20201029-p8157_hw3_jsg2145

Jared Garfinkel

10/29/2020

Read in data

```
data = read_delim(file = "./data/toenail.txt", delim = " ", col_names = c("id", "y", "treatment", "month

data
```

```
## # A tibble: 1,908 x 5
##        id     y treatment  month visit
##     <dbl> <dbl>     <dbl>  <dbl> <dbl>
##  1     1     1         1  0          1
##  2     1     1         1  0.857      2
##  3     1     1         1  3.54       3
##  4     1     0         1  4.54       4
##  5     1     0         1  7.54       5
##  6     1     0         1 10.0        6
##  7     1     0         1 13.1        7
##  8     2     0         0  0          1
##  9     2     0         0  0.964      2
## 10     2     1         0  2          3
## # ... with 1,898 more rows
```

Question 1

Part 1

Setup the random effects model.

```
lme1 = lmer(y ~ treatment + (month|id), data = data)

summary(lme1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ treatment + (month | id)
##    Data: data
##
## REML criterion at convergence: 877.5
##
## Scaled residuals:
```

```
##      Min      1Q  Median      3Q     Max
## -3.0795 -0.1645 -0.0384  0.0107  3.6056
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.228845 0.47838
##           month       0.002163 0.04651  -0.91
##  Residual             0.052686 0.22953
## Number of obs: 1908, groups:  id, 294
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  0.14131    0.01976   7.152
## treatment   -0.04290    0.02780  -1.543
##
## Correlation of Fixed Effects:
##           (Intr)
## treatment -0.711
```

Part 2

The treatment is associated with a 0.04290 times reduction in the presence of moderate to severe onycholysis under the random effects model

Part 4

compared to the generalized estimating equation model in HW2, the sign of the association changes. That is, the treatment appears to work when the random effects are included in the model, whereas the treatment appears to not work under the fixed model.

Question 2

Read in data

```
df = read_delim(file = "./data/skin.txt", delim = " ", col_names = c("id", "center", "age", "skin", "gen
  mutate(tr = as_factor(tr),
         year = as.numeric(year),
         y = as.numeric(y),
         age = as.numeric(age),
         gender = as_factor(gender),
         skin = as_factor(skin),
         exposure = as.numeric(exposure),
         current_age = age + year - 1)

df
```

```
## # A tibble: 7,081 x 10
##    id      center    age skin   gender exposure     y tr     year current_age
##    <chr>   <chr>   <dbl> <fct>  <fct>     <dbl> <dbl> <fct> <dbl>       <dbl>
```

```
## 1  "    10~ "        ~      51 "      ~ "      ~          4     0 "    ~       1          51
## 2  "    10~ "        ~      51 "      ~ "      ~          4     1 "    ~       2          52
## 3  "    10~ "        ~      51 "      ~ "      ~          4     1 "    ~       3          53
## 4  "    10~ "        ~      51 "      ~ "      ~          4     1 "    ~       4          54
## 5  "    10~ "        ~      51 "      ~ "      ~          4     0 "    ~       5          55
## 6  "    10~ "        ~      68 "      ~ "      ~          2     0 "    ~       1          68
## 7  "    10~ "        ~      68 "      ~ "      ~          2     0 "    ~       2          69
## 8  "    10~ "        ~      68 "      ~ "      ~          2     0 "    ~       3          70
## 9  "    10~ "        ~      68 "      ~ "      ~          2     0 "    ~       4          71
## 10 "    10~ "        ~      68 "      ~ "      ~          2     0 "    ~       5          72
## # ... with 7,071 more rows
```

Part 1

```
lme2 = lmer(y ~ tr + year + (year|id), data = df)
```

Part 2

```
summary(lme2)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ tr + year + (year | id)
##    Data: df
##
## REML criterion at convergence: 16725.5
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -8.3079 -0.2497 -0.1310 -0.0770 19.3632
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.18547  0.4307
##           year        0.01128  0.1062   -0.04
##  Residual             0.45260  0.6728
## Number of obs: 7081, groups:  id, 1683
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) 0.244007   0.026033   9.373
## tr    1     0.036371   0.029245   1.244
## year        0.004954   0.006871   0.721
##
## Correlation of Fixed Effects:
##         (Intr) tr1
## tr   1 -0.574
## year   -0.600  0.006
## convergence code: 0
## Model failed to converge with max|grad| = 0.00230329 (tol = 0.002, component 1)
```

The fixed effects coefficients show that there is an increase in skin cancer on average among all patients of 0.00495 for each 1 unit increase in year.

There is an increase in number of skin cancers of 0.0364 for those in treatment 1 compared to treatment 0, indicating the treatment may not be beneficial.

The random effect parameter shows that $\sigma_{b2} = 0.106$. This means that a 95% confidence interval of the effect of year on the outcome can be calculated as $(B_1 \pm 1.96 * \sigma_{b2}) = (-0.203, 0.213)$.

Since this includes 0 in the range, we accept the null that year is not associated with the outcome of interest.

Part 3

Based on these results, this study is not evidence that beta carotene is beneficial to skin health. Using both a GEE model and a random effects model, treatment with beta carotene could not be significantly associated with a change in the number of new cancer growths in patients.

Part 4

```
lme3 = lmer(y ~ tr + year + skin + age + exposure + (year|id), data = df)
```

```
summary(lme3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: y ~ tr + year + skin + age + exposure + (year | id)
##    Data: df
##
## REML criterion at convergence: 16224.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -7.9279 -0.2658 -0.0878 -0.0399 19.5899
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept) 0.09338  0.3056
##           year        0.01150  0.1072   -0.09
##  Residual             0.45191  0.6722
## Number of obs: 7081, groups:  id, 1683
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept) -0.172554   0.084545  -2.041
## tr     1     0.021745   0.025018   0.869
## year         0.005595   0.006861   0.815
## skin    0   -0.031502   0.025343  -1.243
## age          0.002856   0.001264   2.259
## exposure     0.090249   0.003747  24.087
##
## Correlation of Fixed Effects:
##            (Intr) tr1    year   skin0  age
## tr     1   -0.158
```

```
## year         -0.186  0.007
## skin      0 -0.188  0.012 -0.005
## age         -0.934  0.008  0.001  0.007
## exposure    -0.072 -0.022  0.006  0.132 -0.079
## convergence code: 0
## Model failed to converge with max|grad| = 0.0155789 (tol = 0.002, component 1)
```

Part 5

After adjusting for skin type, age, and number of previous skin cancers recorded, there is still no evidence
of a significant association between the treatment and outcome of interest.

Part 6

This is similar to the results in HW2 since it could not be shown that there was an association between beta
carotene treatment and skin cancer growth.