Université Paris Dauphine — Année 2001-2002

U.F.R. Mathématiques de la Décision

MD4 - Modèles Linéaires et généralisations

Examen 17/12/2001

Les 3 exercices sont indépendants. Seules les réponses soigneusement justifiées seront prises en compte.

Exercice 1 Nous considérons le modèle de régression linéaire tel que pour tout $i \in \{1, \ldots, n\}$, nous avons

$$Y_i = \alpha x_i + E_i$$

où α est un paramètre inconnu et E_i est un variable aléatoire réelle telle que : $\mathbb{E}(E_i) = 0$, $\mathbb{V}(E_i) = \sigma^2 k_i$ où $k_i > 0$ pour tout $i \in \{1, \ldots, n\}$ et $cov(E_i, E_j) = 0$ lorsque $i \neq j$.

Posons $Y = (Y_1, \ldots, Y_n)$, $X = (x_1, \ldots, x_n)$, $E = (E_1, \ldots, E_n)$ et $\mathbb{V}(E) = \Sigma$.

1 Donner l'expression de l'estimateur des Moindres Carrés Ordinaires $\widehat{\alpha}$ de α. Calculer sa variance.

2 Calculer l'expression de l'estimateur des Moindres Carrés Généralisés $\widetilde{\alpha}$ de α. Calculer sa variance.

3 Comparer $\mathbb{V}(\widehat{\alpha})$ et $\mathbb{V}(\widetilde{\alpha})$.

4 Nous supposons que Y est un vecteur gaussien. Nous avons $n = 456$, $\sum_{i=1}^{456} \left(\frac{x_i y_i}{k_i} \right) = 15$, $\sum_{i=1}^{456} \left(\frac{x_i^2}{k_i} \right) = 25$, $\sum_{i=1}^{456} (y_i - a x_i)^2 = 6$ où a est la réalisation de $\widetilde{\alpha}$. Proposer un intervalle de confiance de niveau 95 % pour le paramètre α. Nous rappelons que $F^{-1}_{T(455)}(0.975) = 1.965$.

Exercice 2 Nous considérons un échantillon i.i.d $X_1, \dots, X_n$ de loi de probabilité $\mathcal{N}(\mu_X, \sigma^2)$ et un échantillon i.i.d. $Y_1, \dots, Y_n$ indépendants du précédent de loi de probabilité $\mathcal{N}(\mu_Y, \sigma^2)$.

1 Donner les expressions des estimateurs du maximum de vraisemblance $\widehat{\mu_X}$, $\widehat{\mu_Y}$ et $\widehat{\sigma^2}$ des paramètres μ_X, μ_Y et σ^2.

2 Calculer la matrice d'information de Fisher $\mathrm{I}(\mu_X, \mu_Y, \sigma^2)$.

3 Nous souhaitons tester l'hypothèse H_0 selon laquelle $\mu_X = \mu_Y$ contre l'hypothèse H_1 selon laquelle $\mu_X \neq \mu_Y$. Construire les statistiques des tests de Wald et du rapport de vraisemblance.

4 D'après le test de Wald, pour un risque de première espèce asymptotiquement égal à 5 %, quelle décision prenons-nous lorsque $n = 20$, $\overline{x} = 3$, $\overline{y} = 4.2$, $\left(\sum_{i=1}^{20}(x_i - \overline{x})^2 + \sum_{i=1}^{20}(y_i - \overline{y})^2\right) = 80$?
Nous rappelons que $F^{-1}_{\chi^2(1)}(0.95) = 3.841$.

Exercice 3 Nous considérons n variables aléatoires indépendantes de loi de Poisson de paramètre $\mu_i = \mathbb{E}(Y_i)$. Nous supposons que $\log(\mu_i) = \alpha + \beta x_i$ où α et β sont des paramètres inconnus et pour tout $i \in \{1, \dots, n\}$ x_i est un réel connu.

1 Donner les équations de vraisemblance. Pouvons-nous calculer les expressions analytiques des estimateurs du maximum de vraisemblance de α et β ?

2 Calculer la matrice d'information de Fisher $\mathrm{I}(\alpha, \beta)$.

3 Notons $\widehat{\beta}$ l'estimateur du maximum de vraisemblance de β. Quelle est la loi de probabilité asymptotique de la variable aléatoire $\sqrt{n}\left(\widehat{\beta} - \beta\right)$?

4 Pour un risque de première espèce asymptotiquement égal à 5%, proposer un test de l'hypothèse H_0 selon laquelle $\beta = 0$ contre l'hypothèse H_1 selon laquelle $\beta \neq 0$.