Devoir à rendre pour le 19 Décembre 2006

Statistiques - 2006-2007

Dans tout ce qui suit, les intervalles de confiance demandés sont bilatères.

Exercice 1
Soient $Y_1, ..., Y_n$ n varibles aléatoires réelles i.i.d. $\mathcal{N}(m, \sigma^2)$. On suppose que m et σ^2 sont inconnues. Donner, en le justifiant, un intervalle de confiance exact sur m au niveau α.

Exercice 2 Soit $X_1, \ldots, X_n$ un n-échantillon de loi $\mathcal{N}(\mu_X, \sigma_X^2)$. Soit $Y_1, \ldots, Y_m$ un m-échantillon de loi $\mathcal{N}(\mu_Y, \sigma_Y^2)$. On suppose que $\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2$ sont inconnus.

1. Supposons que $\sigma_X^2 = \sigma_Y^2$. Donner, en le justifiant, un intervalle de confiance sur $\mu_X - \mu_Y$ au niveau α.
2. On ne suppose plus que $\sigma_X^2 = \sigma_Y^2$. Donner, en le justifiant, un intervalle de confiance asymptotique sur $\mu_X - \mu_Y$ au niveau α (ie n et m tendent vers $+\infty$).
3. Construire un test de niveau α (qui ne soit pas le test trivial $\phi = 0$) de "H_0 : $\sigma_X^2 = \sigma_Y^2$" contre "H_1 : $\sigma_X^2 \neq \sigma_Y^2$" et donner sa puissance à n et m fixés.

Exercice 3 Soit Y un vecteur Gaussien de dimension n tel que $Y = X\theta + \varepsilon$ où $X \in M_{n,p}$ est injectif et connu, $\theta \in \mathbb{R}^p$ est inconnu, $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$ avec $\sigma^2 \geq 0$ inconnu. Soient $a_1, \ldots, a_\ell$ des vecteurs déterministes de $\mathbb{R}^p$ linéairement indépendants. On cherche des intervalles de confiances simultanés sur tous les $a_i^* \theta$ au niveau α , ie on cherche ℓ intervalles de confiance $I_i(\alpha)$ tels que

$$P_{\theta,\sigma^2}\left(\forall i \in \{1, ..., \ell\}, a_i^* \theta \in I_i(\alpha)\right) \geq 1 - \alpha.$$

1. (Méthode de Bonferroni) Soit i fixé. Construire un intervalle de confiance exact au niveau α, $\tilde{I}_i(\alpha)$, pour $a_i^* \theta$. Montrer que le choix $I_i(\alpha) = \tilde{I}_i(\alpha/\ell)$ permet de réaliser ℓ intervalles de confiance simultanés.
2. Pour tout i soit $Z_i = a_i^*(X^*X)^{-1}X^*Y$ et soit Z tel que $Z^* = (Z_1, ..., Z_\ell)$. Montrer que Z est un vecteur Gaussien dont on donnera la moyenne μ et la matrice de covariance Λ. Montrer que Λ est définie positive.
3. Donner la loi de $\|\Lambda^{-1/2}(Z - \mu)\|^2$.
4. En déduire une région de confiance exacte pour μ, B, au niveau α , ie telle que
$$P_{\theta,\sigma^2}\left(\mu \in B\right) \geq 1 - \alpha.$$
Indication : on pourra exprimer B comme un ellipsoïde que l'on précisera.
5. En déduire un choix d'intervalles $I_i(\alpha)$.
6. Pour simplifier on suppose que $a_i^*(X^*X)^{-1}a_j$ vaut 1 si $i = j$ et 0 sinon. Montrer que les intervalles construits en 5 sont plus fins que ceux construits en 1 quand α tend vers 0.