

Conservatoire National des Arts et Métiers : STA 112
Examen
29 juin 2012. 18h00-20h00

Documents (papier ou électronique) et machines à calculer autorisés

Exercice 1

On mesure chez 50 athlètes, leur taux de glucocorticoïdes à 10 reprises.

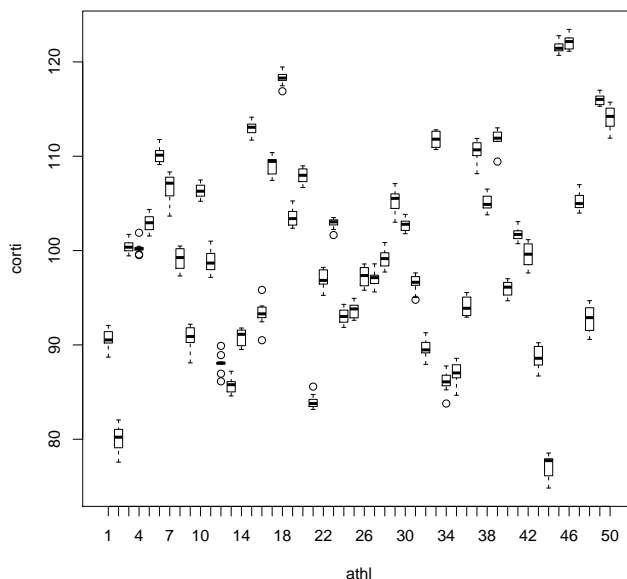


Figure 1: Taux de glucocorticoïdes moyen par athlète

1. A partir de la Figure 1, précisez les sources de variabilité du taux de glucocorticoïdes Y_{ij} chez le sportif i à la mesure j .

Il faut évoquer les types de variabilité : intra individus (les 10 mesures de chaque athlète sont différentes) et inter individus. Plus précisément, On observe une dispersion intra-individuelle faible (les box-plot par athlètes ne sont pas large) et une forte variation inter-individuelle. les concentrations semblent groupées autour de 100. Cependant, certains athlètes ont des concentrations qui s'éloignent de cette valeur.

2. On se propose d'utiliser le modèle suivant $Y_{ij} = \mu + A_i + \varepsilon_{ij}$. On supposera que : $A_i \sim N(0, \sigma_A^2)$, $\varepsilon_{ij} \sim N(0, \sigma^2)$ et A_i et ε_{ij} sont indépendants.

Justifiez son utilisation et commentez les estimations des paramètres $(\hat{\mu}, \hat{\sigma}_A, \hat{\sigma})$.

Ce modèle est approprié car la seule source de variabilité est le facteur Athlète. La normalité des effets aléatoires devra être testée. L'estimation de l'intercept proche de 100 confirme l'inspection visuelle.

$\hat{\mu}$	$\hat{\sigma}_A$	$\hat{\sigma}$
99.91	10.50	0.98

3. Calculez un intervalle de confiance à 95% pour le taux de glucocorticoïdes. Indication : le modèle de régression vous fournit la loi de Y . La réponse moyenne suit une loi normale $N(\mu, \sigma_A^2 + \sigma^2)$. L'intervalle à 95% est donc $\mu \pm 2\sqrt{\sigma_A^2 + \sigma^2}$. Soit [78.9; 121]

Exercice2

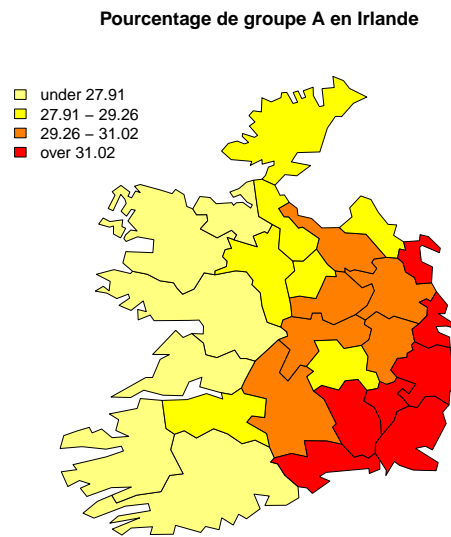


Figure 2: Pourcentage des individus de groupe sanguin A dans les 26 comtés d'Irlande

Le calcul des indices de Moran et Geary nous fournit les résultats suivants :

	Indice	Espérance
Moran	0.55	-0.038
Geary	0.38	1

1. Qu'en concluez-vous ? Les Indices de Moran et Geary sont des indices d'autocorrélation spatiale. L'indice de Geary étant plus petit que 1, il y a une corrélation spatiale (positive).

L'indice de Moran fournit une conclusion similaire. La répartition du groupe sanguin A en Irlande possède une structure spatiale

2. Proposez un calcul de l'espérance du Moran. $E(I) = -\frac{1}{n-1} = -0.04$ sous l'hypothèse d'indépendance spatiale !

Exercice 3

On considère les séquences suivantes $S_1 = 0123$ et $S_2 = 0000$

1. Dénombrer et lister les substitutions possibles. Les 2 séquences sont de longueur $k = 4$. Les substitutions compatibles avec la nature des données sont $(0, 1), (0, 2), (0, 3), (1, 2), (1, 3), (2, 3)$.
2. On précise maintenant que ces 2 séquences retracent le nombre d'enfants de 2 femmes. Quels coûts associés aux substitutions précédentes proposeriez-vous ? Il faudrait prendre en compte le fait que le passage de 0 à 3 enfants, à un coût supérieur au passage de 0 à 1. D'autre part les coûts ne seraient pas symétriques.