Modèles Linéaires à Effets Mixtes

Pr Roch Giorgi

roch.giorgi@univ-amu.fr

SESSTIM, Faculté de Médecine, Aix-Marseille Université, Marseille, France

http://sesstim-orspaca.org

http://optim-sesstim.univ-amu.fr/

Exemple Introductif (1)

- Etude de 15 visages de blancs et 15 visages de noirs
- Critère de jugement : échelle de -20 à +20
- 30 individus sollicités
- Question
  - ✓ Y a-t-il une différence dans la valeur du critère de jugement selon le visage ?
- Comment analyser ces données ?
  - ✓ Pour chaque individus, calcul du score de différence
  - ✓ Régression sur le score de différence

Exemple Introductif (2)

$$y_i = \alpha_0 + \varepsilon_i$$

y_i = score de différence de l'individus i
α_0 = intercept du modèle
ε_i = résidus de l'individus i

- Modèle linéaire simple
- Elimine problème de non-indépendance des résidus
- Fonction dans R
  - ✓ `lm(y ~ 1, data)`

Exemple Introductif (3)

- Remarques
  - ✓ Similaire à une ANOVA intra à 2 modalités
  - ✓ Faible possibilité de généraliser
    - Plus de 2 mesures ?
    - Effets individuels ?
- Autre solution statistique
  - ✓ Modéliser / Estimer les sources de non-indépendance (variabilité)
  - ⇒ Modèle mixte

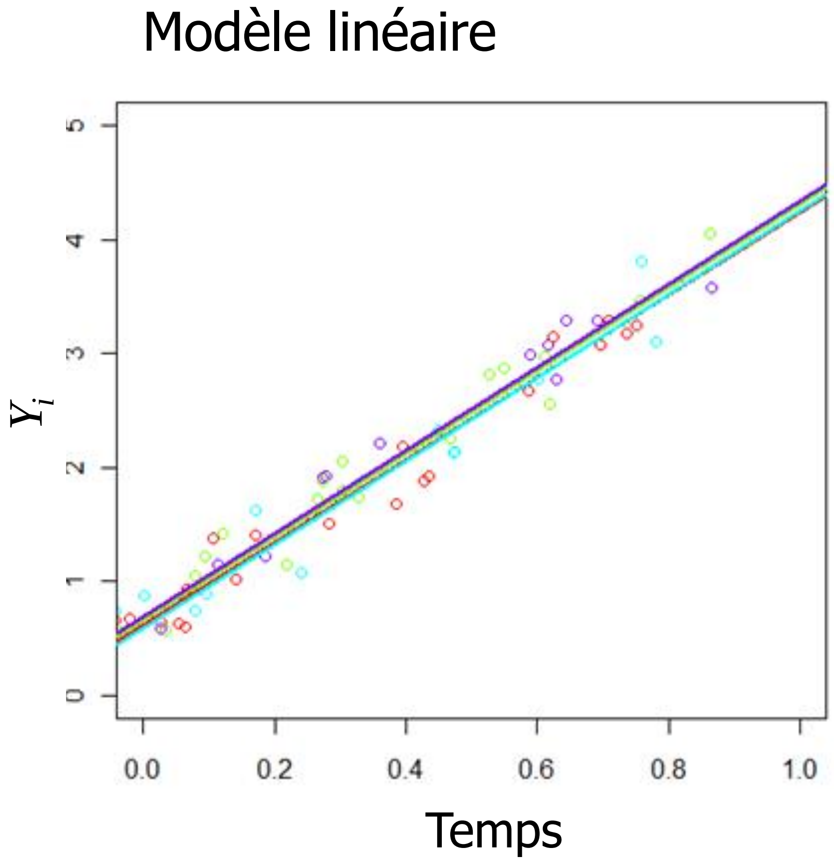
Modèle Linéaire (1)

$$Y_i = \beta X_i + \varepsilon_i$$

Y_i = n x 1 mesures du sujet i (continue)
X_i = n x p matrice de vecteur de covariables
β = p x 1 paramètres pour les covariables
ε_i = n x 1 résidus de l'individus i

- Hypothèses
  - ✓ Observations indépendantes entre les sujets
  - ✓ Observations non-indépendantes pour un même sujet
  - ✓ Y suit une loi Normale de variance-covariance Σ constante pour tout X (homoscédasticité) : Var(ε_i)=σ^2

Modèle Linéaire (2)



A $t=0$, il y a un point de départ moyen.
Le même pour tous les sujets

Modèle Linéaire – Estimation (1)

$$\underset{\beta}{Min} \sum_{i=1}^{n} \left[y_i - \sum_{k=1}^{p} \beta_k x_{ik} \right]^2$$

- Moindres carrés ordinaires : observations indépendantes

$$\hat{\beta} = \left(X^{'} X \right)^{-1} X^{'} Y$$

$$\operatorname{var}\left(\hat{\beta}\right) = \left(X^{'} X \right)^{-1} X^{'} \operatorname{var}\left(Y\right) X \left(X^{'} X \right)^{-1}$$

Indépendance des observation

$$\operatorname{var}\left(\hat{\beta}\right) = \sigma_{\varepsilon}^{2} \left(X^{'} X \right)^{-1}$$ *meilleur estimateur linéaire non biaisé*

Modèle Linéaire – Estimation (2)

$$\underset{\beta}{Min}\sum_{i=1}^{n}\left[y_i-\sum_{k=1}^{p}\beta_k x_{ik}\right]^2$$

- Moindres carrés généralisés : observations non-indépendantes

$$\hat{\beta}=\left(X'\Sigma^{-1}X\right)^{-1}X'\Sigma^{-1}Y$$

$$\operatorname{var}\left(\hat{\beta}\right)=\left(X'\Sigma^{-1}X\right)^{-1}$$

où

$\operatorname{var}(Y)=\Sigma$ a une structure reflétant la corrélation

Modèle Linéaire

- Ne permet pas de modéliser de manière appropriée l'évolution d'une mesure au cours du temps (>2)
- Ne prend pas en compte variabilité individuelle
  - ✓ Par rapport à l'évolution moyenne
  - ✓ Selon le(s) facteur(s) étudié(s)
  - ✓ Au niveau des résidus

Modèle Mixte – Principe (1)

1. Existence d'effets identiques pour tous les sujets de la population (tendance moyenne)
2. Chaque sujet a un certain niveau de réponse par rapport à la population
   - ✓ Avec 2 sources de variations aléatoires
     - Entre les sujets
     - Variation des mesures par sujet

1. Effets fixes
2. Effets aléatoires
   - ✓ Permettent d'expliciter les différences entre les individus (les corrélations) sans observer les déterminants de cette variabilité interindividuelle

Modèle Mixte – Principe (2)

- Effets fixes
  - ✓ Effet d'une variable dont les niveaux peuvent être reproduits et forment un ensemble fini

- Effets aléatoires
  - ✓ Effet d'une variable dont les niveaux sont échantillonnés aléatoirement à partir d'une population plus importante

Modèle Mixte *vs* MANOVA

- Modélisation explicite du changement individuel au cours du temps
- Nombre de mesures différents par sujet
- Temps de mesures différents par sujet
- Temps continu
- Spécification flexible de la structure de covariance
- Covariables dépendantes ou non du temps
- Données manquantes
- Domaine des modèles multi-niveaux
- ...

Modèle Mixte – Formulation (1)

$$Y_i = \beta X_i + \gamma_i Z_i + \varepsilon_i$$

Y_i = n_i x 1 mesures du sujet i (continue)
X_i = n_i x p matrice de vecteur pour les effets fixes
β = p x 1 paramètres pour les effets fixes
Z_i = n_i x p matrice de vecteur pour les effets aléatoires
γ_i = r x 1 paramètres pour les effets aléatoires
ε_i = n_i x 1 résidus de l'individus i

Modèle Mixte – Formulation (2)

Evolution moyenne au cours du temps

Déviations individuelles par rapport à la moyenne

$$Y_i = \beta X_i + \gamma_i Z_i + \varepsilon_i$$

« Vraie valeur » (non observée) dépourvue d'erreur de mesure

Modèle Mixte – Hypothèses (1)

$$Y_i = \beta X_i + \gamma_i Z_i + \varepsilon_i$$

Effets aléatoires et erreurs sont indépendants d'un sujet à l'autre

$$\gamma_i \perp \gamma_j, \forall i, j = 1, \ldots, N$$

$$\varepsilon_i \perp \varepsilon_j, \forall i, j = 1, \ldots, N$$

Erreurs sont indépendantes d'une mesure à l'autre chez un même sujet

$$\varepsilon_{ij} \perp \varepsilon_{ik}, \forall j, k = 1, \ldots, n_i$$

Modèle Mixte – Hypothèses (2)

$$Y_i = \beta X_i + \gamma_i Z_i + \varepsilon_i$$

Effets aléatoires γ_i non estimés

$$\gamma_i \sim N\left(0, \Sigma_\gamma\right)$$

Erreur de mesure indépendante pour chaque observation

$$\varepsilon_i \sim N\left(0, \Sigma_i\right)$$

Estimation

- Conditionnellement au vecteur des effets aléatoires γ_i les mesures répétées sont indépendantes entre elles
- En supposant l'indépendance entre les sujets
- Ecriture de la vraisemblance du modèle
- Estimation du modèle (fixe, aléatoire) par la méthode du maximum de vraisemblance
  - ✓ Existe d'autres possibilités

Prédictions – Résidus

- Effets aléatoires individuels peuvent être estimés
- Marginalement aux effets aléatoires
  - ✓ Prédictions $\hat{Y}_i = E\left(Y\right) = \hat{\beta} X_i$
  - ✓ Résidus $Y - \hat{\beta} X_i$
- Spécifiquement aux effets aléatoires
  - ✓ Prédictions $\hat{Y}_i | \hat{\gamma}_i = E\left(Y | \hat{\gamma}_i\right) = \hat{\beta} X_i + \hat{\gamma}_i Z_i$
  - ✓ Résidus $Y_i - \hat{\beta} X_i - \hat{\gamma}_i Z_i$

Remarques

- Prédictions individuelles
  - ✓ Permet de repérer individus particuliers (écart entre profils individuels et profil moyen)
- Tests / Comparaisons de modèles
  - ✓ Rapport de vraisemblance, si modèles emboîtés

Modèle Mixte à Intercept Aléatoire (1)

Modèle linéaire

$$Y_i = \beta X_i + \varepsilon_i \quad \Rightarrow \quad y_i = \beta_0 + \beta t_i + \varepsilon_i \quad \text{avec} \quad \varepsilon_i \sim N\left(0, \sigma^2\right)$$

$$y_i = \left(\beta_0 + \gamma_{0i}\right) + \beta t_i + \varepsilon_i$$

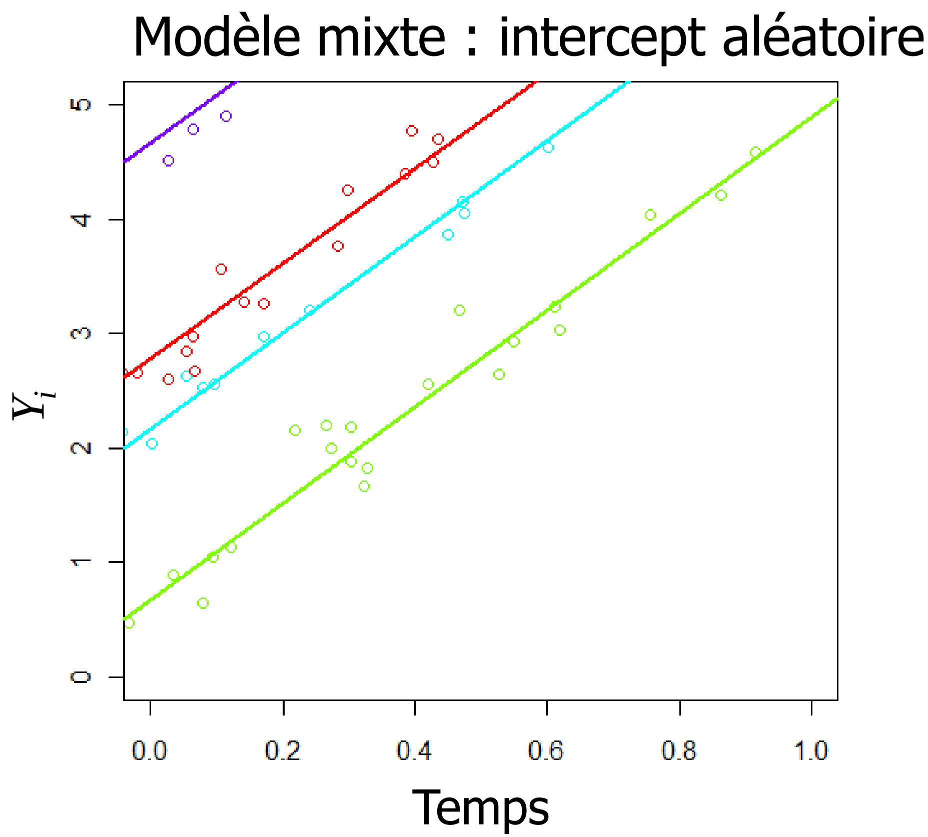
$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$

$$\gamma_{0i} \sim N\left(0, \sigma^2_{\gamma_0}\right)$$

- β_0 et β = effets fixes
  γ_{0i} = effets aléatoires
  ε_i = résidus de l'individus i

- β_0, β, σ, *et* σ_{γ_0} sont à estimer

Modèle Mixte à Intercept Aléatoire (2)



A $t=0$, il y a un point de départ moyen de β_0

Chaque sujet peut varier de $\gamma_{0i} \sim N\left(0, \sigma^2_{\gamma_0}\right)$

sans déviation par rapport à l'évolution moyenne

Modèle Mixte à Pente Aléatoire (1)

Modèle linéaire

$Y_i = \beta X_i + \varepsilon_i \;\Rightarrow\; y_i = \beta_0 + \beta t + \varepsilon_i$ avec $\varepsilon_i \sim N\left(0, \sigma^2\right)$

$$y_i = \beta_0 + \left(\beta + \gamma_i\right) t_i + \varepsilon_i$$

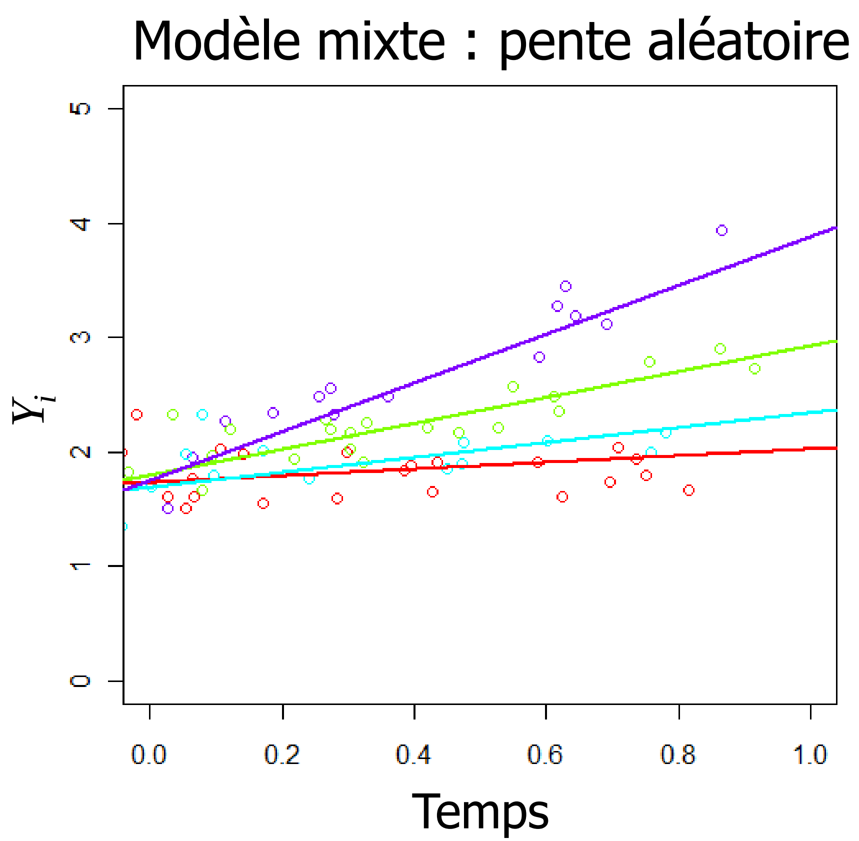
$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$

$$\gamma_i \sim N\left(0, \sigma_\gamma^2\right)$$

- β_0 et β = effets fixes
  γ_i = effets aléatoires
  ε_i = résidus de l'individus i

- β_0, β , σ, *et* σ_γ sont à estimer

Modèle Mixte à Pente Aléatoire (2)



A $t=0$, il y a un point de départ moyen de β_0

qui est le même pour tous les sujets

L'évolution de chaque sujet peut varier de $\gamma_i \sim N\left(0, \sigma_\gamma^2\right)$
avec une déviation par rapport à l'évolution moyenne

Modèle Mixte à Intercept et Pente Aléatoire

Modèle linéaire

$Y_i = \beta X_i + \varepsilon_i \Rightarrow y_i = \beta_0 + \beta t + \varepsilon_i$ avec $\varepsilon_i \sim N\left(0, \sigma^2\right)$

$$y_i = \left(\beta_0 + \gamma_{0i}\right) + \left(\beta + \gamma_i\right) t_i + \varepsilon_i$$

$$\varepsilon_i \sim N\left(0, \sigma^2\right)$$

$$\gamma_i \sim N\left(0, \sigma_\gamma^2\right) \quad \gamma_{0i} \sim N\left(0, \sigma_{\gamma_0}^2\right)$$

☞ β_0 et β = effets fixes
γ_{0i} et γ_i = effets aléatoires
ε_i = résidus de l'individus i

☞ β_0, β , σ, *et* σ_γ sont à estimer

Effets Fixes : Interprétation

$$\boxed{y_i = (\beta_0 + \gamma_{0i}) + (\beta + \gamma_i) t_i + \varepsilon_i}$$

$$y_i = \underbrace{\beta_0 + \beta t_i}_{\text{effets fixes}} + \underbrace{\gamma_{0i} + \gamma_i t_i + \varepsilon_i}_{\text{effets aléatoires}}$$

- β_0 = Intercept (valeur estimée à $t=0$) ; niveau moyen
- β = Pente (changement de l'estimation quand augmentation d'une unité)
- Tout 2 peuvent être testés
  - ✓ $H_0 : \beta_0 = 0$
  - ✓ $H_0 : \beta = 0$

Effets Aléatoires : Interprétation

$$\boxed{y_i = (\beta_0 + \gamma_{0i}) + (\beta + \gamma_i) t_i + \varepsilon_i}$$

$$y_i = \underbrace{\beta_0 + \beta t_i}_{\text{effets fixes}} + \underbrace{\gamma_{0i} + \gamma_i t_i + \varepsilon_i}_{\text{effets aléatoires}}$$

- Pas d'estimation des paramètres
- Estimation de la variance des effets aléatoires
  - ✓ Pour l'intercept γ_{0i}
  - ✓ Pour la pente γ_i
  - ✓ Pour l'erreur résiduelle ε_i

Exemple : Orthodont

- Evolution de mesures orthodontic au cours du temps chez des enfants
- Mesure tous les 2 ans entre 8 et 14 ans de la distance (en millimètre, sur cliché radio) entre l'hypophyse et la fente pterygomaxillaire
- *Source : package* `nlme` *de R*

Descriptif (1)

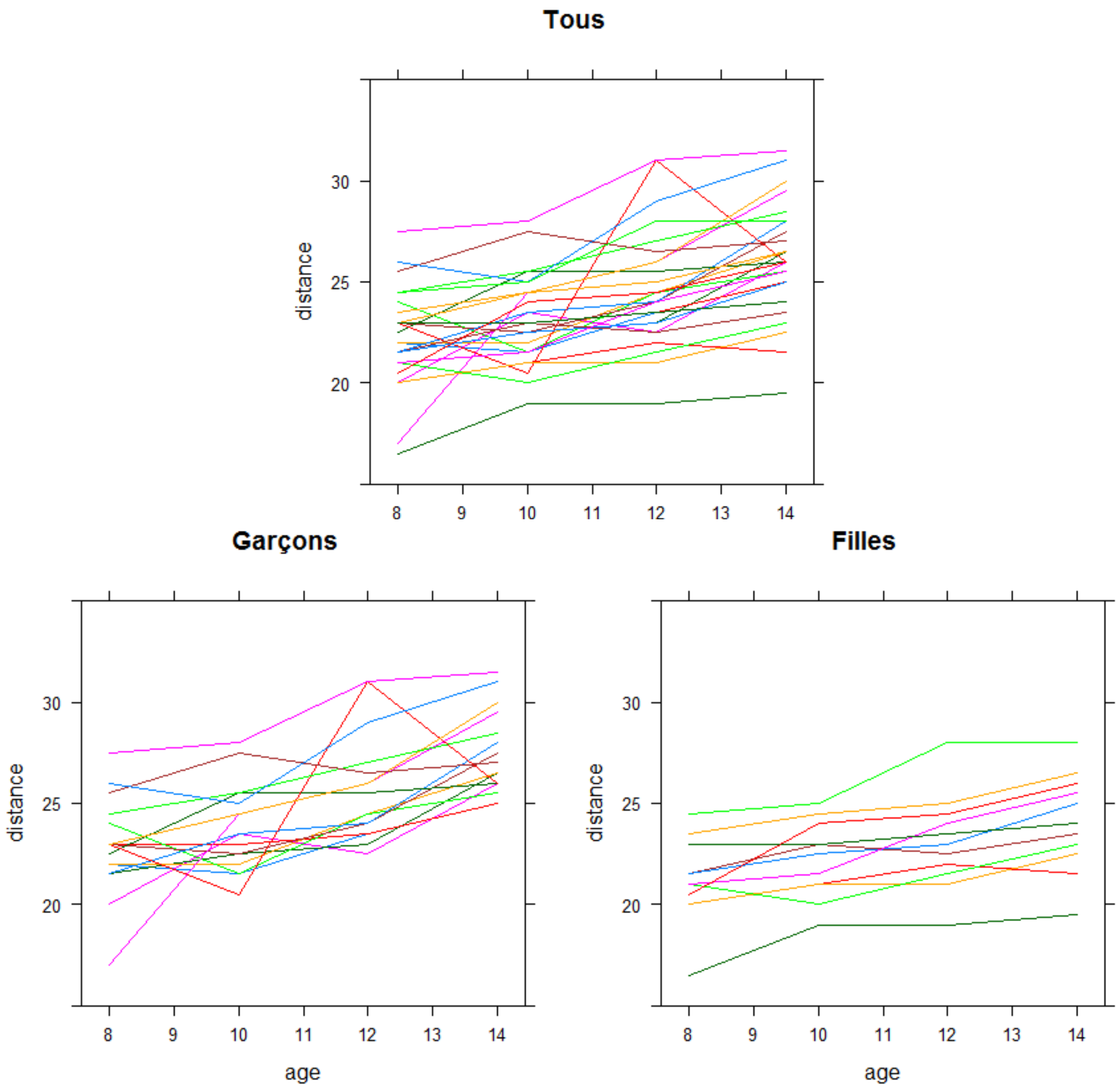
```
> with(Orthodont, tapply(distance, age, summary))
$`8`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  16.50   21.00   22.00   22.19   23.25   27.50

$`10`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  19.00   21.50   23.00   23.17   24.50   28.00

$`12`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  19.00   23.00   24.00   24.65   26.00   31.00

$`14`
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  19.50   25.00   26.00   26.09   27.75   31.50
```

Descriptif (2)



Modèle Mixte (`lme`)

$$y_i = \underbrace{\beta_0 + \beta t_i}_{\text{effets fixes}} + \underbrace{\gamma_{0i} + \gamma_i t_i}_{\text{effets aléatoires}} + \varepsilon_i$$

```
lme(distance ~ 1 + agec,
              random= ~ 1 + agec|Subject, data=Orthodont)
```

Identique à

```
lme(distance ~ agec, random=~ agec|Subject,data=Orthodont)
```

Remarque : `agec` *âge centré sur âge médian (11 ans)*

Résultats (`lme`) (1)

```
> summary(lme1)
Linear mixed-effects model fit by REML
 Data: Orthodont
       AIC      BIC    logLik
  454.6367 470.6173 -221.3183

Random effects:
 Formula: ~agec | Subject
 Structure: General positive-definite, Log-Cholesky
parametrization
            StdDev    Corr
(Intercept) 2.1343289 (Intr)
agec        0.2264278 0.503
Residual    1.3100402
…
```

Méthode d'estimation

Résultats (`lme`) (2)

```
> summary(lme1)
Linear mixed-effects model fit by REML
 Data: Orthodont
       AIC      BIC    logLik
  454.6367 470.6173 -221.3183
```

Indices d'ajustement du modèle

```
Random effects:
 Formula: ~agec | Subject
 Structure: General positive-definite, Log-Cholesky
parametrization
            StdDev    Corr
(Intercept) 2.1343289 (Intr)
agec        0.2264278 0.503
Residual    1.3100402
…
```

Résultats (`lme`) (3)

```
> summary(lme1)
Linear mixed-effects model fit by REML
 Data: Orthodont
        AIC      BIC    logLik
  454.6367 470.6173 -221.3183

Random effects:
 Formula: ~agec | Subject
 Structure: General positive-definite, Log-Cholesky
parametrization
            StdDev    Corr
(Intercept) 2.1343289 (Intr)
agec        0.2264278 0.503
Residual    1.3100402
…
```

Effets aléatoires

Déviations standards estimées des effets aléatoires σ.
(les moyennes = 0)

Intercept
Pente
Résiduelle

≠ ajustements appliqués aux observations γ_{0i}, γ_i

Résultats (`lme`) (4)

```
> ranef(fm1)
     (Intercept)         age
M16    0.1888889 -0.11018519
M05   -3.1111111  0.18981481
M02   -1.9111111  0.11481481
M11    3.2888889 -0.33518519
M07   -1.8111111  0.13981481
…            …            …
```

ajustements appliqués aux observations γ_{0i}, γ_i
(Best Linear Unbiased Predictors BLUPS)

Pour observation M11, ajustement pour

intercept pente

Résultats (`lme`) (5)

```
> summary(lme1)
Linear mixed-effects model fit by REML
 Data: Orthodont
      AIC      BIC    logLik
  454.6367 470.6173 -221.3183

Random effects:
 Formula: ~agec | Subject
 Structure: General positive-definite, Log-Cholesky
parametrization
            StdDev    Corr
(Intercept) 2.1343289 (Intr)
agec        0.2264278 0.503
Residual    1.3100402
…
```

Corrélation entre l'intercept aléatoire et la pente aléatoire

Résultats (lme) (6)

```
> summary(lme1)
…
Fixed effects: distance ~ agec
                 Value Std.Error DF  t-value p-value
(Intercept) 24.023148 0.4296601 80 55.91198       0
agec         0.660185 0.0712533 80  9.26533       0
 Correlation:
     (Intr)
agec 0.294

Standardized Within-Group Residuals:
     Min        Q1       Med        Q3       Max
-3.22310 -0.49376  0.00731  0.47215  3.91603

Number of Observations: 108
Number of Groups: 27
```

Effets fixes

Prédiction pour une valeur 0 de agec

Pente : changement de la prédiction quand agec augmente de 1 unité

Résultats (`lme`) (7)

```
> summary(lme1)
…
Fixed effects: distance ~ agec
                Value Std.Error DF  t-value p-value
(Intercept) 24.023148 0.4296601 80 55.91198       0
agec         0.660185 0.0712533 80  9.26533       0
 Correlation:
     (Intr)
agec 0.294

Standardized Within-Group Residuals:
     Min        Q1       Med        Q3       Max
-3.22310 -0.49376  0.00731  0.47215  3.91603

Number of Observations: 108
Number of Groups: 27
```

Test de l'effet de `agec`

Résultats (`lme`) (8)

- Test des effets aléatoires : modèle à intercept et pente aléatoires versus
  - ✓ Modèle sans intercept
  - ✓ Modèle sans effet pente
- Comparés par une ANOVA

```
> # Test des effets aléatoires : intercept
> lme.1b <- update(lme1, random=~ -1 + agec|Subject)
> anova(lme1, lme.1b)
       Model df      AIC      BIC    logLik   Test  L.Ratio p-value
lme1       1  6 454.6367 470.6173 -221.3183
lme.1b     2  4 517.1695 527.8233 -254.5848 1 vs 2 66.53281  <.0001
> # Test des effets aléatoires : pente
> lme.1a <- update(lme1, random=~ 1|Subject)
> anova(lme1, lme.1a)
       Model df      AIC      BIC    logLik   Test L.Ratio p-value
lme1       1  6 454.6367 470.6173 -221.3183
lme.1a     2  4 455.0025 465.6563 -223.5013 1 vs 2 4.36583  0.1127
```

Extension : Modèle Linéaire $\rightarrow$ Généralisé (1)

- Modèle linéaire

$$Y_i = \beta X_i + \varepsilon_i$$

ou $$E(Y_i) = \mu_i = \beta X_i$$

- Modèle linéaire généralisé (GLM)

$$g\left(E\left(Y_i\right)\right) = g\left(\mu_i\right) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip} = \beta X_i$$

où g est appelée fonction de lien

partie droite de l'équation : prédicteur linéaire

Source : E. Cantoni. Colloque d'Epidémiologie Clinique et Biostatistique. 9 Juin 2009

Extension : Modèle Linéaire → Généralisé (2)

- Hypothèses
  - ✓ Y_i suit une distribution discrète ou continue dans la « famille exponentielle »
  - ✓ Distributions : binomiale, Poisson, exponentielle, Gamma
  - ✓ avec $E(Y_i) = \mu_i$ et $Var(Y_i) = \upsilon(\mu_i)$
- Estimation
  - ✓ Maximum de vraisemblance
- Inférence basée sur la vraisemblance (ou déviance)
  - ✓ Distribution asymptotique de l'estimateur (et donc de la variance)
  - ✓ Comparaison de modèles emboîtés (LRT)
  - ✓ Critère d'Akaike (AIC)

Source : E. Cantoni. Colloque d'Epidémiologie Clinique et Biostatistique. 9 Juin 2009

GLM pour Réponses Binaires

- Si Y_i = 0 ou 1
- Alors $E(Y_i) = \mu_i = p_i = P(Y_i = 1)$

$$Var(Y_i) = p_i(1 - p_i)$$

- Différentes possibilités pour g

Lien logit	$\ln\left(p_i \big/ 1 - p_i \right) = \beta X_i$
Lien probit	$\phi(p_i) = \beta X_i$
Lien c-loglog	$\log(-\log(1 - p_i)) = \beta X_i$

- Si Y_i est $Bin(n_i, p_i)$, on travaille avec les proportions Y_i/n_i

Source : E. Cantoni. Colloque d'Epidémiologie Clinique et Biostatistique. 9 Juin 2009

GLM pour Réponses Poisson

- Si $Y_i \sim \mathrm{P}(\lambda_i)$
- Alors $E(Y_i) = \lambda_i$

  $Var(Y_i) = \lambda_i$
- Fonction de lien logarithmique g : $\log(\lambda_i) = \beta X_i$

Source : E. Cantoni. Colloque d'Epidémiologie Clinique et Biostatistique. 9 Juin 2009

GLM pour Réponses Gamma

- Si $Y_i \sim \text{Gamma}(\mu_i, \phi)$
- Alors $E(Y_i) = \mu_i$

  $Var(Y_i) = \mu_i^2 / \phi$
- Fonction de lien logarithmique g : $\log(\lambda_i) = \beta X_i$

Source : E. Cantoni. Colloque d'Epidémiologie Clinique et Biostatistique. 9 Juin 2009

Modèle Linéaire Mixte Généralisé (1)

- Modèle linéaire mixte

$$Y_i = \beta X_i + \gamma_i Z_i + \varepsilon_i$$

  - ✓ γ_i indépendants selon $N(0, \Sigma_\gamma)$
  - ✓ $Y_i \mid \gamma_i$ indépendants selon une distribution de la famille exponentielle
  - ✓ Avec $E\left(Y_i \mid \gamma_i\right) = \mu_i$ et $Var\left(Y_i \mid \gamma_i\right) = \phi\upsilon\left(\mu_i\right)$

- Modèle linéaire mixte généralisé (GLMM)

$$g\left(E\left(Y_i \mid \gamma_i\right)\right) = g\left(\mu_i\right) = \beta X_i + \gamma_i Z_i$$

où g est appelée fonction de lien

Source : E. Cantoni. Colloque d'Epidémiologie Clinique et Biostatistique. 9 Juin 2009

Modèle Linéaire Mixte Généralisé (2)

- Les effets aléatoires sont inclus dans le prédicteur linéaire
- Inférence/Interprétation de type spécifique aux sujets
- Nécessite l'utilisation de routines d'optimisation non-linéaire et d'algorithmes d'intégration numérique pour ajuster ces modèles

GLMM avec R

- Fonction `glmer` du package `lme4` de R

```
glmer(formula, data, familly)
```

Contient la partie aléatoire (…|…)

Binomial, gaussian, poisson,…

Sources

- Pinheiro JC, Bates DM. Mixted-effects models in S and S-Plus. Springer 2000.
- Cantoni E. Modèles de régression pour variables réponses discrètes et continues asymétriques. Colloque d'Epidémiologie Clinique et Biostatistique. 9 Juin 2009 (www.unige.ch/ses/metri/cantoni)