

Modèles mixtes pour l'analyse de données longitudinales : TP

Partie 1 : modèles linéaires mixtes en R (fonction lme)

La fonction `lme` de la librairie `nlme` permet d'analyser un modèle linéaire à effets mixtes par maximum de vraisemblance. Il est plus facile d'organiser les données en `data.frame` (vous pouvez taper `?data.frame` dans la console R pour accéder à la page d'aide). Le jeu de données `Orthodont` de la librairie `nlme` est utilisé en exemple. `Orthodont` contient des données provenant d'une étude dentaire (Pothoff et Roy, 1964). Il s'agit de données répétées, il y a quatre mesures pour chaque individu.

Un appel de la fonction `lme` se fait par : `lme(fixed,data,random)`. Le premier argument est une formule spécifiant le modèle des effets fixes, le second argument correspond au nom du jeu de données étudié et le troisième argument est une formule indiquant les effets aléatoires et leur structure. Par exemple,

```
> fit.lme<- lme(distance~age,data=Orthodont,random=~age)
```

- Pour obtenir les effets aléatoires individuels estimés (post-hoc), on utilise la fonction `ranef`.
- On peut obtenir les résidus par la fonction `resid` et son option `level`. Pour obtenir les résidus standardisés (ou de Pearson) individuels, c'est-à-dire les résidus individuels bruts divisés par leur déviation standard, on utilise `resid(fit.lme,level=1,type= "pearson")`.
- On peut obtenir les intervalles de confiance de chaque paramètre par la fonction `intervals`.
- Par défaut, la fonction `lme` utilise la méthode de vraisemblance restreinte (REML). Pour utiliser le maximum de vraisemblance, on utilise l'option `method= "ML"` dans l'appel de `lme`.
- La fonction `update` permet de modifier un objet `lme` préexistant. Par exemple,
`update(fit.lme, method= ``ML'')`

Exercice 1 : courbe de croissance d'une mesure orthodontique

1. Chargez le jeu de données `Orthodont` de la librairie `nlme`. De quoi est constitué ce jeu de données ? Comment sont regroupées les données ? Pourquoi utiliser un modèle mixte pour étudier ces données ?
2. Tracez le “spaghetti plot” des données.
3. Etudiez la relation entre la variable `distance` et la variable `age` par un modèle linéaire mixte (avec effet aléatoire sur la variable `age` et l’intercept).
4. Comparez les résultats obtenus par les méthodes d’estimation ML et REML.
5. Tracez le “spaghetti plot” des données pour chaque sexe. Qu’observez-vous ? Testez si le sexe a une influence sur le modèle précédent.
6. Choisissez la structure de la matrice de variance-covariance des effets aléatoires.
7. Comparez le modèle d’erreur homoscedastique à deux modèles hétéroscedastiques couramment utilisés. Concluez.
8. Tracez les graphes de qualité d’ajustement du modèle retenu.