P8157 HW 3

Sabrina Lin stl2137

11/2/2020

Question 1

```
### Import data

toenail_dat <- read.delim("/Users/SabrinaLin/Documents/Fall_2020_Longitudinal/HW2/toenail.txt", header =
  janitor::clean_names() %>%
  mutate(
    treatment = as.factor(treatment),
    visit = as.factor(visit)
  )
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote = quote, :
## header and 'col.names' are of different lengths
```

Part 1

Consider a random effects model with a random intercept for the log odds of moderate or severe onycholysis. Assuming linear trends and month as the time variable.

```
glme_toenail <- glmer(y ~ treatment + month + treatment*month + (1|id), data = toenail_dat, family = bi

sum_glme_toenail <- summary(glme_toenail)
```

Part 2

Provide Interpretations for the fixed effects coefficients in your model. Interpret the random effect parameter.

- β_0: The log-odds of moderate or severe onycholysis for a typical individual in the control group at baseline is -2.5098601, holding all other variables constant.
- β_1: The log-odds ratio of moderate or severe onycholysis comparing an individual in treatment group 1 to an individual in treatment group 0 at baseline who have an identical propensity for moderate or severe onycholysis is -0.3048381, holding all other variables constant.
- β_2: The log-odds ratio of moderate or sever onycholysis for every month increase for a given individual in treatment group 0 is -0.3997336, holding all other variables constant.
- β_3: The difference in the log-odds ratio of moderate or severe onycholysis for every month increase between treatment group 1 and treatment group 0 is -0.1371361, holding all other variables constant.

- Random effect parameter: The variance of the random effects σ_b^2 is 20.76, which is not small, suggesting that there is variation in the random effects in the underlying propensity for response between the individuals.

Part 3

From the results of your analysis what conclusions do you draw about the effect of treatment on changes in the severity of onycholysis over time? Provide results that support your conclusions.

- From the model constructed, we see that with a p-value of 0.65728, the treatment beta value is not significant. Although the interaction term between treatment and months is significant with a p-value of 0.04846, this significance should be noted as borderline. The interaction significance may be driven by the significance of the beta value of month, where the p-value $<$ 2e-16. This then implies that from our model, time is the covariate that has a significant change in the severity of onycholysis and not necessarily the treatment.

Part 4

How are the interpretations different from the model in HW2.

- The interpretations for the random intercept model differ from the interpretations for the GEE models by stating that the beta coefficient is for a given individual rather than for the general population. We are also able to look at the between individual variability in the random intercept model, which we were unable to distinguish in the GEE model.

Question 2

```
skin_dat <- read.delim("/Users/SabrinaLin/Documents/Fall_2020_Longitudinal/HW2/skin.txt", header = TRUE
  janitor::clean_names() %>%
  mutate(
    treatment = as.factor(treatment),
    gender = as.factor(gender),
    skin = as.factor(skin)
  )
```

```
## Warning in read.table(file = file, header = header, sep = sep, quote = quote, :
## header and 'col.names' are of different lengths
```

Part 1

Set up a suitable random effects (random intercept) model for rate of skin cancers with Treatment and Year as covariates.

```
glme_skin <- glmer(y ~ treatment + year + treatment*year + (1|id), data = skin_dat, family = poisson)

summary(glme_skin)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: y ~ treatment + year + treatment * year + (1 | id)
```

```
##    Data: skin_dat
##
##      AIC      BIC   logLik deviance df.resid
##   8427.5   8461.8  -4208.8   8417.5     7076
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.6779 -0.3890 -0.2394 -0.2260  6.4430
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  id     (Intercept) 2.189    1.48
## Number of obs: 7081, groups:  id, 1683
##
## Fixed effects:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -2.4149720  0.1108829 -21.779   <2e-16 ***
## treatment1       0.0798434  0.1383237   0.577    0.564
## year            -0.0002495  0.0261529  -0.010    0.992
## treatment1:year  0.0348951  0.0358871   0.972    0.331
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##             (Intr) trtmn1 year
## treatment1  -0.660
## year        -0.638  0.505
## tretmnt1:yr  0.457 -0.693 -0.728
```

- the interaction term in the above model is not significant, thus for the following models, the interaction term will be taken out.

```
glme_skin <- glmer(y ~ treatment + year + (1|id), data = skin_dat, family = poisson(link = "log"))

sum_glme_skin <- summary(glme_skin)
```

Part 2

Provide Interpretations for the fixed effects coefficients in your model. Interpret the random effect parameter.

- β_0: The log-rate of skin cancer for a typical individual in the control group at baseline is -2.464608.
- β_1: The log-rate ratio comparing a patient in treatment group 1 to a patient in treatment group 0 at baseline who have an identical underlying propensity for response is 0.1730909, holding all other variables constant.
- β_2: The log-rate ratio for every year increase for a typical individual in treatment group 0 who have an identical underlying propensity for response is 0.0182672, holding all other variables constant.
- Random effect parameter: The variance of the random effects σ_b^2 is 2.189, which is not very large, suggesting that there is not too much variation in the random effects in the underlying propensity for response between the individuals.

Part 3

From the results of your analysis what conclusions do you draw about the effect of beta carotene on the rate of skin cancers? Provide results that support your conclusions.

- The treatment positively affects the outcome on the rate of skin cancers, thus compared to placebo, beta carotene appears to increase the rate of skin cancers. That being said, the treatment variable in the model is not significant (p-value = 0.0828), thus the beta crotene treatment does not significantly effect the rate of skin cancers.

Part 4

Repeat the above analysis adjusting for skin type, age, and the count of the number of previous skin cancers. What conclusions do you draw about the effect of beta carotene on the adjusted rate of skin cancers?

```
mod_adjust_skin <- glmer(y ~ treatment + year + skin + age + exposure + (1 |id), data = skin_dat, family
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, :
## Model failed to converge with max|grad| = 0.00258943 (tol = 0.002, component 1)
```

```
## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unider
##  - Rescale variables?
```

```
summary(mod_adjust_skin)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##  Family: poisson  ( log )
## Formula: y ~ treatment + year + skin + age + exposure + (1 | id)
##    Data: skin_dat
##
##      AIC      BIC   logLik deviance df.resid
##   8071.6   8119.6  -4028.8   8057.6     7074
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -2.7420 -0.3471 -0.2456 -0.2090  6.4933
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  id     (Intercept) 1.375    1.173
## Number of obs: 7081, groups:  id, 1683
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -4.219122   0.316609 -13.326  < 2e-16 ***
## treatment1   0.125507   0.087063   1.442 0.149426
## year         0.020428   0.017870   1.143 0.252968
## skin1        0.328740   0.087852   3.742 0.000183 ***
## age          0.018423   0.004632   3.978 6.95e-05 ***
## exposure     0.188588   0.010675  17.666  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
```

```
##            (Intr) trtmn1 year   skin1  age
## treatment1 -0.146
## year       -0.164  0.011
## skin1      -0.141 -0.003  0.016
## age        -0.938  0.003  0.003  0.006
## exposure   -0.105 -0.018  0.021 -0.106 -0.044
## optimizer (Nelder_Mead) convergence code: 0 (OK)
## Model failed to converge with max|grad| = 0.00258943 (tol = 0.002, component 1)
## Model is nearly unidentifiable: very large eigenvalue
##  - Rescale variables?
```

- With skin type, age, and the count of the number of previous skin cancers included in this model, the treatment of beta carotene also appears to increase the rate of skin cancers. Like the previous model, treatment is still insignificant in the model with a p-value of 0.149426. Skin type (burns), age and the count of the number of previous skin cancers do significantly effect the outcome on the rate of skin cancers however, with respective p-values of 0.000183, 6.95e-05 and <2e-16.

Part 5

How are the interpretations different from the model in HW2.

- The interpretations for the random intercept model differ from the interpretations for the GEE models by stating that the beta coefficient is for a given individual rather than for the general population. We are also able to look at the between individual variability in the random intercept model, which we were unable to distinguish in the GEE model.