

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE ESTATÍSTICA
MAT02035 - Modelos para dados correlacionados

Avaliação pontual 02

Exercício 1. (10 pontos) Em um ensaio clínico de pacientes com doenças respiratórias, 111 pacientes de duas clínicas diferentes foram aleatorizados para receber placebo ou um tratamento ativo. Os pacientes foram examinados na linha de base e em quatro visitas durante o tratamento. Em cada exame, o estado respiratório (categorizado como 1 = bom, 0 = ruim) foi determinado. O principal objetivo das análises é compreender os efeitos conjuntos do tratamento e do tempo na probabilidade de que o estado respiratório seja classificado como bom. Também é de interesse determinar se o efeito do tratamento é o mesmo para os pacientes das duas clínicas.

Os dados brutos são armazenados em um arquivo externo: `respir.dta`.

Cada linha do conjunto de dados contém as oito variáveis a seguir:

- `yd, clinic, trt, y0, y1, y2, y3, y4`.

Nota: A variável resposta “estado respiratório”, Y_j , é codificada como 1 = bom e 0 = ruim, na j -ésima ocasião. A variável categórica Tratamento (`trt`) é codificada A = Droga ativa, P = Placebo. A variável categórica Clínica (`clinic`) é codificada 1 = clínica 1, 2 = clínica 2.

1. Ignorando a variável clínica, considere um modelo para o log-chances de que o estado respiratório seja classificado como bom (modelo logístico), incluindo os principais efeitos do tratamento e tempo (em que o tempo é considerado uma variável categórica com cinco níveis) e sua interação. Use equações de estimativa generalizadas (GEE), assumindo log-razões de chances separadas (**ou seja, *odds ratio* não estruturado**; ou correlações duas-a-duas separadas, se o *software* disponível não permitir que a associação dentro do indivíduo seja parametrizada em termos de log-razões de chances) entre as cinco respostas binárias. Construa um teste da hipótese nula de nenhum efeito do tratamento sobre as mudanças na *log*-chances de que o estado respiratório seja classificado como bom com base nos erros padrão empíricos (**estimador sanduíche**).

Vamos considerar o seguinte modelo

$$\begin{aligned}\text{logit}(\mu_{ij}) = & \beta_0 + \beta_1 I(\text{tempo} = 1) + \beta_2 I(\text{tempo} = 2) + \beta_3 I(\text{tempo} = 3) + \beta_4 I(\text{tempo} = 4) \\ & + \beta_5 I(\text{trt} = A) + \beta_6 [I(\text{tempo} = 1) \times I(\text{trt} = A)] + \beta_7 [I(\text{tempo} = 2) \times I(\text{trt} = A)] \\ & + \beta_8 [I(\text{tempo} = 3) \times I(\text{trt} = A)] + \beta_9 [I(\text{tempo} = 4) \times I(\text{trt} = A)],\end{aligned}$$

em que, $\text{logit}(\mu_{ij}) = \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right)$, β_0 (excepcionalmente) é um termo de intercepto e $I(\cdot)$ representa uma variável indicadora. Além disso, vamos assumir que: a. $\text{Var}(Y_{ij}|X_{ij}) = \mu_{ij}(1 - \mu_{ij})$ (variância Bernoulli). b. $\text{OR}(Y_{ij}, Y_{ik}) = \alpha_{jk}$ (ou seja, *odds ratio* não estruturado), em que

$$\text{OR}(Y_j, Y_k) = \frac{\Pr(Y_j = 1, Y_k = 1) \Pr(Y_j = 0, Y_k = 0)}{\Pr(Y_j = 1, Y_k = 0) \Pr(Y_j = 0, Y_k = 1)}.$$

```
# -----
# Carregando pacotes em R
# -----

library(haven)
library(tidyr)
library(dplyr)
library(geepack)
library(doBy)

# -----
# Carregando dados
# -----

df <- read_dta(file = here::here("data", "respir.dta"))
df

## # A tibble: 111 x 8
##   clinic    id trt      y0      y1      y2      y3      y4
##   <dbl> <dbl> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1      1      1  P      0      0      0      0      0
## 2      1      2  P      0      0      0      0      0
## 3      1      3  A      1      1      1      1      1
## 4      1      4  P      1      1      1      1      0
## 5      1      5  P      1      1      1      1      1
## 6      1      6  A      0      0      0      0      0
## 7      1      7  P      0      1      0      1      1
## 8      1      8  A      0      0      0      0      0
## 9      1      9  A      1      1      1      1      1
## 10     1     10  P      1      0      1      1      0
## # ... with 101 more rows
```

```
# -----
# Reformatando dados: de largo para longo
# -----

df.longo <- df %>%
  gather(key = "tempo", value = "y", -id, -trt, -clinic)
df.longo
```

```
## # A tibble: 555 x 5
##   clinic    id trt   tempo     y
##   <dbl> <dbl> <chr> <chr> <dbl>
## 1      1      1 1 P     y0      0
## 2      1      2 2 P     y0      0
## 3      1      3 3 A     y0      1
## 4      1      4 4 P     y0      1
## 5      1      5 5 P     y0      1
## 6      1      6 6 A     y0      0
## 7      1      7 7 P     y0      0
## 8      1      8 8 A     y0      0
## 9      1      9 9 A     y0      1
## 10     1     10 10 P     y0      1
## # ... with 545 more rows
```

```
# -----
# Formatando variáveis
# -----

df.longo$clinic <- factor(df.longo$clinic)

df.longo$tempo.fac <- factor(df.longo$tempo,
                             levels = paste0("y", 0:4),
                             labels = 0:4)

df.longo$trt <- factor(df.longo$trt)
df.longo$trt <- relevel(df.longo$trt, ref = "P")
```

```
# -----
# Ordenando os dados
# Este passo é importante para o correto
# ajuste dos modelos marginais via função
# geeglm (para mais informações veja a
# vinheta e do pacote e o help da função)
# -----

df.longo <- df.longo %>%
```

```

  arrange(clinic, trt, id, tempo)
df.longo

```

```

## # A tibble: 555 x 6
##   clinic    id trt   tempo     y tempo.fac
##   <fct>   <dbl> <fct> <chr> <dbl> <fct>
## 1 1      1      1 P     y0      0 0
## 2 1      1      1 P     y1      0 1
## 3 1      1      1 P     y2      0 2
## 4 1      1      1 P     y3      0 3
## 5 1      1      1 P     y4      0 4
## 6 1      2      2 P     y0      0 0
## 7 1      2      2 P     y1      0 1
## 8 1      2      2 P     y2      0 2
## 9 1      2      2 P     y3      0 3
## 10 1     2      2 P     y4      0 4
## # ... with 545 more rows

```

```

# -----
# Ajustando modelo marginal com efeitos
# principais de tempo, tratamento e efeito
# de interação (GEE)
# -----

```

```

fit1 <- geeglm(y ~ tempo.fac*trt,
              id = id,
              family = binomial(link = "logit"),
              data = df.longo,
              corstr = "unstructured",
              std.err = "san.se")
summary(fit1)

```

```

##
## Call:
## geeglm(formula = y ~ tempo.fac * trt, family = binomial(link = "logit"),
##       data = df.longo, id = id, corstr = "unstructured", std.err = "san.se")
##
## Coefficients:
##              Estimate      Std.err  Wald Pr(>|W|)
## (Intercept)  -1.759e-01  2.659e-01  0.437  0.50835
## tempo.fac1    1.408e-01  2.812e-01  0.251  0.61659
## tempo.fac2   -2.884e-01  3.046e-01  0.897  0.34366
## tempo.fac3   -1.487e-16  2.829e-01  0.000  1.00000
## tempo.fac4   -7.097e-02  2.748e-01  0.067  0.79618
## trtA         -4.725e-02  3.817e-01  0.015  0.90149

```

```
## tempo.fac1:trtA  8.600e-01  3.777e-01  5.186  0.02277 *
## tempo.fac2:trtA  1.377e+00  4.492e-01  9.390  0.00218 **
## tempo.fac3:trtA  1.179e+00  4.109e-01  8.227  0.00413 **
## tempo.fac4:trtA  7.461e-01  4.216e-01  3.131  0.07680 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = unstructured
## Estimated Scale Parameters:
##
##             Estimate Std.err
## (Intercept)      1 0.03922
## Link = identity
##
## Estimated Correlation Parameters:
##             Estimate Std.err
## alpha.1:2      0.5203 0.07653
## alpha.1:3      0.3477 0.08875
## alpha.1:4      0.4524 0.08186
## alpha.1:5      0.4016 0.09024
## alpha.2:3      0.4771 0.08997
## alpha.2:4      0.4177 0.09318
## alpha.2:5      0.4948 0.08878
## alpha.3:4      0.5475 0.08793
## alpha.3:5      0.5112 0.08653
## alpha.4:5      0.5630 0.08535
## Number of clusters: 111 Maximum cluster size: 5
```

Vamos avaliar a hipótese $H_0 : \beta_6 = \beta_7 = \beta_8 = \beta_9 = 0$. Esta pode ser reescrita como $H_0 : L\beta = 0$, em que β é o vetor de coeficientes de regressão e

$$L = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

É possível utilizar uma estatística de Wald multivariada para testar H_0 .

```
L <- rbind(c(rep(0, 6), 1, 0, 0, 0),
           c(rep(0, 6), 0, 1, 0, 0),
           c(rep(0, 6), 0, 0, 1, 0),
           c(rep(0, 6), 0, 0, 0, 1))

esticon(fit1, L, joint.test = TRUE)
```

```
## X2.stat DF Pr(>|X^2|)
```

```
## 1    11.99  4      0.0174
```

Este teste produz uma estatística Wald, $W^2 = 11.99$, com 4 graus de liberdade ($p < 0.05$), e a rejeita-se a hipótese nula ao nível de significância de 0.05. Assim, um modelo de regressão logística marginal sem as interações *tempo* × *trt* não é defensável.

Este mesmo teste poderia ser realizado com a função `anova`. Para isso, é preciso ajustar um modelo sem interações e realizar uma comparação entre estes modelos.

```
fit2 <- geeglm(y ~ tempo.fac + trt,
              id = id,
              family = binomial(link = "logit"),
              data = df.longo,
              corstr = "unstructured",
              std.err = "san.se")
summary(fit2)

##
## Call:
## geeglm(formula = y ~ tempo.fac + trt, family = binomial(link = "logit"),
##       data = df.longo, id = id, corstr = "unstructured", std.err = "san.se")
##
## Coefficients:
##              Estimate Std.err Wald Pr(>|W|)
## (Intercept)  -0.570    0.250  5.22   0.0223 *
## tempo.fac1    0.559    0.197  8.01   0.0046 **
## tempo.fac2    0.375    0.229  2.69   0.1009
## tempo.fac3    0.562    0.211  7.09   0.0078 **
## tempo.fac4    0.297    0.216  1.89   0.1691
## trtA          0.723    0.292  6.12   0.0134 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = unstructured
## Estimated Scale Parameters:
##
##              Estimate Std.err
## (Intercept)    0.994  0.0262
## Link = identity
##
## Estimated Correlation Parameters:
##              Estimate Std.err
## alpha.1:2      0.523  0.0843
## alpha.1:3      0.315  0.0927
## alpha.1:4      0.432  0.0872
## alpha.1:5      0.414  0.0906
```

```
## alpha.2:3      0.463  0.0868
## alpha.2:4      0.408  0.0901
## alpha.2:5      0.495  0.0857
## alpha.3:4      0.529  0.0845
## alpha.3:5      0.493  0.0862
## alpha.4:5      0.555  0.0830
## Number of clusters: 111 Maximum cluster size: 5
```

```
anova(fit1, fit2)
```

```
## Analysis of 'Wald statistic' Table
##
## Model 1 y ~ tempo.fac * trt
## Model 2 y ~ tempo.fac + trt
##   Df X2 P(>|Chi|)
## 1  4 12    0.017 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Perceba que o resultado é idêntico ao resultado da função `esticon`.

2. Que conclusões você tira sobre o efeito do tratamento nas mudanças na log-chances? Forneça resultados que suportem suas conclusões.

Na tabela a seguir são apresentadas as razões de chances de que o estado respiratório seja classificado como bom (ao longo do tempo comprado com a linha de base) para os grupos de tratamento Droga Ativa (Droga A) e Placebo. O grupo que recebeu o tratamento com a Droga Ativa apresentou de 96.4% a 225% vezes as chances de que o estado respiratório seja classificado como bom em relação às chances na linha de base. O grupo placebo não apresenta mudanças significativas nas chances em comparação com a linha de base ($OR \approx 1$).

Table 1: Mudanças (ao longo do tempo em comparação com a linha de base) na razão de chances (odds ratio, OR) de que o estado respiratório seja classificado como bom.

	Tempo 1	Tempo 2	Tempo 3	Tempo 4
Placebo	1.15	0.75	1.00	0.93
Droga A	2.72	2.97	3.25	1.96

3. Os pacientes neste estudo foram retirados de duas clínicas separadas. Repita a análise para o Item 1, permitindo que os efeitos do tratamento (e, possivelmente, o tempo) dependam da clínica.
 - a. O efeito do tratamento é o mesmo nas duas clínicas? Apresente resultados para apoiar sua conclusão.
 - b. Encontre um modelo parcimonioso que descreva os efeitos da clínica, do tratamento e do tempo nas *log*-chances de que o estado respiratório seja classificado como

bom. Para o modelo selecionado, forneça uma interpretação clara dos parâmetros de regressão estimados para o modelo final selecionado.

Foram ajustado diferentes modelos para avaliar o efeito de clínica. Estes modelos foram comparados utilizando o teste de Wald multivariado (função `anova`)

```
# Modelo com efeitos principais de
# tempo, tratamento e clínica, e
# efeitos de interação duas-a-duas
# e entre tempo, tratamento e clínica

fit3 <- geeglm(y ~ tempo.fac*trt*clinic,
              id = id,
              family = binomial(link = "logit"),
              data = df.longo,
              corstr = "unstructured",
              std.err = "san.se")

summary(fit3)
```

```
##
## Call:
## geeglm(formula = y ~ tempo.fac * trt * clinic, family = binomial(link = "logit"),
##       data = df.longo, id = id, corstr = "unstructured", std.err = "san.se")
##
## Coefficients:
##
##              Estimate   Std.err Wald Pr(>|W|)
## (Intercept)   -7.99e-01  4.01e-01  3.96   0.047 *
## tempo.fac1     4.50e-01  3.28e-01  1.89   0.170
## tempo.fac2     1.57e-01  3.49e-01  0.20   0.654
## tempo.fac3     4.50e-01  3.28e-01  1.89   0.170
## tempo.fac4    -1.21e-16  3.95e-01  0.00   1.000
## trtA           1.05e-01  5.73e-01  0.03   0.854
## clinic2        1.23e+00  5.58e-01  4.90   0.027 *
## tempo.fac1:trtA  3.17e-01  4.56e-01  0.48   0.487
## tempo.fac2:trtA  9.11e-01  5.53e-01  2.71   0.099 .
## tempo.fac3:trtA  7.74e-01  5.06e-01  2.33   0.127
## tempo.fac4:trtA  4.70e-01  6.10e-01  0.59   0.441
## tempo.fac1:clinic2 -5.98e-01  5.89e-01  1.03   0.310
## tempo.fac2:clinic2 -8.80e-01  6.22e-01  2.00   0.157
## tempo.fac3:clinic2 -8.86e-01  5.80e-01  2.33   0.126
## tempo.fac4:clinic2 -1.48e-01  5.93e-01  0.06   0.803
## trtA:clinic2     -3.18e-01  7.92e-01  0.16   0.689
## tempo.fac1:trtA:clinic2 1.36e+00  8.34e-01  2.65   0.104
## tempo.fac2:trtA:clinic2 1.07e+00  9.53e-01  1.26   0.261
## tempo.fac3:trtA:clinic2 9.20e-01  8.66e-01  1.13   0.288
## tempo.fac4:trtA:clinic2 7.07e-01  9.20e-01  0.59   0.442
```



```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = unstructured
## Estimated Scale Parameters:
##
##             Estimate Std.err
## (Intercept)      1 0.0825
##   Link = identity
##
## Estimated Correlation Parameters:
##             Estimate Std.err
## alpha.1:2      0.496 0.0912
## alpha.1:3      0.326 0.0978
## alpha.1:4      0.439 0.0923
## alpha.1:5      0.355 0.1025
## alpha.2:3      0.446 0.1043
## alpha.2:4      0.387 0.1055
## alpha.2:5      0.467 0.1038
## alpha.3:4      0.552 0.1020
## alpha.3:5      0.479 0.1034
## alpha.4:5      0.548 0.1029
## Number of clusters: 111 Maximum cluster size: 5

# Modelo com efeitos principais de
# tempo, tratamento e clínica, e
# efeitos de interação duas-a-duas

fit4 <- geeglm(y ~ tempo.fac*trt + clinic*tempo.fac + clinic*trt,
              id = id,
              family = binomial(link = "logit"),
              data = df.longo,
              corstr = "unstructured",
              std.err = "san.se")
summary(fit4)

##
## Call:
## geeglm(formula = y ~ tempo.fac * trt + clinic * tempo.fac + clinic *
##       trt, family = binomial(link = "logit"), data = df.longo,
##       id = id, corstr = "unstructured", std.err = "san.se")
##
## Coefficients:
##             Estimate Std.err Wald Pr(>|W|)
## (Intercept)   -0.61372 0.37470 2.68  0.1014

```

```

## tempo.fac1      0.14529  0.30310  0.23   0.6317
## tempo.fac2     -0.09944  0.33622  0.09   0.7674
## tempo.fac3      0.23140  0.30685  0.57   0.4508
## tempo.fac4     -0.17324  0.35444  0.24   0.6250
## trtA            -0.29477  0.54020  0.30   0.5853
## clinic2         0.88014  0.48320  3.32   0.0685 .
## tempo.fac1:trtA  0.95192  0.40172  5.62   0.0178 *
## tempo.fac2:trtA  1.45010  0.47334  9.39   0.0022 **
## tempo.fac3:trtA  1.23881  0.42881  8.35   0.0039 **
## tempo.fac4:trtA  0.84428  0.46580  3.29   0.0699 .
## tempo.fac1:clinic2 0.00533  0.41145  0.00   0.9897
## tempo.fac2:clinic2 -0.38107  0.46765  0.66   0.4151
## tempo.fac3:clinic2 -0.45821  0.42696  1.15   0.2832
## tempo.fac4:clinic2 0.19094  0.45655  0.17   0.6758
## trtA:clinic2     0.38284  0.60141  0.41   0.5244
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = unstructured
## Estimated Scale Parameters:
##
##           Estimate Std.err
## (Intercept)   0.991  0.0551
## Link = identity
##
## Estimated Correlation Parameters:
##           Estimate Std.err
## alpha.1:2     0.489  0.0856
## alpha.1:3     0.326  0.0942
## alpha.1:4     0.440  0.0874
## alpha.1:5     0.356  0.0985
## alpha.2:3     0.447  0.0967
## alpha.2:4     0.386  0.0995
## alpha.2:5     0.462  0.0981
## alpha.3:4     0.547  0.0946
## alpha.3:5     0.479  0.0954
## alpha.4:5     0.549  0.0950
## Number of clusters: 111 Maximum cluster size: 5

# Modelo com efeitos principais de
# tempo, tratamento e clínica, e
# efeitos de interação
# tempo e tratamento
# e clínica e tratamento

```

```
fit5 <- geeglm(y ~ tempo.fac*trt + clinic*trt,
              id = id,
              family = binomial(link = "logit"),
              data = df.longo,
              corstr = "unstructured",
              std.err = "san.se")
summary(fit5)
```

```
##
## Call:
## geeglm(formula = y ~ tempo.fac * trt + clinic * trt, family = binomial(link = "logit",
##      data = df.longo, id = id, corstr = "unstructured", std.err = "san.se")
##
## Coefficients:
##              Estimate Std.err Wald Pr(>|W|)
## (Intercept)    -0.56967  0.36607  2.42   0.1197
## tempo.fac1       0.14642  0.29247  0.25   0.6166
## tempo.fac2     -0.29694  0.31493  0.89   0.3457
## tempo.fac3     -0.00116  0.29414  0.00   0.9969
## tempo.fac4     -0.07372  0.28525  0.07   0.7961
## trtA           -0.29208  0.53268  0.30   0.5835
## clinic2         0.78917  0.41037  3.70   0.0545 .
## tempo.fac1:trtA  0.94344  0.40153  5.52   0.0188 *
## tempo.fac2:trtA  1.48889  0.48016  9.62   0.0019 **
## tempo.fac3:trtA  1.28465  0.44109  8.48   0.0036 **
## tempo.fac4:trtA  0.81738  0.45476  3.23   0.0723 .
## trtA:clinic2     0.41833  0.60133  0.48   0.4866
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = unstructured
## Estimated Scale Parameters:
##
##              Estimate Std.err
## (Intercept)         1  0.0724
## Link = identity
##
## Estimated Correlation Parameters:
##              Estimate Std.err
## alpha.1:2       0.481  0.0886
## alpha.1:3       0.324  0.0958
## alpha.1:4       0.439  0.0927
## alpha.1:5       0.346  0.0980
## alpha.2:3       0.446  0.1025
```

```
## alpha.2:4      0.386  0.1036
## alpha.2:5      0.447  0.1010
## alpha.3:4      0.579  0.1020
## alpha.3:5      0.468  0.1008
## alpha.4:5      0.540  0.1012
## Number of clusters: 111 Maximum cluster size: 5
```

```
# Modelo com efeitos principais de
# tempo, tratamento e clínica, e
# efeito de interação
# tempo e tratamento
```

```
fit6 <- geeglm(y ~ tempo.fac*trt + clinic,
               id = id,
               family = binomial(link = "logit"),
               data = df.longo,
               corstr = "unstructured",
               std.err = "san.se")
```

```
summary(fit6)
```

```
##
## Call:
## geeglm(formula = y ~ tempo.fac * trt + clinic, family = binomial(link = "logit"),
##       data = df.longo, id = id, corstr = "unstructured", std.err = "san.se")
##
## Coefficients:
##              Estimate Std.err Wald Pr(>|W|)
## (Intercept)  -0.66775  0.32877  4.13  0.0423 *
## tempo.fac1     0.14785  0.29867  0.25  0.6206
## tempo.fac2    -0.29985  0.32122  0.87  0.3506
## tempo.fac3    -0.00179  0.30037  0.00  0.9952
## tempo.fac4    -0.07474  0.29134  0.07  0.7975
## trtA          -0.07564  0.39366  0.04  0.8476
## clinic2        0.98107  0.29984 10.71  0.0011 **
## tempo.fac1:trtA 0.92134  0.39891  5.33  0.0209 *
## tempo.fac2:trtA 1.46659  0.47462  9.55  0.0020 **
## tempo.fac3:trtA 1.26147  0.43522  8.40  0.0037 **
## tempo.fac4:trtA 0.80116  0.44750  3.21  0.0734 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = unstructured
## Estimated Scale Parameters:
##
##              Estimate Std.err
```

```
## (Intercept)          1  0.0634
##   Link = identity
##
## Estimated Correlation Parameters:
##           Estimate Std.err
## alpha.1:2    0.480  0.0861
## alpha.1:3    0.327  0.0945
## alpha.1:4    0.438  0.0905
## alpha.1:5    0.345  0.0964
## alpha.2:3    0.456  0.0987
## alpha.2:4    0.394  0.1007
## alpha.2:5    0.452  0.0982
## alpha.3:4    0.577  0.0978
## alpha.3:5    0.470  0.0975
## alpha.4:5    0.544  0.0968
## Number of clusters:  111 Maximum cluster size: 5
```

```
anova(fit1, fit3)
```

```
## Analysis of 'Wald statistic' Table
##
## Model 1 y ~ tempo.fac * trt * clinic
## Model 2 y ~ tempo.fac * trt
##   Df   X2 P(>|Chi|)
## 1 10 16.7   0.081 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(fit1, fit4)
```

```
## Analysis of 'Wald statistic' Table
##
## Model 1 y ~ tempo.fac * trt + clinic * tempo.fac + clinic * trt
## Model 2 y ~ tempo.fac * trt
##   Df   X2 P(>|Chi|)
## 1   6 14.1   0.028 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(fit4, fit5)
```

```
## Analysis of 'Wald statistic' Table
##
## Model 1 y ~ tempo.fac * trt + clinic * tempo.fac + clinic * trt
## Model 2 y ~ tempo.fac * trt + clinic * trt
##   Df   X2 P(>|Chi|)
## 1   4  3.6   0.46
```

```
anova(fit5, fit6)
```

```
## Analysis of 'Wald statistic' Table
##
## Model 1 y ~ tempo.fac * trt + clinic * trt
## Model 2 y ~ tempo.fac * trt + clinic
##   Df    X2 P(>|Chi|)
## 1   1 0.484      0.49
```

Ao nível de significância de 0.05, utilizando um teste de Wald multivariado, é defensável um modelo com efeitos principais para tempo, tratamento e clínica e um efeito de interação entre tempo e tratamento. Ou seja, o efeito de tratamento é o mesmo nas duas clínicas. O efeito de clínica descreve a heterogeneidade dos participantes entre clínicas.

	Estimativa	EP	Z
(Intercept)	-0.67	0.329	4.1250
tempo.fac1	0.15	0.299	0.2451
tempo.fac2	-0.30	0.321	0.8714
tempo.fac3	0.00	0.300	0.0000
tempo.fac4	-0.07	0.291	0.0658
trtA	-0.08	0.394	0.0369
clinic2	0.98	0.300	10.7055
tempo.fac1:trtA	0.92	0.399	5.3345
tempo.fac2:trtA	1.47	0.475	9.5482
tempo.fac3:trtA	1.26	0.435	8.4011
tempo.fac4:trtA	0.80	0.447	3.2052

4. Para o modelo final selecionado no Item 3, construa uma tabela das probabilidades estimadas de que o estado respiratório seja classificado como bom em função do tempo e do grupo de tratamento (e, possivelmente, da clínica). O que você conclui dessa tabela?

A gráfico a seguir (análogo a tabela) apresenta as probabilidades estimadas para cada tempo do estudo e clínica/tratamento. As probabilidades de classificação de estado respiratório bom se mantém a mesma ao longo do tempo para o grupo placebo (≈ 0.33 para os participantes da clínica 1, ≈ 0.57 para os participantes da clínica 2). Já para o grupo que recebeu a Droga Ativa, as probabilidades de classificação de estado respiratório bom aumentam ao longo do tempo (tendência quadrática) em relação a linha de base. A diferença entre clínicas representa uma diferença (heterogeneidade) no estado respiratório dos participantes das duas clínicas.

