Statistique/Modèle linéaire

Partiel du 14 novembre 2001

Avec documents (durée deux heures). Deux exercices indépendants.

Exercice 1

On introduit une fonction de coût appelée LINEX,

$$\mathrm{L}(\theta, d) = e^{c(d-\theta)} - c(d-\theta) - 1, \qquad c > 0$$

1. Montrer que $\mathrm{L}(\theta, d)$ est toujours positive et ne s'annule que pour $d = \theta$.

2. Montrer que
$$\lim_{c \to 0} \frac{\mathrm{L}(\theta, d)}{c^2} = \frac{(d-\theta)^2}{2}\,.$$

3. Montrer que, pour une observation x de loi f_θ et une loi a priori $\pi(\theta)$, l'estimateur de Bayes sous le coût LINEX est
$$\delta^\pi(x) = \frac{-1}{c} \log \left\{ \mathbb{E}^{\pi(\theta|X=x)} [e^{-c\theta} | X = x] \right\}\,. \tag{1}$$

4. Déterminer l'estimateur de Bayes δ^π donné par (1) dans le cas où X suit une loi $\mathcal{N}(\theta, 1)$ et où π est la loi non-informative $\pi(\theta) = 1$.

5. Dans le cadre de la question 4, comparer le risque classique $R(\theta, \delta^\pi) = \mathbb{E}_\theta \left[\mathrm{L}(\theta, \delta^\pi(X)) \right]$ de l'estimateur de Bayes ainsi obtenu avec celui de $\delta_0(x) = x$, sous le coût LINEX.

6. Déterminer l'estimateur de Bayes δ^π donné par (1) lorsque x suit une loi $\mathcal{G}(\alpha, \theta)$ et lorsque π est la loi non-informative $\pi(\theta) = 1/\theta$ sur $\mathbb{R}_+$. On rappelle que la loi gamma $\mathcal{G}(\alpha, \theta)$ a pour densité
$$f(x|\alpha, \theta) = \frac{\theta^\alpha x^{\alpha-1}}{\Gamma(\alpha)}\, e^{-\theta x}\,, \qquad x \in \mathbb{R}_+$$

Exercice 2

On considère $X_1, \dots, X_n$ échantillon iid $\mathcal{N}(\theta, \sigma^2)$, avec σ connu.

1. Donner la vraisemblance de cet échantillon et montrer qu'elle factorise en $\overline{x}$, moyenne empirique des x_i, réalisations des variables X_i.

2. Si on teste $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$, avec $\theta_0 < \theta_1$, montrer que le lemme de Neyman–Pearson conduit à la procédure de test optimale suivante: accepter H_0 si et seulement si
$$\overline{x} < \frac{(2\sigma^2 \log k) - \theta_0^2 + \theta_1^2}{2(\theta_1 - \theta_0)}, \tag{2}$$
où k est une constante déterminée par le niveau du test.

3. Montrer que, malgré la dépendance formelle de (2) à θ_1 et θ_0, le choix d'un niveau α pour le test de H_0 contre H_1 conduit à une borne (2) indépendante de θ_1. (On donnera cette borne en fonction des quantiles de la loi normale $\mathcal{N}(0, 1)$.)

4. Déduire de la question 3 que le test précédent est aussi optimal (c'est à dire plus puissant) pour tester $H_0 : \theta \leq \theta_0$ contre $H_1 : \theta > \theta_1$ en montrant qu'il est bien au niveau α pour tous les $\theta \leq \theta_0$.

5. Pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta \neq \theta_0$, montrer, à partir des questions précédentes, qu'on peut construire deux tests au niveau α dont les fonctions puissances ne sont pas comparables. En déduire qu'il n'existe pas de test uniformément plus puissant pour ce problème.

6. On considère à présent le problème de test inverse, soit
$$H_0 : \theta \neq \theta_0 \quad \text{contre} \quad H_1 : \theta = \theta_0 \,.$$
Montrer qu'une procédure de test φ uniformément plus puissante au niveau α, maximisant la probabilité de rejet de H_0, $P_{\theta_0}(\varphi(X_1, \dots, X_n) = 0)$, sous la contrainte
$$P_\theta(\varphi(X_1, \dots, X_n) = 0) \leq \alpha, \quad \theta \neq \theta_0 \,,$$
est telle que
$$\varphi^*(x_1, \dots, x_n) = \begin{cases} 1 & \text{si } |\overline{x} - \theta_0| > c, \\ 0 & \text{si } |\overline{x} - \theta_0| < c, \end{cases}$$
où c est déterminé par α.