Modelos Marginais: Estimadores GLS e GEE

José Luiz Padilha da Silva

11 de setembro de 2019

Exemplo 1: Crianças Expostas ao Chumbo

O estudo *Tratamento de Crianças Expostas ao Chumbo (TLC)* foi um ensaio clínico aleatorizado envolvendo crianças com níveis de chumbo no sangue entre $20 - 44$ microgramas/dL. Os grupos de comparação são placebo e um tratamento ativo. Os dados consistem de quatro medidas repetidas de níveis de chumbo no sangue obtidos na linha de base (semana 0), semana 1, semana 4 e semana 6 em 100 crianças aleatoriamente alocadas entre os dois grupos.

Análise Exploratória

```
library(reshape); library(plyr); library(nlme); library(ggplot2); library(geepack)
datawide <- read.table("https://docs.ufpr.br/~jlpadilha/CE075/Datasets/chumbo.txt",
                       header=TRUE)
head(datawide)
```

```
##   ID Grupo Sem0 Sem1 Sem4 Sem6
## 1  1     P 30.8 26.9 25.8 23.8
## 2  2     A 26.5 14.8 19.5 21.0
## 3  3     A 25.8 23.0 19.1 23.2
## 4  4     P 24.7 24.5 22.0 22.5
## 5  5     A 20.4  2.8  3.2  9.4
## 6  6     A 20.4  5.4  4.5 11.9
```

Os dados estão no formato *largo*, no qual cada indivíduo é representado por uma linha e as medidas repetidas são apresentadas por colunas. A seguir um resumo das observações por grupo.

```
summary(subset(datawide, Grupo=="A")[,3:6]) #Grupo tratamento
```

```
##       Sem0            Sem1             Sem4             Sem6
##  Min.   :19.70   Min.   : 2.800   Min.   : 3.000   Min.   : 4.10
##  1st Qu.:22.12   1st Qu.: 7.225   1st Qu.: 9.125   1st Qu.:15.40
##  Median :26.20   Median :12.250   Median :15.350   Median :18.85
##  Mean   :26.54   Mean   :13.522   Mean   :15.514   Mean   :20.76
##  3rd Qu.:29.55   3rd Qu.:17.500   3rd Qu.:19.725   3rd Qu.:23.75
##  Max.   :41.10   Max.   :39.000   Max.   :40.400   Max.   :63.90
```

```
summary(subset(datawide, Grupo=="P")[,3:6]) #Grupo controle
```

```
##       Sem0            Sem1            Sem4            Sem6
##  Min.   :19.70   Min.   :14.90   Min.   :15.30   Min.   :13.50
##  1st Qu.:21.88   1st Qu.:20.93   1st Qu.:19.82   1st Qu.:19.95
##  Median :25.25   Median :24.10   Median :22.45   Median :22.35
##  Mean   :26.27   Mean   :24.66   Mean   :24.07   Mean   :23.65
##  3rd Qu.:29.73   3rd Qu.:27.82   3rd Qu.:27.45   3rd Qu.:27.50
##  Max.   :38.10   Max.   :40.80   Max.   :38.60   Max.   :43.30
```

Nota-se uma pequena diminuição das médias dos níveis de chumbo no grupo controle ao longo do tempo. Para o grupo tratamento há um grande decréscimo do baseline para a primeira semana e subsequentes aumentos nas semanas seguintes.

Temos as seguintes estimativas para as correlações:

```r
round(cor(datawide[,3:6]),2) #Todos os indivídos
```

```
##      Sem0 Sem1 Sem4 Sem6
## Sem0 1.00 0.42 0.47 0.56
## Sem1 0.42 1.00 0.84 0.56
## Sem4 0.47 0.84 1.00 0.58
## Sem6 0.56 0.56 0.58 1.00
```

```r
round(cor(subset(datawide, Grupo=="A")[,3:6]),2) # Grupo tratamento
```

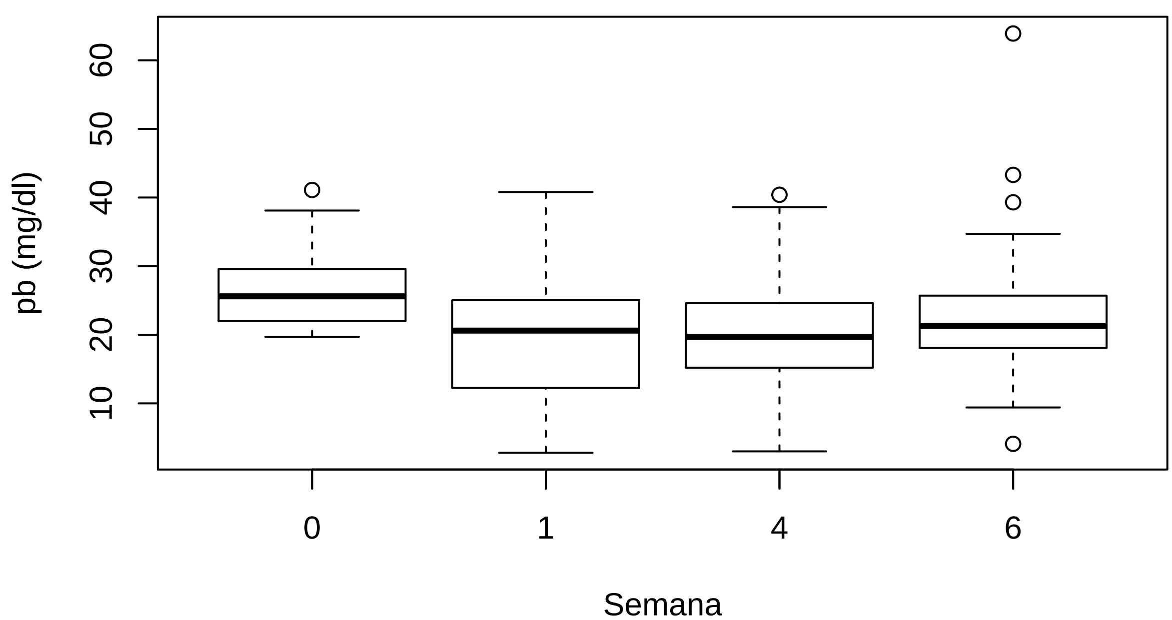
```
##      Sem0 Sem1 Sem4 Sem6
## Sem0 1.00 0.40 0.38 0.50
## Sem1 0.40 1.00 0.73 0.51
## Sem4 0.38 0.73 1.00 0.45
## Sem6 0.50 0.51 0.45 1.00
```

```r
round(cor(subset(datawide, Grupo=="P")[,3:6]),2) # Grupo controle
```

```
##      Sem0 Sem1 Sem4 Sem6
## Sem0 1.00 0.83 0.84 0.76
## Sem1 0.83 1.00 0.86 0.76
## Sem4 0.84 0.86 1.00 0.87
## Sem6 0.76 0.76 0.87 1.00
```

A seguir os boxplots marginais

```r
with(datawide, boxplot(Sem0,Sem1,Sem4,Sem6,ylab="pb (mg/dl)",xlab="Semana"))
axis(1, 1:4, c(0,1,4,6))
```



Devemos ter cuidado pois o boxplot não considera a estrutura longitudinal dos dados.

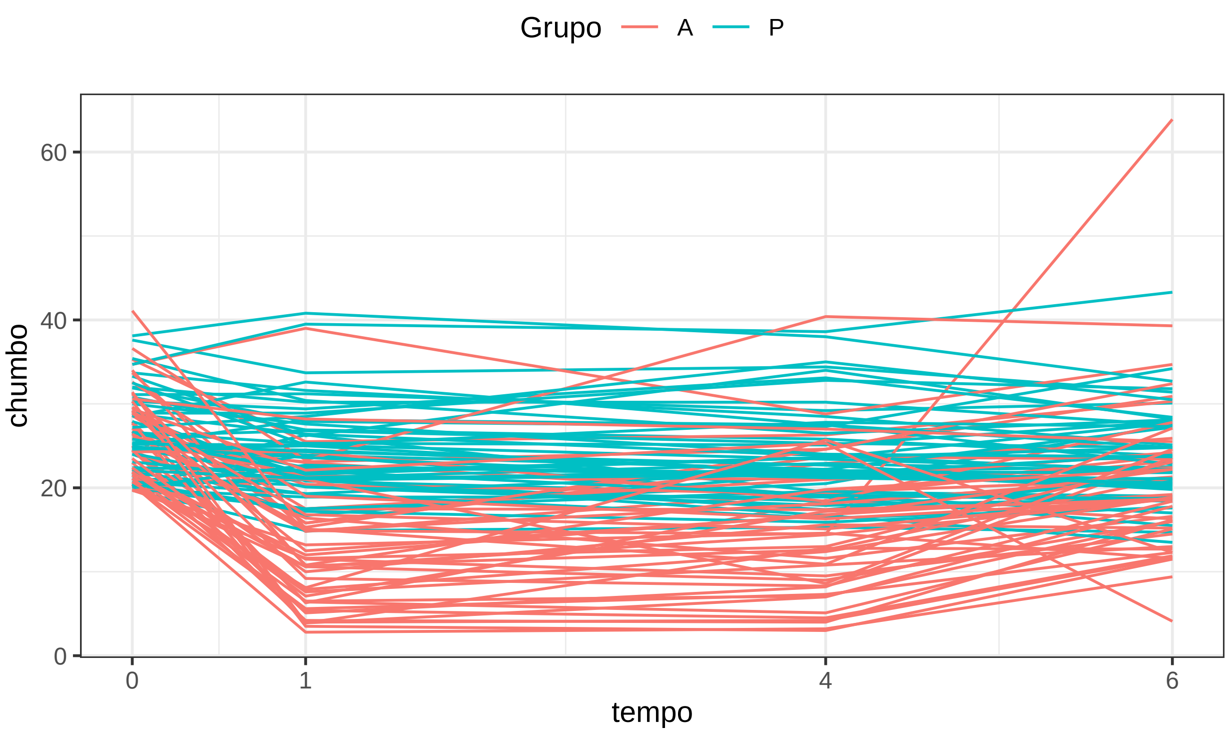
É mais conveniente trabalharmos com os dados no formato *longo*, no qual cada variável é representada por uma coluna e há uma linha para cada medida repetida do indivíduo. Vamos usar a função `reshape` do pacote de mesmo nome.

```
datalong <- reshape(data=datawide,direction="long", idvar="ID", v.names="chumbo",
                    varying = list(names(datawide)[3:6]), time= c(0,1,4,6), timevar="tempo")
datalong <- arrange(datalong, ID); head(datalong, 8)
```

```
##   ID Grupo tempo chumbo
## 1  1     P     0   30.8
## 2  1     P     1   26.9
## 3  1     P     4   25.8
## 4  1     P     6   23.8
## 5  2     A     0   26.5
## 6  2     A     1   14.8
## 7  2     A     4   19.5
## 8  2     A     6   21.0
```

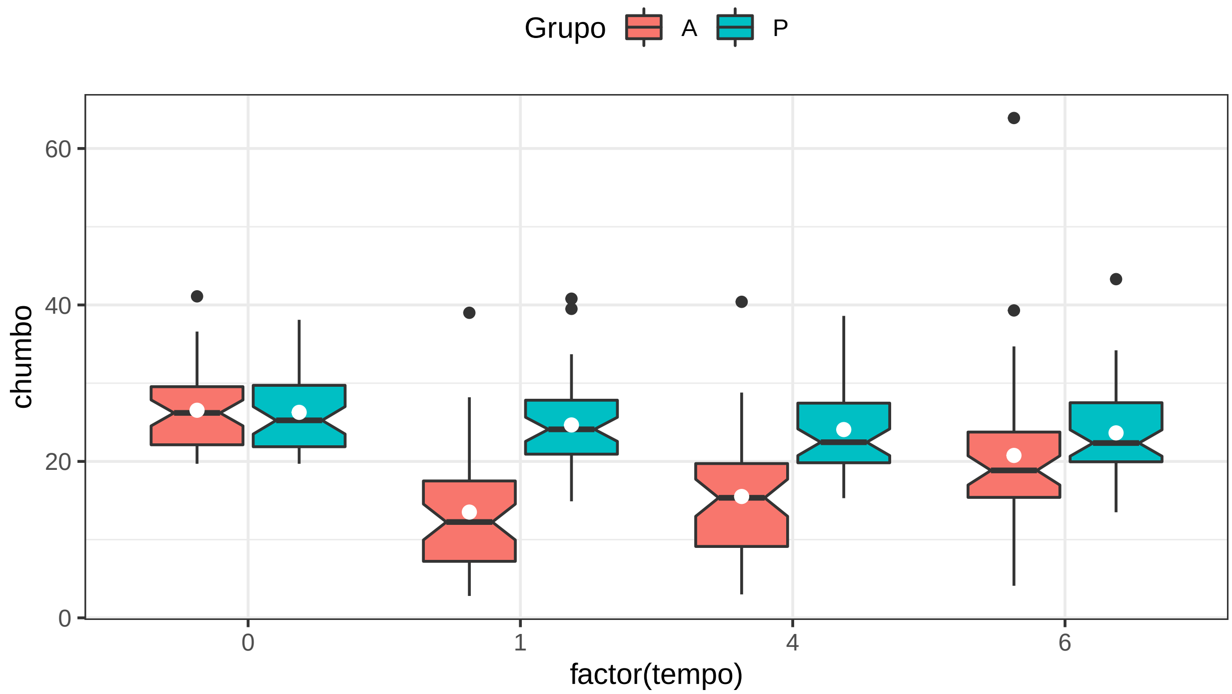
A representação gráfica mais interessante para as respostas em nível individual é o gráfico de perfis:

```
p1 <- ggplot(datalong, aes(x=tempo, y=chumbo, color=Grupo)) +   theme_bw() +
      geom_line(aes(group=ID)) + theme(legend.position="top") +
      scale_x_continuous(breaks=unique(datalong$tempo))
p1
```



Podemos examinar as diferenças dentro de cada tempo por meio de boxplots:

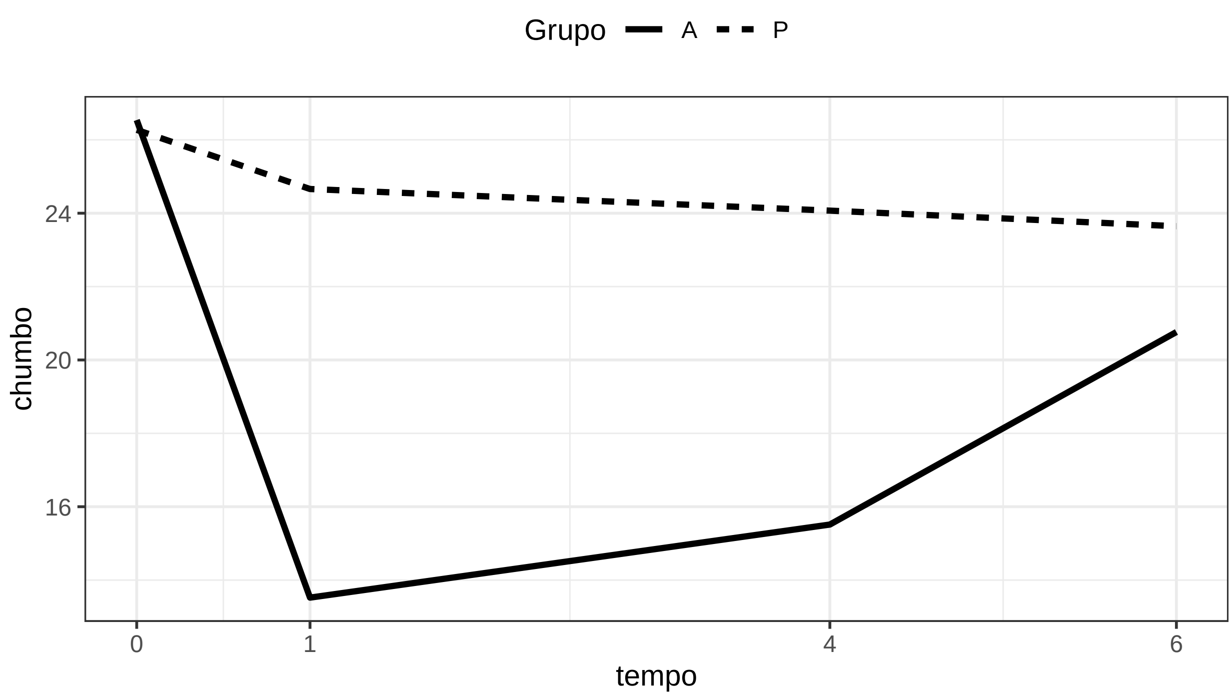
```
p2 <- ggplot(datalong, aes(x=factor(tempo), y=chumbo, fill=Grupo)) + geom_boxplot(notch=TRUE) +
      theme_bw() + theme(legend.position="top") + stat_summary(fun.y="mean", geom="point", size=2,
      position=position_dodge(width=0.75), color="white", show.legend=FALSE)
p2
```



Da ajuda do `geom_boxplot`: *In a notched box plot, the notches extend '1.58IQR / sqrt(n)'. This gives a roughly 95% confidence interval for comparing medians. See McGill et al. (1978) for more details.*

Os valores centrais em branco representam as médias.

```
p3 <- ggplot(datalong, aes(x=tempo, y=chumbo, group = Grupo, shape = Grupo)) + theme_bw() +
      stat_summary(fun.y="mean",geom="line", size=1.1, aes(linetype = Grupo)) +
      theme(legend.position="top") + scale_x_continuous(breaks=unique(datalong$tempo)); p3
```



Como vimos, as maiores diferenças ocorrem no tempo 1 e vão diminuindo ao longo das semanas. Passaremos aos ajustes dos modelos marginais por mínimos quadrados generalizados. Consideraremos as estruturas de correlação do tipo *independente*, *simetria composta*, *AR(1)* e *não estruturada*.

Formulação 1: Modelo linear de efeitos fixos (com intercepto)

O primeiro modelo é da forma `chumbo ~ tempo*Grupo` e tem a seguinte representação:

$$
\begin{aligned}
E(Y_{ij}) = &\beta_1 + \beta_2 I(tempo_j = 1) + \beta_3 I(tempo_j = 4) + \beta_4 I(tempo_j = 6) + \beta_5 I(Grupo_i = P) + \\
&\beta_6 I(Grupo_i = P) \times I(tempo_j = 1) + \beta_7 I(Grupo_i = P) \times I(tempo_j = 4) + \\
&\beta_8 I(Grupo_i = P) \times I(tempo_j = 6).
\end{aligned}
$$

Estimador GEE

Para fazer o ajuste GEE no `R` podemos utilizar os seguintes comandos:

```
datalong$tempo=as.factor(datalong$tempo); datalong$Grupo=as.factor(datalong$Grupo)
gee1a.ind<-geeglm(chumbo ~ tempo*Grupo, corstr="independence", id=ID,
                  family="gaussian", data=datalong) #Independente
gee1a.exch<-geeglm(chumbo ~ tempo*Grupo, corstr="exchangeable", id=ID,
                   family="gaussian", data=datalong) #Simetria composta
gee1a.ar1<-geeglm(chumbo ~ tempo*Grupo, corstr="ar1", id=ID,
                  data=datalong) #AR(1)
gee1a.unst<-geeglm(chumbo ~ tempo*Grupo, corstr="unstructured", id=ID,
                   data=datalong) #Não estruturada
```

Os resultados dos ajustes são:

```
# Independente
round(coef(summary(gee1a.ind)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## (Intercept)      26.540   0.703 1425.530    0.000
## tempo1          -13.018   1.021  162.690    0.000
## tempo4          -11.026   1.053  109.606    0.000
## tempo6           -5.778   1.126   26.313    0.000
## GrupoP           -0.268   0.994    0.073    0.788
## tempo1:GrupoP    11.406   1.109  105.840    0.000
## tempo4:GrupoP     8.824   1.141   59.820    0.000
## tempo6:GrupoP     3.152   1.244    6.421    0.011
```

```
# Simetria composta
round(coef(summary(gee1a.exch)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## (Intercept)      26.540   0.703 1425.530    0.000
## tempo1          -13.018   1.021  162.690    0.000
## tempo4          -11.026   1.053  109.606    0.000
## tempo6           -5.778   1.126   26.313    0.000
## GrupoP           -0.268   0.994    0.073    0.788
## tempo1:GrupoP    11.406   1.109  105.840    0.000
## tempo4:GrupoP     8.824   1.141   59.820    0.000
## tempo6:GrupoP     3.152   1.244    6.421    0.011
```

```
# AR(1)
round(coef(summary(gee1a.ar1)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## (Intercept)     26.540   0.703 1425.530    0.000
## tempo1         -13.018   1.021  162.690    0.000
## tempo4         -11.026   1.053  109.606    0.000
## tempo6          -5.778   1.126   26.313    0.000
## GrupoP          -0.268   0.994    0.073    0.788
## tempo1:GrupoP   11.406   1.109  105.840    0.000
## tempo4:GrupoP    8.824   1.141   59.820    0.000
## tempo6:GrupoP    3.152   1.244    6.421    0.011
```

```r
# Não estruturada
round(coef(summary(gee1a.unst)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## (Intercept)     26.540   0.703 1425.530    0.000
## tempo1         -13.018   1.021  162.690    0.000
## tempo4         -11.026   1.053  109.606    0.000
## tempo6          -5.778   1.126   26.313    0.000
## GrupoP          -0.268   0.994    0.073    0.788
## tempo1:GrupoP   11.406   1.109  105.840    0.000
## tempo4:GrupoP    8.824   1.141   59.820    0.000
## tempo6:GrupoP    3.152   1.244    6.421    0.011
```

Note a concordância das inferências mesmo considerando estruturas de correlação de trabalho bastante distintas.

```r
round(summary(gee1a.ind)$corr,3)
```

```
## [1] Estimate Std.err
## <0 rows> (or 0-length row.names)
```

```r
round(summary(gee1a.exch)$corr,3)
```

```
##       Estimate Std.err
## alpha    0.595   0.072
```

```r
round(summary(gee1a.ar1)$corr,3)
```

```
##       Estimate Std.err
## alpha    0.733   0.052
```

```r
round(summary(gee1a.unst)$corr,3)
```

```
##           Estimate Std.err
## alpha.1:2    0.435   0.076
## alpha.1:3    0.449   0.085
## alpha.1:4    0.506   0.058
## alpha.2:3    0.809   0.112
## alpha.2:4    0.676   0.104
## alpha.3:4    0.698   0.137
```

Vamos comparar as estimativas do modelo marginal dadas pelo GEE com aquelas do estimador GLS para as mesmas estruturas de correlação.

Estimador GLS

Fazemos o ajuste no `R` através dos seguintes comandos:

```r
gls1a.ind<-gls(chumbo ~ tempo*Grupo, data=datalong) #Independente
gls1a.exch<-gls(chumbo ~ tempo*Grupo, correlation=corCompSymm(form=~1|ID),
               data=datalong) #Simetria composta
gls1a.ar1<-gls(chumbo ~ tempo*Grupo, correlation=corAR1(form=~1|ID),
              data=datalong) #AR(1)
gls1a.unst<-gls(chumbo ~ tempo*Grupo, correlation=corSymm(form=~1|ID),
               data=datalong) #Não estruturada
```

Os resultados são:

```r
# Independente
round(coef(summary(gls1a.ind)),3)
```

```
##                 Value Std.Error t-value p-value
## (Intercept)    26.540     0.937  28.324   0.000
## tempo1        -13.018     1.325  -9.824   0.000
## tempo4        -11.026     1.325  -8.321   0.000
## tempo6         -5.778     1.325  -4.360   0.000
## GrupoP         -0.268     1.325  -0.202   0.840
## tempo1:GrupoP  11.406     1.874   6.086   0.000
## tempo4:GrupoP   8.824     1.874   4.709   0.000
## tempo6:GrupoP   3.152     1.874   1.682   0.093
```

```r
# Simetria composta
round(coef(summary(gls1a.exch)),3)
```

```
##                 Value Std.Error t-value p-value
## (Intercept)    26.540     0.937  28.324   0.000
## tempo1        -13.018     0.843 -15.445   0.000
## tempo4        -11.026     0.843 -13.082   0.000
## tempo6         -5.778     0.843  -6.855   0.000
## GrupoP         -0.268     1.325  -0.202   0.840
## tempo1:GrupoP  11.406     1.192   9.569   0.000
## tempo4:GrupoP   8.824     1.192   7.403   0.000
## tempo6:GrupoP   3.152     1.192   2.644   0.009
```

```r
# AR(1)
round(coef(summary(gls1a.ar1)),3)
```

```
##                 Value Std.Error t-value p-value
## (Intercept)    26.540     0.932  28.482   0.000
## tempo1        -13.018     0.801 -16.261   0.000
## tempo4        -11.026     1.022 -10.785   0.000
## tempo6         -5.778     1.140  -5.067   0.000
## GrupoP         -0.268     1.318  -0.203   0.839
## tempo1:GrupoP  11.406     1.132  10.075   0.000
## tempo4:GrupoP   8.824     1.446   6.103   0.000
## tempo6:GrupoP   3.152     1.613   1.955   0.051
```

```r
# Não estruturada
round(coef(summary(gls1a.unst)),3)
```

```
##                 Value Std.Error t-value p-value
## (Intercept)    26.540     0.937  28.310   0.000
## tempo1        -13.018     0.843 -15.450   0.000
## tempo4        -11.026     0.858 -12.856   0.000
## tempo6         -5.778     0.903  -6.396   0.000
```

```
## GrupoP         -0.268     1.326  -0.202   0.840
## tempo1:GrupoP  11.406     1.192   9.572   0.000
## tempo4:GrupoP   8.824     1.213   7.275   0.000
## tempo6:GrupoP   3.152     1.278   2.467   0.014
```

Observe como os erros padrões estimados dependem da escolha da estrutura de correlação. Isso ocorre apesar da coincidência nas estimativas pontuais dos parâmetros de média.

```r
gls1a.ind$modelStruct$corStruct
```

```
## NULL
```

```r
gls1a.exch$modelStruct$corStruct
```

```
## Correlation structure of class corCompSymm representing
##       Rho
## 0.5954401
```

```r
gls1a.ar1$modelStruct$corStruct
```

```
## Correlation structure of class corAR1 representing
##       Phi
## 0.6309418
```

```r
gls1a.unst$modelStruct$corStruct
```

```
## Correlation structure of class corSymm representing
##  Correlation:
##   1     2     3
## 2 0.596
## 3 0.582 0.769
## 4 0.536 0.552 0.551
```

Podemos comparar as diferentes estruturas de correlação por meio dos critérios AIC e BIC para estruturas gerais (não encaixadas) e via teste de razão de verossimilhanças para modelos encaixados.

```r
anova(gls1a.unst, gls1a.exch)
```

```
##            Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## gls1a.unst     1 15 2471.632 2531.200 -1220.816
## gls1a.exch     2 10 2480.621 2520.334 -1230.311 1 vs 2 18.98944  0.0019
```

```r
anova(gls1a.unst, gls1a.ar1)
```

```
##            Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## gls1a.unst     1 15 2471.632 2531.200 -1220.816
## gls1a.ar1      2 10 2492.631 2532.343 -1236.315 1 vs 2 30.99904  <.0001
```

```r
anova(gls1a.exch, gls1a.ar1)
```

```
##            Model df      AIC      BIC    logLik
## gls1a.exch     1 10 2480.621 2520.334 -1230.311
## gls1a.ar1      2 10 2492.631 2532.343 -1236.315
```

Pelo teste da razão de verossimilhanças escolhemos o modelo com correlação não estruturada. O critério AIC também identifica a correlação não estruturada como a melhor escolha; já o BIC elege a correlação simetria composta.

Como consideramos o tempo como fator e efeito de interação tempo e grupo, podemos testar a diferença entre os grupos para um tempo especificado. Seja $\delta_j,\ j = 1, \ldots, 4$ a verdadeira diferença entre os grupos tratamento e controle no tempo j. Para $j = 4$ (seis semanas), por exemplo, o modelo especifica:

$$\hat{E}(Y_{i4}|grupo_i = P) = \hat{\beta}_1 + \hat{\beta}_4,$$
$$\hat{E}(Y_{i4}|grupo_i = A) = \hat{\beta}_1 + \hat{\beta}_4 + \hat{\beta}_5 + \hat{\beta}_8.$$

Assumindo correlação de trabalho não estruturada e estimador GLS, temos

$$\hat{\delta}_4 = \hat{\beta}_5 + \hat{\beta}_8$$

```
delta4 <- as.numeric(coef(gls1a.unst)[5]+coef(gls1a.unst)[8])
delta4
```

```
## [1] 2.884
```

A variância estimada de $\hat{\delta}$ é dada por:

$$\widehat{Var}(\hat{\delta}_4) = \widehat{Var}(\hat{\beta}_5) + \widehat{Var}(\hat{\beta}_8) + 2\widehat{Cov}(\hat{\beta}_5, \hat{\beta}_8).$$

```
var.beta <- vcov(gls1a.unst)
round(var.beta,2)
```

```
##               (Intercept) tempo1 tempo4 tempo6 GrupoP tempo1:GrupoP
## (Intercept)          0.88  -0.35  -0.37  -0.41  -0.88          0.35
## tempo1              -0.35   0.71   0.52   0.37   0.35         -0.71
## tempo4              -0.37   0.52   0.74   0.38   0.37         -0.52
## tempo6              -0.41   0.37   0.38   0.82   0.41         -0.37
## GrupoP              -0.88   0.35   0.37   0.41   1.76         -0.71
## tempo1:GrupoP        0.35  -0.71  -0.52  -0.37  -0.71          1.42
## tempo4:GrupoP        0.37  -0.52  -0.74  -0.38  -0.74          1.04
## tempo6:GrupoP        0.41  -0.37  -0.38  -0.82  -0.82          0.74
##               tempo4:GrupoP tempo6:GrupoP
## (Intercept)            0.37          0.41
## tempo1                -0.52         -0.37
## tempo4                -0.74         -0.38
## tempo6                -0.38         -0.82
## GrupoP                -0.74         -0.82
## tempo1:GrupoP          1.04          0.74
## tempo4:GrupoP          1.47          0.76
## tempo6:GrupoP          0.76          1.63
```

Logo, temos:

```
var.delta4 <- as.numeric(diag(var.beta)[5]+diag(var.beta)[8]+2*var.beta[5,8])
var.delta4
```

```
## [1] 1.757711
```

```
sqrt(var.delta4)
```

```
## [1] 1.325787
```

A estatística de teste é:

```
delta4/sqrt(var.delta4)
```

```
## [1] 2.175312
```

Observamos uma diferença estatisticamente significativa entre os grupos. Contudo, é mais direto realizar comparações como essa por meio de outra especificação do preditor linear. O modelo permanecerá o mesmo mas as comparações dentro de cada tempo serão mais diretas, pois envolverão apenas um parâmetro.

Formulação 2: Modelo linear de efeitos fixos (sem intercepto)

O segundo modelo é da forma `chumbo ~ tempo*Grupo - Grupo - 1` e tem a seguinte representação:

$$
\begin{aligned}
E(Y_{ij}) =& \beta_1 I(tempo_j = 0) + \beta_2 I(tempo_j = 1) + \beta_3 I(tempo_j = 4) + \beta_4 I(tempo_j = 6) + \\
& \beta_5 I(Grupo_i = P) \times I(tempo_j = 0) + \beta_6 I(Grupo_i = P) \times I(tempo_j = 1) + \\
& \beta_7 I(Grupo_i = P) \times I(tempo_j = 4) + \beta_8 I(Grupo_i = P) \times I(tempo_j = 6).
\end{aligned}
$$

Os quatro últimos parâmetros permitem a comparação entre os dois grupos. Eles representam diretamente as estimativas de $\delta_j,\ j = 1, \ldots, 4$.

Estimador GEE

```
gee1b.ind<-geeglm(chumbo ~ tempo*Grupo - Grupo - 1, corstr="independence", id=ID,
                 family="gaussian", data=datalong) #Independente
gee1b.exch<-geeglm(chumbo ~ tempo*Grupo - Grupo - 1, corstr="exchangeable", id=ID,
                  family="gaussian", data=datalong) #Simetria composta
gee1b.ar1<-geeglm(chumbo ~ tempo*Grupo - Grupo - 1, corstr="ar1", id=ID, data=datalong) #AR(1)
gee1b.unst<-geeglm(chumbo ~ tempo*Grupo - Grupo - 1, corstr="unstructured", id=ID,
                  data=datalong) #Não estruturada
```

As estimativas são dadas por:

```
# Independente
round(coef(summary(gee1b.ind)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## tempo0           26.540   0.703 1425.530    0.000
## tempo1           13.522   1.074  158.472    0.000
## tempo4           15.514   1.099  199.163    0.000
## tempo6           20.762   1.294  257.243    0.000
## tempo0:GrupoP    -0.268   0.994    0.073    0.788
## tempo1:GrupoP    11.138   1.318   71.363    0.000
## tempo4:GrupoP     8.556   1.363   39.417    0.000
## tempo6:GrupoP     2.884   1.516    3.618    0.057
```

```
# Simetria composta
round(coef(summary(gee1b.exch)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## tempo0           26.540   0.703 1425.530    0.000
## tempo1           13.522   1.074  158.472    0.000
## tempo4           15.514   1.099  199.163    0.000
## tempo6           20.762   1.294  257.243    0.000
## tempo0:GrupoP    -0.268   0.994    0.073    0.788
## tempo1:GrupoP    11.138   1.318   71.363    0.000
## tempo4:GrupoP     8.556   1.363   39.417    0.000
## tempo6:GrupoP     2.884   1.516    3.618    0.057
```

```
# AR(1)
round(coef(summary(gee1b.ar1)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## tempo0           26.540   0.703 1425.530    0.000
## tempo1           13.522   1.074  158.472    0.000
## tempo4           15.514   1.099  199.163    0.000
## tempo6           20.762   1.294  257.243    0.000
```

```
## tempo0:GrupoP   -0.268   0.994    0.073    0.788
## tempo1:GrupoP   11.138   1.318   71.363    0.000
## tempo4:GrupoP    8.556   1.363   39.417    0.000
## tempo6:GrupoP    2.884   1.516    3.618    0.057
```

```
# Não estruturada
round(coef(summary(gee1b.unst)),3)
```

```
##                Estimate Std.err     Wald Pr(>|W|)
## tempo0           26.540   0.703 1425.530    0.000
## tempo1           13.522   1.074  158.472    0.000
## tempo4           15.514   1.099  199.163    0.000
## tempo6           20.762   1.294  257.243    0.000
## tempo0:GrupoP    -0.268   0.994    0.073    0.788
## tempo1:GrupoP    11.138   1.318   71.363    0.000
## tempo4:GrupoP     8.556   1.363   39.417    0.000
## tempo6:GrupoP     2.884   1.516    3.618    0.057
```

Estimador GLS

```
gls1b.ind<-gls(chumbo ~ tempo*Grupo - Grupo - 1, data=datalong) #Independente
gls1b.exch<-gls(chumbo ~ tempo*Grupo - Grupo - 1,
               correlation=corCompSymm (form=~1|ID), data=datalong) #Simetria Composta
gls1b.ar1<-gls(chumbo ~ tempo*Grupo - Grupo - 1,
              correlation=corAR1 (form=~1|ID), data=datalong) #AR1
gls1b.unst<-gls(chumbo ~ tempo*Grupo - Grupo - 1,
               correlation=corSymm (form=~1|ID), data=datalong) #Não estruturada
```

```
# Independente
round(coef(summary(gls1b.ind)),3)
```

```
##                 Value Std.Error t-value p-value
## tempo0         26.540     0.937  28.324    0.00
## tempo1         13.522     0.937  14.431    0.00
## tempo4         15.514     0.937  16.557    0.00
## tempo6         20.762     0.937  22.158    0.00
## tempo0:GrupoP -0.268     1.325  -0.202    0.84
## tempo1:GrupoP 11.138     1.325   8.405    0.00
## tempo4:GrupoP  8.556     1.325   6.457    0.00
## tempo6:GrupoP  2.884     1.325   2.176    0.03
```

```
# Simetria composta
round(coef(summary(gls1b.exch)),3)
```

```
##                 Value Std.Error t-value p-value
## tempo0         26.540     0.937  28.324    0.00
## tempo1         13.522     0.937  14.431    0.00
## tempo4         15.514     0.937  16.557    0.00
## tempo6         20.762     0.937  22.158    0.00
## tempo0:GrupoP -0.268     1.325  -0.202    0.84
## tempo1:GrupoP 11.138     1.325   8.405    0.00
## tempo4:GrupoP  8.556     1.325   6.457    0.00
## tempo6:GrupoP  2.884     1.325   2.176    0.03
```

```
# AR(1)
round(coef(summary(gls1b.ar1)),3)
```

```
##                Value Std.Error t-value p-value
## tempo0         26.540     0.932  28.482   0.000
## tempo1         13.522     0.932  14.512   0.000
## tempo4         15.514     0.932  16.649   0.000
## tempo6         20.762     0.932  22.282   0.000
## tempo0:GrupoP -0.268     1.318  -0.203   0.839
## tempo1:GrupoP 11.138     1.318   8.452   0.000
## tempo4:GrupoP  8.556     1.318   6.493   0.000
## tempo6:GrupoP  2.884     1.318   2.189   0.029
```

```
# Não estruturada
round(coef(summary(gls1b.unst)),3)
```

```
##                Value Std.Error t-value p-value
## tempo0         26.540     0.937  28.310    0.00
## tempo1         13.522     0.937  14.424    0.00
## tempo4         15.514     0.937  16.549    0.00
## tempo6         20.762     0.937  22.147    0.00
## tempo0:GrupoP -0.268     1.326  -0.202    0.84
## tempo1:GrupoP 11.138     1.326   8.401    0.00
## tempo4:GrupoP  8.556     1.326   6.454    0.00
## tempo6:GrupoP  2.884     1.326   2.175    0.03
```

Comparando os estimadores GLS e GEE notamos que o primeiro apresenta erros padrões menores. Assim, assumindo um nível de significância $\alpha = 0.05$, o efeito de grupo no GEE não é significativo em seis semanas, embora o seja na primeira e na quarta. Na aleatorização (semana zero) os grupos são estatisticamente iguais, como esperado. As inferências via GLS coincidem com as do GEE, exceto para a semana seis, para a qual a diferença encontrada é estatisticamente significativa.

Observação: o modelo GLS acima é homocedástico pois há apenas uma variância residual. É possível, contudo, estimar variâncias diferentes por grupo por meio do argumento `weights`. Por exemplo, a opção `weights = varIdent(form = ~ 1 | Grupo)` especifica variâncias diferentes por grupo. De forma similar, é possível especificar variâncias heterogêneas por tempo. Como exercício, avalie a necessidade de considerar variâncias distintas por grupo ou tempo.