TD DE MODÈLES LINÉAIRES I - SÉRIE 4

Exercice 1. Soit X de loi $\mathcal{N}(0, I_n)$. Soit P une projection orthogonale de rang r. Montrer que $X'PX$ suit un χ^2_r.

Exercice 2. Soit X de loi $\mathcal{N}(0, I_n)$ et δ un vecteur fixé de $\mathbb{R}^n$. On définit la variable aléatoire réelle Y par

$$Y = \| X + \delta \|^2 = \sum_{i=1}^{n} (X_i + \delta_i)^2.$$

On dit que Y suit un $\chi^2(n, \| \delta \|^2)$. Supposons $\delta \neq 0$. Calculer l'espérance et la variance de Y.

Exercice 3. On considère les trois modèles de régression suivants

(1) $Y_k = b_1 + b_2 \ln X_k + \varepsilon_k$, $k = 1, ..., n$,

(2) $Y_k = b_1 + b_2 X_k + \varepsilon_k$, $k = 1, ..., n$,

(3) $Y_k = b_1 + b_2 X_k + b_3 X_k^2 + \varepsilon_k$, $k = 1, ..., n$,

où les ε_k sont des $\mathcal{N}(0, \sigma^2)$ indépendantes, $k = 1, ..., n$ et les $X_k \in \mathbb{R}$ sont déterministes.

(1) Calculer les estimateurs linéaires sans biais de variance minimale (ESBVM) pour chacun des trois modèles.

(2) Donner des estimateurs de σ^2 et leurs propriétés (ne pas expliciter les calculs).

Exercice 4. Soit le modèle de régression linéaire simple

$$Y_i = \theta_1 + \theta_2 X_i + \varepsilon_i, \quad i = 1, \ldots, n,$$

où ε_i sont des variables aléatoires gaussiennes indépendantes, de moyenne 0 et de variance $\sigma^2 > 0$ inconnue, $X_i \in \mathbb{R}$ sont des valeurs déterministes et θ_1 et θ_2 sont des paramètres réels inconnus. On note

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i,$$

et on suppose dans la suite que

$$S_X^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 > 0.$$

1) Expliciter $\hat{\theta}_1$ et $\hat{\theta}_2$, les estimateurs des moindres carrés de θ_1 et θ_2 respectivement. Expliciter également l'estimateur $\hat{\sigma}_n^2$ de σ^2.

2) Trouver les variances de $\hat{\theta}_1$, $\hat{\theta}_2$, ainsi que les covariances $Cov(\hat{\theta}_1, \hat{\theta}_2)$ et $Cov(\bar{Y}, \hat{\theta}_2)$. Montrer que $\hat{\theta}_1 \perp\!\!\!\perp \hat{\theta}_2$ si et seulement si $\bar{X} = 0$.

3) Donner la loi de la statistique

$$\frac{\sqrt{n}\, S_X(\hat{\theta}_2 - \theta_2)}{\hat{\sigma}_n}.$$

4) Soit $0 < \alpha < 1$. Proposer un test de niveau (exact) α de l'hypothèse $H_0 : \ \theta_2 = 0$ contre l'alternative $H_1 : \ \theta_2 \neq 0$.

Problème (janvier 1988). On considère un vecteur aléatoire normal $X = (X_1, ..., X_n)'$ de loi $\mathcal{N}(M, I_n)$, où $M = (\mu_1, ..., \mu_n)', n \geq 2$.

On suppose que M vérifie les équations

$$\sum_{i=1}^{n} \mu_i = 0 \qquad \text{et} \qquad \sum_{i=1}^{n} (-1)^i \mu_i = 0.$$

(1) Déterminer l'ESBMV de M.

(2) Tester l'hypothèse M = 0 au seuil α.

(3) A.N. $n = 5, X_1 = 2, X_2 = 0, X_3 = 4, X_4 = -1, X_5 = -2, \alpha = 0.05$.