Linear Mixed Model Example: Dental Growth

Mila Sun

Winter 2022

1 Install and load required packages

```
#uncomment and run the following line if you did have these packages installed
#install.packages(c("nlme","lme4","mlmRev","ggplot2","lattice","gridExtra"))

# load packages
library(lme4)
```

```
## Warning: package 'lme4' was built under R version 3.6.2
```

```
library(nlme)
library(mlmRev)  #for Exam data
```

```
## Warning: package 'mlmRev' was built under R version 3.6.2
```

```
# for data visualization
library(ggplot2)
library(lattice)
library(gridExtra)
```

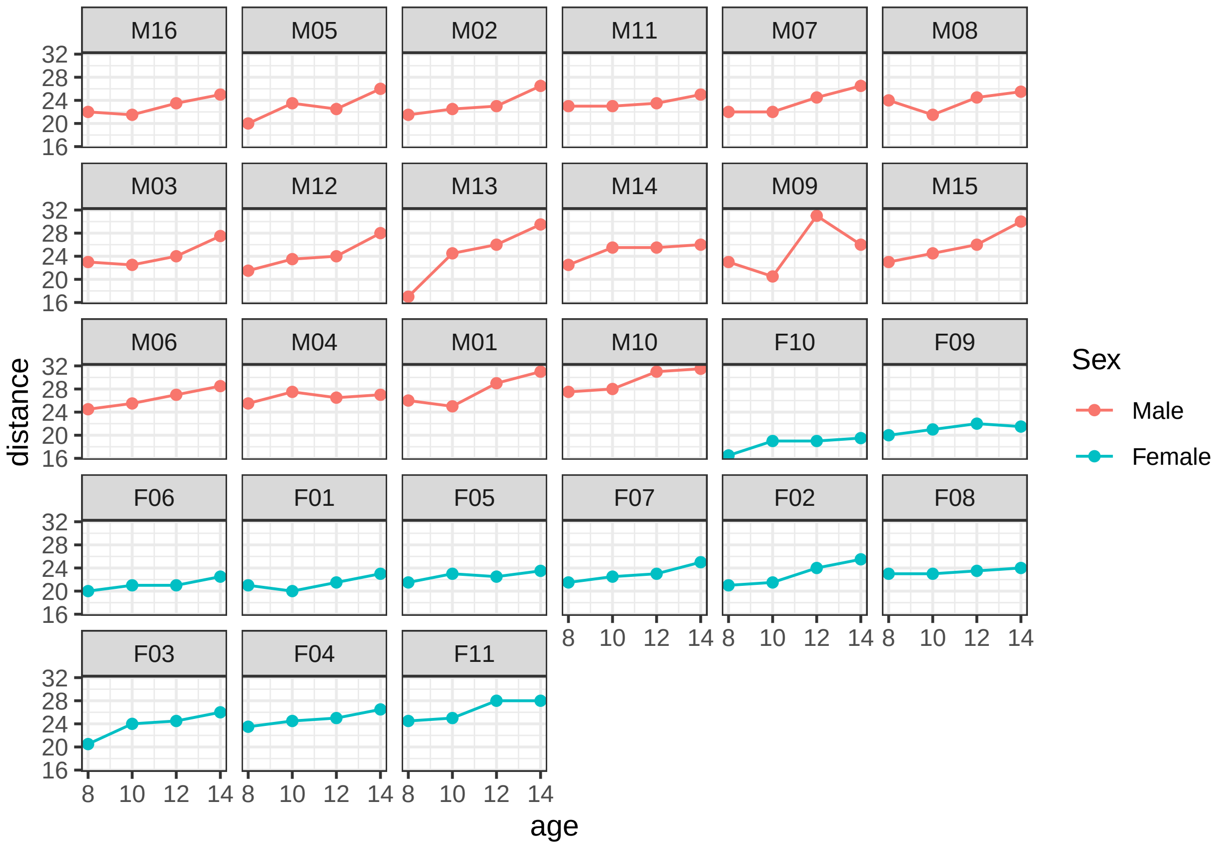
2 Goal:

Characterize dental growth among children, ages 8 to 14 years:

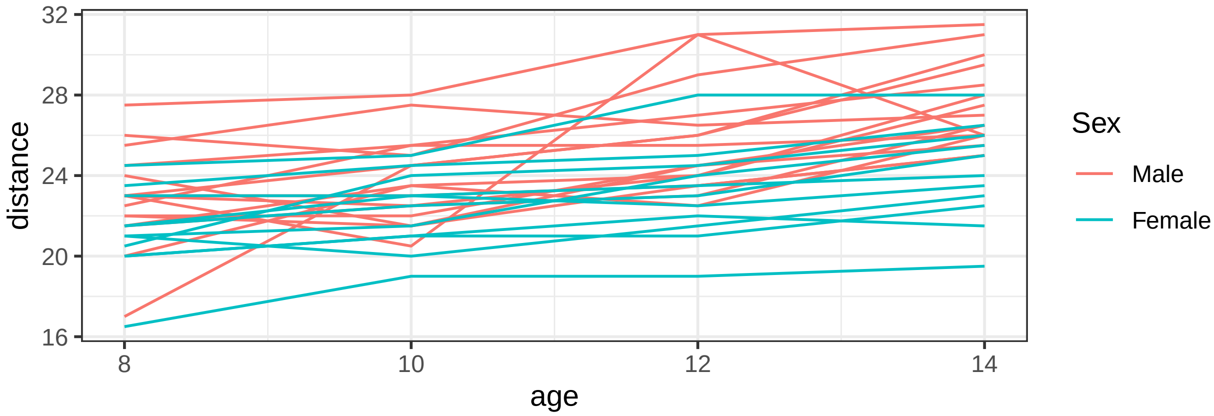
- Estimate the average growth curve among all children
- Estimate the growth curve for individual children
- Characterize the degree of heterogeneity across children
- Identify factors that predict growth

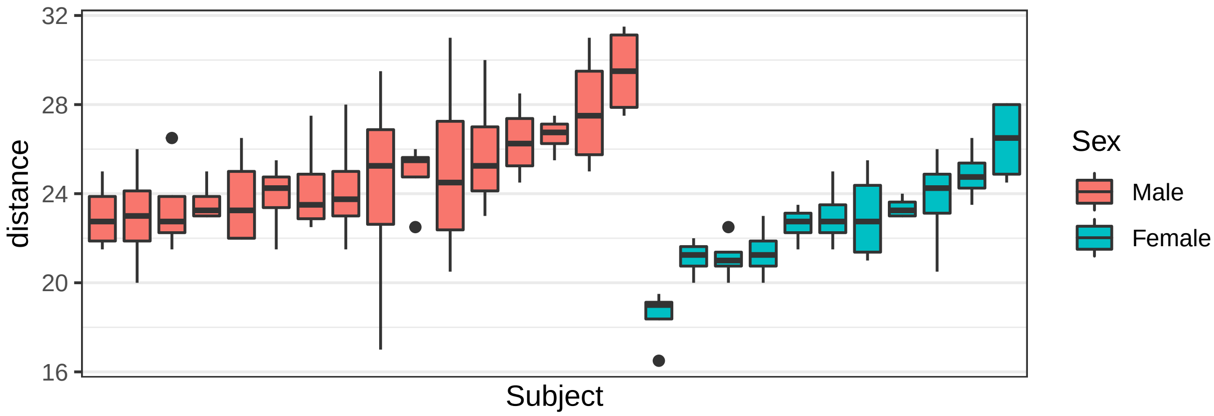
```
#---- load Dental growth data ----
data(Orthodont)

ggplot(Orthodont, aes(x=age, y=distance, color=Sex)) +
  geom_point() + geom_line() +
  facet_wrap(~Subject) + theme_bw()
```




```
pl1 = ggplot(Orthodont, aes(x=age, y=distance, color=Sex)) +
  geom_line(aes(group=Subject)) + theme_bw()
pl2 = ggplot(Orthodont, aes(x=Subject, y=distance, fill=Sex)) +
  geom_boxplot()+ scale_x_discrete(breaks = NULL) + theme_bw()
grid.arrange(pl1, pl2, ncol=1)
```





3 Practice

Characterize dental growth among males and females, ages 8 to 14 years

$$E[Y_{ij}] = \beta_0 + \beta_1(Age_{ij} - 8) + \beta_2 Gender_i + \beta_3(Age_{ij} - 8) \times Gender_i$$

Consider various specifications for the random effects structure

- Random intercepts
- Random intercepts and slopes (for age)

Note: In practice, selection of a random effects structure should be guided by a prior knowledge and/or exploratory analysis, or specified as relevant to the scientific question of interest.

```
# random intercepts
m_ri <- lmer(distance ~  I(age-8)*Sex + (1 | Subject), data=Orthodont)
summary(m_ri)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: distance ~ I(age - 8) * Sex + (1 | Subject)
##    Data: Orthodont
##
## REML criterion at convergence: 433.8
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -3.5980 -0.4546  0.0158  0.5024  3.6862
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  Subject  (Intercept) 3.299    1.816
##  Residual             1.922    1.386
## Number of obs: 108, groups:  Subject, 27
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)            22.6156     0.5387  41.978
## I(age - 8)              0.7844     0.0775  10.121
## SexFemale              -1.4065     0.8441  -1.666
## I(age - 8):SexFemale  -0.3048     0.1214  -2.511
##
## Correlation of Fixed Effects:
##             (Intr) I(g-8) SexFml
## I(age - 8)  -0.432
## SexFemale   -0.638  0.275
## I(g-8):SxFm  0.275 -0.638 -0.432
```

```r
# random intercepts and slopes
m_rs <- lmer(distance ~  I(age-8)*Sex + (I(age-8) | Subject), data=Orthodont)
summary(m_rs)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: distance ~ I(age - 8) * Sex + (I(age - 8) | Subject)
##    Data: Orthodont
##
## REML criterion at convergence: 432.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1681 -0.3859  0.0071  0.4452  3.8495
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  Subject  (Intercept) 3.23394  1.7983
##           I(age - 8)  0.03252  0.1803   -0.09
##  Residual             1.71621  1.3100
## Number of obs: 108, groups:  Subject, 27
##
## Fixed effects:
##                       Estimate Std. Error t value
## (Intercept)            22.6156     0.5265  42.954
## I(age - 8)              0.7844     0.0860   9.121
## SexFemale              -1.4065     0.8249  -1.705
## I(age - 8):SexFemale  -0.3048     0.1347  -2.262
##
## Correlation of Fixed Effects:
##             (Intr) I(g-8) SexFml
## I(age - 8)  -0.396
## SexFemale   -0.638  0.253
## I(g-8):SxFm  0.253 -0.638 -0.396
```

```r
# compare models
anova(m_ri, m_rs)
```

```
## refitting model(s) with ML (instead of REML)
```

```
## Data: Orthodont
## Models:
## m_ri: distance ~ I(age - 8) * Sex + (1 | Subject)
## m_rs: distance ~ I(age - 8) * Sex + (I(age - 8) | Subject)
##      npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
## m_ri    6 440.64 456.73 -214.32   428.64
## m_rs    8 443.81 465.26 -213.90   427.81 0.8331  2     0.6593
```

4 Results interpretation

- For random coefficients model:
  - $Corr(u_0, u_1) = -0.09 < 0$ indicates subjects with low rate of growth have high distance
  - $\hat{\sigma}^2_{u_0} = 3.23$ indicates mild variability in level of dental length
  - $\hat{\sigma}^2_{u_1} = 0.03$ indicates mild variability in change in length over time
- AIC and LR indicate model 1 is a reasonable fit to the data
- $\hat{\beta}_3 = -0.30 < 0$ indicates increase in average dental length is larger for males