i=1}^n X_i$ et on pose $c_\alpha = \sqrt{2\ln(1/\alpha)/n}$.

1. Montrer que le test $\phi = \mathbf{1}_{\{\hat{\theta} > c_\alpha\}}$ est de niveau α.
2. Montrer que, pour tout $\theta > c_\alpha$, $\beta(\theta) > 1 - e^{-n(\theta - c_\alpha)^2/2}$.

Exercice 3. Soient n et p deux entiers tels que $p \leq n$ et soit X une matrice déterministe fixée à n lignes et p colonnes, de rang p. Soit $\mathbf{Y} = (Y_1, \dots, Y_n)$ une variable aléatoire de loi $\mathcal{N}(X\beta, \sigma^2 I)$ où $\beta \in \mathbb{R}^p$ et $\sigma^2 > 0$ sont inconnus. Soit $0 \leq r < p$. Dans cet exercice, on étudie le test du rapport de vraisemblance dans le cas où $H_0 : \beta_{r+1} = \cdots = \beta_p = 0$.

1. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}$ de $\theta = (\beta, \sigma^2)$ et en déduire que

$$\sup_{\theta \in \Theta} L(\theta, Y) = \frac{1}{(2\pi)^{n/2} \hat{\sigma}^2} e^{-n/2}.$$

2. Déterminer l'estimateur du maximum de vraisemblance $\hat{\theta}_0 de \theta = (\beta, \sigma^2)$ sous la contrainte que $\beta_{r+1} = \cdots = \beta_p = 0$. En posant

$$\mathbf{M} = X(X^\top X)^{-1} X^\top - X_0 (X_0^\top X_0)^{-1} X_0^\top = \Pi_X - \Pi_{X_0},$$

où X_0 est la matrice extraite de X, de taille $n \times r$ contenant les r premières colonnes de X, en déduire que

$$\sup_{\theta \in \Theta_0} L(\theta, \mathbf{Y}) = \frac{1}{(2\pi)^{n/2} (\hat{\sigma}^2 + \frac{1}{n} \mathbf{Y}^\top \mathbf{M} \mathbf{Y})^{n/2}} e^{-n/2}.$$

3. En déduire que le test du rapport de vraisemblance rejette H_0 si

$$F = \frac{n-p}{p-r} \frac{\mathbf{Y}^\top \mathbf{M} \mathbf{Y}}{n \hat{\sigma}^2} > k_\alpha,$$

pour un certain seuil k_α.

4. Montrer que, si $\theta \in \Theta_0$, F suit une loi de Fisher dont on précisera les paramètres.

Exercice 4. (*)

1. Rappeler la définition d'une matrice symétrique définie positive. D'une matrice symétrique semi-définie positive.
2. Les matrices suivantes sont-elles symétriques définies positives ? Symétriques semi-définies positives ?

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & -2 \\ -2 & 2 \end{pmatrix}$$

Exercice 5. (*) Corrigez tout ce qui est faux dans le raisonnement suivant :

≪ *Soient X et Y deux v.a. gaussiennes d'espérance 1 telles que $Var(X) = 1$, $Var(Y) = 2$ et $Cov(X, Y) = -2$. Alors (X, Y) est un vecteur gaussien d'espérance $(1, 1)$ et de matrice de covariance*

$$\begin{pmatrix} 2 & -2 \\ -2 & 1 \end{pmatrix}$$

donc le vecteur $(X + Y, X - Y)$ est un vecteur gaussien d'espérance $(2, 0)$ et de matrice de covariance

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 4 & -3 \end{pmatrix} \gg .$$

(*) De telles variables existent-elles ?