Trabalho 2 - Lista de exercícios

Carolina Bercki (GRR20137542); Ketlin Padilha (GRR20137564); Marcelo Maceno (GRR20165678);

20 de setembro de 2017

```
dados <- read.dta("toenail.dta")
dados$trt <- as.factor(dados$trt)
summary(dados)
```

```
##        id              y            trt          month            visit      
##  Min.   :  1.0   Min.   :0.0000   0:937   Min.   : 0.000   Min.   :1.000  
##  1st Qu.:101.8   1st Qu.:0.0000   1:971   1st Qu.: 1.000   1st Qu.:2.000  
##  Median :192.0   Median :0.0000           Median : 3.000   Median :4.000  
##  Mean   :189.8   Mean   :0.2138           Mean   : 4.691   Mean   :3.896  
##  3rd Qu.:276.2   3rd Qu.:0.0000           3rd Qu.: 8.893   3rd Qu.:6.000  
##  Max.   :383.0   Max.   :1.0000           Max.   :18.500   Max.   :7.000
```

12.1.1

Modelo:

$logitE(Y_i) = \beta_1 + \beta_2 * month_{i,j} + \beta_3 * treatment_i * month_{i,j}$

```
gee <- geeglm(y~month+month*trt-trt,
      family=binomial(link="logit"), data=dados, id=id,
      corstr ="exchangeable", std.err="san.se")
summary(gee)
```

```
## 
## Call:
## geeglm(formula = y ~ month + month * trt - trt, family = binomial(link = "logit"), 
##     data = dados, id = id, corstr = "exchangeable", std.err = "san.se")
## 
##  Coefficients:
##             Estimate  Std.err   Wald Pr(>|W|)    
## (Intercept) -0.57822  0.13041 19.661 9.25e-06 ***
## month       -0.17132  0.02957 33.574 6.86e-09 ***
## month:trt1  -0.07770  0.05379  2.086    0.149    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Estimated Scale Parameters:
##             Estimate Std.err
## (Intercept)    1.088  0.5265
## 
## Correlation: Structure = exchangeable  Link = identity 
## 
## Estimated Correlation Parameters:
##       Estimate Std.err
## alpha   0.4217  0.2203
## Number of clusters:   294   Maximum cluster size: 7
```

12.1.2

Interpretação do parâmetro β_2 do modelo (correspondente ao mês)

```
coef(gee)[2]
```

```
##   month
## -0.1713
```

Estima-se que para a população de indivíduos que estão sob o tratamento B a chance de a oconólise evoluir a um estado moderado ou severo diminui em 15.74% $(1 - exp(\beta_2))$ após 1 mês de tratamento. Já para a população sob mesma condição, porém, após 2 meses de tratamento, a chance diminui em 29.01% $(1 - exp(\beta_2 * 2))$.
Já para a população que está submetida sob o tratamento A estima-se que a chance de a oconólise evoluir a um estado moderado ou severo diminui em 22.04% $(1 - exp(\beta_2 + \beta_3))$ após 1 mês de tratamento. Já para a população sob mesma condição, porém, após 2 meses de tratamento, a chance diminui em 34.32% $(1 - exp(\beta_2 * 2 + \beta_3))$.

12.1.3

Interpretação do parâmetro β_3 do modelo (correspondente à interação tratamento.mês)

```
coef(gee)[3]
```

```
## month:trt1
##    -0.0777
```

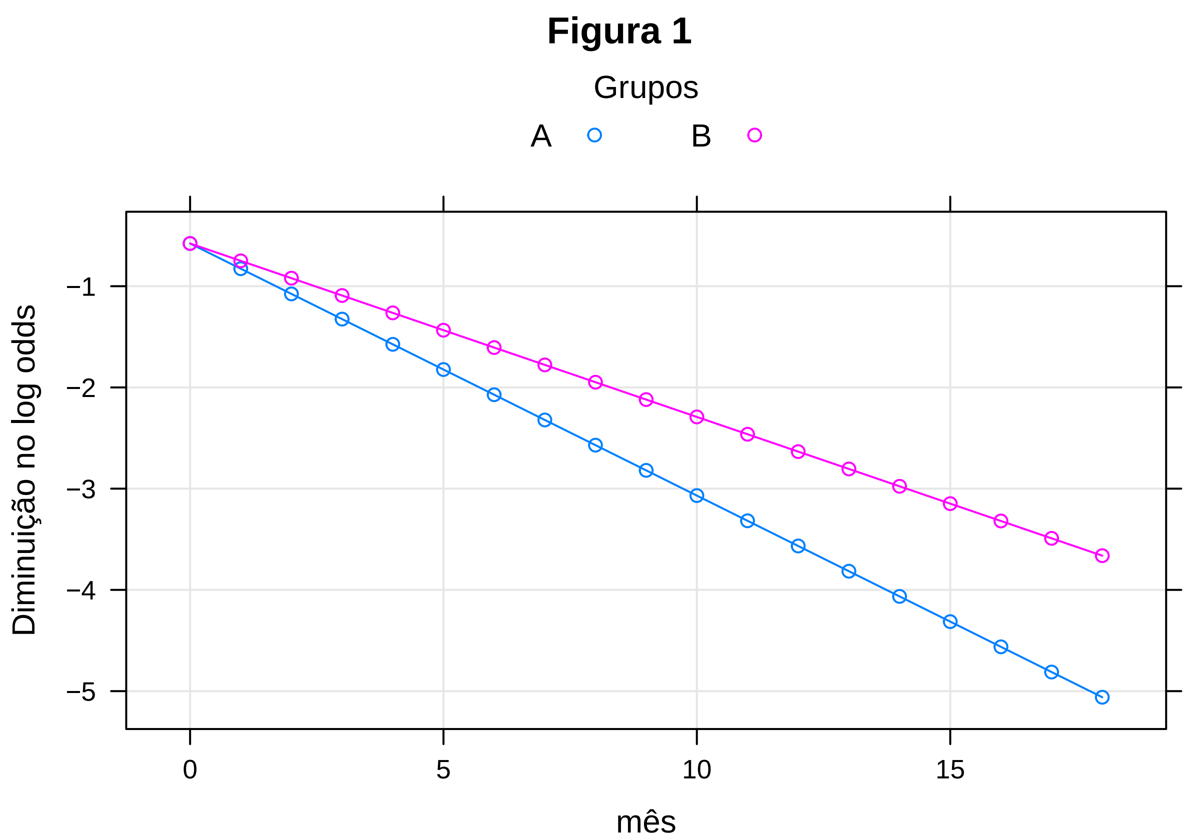
Estima-se que para a população de indivíduos que estão submetidos ao tratamento A tem-se uma chance de 7.48% $(1 - exp(\beta_3)$ de ocorrer onicólise moderada ou severa menor do que o grupo que está submetido ao tratamento B.

12.1.4

Gráfico de log(Odd) dos tratamentos em função do mês.

```
tabela1 <- data.frame(grupo=rep(c("A","B"),each=19),
                      mês=rep(c(0:18),times=2),
dimlogodd=c((coef(gee)[1]+coef(gee)[2]*0:18+coef(gee)[3]*0:18),
           (coef(gee)[1]+coef(gee)[2]*0:18)))

xyplot(dimlogodd ~ mês,
       groups=grupo,
       type=c("p","l"),
       ylab=list("Diminuição no log odds"),
       grid=T,
       main="Figura 1",
       auto.key=list(columns=2, cex.title=1,
       title=expression('Grupos')),
       data=tabela1)
```



Pela Figura 1 tem-se que há uma tendência de diminuição no logodds para ambos os grupos. Para o grupo A a diminuição é mais acentuada do que para o grupo B. Assim, para o grupo A tem-se que o efeito do tratamento no tempo é maior do que para o grupo B.

12.1.5

Modelo misto com intercepto aleatório:

$logitE(Y_i) = (\beta_1 + \_i) + \beta_2 * month_{i,j} + \beta_3 * treatment_i * month_{i,j}$

```
m1 <- glmer(y~(1|id)+month+month*trt-trt,
            family=binomial,
            nAGQ=20,
            data=dados)
summary(m1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 20) [glmerMod]
##  Family: binomial  ( logit )
## Formula: y ~ (1 | id) + month + month * trt - trt
##    Data: dados
##
##      AIC      BIC   logLik deviance df.resid
##   1258.8   1281.0   -625.4   1250.8     1904
##
## Scaled residuals:
```

```
##    Min     1Q Median     3Q    Max
##  -3.18  -0.19  -0.09  -0.01  44.33
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  id     (Intercept) 16       4
## Number of obs: 1908, groups:  id, 294
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.6972     0.3284   -5.17  2.4e-07 ***
## month        -0.3883     0.0432   -8.98  < 2e-16 ***
## month:trt1   -0.1424     0.0649   -2.19    0.028 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) month
## month      -0.040
## month:trt1 -0.001 -0.541
```

12.1.6

Estimativa de σ_b^2

```
VarCorr(m1)
```

```
##  Groups Name        Std.Dev.
##  id     (Intercept) 4
```

Verifica-se que a variância do intercepto aleatório é igual a 16, um valor elevado. Assim, há uma grande variabilidade na propensão de experimentar maior grau de infecção de unha do pé. Tem-se o seguinte intervalo de 95 % de confiança para a propensão:

0 - 1

Praticamente uma variação de 0 a 100 % de propensão.

12.1.7

O parâmetro β_2 é o seguinte:

```
m1@beta[2]
```

```
## [1] -0.3883
```

Estima-se que para um indivíduo do grupo que está sob o tratamento B a chance de a oconólise evoluir a um estado moderado ou severo diminui em 32.18% $(1 - exp(\beta_2))$ após 1 mês de tratamento. Já após 2 meses de tratamento, a chance diminui em 54% $(1 - exp(\beta_2 * 2))$.
Já para um indivíduo que está submetida sob o tratamento A estima-se que a chance de a oconólise evoluir a um estado moderado ou severo diminui em NA% $(1 - exp(\beta_2 + \beta_3))$ após 1 mês de tratamento. Já após 2 meses de tratamento, a chance diminui em 60.11% $(1 - exp(\beta_2 * 2 + \beta_3))$. Pode interpretar também em função do log Odds, sendo que para um indivíduo pertencente ao grupo B estima-se que o log Odds diminui linearmente em -0.3883 após 1 mês e para um indivíduo pertencente ao grupo A diminui linearmente em -0.5307 após 1 mês.

12.1.8

O parâmetro β_3 é o seguinte:

```
m1@beta[3]
```

```
## [1] -0.1424
```

Estima-se que para um indivíduo que está submetidos ao tratamento A tem-se uma chance de 13.27 $(1 - exp(\beta_3)\%$ menor de ocorrer onicólise moderada ou severa menor do que o grupo que um indivíduo que está submetido ao tratamento B, sendo que este indivíduo possui o mesmo risco de experimentar maior grau de infecção de unha do pé quando da aleatorização.

12.1.9

Para os dois modelos as estimativas de β_3 são as seguintes:

```
coef(gee)[3] #Modelo marginal
```

```
## month:trt1
##    -0.0777
```

```
m1@beta[3] #Modelo misto de efeito aleatório
```

```
## [1] -0.1424
```

Conforme verificado a a chance de a oconólise evoluir a um estado moderado ou severo em relação ao grupo de tratamento B é de 7.48% menor utilizando o modelo marginal, já utilizando um modelo misto de efeito aleatório é de 13.27% menor. Isto ocorreu pois há uma diferença na interpretação do β_3 para os dois modelos. Para o modelo de efeitos mistos o parâmetro se refere ao efeito do tratamento na diminuição da chance de evolução da oconólise em um determinado indivíduo. Já no caso do modelo marginal refere-se à prevalência de indivíduos com oconólise moderada ou severa na população em relação ao tratamento B contra o A.

12.1.10

Variando os pontos de quadratura do modelo:

```
#Número de pontos de quadratura = 2
m12 <- glmer(y~(1|id)+month+month*trt-trt,
            family=binomial,
            nAGQ=2,
            data=dados)

#Número de pontos de quadratura = 5
m15 <- glmer(y~(1|id)+month+month*trt-trt,
            family=binomial,
            nAGQ=5,
            data=dados)

#Número de pontos de quadratura = 10
m110 <- glmer(y~(1|id)+month+month*trt-trt,
```

```r
            family=binomial,
            nAGQ=10,
            data=dados)

#Número de pontos de quadratura = 20
m120 <- glmer(y~(1|id)+month+month*trt-trt,
            family=binomial,
            nAGQ=20,
            data=dados)

#Número de pontos de quadratura = 30
m130 <- glmer(y~(1|id)+month+month*trt-trt,
            family=binomial,
            nAGQ=30,
            data=dados)

#Número de pontos de quadratura = 50
m150 <- glmer(y~(1|id)+month+month*trt-trt,
            family=binomial,
            nAGQ=30,
            data=dados)


summary(m150)
```

```
## Generalized linear mixed model fit by maximum likelihood (Adaptive
##   Gauss-Hermite Quadrature, nAGQ = 30) [glmerMod]
##  Family: binomial  ( logit )
## Formula: y ~ (1 | id) + month + month * trt - trt
##    Data: dados
##
##      AIC      BIC   logLik deviance df.resid
##   1258.9   1281.1   -625.4   1250.9     1904
##
## Scaled residuals:
##    Min     1Q Median     3Q    Max
##  -3.18  -0.19  -0.09  -0.01  44.41
##
## Random effects:
##  Groups Name        Variance Std.Dev.
##  id     (Intercept) 16       4.01
## Number of obs: 1908, groups:  id, 294
##
## Fixed effects:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -1.6978     0.3304   -5.14  2.8e-07 ***
## month        -0.3885     0.0433   -8.97  < 2e-16 ***
## month:trt1   -0.1424     0.0649   -2.19    0.028 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##            (Intr) month
```

```
## month      -0.035
## month:trt1  0.001 -0.540
```

```r
#Estimativas dos parâmetros utilizando o número de pontos de quadratura descritos nas colunas.
t(data.frame("2 pontos"=summary(m12)$coefficients[,1:2],
           "5 pontos"=summary(m15)$coefficients[,1:2],
           "10 pontos"=summary(m110)$coefficients[,1:2],
           "20 pontos"=summary(m120)$coefficients[,1:2],
           "30 pontos"=summary(m130)$coefficients[,1:2],
           "50 pontos"=summary(m150)$coefficients[,1:2],
           check.names=FALSE))
```

```
##                       (Intercept)    month month:trt1
## 2 pontos.Estimate         -1.4916 -0.36063   -0.13034
## 2 pontos.Std. Error        0.2723  0.03961    0.05893
## 5 pontos.Estimate         -1.5216 -0.37986   -0.13866
## 5 pontos.Std. Error        0.2946  0.04225    0.06288
## 10 pontos.Estimate        -1.7190 -0.39065   -0.14320
## 10 pontos.Std. Error       0.3476  0.04380    0.06535
## 20 pontos.Estimate        -1.6972 -0.38832   -0.14236
## 20 pontos.Std. Error       0.3284  0.04325    0.06490
## 30 pontos.Estimate        -1.6978 -0.38853   -0.14244
## 30 pontos.Std. Error       0.3304  0.04332    0.06494
## 50 pontos.Estimate        -1.6978 -0.38853   -0.14244
## 50 pontos.Std. Error       0.3304  0.04332    0.06494
```

Verifica-se que para estimar bem os parâmetros necessita-se acima de 20 pontos de quadratura, tanto para a estimativas do valor pontual quanto para o erro padrão dos parâmetros.

```r
#Estimativa do erro padrão do intercepto aleatório
data.frame("2 pontos"=3.2452,
           "5 pontos"=3.6903,
           "10 pontos"=4.0606,
           "20 pontos"=4.0017,
           "30 pontos"=4.0058,
           "50 pontos"=4.0058,
           check.names=FALSE)
```

```
##   2 pontos 5 pontos 10 pontos 20 pontos 30 pontos 50 pontos
## 1    3.245     3.69     4.061     4.002     4.006     4.006
```

Para estimar bem o erro padrão do intercepto aleatório é necessário uma quadratura acima de 20 pontos de quadratura. Conclui-se que os resultados dependem do número de pontos de quadratura.