STA112 : Exemple

A. Latouche

Sept 2013 : Ex 2

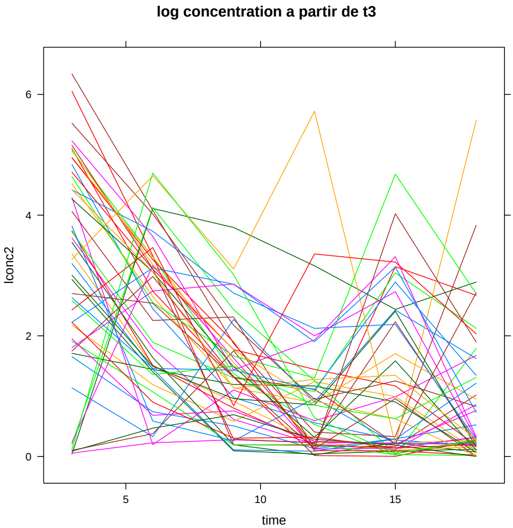
Soit Y_i le nombre observé de cas dans l'unité géographique i, E_i le nombre attendu de cas. On suppose que $Y_i \sim Poisson(\theta_i E_i)$.

1. Calculer la variance de $\hat{\theta}_i$ (estimateur du maximum de vraisemblance).
2. Sur les données de cas de cancers oral en écosse (56 unités géographiques), on obtient les résultats suivants la valeur moyenne du SMR vaut 1.57, la variance des SMR vaut 1.71. Le modèle proposé pour les Y_i est-il compatible avec ces estimations ?
3. Quel modèle proposeriez vous pour prendre en compte la nature spatiale des données ?

Sept 2013 : Ex 3

- La réponse anticorps de l'enfant se mesure en concentration d'anticorps (mg/l).
- On souhaite expliquer le comportement immunitaire de 50 nouveaux nés de 0 à 18 mois.
- On note $Y_{i,t}$ la concentration d'anticorps de l'individu i au temps t (t variant de 0 à 18 mois).

Le tracé des log-concentrations de l'échantillon d'enfants est représenté en Figure (1).



1. On note le modèle de régression

$$M_1 : Y_{i,t} = \alpha + \alpha_i + \beta t + \varepsilon_{i,t}$$

2. On note le modèle de régression

$$M_2 : Y_{i,t} = \alpha + \alpha_i + (\beta + \beta_i)t + \varepsilon_{i,t}$$

On notera $\varepsilon_{i,t} \sim N(0, \sigma^2)$ le terme d'erreur aléatoire,
$\alpha_i \sim N(0, \mu^2)$ et $\beta_i \sim N(0, \nu^2)$
L'estimation des paramètres du modèle M_1 et M_2 nous fournit :

Modèle	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\beta}$	$\hat{\nu}$	$\hat{\sigma}$
M_1	3.15	0.30	-0.14		1.41
M_2	3.14	1.20	-0.14	0.01	1.20

Table : Estimation des effets fixes et des effets aléatoires

Id	(Intercept)	time
1	-0.41	-0.06
2	0.84	-0.08

Table : Prédiction des effets aléatoires pour le modèle M_2

1. Commenter les différences entre les 2 modèles.
2. Calculer un intervalle de prédiction à 95% de la log-concentration pour les enfants 1 et 2.

La réponse conditionnelle aux effets aléatoires suit une loi

$$N(\alpha + \alpha_i + (\beta + \beta_i)\ t, \sigma^2)$$

Pour un temps donné, t^*, un intervalle de prédiction à 95 % pour l'individu i s'écrit sous la forme

$$\alpha + \alpha_i + (\beta + \beta_i)\ t^* \pm 2\ \sigma$$

3. Calculer un intervalle de prédiction à 95% de la log-concentration pour l'échantillon de 50 enfants.

La réponse moyenne suit une loi

$$N(\alpha + \beta \; t, \mu^2 + \nu^2 \; t^2 + \sigma^2)$$

Pour un temps donné, t^*, un intervalle de prédiction à 95 % s'écrit sous la forme

$$\alpha + \beta \; t^* \pm 2\sqrt{\mu^2 + \nu^2 \; (t^*)^2 + \sigma^2)}$$