Assignment 05

Marwin Carmo

library(nlme)  
library(lmerTest)

## Loading required package: lme4

## Loading required package: Matrix

##   
## Attaching package: 'lme4'

## The following object is masked from 'package:nlme':  
##   
## lmList

##   
## Attaching package: 'lmerTest'

## The following object is masked from 'package:lme4':  
##   
## lmer

## The following object is masked from 'package:stats':  
##   
## step

library(ggplot2)

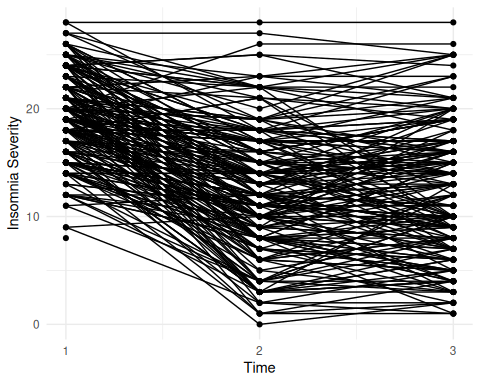
mydata <- read.csv("../data/mydata.csv")  
# removing NAs  
mydata <- mydata[!is.na(mydata$insomnia\_severity), ]  
# convert the time variable to factor  
#mydata$redcap\_event\_name <- factor(mydata$redcap\_event\_name)

## (1) Select a variable in your data for modeling over time. (1 variable, at least 3 occasions). Prepare a long-format data set for use. Use the same variable and data as Assignment 4.

Consistent with last assignments, I will work with the outcome of insomnia severity.

## (2) Plot the raw longitudinal data for all, or some representative subsample of, participants.

mydata |>   
 ggplot(aes( x = redcap\_event\_name, y = insomnia\_severity, group = record\_id)) +  
 geom\_point() +   
 geom\_line() +  
 scale\_x\_continuous(breaks = c(1,2,3)) +  
 labs(x = "Time", y = "Insomnia Severity") +  
 theme\_minimal()



## (3) Unconditional Growth Curve Analysis (polynomial)

### a. Run the multilevel polynomial growth curve models with increasing order (lin, lin+qua, etc.)

## Linear model of change  
  
### random intercept and slope  
mLin <- lmer(insomnia\_severity ~ 1 + redcap\_event\_name +   
 (1+ redcap\_event\_name|record\_id), REML=FALSE, data=mydata)

## boundary (singular) fit: see help('isSingular')

### random intercept only  
mLin\_ri <- lmer(insomnia\_severity ~ 1 + redcap\_event\_name +   
 (1|record\_id), REML=FALSE, data=mydata)  
  
  
  
### random intercept and slope with uncorrelated random effects  
mLin\_nc <- lmer(insomnia\_severity ~ 1 + redcap\_event\_name +   
 (1 + redcap\_event\_name||record\_id), REML=FALSE, data=mydata)  
  
## Linear + Quadratic Model of Change  
# mQuad <- lmer(insomnia\_severity ~ 1 + redcap\_event\_name + I(redcap\_event\_name^2) +  
# (1+ redcap\_event\_name +I(redcap\_event\_name^2)|record\_id),   
# REML=FALSE, data=mydata)  
# quadratic random slope could not be estimated due to fewer observations   
# than random effects in the model  
  
## Linear + Quadratic Model of Change without random quadratic effect  
mQuad\_ri <- lmer(insomnia\_severity ~ 1 + redcap\_event\_name + I(redcap\_event\_name^2) +  
 (1+ redcap\_event\_name |record\_id),   
 REML=FALSE, data=mydata)  
  
## Linear + Quadratic Model of Change with uncorrelated random effects  
mQuad\_nc <- lmer(insomnia\_severity ~ 1 + redcap\_event\_name + I(redcap\_event\_name^2) +  
 (1+ redcap\_event\_name +I(redcap\_event\_name^2)||record\_id),   
 REML=FALSE, data=mydata)

## boundary (singular) fit: see help('isSingular')

### b. Determine the best model representing the data (use theory and/or fit indices to make your decision)

# Comparing linear models  
anova(mLin, mLin\_ri) # removing the random slope worsened the model fit

## Data: mydata  
## Models:  
## mLin\_ri: insomnia\_severity ~ 1 + redcap\_event\_name + (1 | record\_id)  
## mLin: insomnia\_severity ~ 1 + redcap\_event\_name + (1 + redcap\_event\_name | record\_id)  
## npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)   
## mLin\_ri 4 3755.2 3772.9 -1873.6 3747.2   
## mLin 6 3711.7 3738.3 -1849.9 3699.7 47.514 2 4.813e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

anova(mLin, mLin\_nc) # correlated random effects fit significantly better

## Data: mydata  
## Models:  
## mLin\_nc: insomnia\_severity ~ 1 + redcap\_event\_name + (1 + redcap\_event\_name || record\_id)  
## mLin: insomnia\_severity ~ 1 + redcap\_event\_name + (1 + redcap\_event\_name | record\_id)  
## npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)   
## mLin\_nc 5 3714.5 3736.6 -1852.2 3704.5   
## mLin 6 3711.7 3738.3 -1849.9 3699.7 4.7792 1 0.02881 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Comparing linear to quadratic models  
anova(mLin, mQuad\_ri) # the quadratic model with random slope for time had a better fit

## Data: mydata  
## Models:  
## mLin: insomnia\_severity ~ 1 + redcap\_event\_name + (1 + redcap\_event\_name | record\_id)  
## mQuad\_ri: insomnia\_severity ~ 1 + redcap\_event\_name + I(redcap\_event\_name^2) + (1 + redcap\_event\_name | record\_id)  
## npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)   
## mLin 6 3711.7 3738.3 -1849.9 3699.7   
## mQuad\_ri 7 3586.5 3617.5 -1786.2 3572.5 127.22 1 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Comparing quadratic model with random slope for linear time to quadratic model with  
# uncorrelated random slopes for linear and quadratic time  
anova(mQuad\_ri, mQuad\_nc)

## Data: mydata  
## Models:  
## mQuad\_ri: insomnia\_severity ~ 1 + redcap\_event\_name + I(redcap\_event\_name^2) + (1 + redcap\_event\_name | record\_id)  
## mQuad\_nc: insomnia\_severity ~ 1 + redcap\_event\_name + I(redcap\_event\_name^2) + (1 + redcap\_event\_name + I(redcap\_event\_name^2) || record\_id)  
## npar AIC BIC logLik deviance Chisq Df Pr(>Chisq)  
## mQuad\_ri 7 3586.5 3617.5 -1786.2 3572.5   
## mQuad\_nc 7 3588.8 3619.8 -1787.4 3574.8 0 0

# The Quadratic Model of Change with uncorrelated random effects did not fit  
# significantly better than the previous one

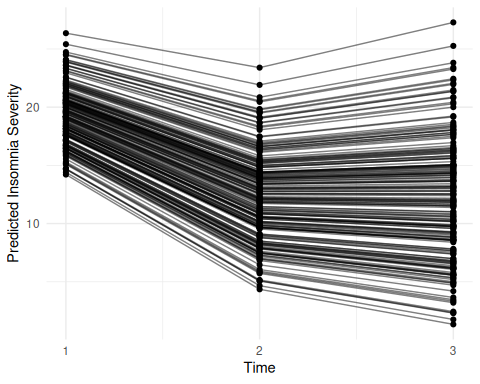
The model with lower fit indices (AIC and BIC) was the model with fixed linear and quadratic terms of change and random effects for the intercept and the linear term of time.

### c. Write out the multilevel equation for the best model

### d. Make a table for the best model as would appear in a paper. Include fixed effects, random effects, and fit indices (log-likelihood, AIC and BIC).

### e. Plot the predicted scores for all, or a representative subset, of individuals.

mydata$predquad <- predict(mQuad\_ri)  
  
mydata |>   
 ggplot(aes( x = redcap\_event\_name, y = predquad, group = record\_id)) +  
 geom\_point() +   
 geom\_line(alpha=.5) +  
 scale\_x\_continuous(breaks = c(1,2,3)) +  
 labs(x = "Time", y = "Predicted Insomnia Severity") +  
 theme\_minimal()



### f. Write a few sentences reporting and interpreting the results

## (4) Growth Curve Analysis with A Time-Invariant Predictor