Examen du 13 mai 2008

Durée : 2 heures. Notes de cours et polycopiés de cours et d'exercices de statistique autorisés

Version très préliminaire !!!

Exercice 1. Soit $n \geq 2$ et $\varepsilon_1, \ldots, \varepsilon_n$ des variables aléatoires gaussiennes indépendantes, centrées et réduites. On observe une réalisation des variables aléatoires $Z_1, \ldots, Z_n$ définies par récurrence :
$Z_1 = \sqrt{\theta}\, \varepsilon_1$ et, pour $i \geq 2$,

$$Z_i = Z_{i-1} + \sqrt{\theta}\, \varepsilon_i.$$

Ici, le paramètre inconnu est $\theta > 0$.

(1) Les variables Z_i sont-elles indépendantes ?

(2) On construit les variables Y_i en posant $Y_1 = 0$ et pour $i \geq 2$,

$$Y_i = (Z_i - Z_{i-1})^2.$$

Ces variables sont-elles observables ? A l'aide des variables Y_i, proposer un estimateur sans biais $\hat{\theta}_n$ de θ qui, de plus, converge vers θ lorsque $n \to \infty$.

(3) Construire un intervalle de confiance de niveau de risque α pour θ.

(4) Proposer un test[1] de l'hypothèse $H_0 \;:\; \theta = 1$ contre $\theta \neq 1$ de niveau α, de sorte que l'erreur de seconde espèce soit arbitrairement petite lorsque le nombre de données augmente.

Exercice 2. On observe une réalisation de n variables aléatoires $X_1, \ldots, X_n$ avec :

$$X_i = \delta_i Z_i^{(1)} + (1 - \delta_i) Z_i^{(2)}$$

où les δ_i sont connus, fixes et appartiennent à la paire $\{0, 1\}$, les $Z_i^{(j)}$ sont toutes indépendants et gaussiennes de variance égale à 1, et

$$\mathbb{E}[Z_i^1] = \mu_1, \quad \mathbb{E}[Z_i^2] = \mu_2.$$

(1) Calculer l'espérance et la variance de X_i. Quelle est la loi des X_i ?

(2) On pose $\beta = (\mu_1, \mu_2)^t$. Montrer que l'on peut écrire le modèle sous la forme d'un modèle linéaire gaussien

$$X = M\beta + \sigma\varepsilon$$

où ε est un vecteur de gaussiennes centrées réduites indépendantes, σ est un réel strictement positif dont on donnera la valeur, et M est une matrice que l'on explicitera en fonction des δ_i.

(3) Proposer un estimateur de β.

(4) Proposer un test de niveau α pour tester l'hypothèse

$$\mu_1 = \mu_2.$$

[1] éventuellement asymptotique

Exercice 4. Soit X le nombre de garçons dans une famille de 4 enfants. Dans un échantillon de 200 familles de 4 enfants, on observe les résultats suivants:

Valeurs de X	0	1	2	3	4
Nombre de familles	13	65	72	35	15

(1) Le test du χ^2 au niveau $0,05$ conduit-il à accepter l'hypothèse que la probabilité p d'avoir un garçon pour chaque naissance est égale à $\frac{1}{2}$?

(2) On se contente maintenant de considérer le nombre total N de garçons nés dans les 200 familles.

A l'aide d'un test de région critique (de rejet) de la forme

$$\mathcal{R}_t = \{|N - \mathbb{E}_0[N]| \geq t\}$$

où $t > 0$ est à choisir, dire si l'on doit rejeter l'hypothèse $p = \frac{1}{2}$ au niveau de risque $0,05$?

Exercice 5. On lance 60 fois un dé et on obtient les résultats suivants :

1	2	3	4	5	6
10	13	8	12	9	8

Au seuil de $0,025$ peut-on conclure que le dé est bien équilibré ?