GEE and GLME

Jay, Alec, Lynn

5/27/2021

```
library(tidyverse)
library(geepack)
library(nlme)
library(glme)
library(lme4)
```

```
aids = read.csv("aids.csv")


aids = aids %>%
  mutate(occasion = ceiling(week),
         gender = factor(gender, level = c("male", "female")),
         treatment = as.factor(treatment))


aids = aids %>%
  mutate(counts = exp(log_cd4) - 1) %>%
  mutate(counts2 = as.integer(round(counts))) %>%
  mutate(occasion_sqr = I(occasion^2))
```

log_cd4 ~ occasion + I(occasion^2) + treatment:occasion + treatment:I(occasion^2) + age

```
model = glmer(counts2 ~ occasion + I(occasion ^ 2) + treatment:occasion + treatment:I(occasion^2) + age
              data = aids,
              family = poisson,
              control = glmerControl(tol = 1e-12),
              nAGQ = 0,
              na.action = na.omit)
summary(model)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## counts2 ~ occasion + I(occasion^2) + treatment:occasion + treatment:I(occasion^2) +
##     age + (1 | id)
##    Data: aids
## Control: glmerControl(tol = 1e-12)
##
##      AIC      BIC   logLik deviance df.resid
##  62503.4  62575.2 -31240.7  62481.4     5025
##
## Scaled residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -12.0587  -1.3875  -0.2441   1.0653  22.9876
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  id     (Intercept) 0.9138   0.9559
## Number of obs: 5036, groups:  id, 1309
##
## Fixed effects:
##                             Estimate Std. Error z value Pr(>|z|)
## (Intercept)                2.478e+00  1.265e-01  19.582  < 2e-16 ***
## occasion                   1.703e-03  1.646e-03   1.035 0.300676
## I(occasion^2)             -4.932e-04  4.777e-05 -10.325  < 2e-16 ***
## age                        1.091e-02  3.273e-03   3.334 0.000856 ***
## occasion:treatment2        1.506e-03  2.223e-03   0.677 0.498096
## occasion:treatment3        2.809e-02  2.149e-03  13.072  < 2e-16 ***
## occasion:treatment4        4.880e-02  2.075e-03  23.521  < 2e-16 ***
## I(occasion^2):treatment2   9.251e-05  6.374e-05   1.451 0.146699
## I(occasion^2):treatment3  -5.108e-04  6.148e-05  -8.309  < 2e-16 ***
## I(occasion^2):treatment4  -7.852e-04  5.878e-05 -13.358  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) occasn I(c^2) age    occs:2 occs:3 occs:4 I(^2):2 I(^2):3
## occasion    -0.015
## I(occasn^2)  0.011 -0.947
## age         -0.977  0.001 -0.001
## occsn:trtm2  0.002 -0.738  0.699 -0.001
## occsn:trtm3  0.002 -0.763  0.723  0.000  0.565
## occsn:trtm4  0.004 -0.791  0.749 -0.002  0.585  0.605
## I(ccsn^2):2 -0.002  0.708 -0.748  0.000 -0.947 -0.542 -0.561
## I(ccsn^2):3 -0.001  0.734 -0.776  0.000 -0.543 -0.947 -0.582  0.581
## I(ccsn^2):4 -0.003  0.768 -0.811  0.001 -0.568 -0.588 -0.947  0.608   0.630
```

We noticed that cannot have occasion^2 as a random variable, which doesn't allow us to test our ideal model. We will instead just use a random intercept and a random occasion slope.

```
model1 = glmer(counts2 ~ occasion + I(occasion ^ 2) + treatment:occasion + treatment:I(occasion^2) + ag
               data = aids,
               family = poisson,
               control = glmerControl(tol = 1e-12),
               nAGQ = 0,
               na.action = na.omit)
summary(model1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 0) [glmerMod]
##  Family: poisson  ( log )
## Formula:
## counts2 ~ occasion + I(occasion^2) + treatment:occasion + treatment:I(occasion^2) +
##     age + (1 + occasion | id)
##    Data: aids
## Control: glmerControl(tol = 1e-12)
##
```

```
##      AIC      BIC   logLik deviance df.resid
##  52553.6  52638.4 -26263.8  52527.6     5023
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -10.0615  -1.0129  -0.0878   0.8022  18.4080
##
## Random effects:
##  Groups Name        Variance  Std.Dev. Corr
##  id     (Intercept) 0.8678324 0.93158
##         occasion    0.0009204 0.03034  -0.15
## Number of obs: 5036, groups:  id, 1309
##
## Fixed effects:
##                              Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 2.543e+00  1.228e-01  20.720  < 2e-16 ***
## occasion                    5.272e-03  2.493e-03   2.115   0.0344 *
## I(occasion^2)              -7.719e-04  5.281e-05 -14.617  < 2e-16 ***
## age                         9.905e-03  3.173e-03   3.122   0.0018 **
## occasion:treatment2         5.524e-03  3.445e-03   1.604   0.1088
## occasion:treatment3         2.737e-02  3.384e-03   8.088 6.05e-16 ***
## occasion:treatment4         4.555e-02  3.326e-03  13.695  < 2e-16 ***
## I(occasion^2):treatment2 -7.764e-05  7.076e-05  -1.097   0.2725
## I(occasion^2):treatment3 -7.229e-04  6.828e-05 -10.588  < 2e-16 ***
## I(occasion^2):treatment4 -8.695e-04  6.585e-05 -13.205  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) occasn I(c^2) age    occs:2 occs:3 occs:4 I(^2):2 I(^2):3
## occasion    -0.025
## I(occasn^2)  0.010 -0.625
## age         -0.976  0.002  0.000
## occsn:trtm2  0.001 -0.716  0.448  0.000
## occsn:trtm3  0.001 -0.729  0.457  0.000  0.527
## occsn:trtm4  0.003 -0.742  0.465 -0.002  0.536  0.546
## I(ccsn^2):2 -0.002  0.463 -0.745  0.001 -0.605 -0.341 -0.347
## I(ccsn^2):3 -0.001  0.480 -0.772 -0.001 -0.347 -0.593 -0.359  0.576
## I(ccsn^2):4 -0.003  0.498 -0.801  0.001 -0.360 -0.366 -0.587  0.597   0.619
```

The best model that we have so far is:

$$log(count) = \beta_0 + \beta 1 * occasion + \beta 2 * occasion^2 + \beta 3 - 5 * treatment : occasion + \beta 6 - 8 * treatment : occasion^2$$

Comparing the Models

```
anova(model, model1)
```

```
## Data: aids
## Models:
## model: counts2 ~ occasion + I(occasion^2) + treatment:occasion + treatment:I(occasion^2) +
## model:     age + (1 | id)
## model1: counts2 ~ occasion + I(occasion^2) + treatment:occasion + treatment:I(occasion^2) +
```

```
## model1:     age + (1 + occasion | id)
##        npar   AIC   BIC logLik deviance  Chisq Df Pr(>Chisq)
## model    11 62503 62575 -31241    62481
## model1   13 52554 52638 -26264    52528 9953.8  2  < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

H_0: the reduced model with only a random intercept is adequate.
Because the p-value $<$ 2.2e-16, is less than $\alpha = 0.05$, we reject the null and conclude that the model with both the random intercept and the random slope fits the data better.

We will also compare these models to the quadratic model previously.

```
model_quadratic2 = lme(log_cd4 ~ occasion + I(occasion^2) +
             treatment:occasion + treatment:I(occasion^2)
               +  age, data = aids,
             random = ~ occasion + I(occasion^2) | id,
             method = "ML")
```

```
summary(model_quadratic2)
```

```
## Linear mixed-effects model fit by maximum likelihood
##   Data: aids
##        AIC      BIC    logLik
##   11888.81 11999.73 -5927.406
##
## Random effects:
##  Formula: ~occasion + I(occasion^2) | id
##  Structure: General positive-definite, Log-Cholesky parametrization
##               StdDev      Corr
## (Intercept)   0.759158207 (Intr) occasn
## occasion      0.041653870  0.327
## I(occasion^2) 0.000946246 -0.393 -0.921
## Residual      0.547194358
##
## Fixed effects:  log_cd4 ~ occasion + I(occasion^2) + treatment:occasion + treatment:I(occasion^2) +
##                               Value  Std.Error   DF   t-value p-value
## (Intercept)                2.5752308 0.11579256 3719 22.240037  0.0000
## occasion                  -0.0121732 0.00490191 3719 -2.483360  0.0131
## I(occasion^2)             -0.0001247 0.00013385 3719 -0.931732  0.3515
## age                        0.0096564 0.00299342 1307  3.225870  0.0013
## occasion:treatment2        0.0073144 0.00686356 3719  1.065682  0.2866
## occasion:treatment3        0.0186455 0.00685665 3719  2.719328  0.0066
## occasion:treatment4        0.0438280 0.00679153 3719  6.453332  0.0000
## I(occasion^2):treatment2 -0.0001423 0.00018672 3719 -0.762371  0.4459
## I(occasion^2):treatment3 -0.0003629 0.00018812 3719 -1.928913  0.0538
## I(occasion^2):treatment4 -0.0008380 0.00018513 3719 -4.526468  0.0000
##  Correlation:
##                           (Intr) occasn I(c^2) age    occs:2 occs:3 occs:4
## occasion                  -0.026
## I(occasion^2)              0.016 -0.939
## age                       -0.975  0.003 -0.002
## occasion:treatment2        0.002 -0.707  0.666 -0.001
## occasion:treatment3        0.002 -0.707  0.667 -0.002  0.505
## occasion:treatment4        0.002 -0.714  0.673 -0.001  0.510  0.510
## I(occasion^2):treatment2 -0.002  0.669 -0.714  0.001 -0.939 -0.478 -0.482
```

```
## I(occasion^2):treatment3 -0.002  0.663 -0.709  0.002 -0.474 -0.939 -0.479
## I(occasion^2):treatment4  0.000  0.674 -0.720 -0.001 -0.481 -0.482 -0.939
##                          I(^2):2 I(^2):3
## occasion
## I(occasion^2)
## age
## occasion:treatment2
## occasion:treatment3
## occasion:treatment4
## I(occasion^2):treatment2
## I(occasion^2):treatment3  0.508
## I(occasion^2):treatment4  0.516   0.512
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -4.48500432 -0.41811557  0.02420634  0.47119729  3.88851463
##
## Number of Observations: 5036
## Number of Groups: 1309
```

Comparing the AIC of the GLME model with our previous LME model:
- GLME AIC = 52553.6
- LME AIC = 11888.81

Because the GLME AIC is greater than the LME AIC, we can conclude that the LME model we created previously is better (fits the data better) than the GLME.