CE075 - Análise de Dados Longitudinais

Silva, J.L.P.

12 de agosto, 2019

Alguns Exemplos Reais

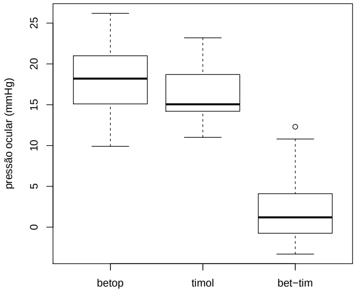
Exemplo 1: Estudo "cross-over"

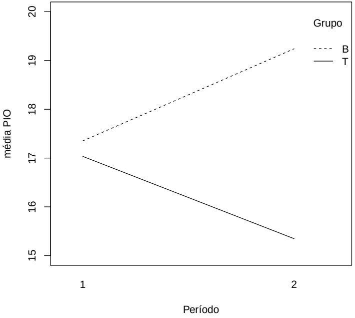
- Um oftalmologista quer comparar o efeito de dois colírios (B: Betoptic, T: Timoptol) redutores da pressão ocular com relação ao fluxo sanguíneo.
- Para tal ele submeteu 32 pacientes aos dois colírios por um período de dois meses com um descanso de igual tamanho. Foram 17 pacientes submetidos ao B e 15 ao T.
- A ordem da aplicação dos colírios foi aleatória.
- Duas medidas de pressão (colírio B e colírio T) foram tomadas ao fim do estudo para cada paciente.
- Eventualmente, a medida de linha de base pode ser útil na análise estatística.

Exemplo 1: Estudo “cross-over”

- Ponto Principal: Existe diferença entre os colírios?
- Existe efeito da ordem?
- Existe efeito de período?
- A medida de linha de base é útil?
- O descanso (“washout") de dois meses foi suficiente?

Exemplo 1: Estudo "cross-over"





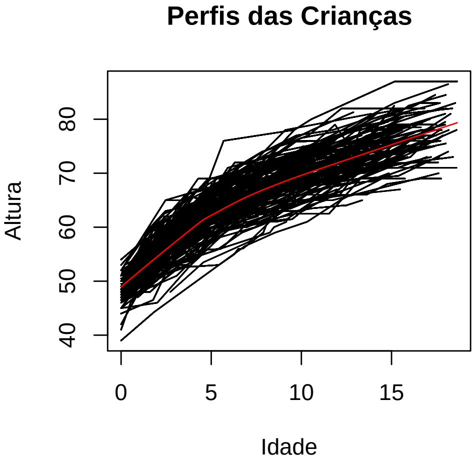
Exemplo 2: Avaliação longitudinal do crescimento de lactentes nascidos de mães infectadas com o HIV-1.

- Comparar longitudinalmente a altura de lactentes infectados e não-infectados nascidos de mães infectadas pelo HIV.
- Uma coorte aberta acompanhada no ambulatório de AIDS pediátrica do Hospital das Clínicas da Universidade Federal de Minas Gerais.
- Período: 1995 a 2003.
- Inclusão: primeiros três meses de vida.
- Grupos: (1) não-infectados: 97; (2) infectados: 42.
- Controlado por sexo.

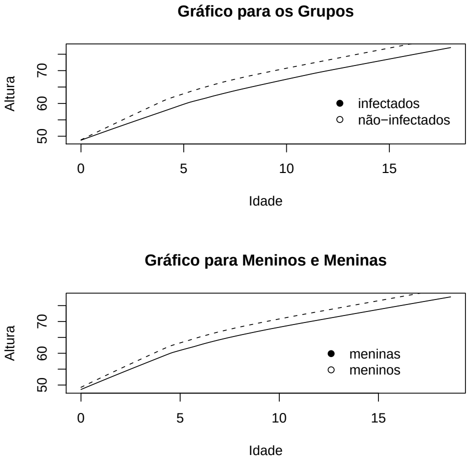
Exemplo 2: Avaliação longitudinal do crescimento de lactentes nascidos de mães infectadas com o HIV-1.

- Visitas regulares ao pediatra.
- Planejado para acompanhamento de 18 meses.
- Tempo: idade da criança.
- Tempo mediano de acompanhamento foi 15 meses (7 a 18).
- Número total de medidas: Não-infectados: 907; Infectados: 411.
- Número médio de visitas por criança: 9,5.
- Delineamento não-balanceado.

Exemplo 2: Avaliação longitudinal do crescimento de lactentes nascidos de mães infectadas com o HIV-1.



Exemplo 2: Avaliação longitudinal do crescimento de lactentes nascidos de mães infectadas com o HIV-1.



Exemplo 3: Marcadores Psicofisiológicos de Proteção e Vulnerabilidade ao Estresse Psicossocial

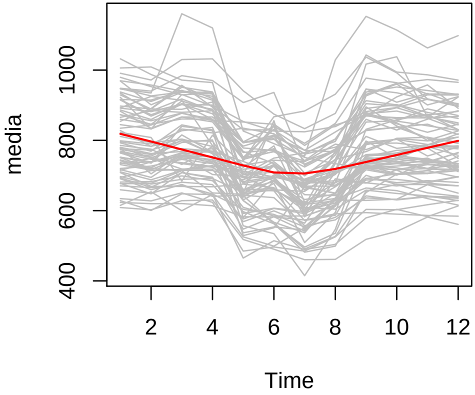
Os objetivos gerais deste estudo são:

- Investigar as reações cardíacas a uma situação de estresse social.
- Investigar a capacidade de regulação dessas respostas em função da afetiva individual (fatores internos) e da indução prévia de um estado de afeto positivo ou negativo (fator externo).

Exemplo 3: Marcadores Psicofisiológicos de Proteção e Vulnerabilidade ao Estresse Psicossocial

- Participaram do experimento 72 estudantes universitários da Universidade de Granada (Espanha) de ambos os sexos, com idade entre 18 a 30 anos.
- Foram utilizadas 40 fotos agradáveis (famílias e bebês) e 40 fotos desagradáveis (pessoas com mutilações) selecionadas do catálogo International Affective Picture System - IAPS.
- O objetivo das fotos é induzir um estado de humor positivo ou negativo, respectivamente.
- Resposta: período cardíaco médio avaliado em 12 momentos.

Exemplo 3: Marcadores Psicofisiológicos de Proteção e Vulnerabilidade ao Estresse Psicossocial



Exemplo 4: Trauma Odontológico

- O indivíduo perde o(s) dente(s) por acidente.
- O dente é reimplantado em um serviço de urgência (Odilon Behrens).
- Em seguida ele é encaminhado ao serviço de trauma da Faculdade de Odontologia da UFMG para tratamento de canal.
- No período entre o reimplante e o canal, existe um processo de reabsorção inflamatória.
- Este processo de reabsorção é medido por um índice.
- A resposta de interesse é a avaliação deste índice, em especial se ele ultrapassou ou não o valor 4 na chegada ao tratamento de canal.

Exemplo 4: Trauma Odontológico

- O objetivo é identificar fatores que aceleram ou desaceleram o crescimento do índice.
- Fatores: período extra-oral, meio de armazenamento, idade, etc.
- Alguns pacientes contribuem com mais de um dente.
- A princípio o estudo é transversal.

Exemplo 5: Dados de Crescimento

Por fim, dados do estudo clássico de crescimento de Potthoff e Roy (1964).

- Se refere à mudança nas medidas ortodônticas ao longo do tempo de 11 meninas e 16 meninos.
- Os indivíduos foram avaliados em quatro períodos de tempo: aos 8, 10, 12 e 14 anos.
- Resposta: distância do centro da pituitária à fissura do maxilar.
- Como objetivo podemos colocar: como essa distância cresce com a idade e testar se há diferença entre os valores para os meninos e meninas e se existe interação entre essas variáveis.

Exemplo 5: Dados de Crescimento

```r
library(mice)
data(potthoffroy)
head(potthoffroy)
```

```
  id sex   d8  d10  d12  d14
1  1   F 21.0 20.0 21.5 23.0
2  2   F 21.0 21.5 24.0 25.5
3  3   F 20.5 24.0 24.5 26.0
4  4   F 23.5 24.5 25.0 26.5
5  5   F 21.5 23.0 22.5 23.5
6  6   F 20.0 21.0 21.0 22.5
```

Exemplo 5: Dados de Crescimento

```
with(potthoffroy,by(potthoffroy[,-c(1,2)],sex,summary,digits=3))
```

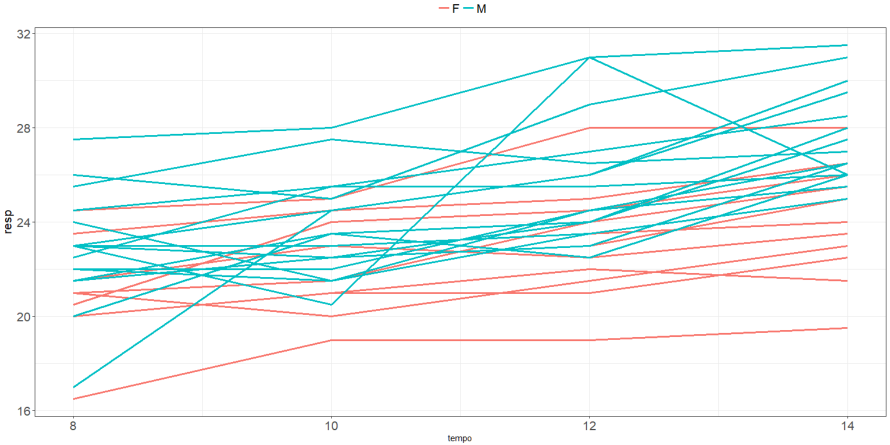
```
sex: F
       d8              d10              d12              d14
 Min.   :16.5   Min.   :19.0   Min.   :19.0   Min.   :19.5
 1st Qu.:20.2   1st Qu.:21.0   1st Qu.:21.8   1st Qu.:22.8
 Median :21.0   Median :22.5   Median :23.0   Median :24.0
 Mean   :21.2   Mean   :22.2   Mean   :23.1   Mean   :24.1
 3rd Qu.:22.2   3rd Qu.:23.5   3rd Qu.:24.2   3rd Qu.:25.8
 Max.   :24.5   Max.   :25.0   Max.   :28.0   Max.   :28.0
--------------------------------------------------------
sex: M
       d8              d10              d12              d14
 Min.   :17.0   Min.   :20.5   Min.   :22.5   Min.   :25.0
 1st Qu.:21.9   1st Qu.:22.4   1st Qu.:23.9   1st Qu.:26.0
 Median :23.0   Median :23.5   Median :25.0   Median :26.8
 Mean   :22.9   Mean   :23.8   Mean   :25.7   Mean   :27.5
 3rd Qu.:24.1   3rd Qu.:25.1   3rd Qu.:26.6   3rd Qu.:28.8
 Max.   :27.5   Max.   :28.0   Max.   :31.0   Max.   :31.5
```

Exemplo 5: Dados de Crescimento

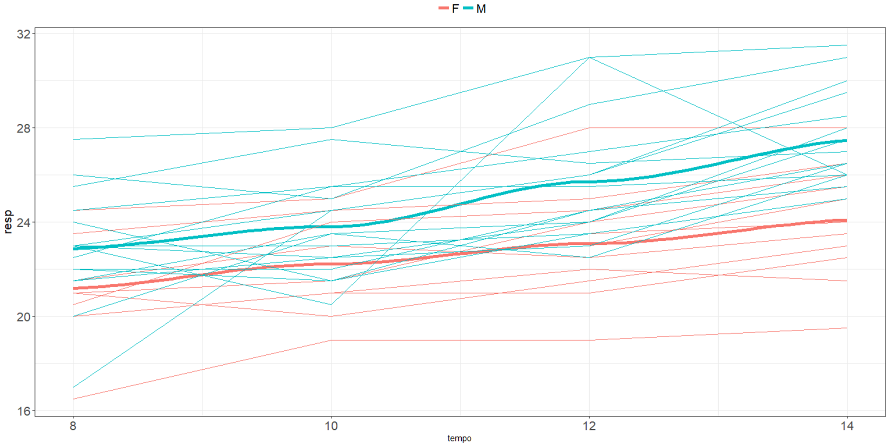
```
with(potthoffroy,by(potthoffroy[,-c(1,2)],sex,cor))
```

```
sex: F
           d8       d10       d12       d14
d8  1.0000000 0.8300900 0.8623146 0.8413558
d10 0.8300900 1.0000000 0.8954156 0.8794236
d12 0.8623146 0.8954156 1.0000000 0.9484070
d14 0.8413558 0.8794236 0.9484070 1.0000000
---------------------------------------------------------
sex: M
           d8       d10       d12       d14
d8  1.0000000 0.4373932 0.5579310 0.3152311
d10 0.4373932 1.0000000 0.3872909 0.6309234
d12 0.5579310 0.3872909 1.0000000 0.5859866
d14 0.3152311 0.6309234 0.5859866 1.0000000
```

Exemplo 5: Dados de Crescimento



Exemplo 5: Dados de Crescimento



Comparação de Médias

Comparação de duas Médias

Vamos retomar a comparação dos colírios:

- Pacientes com pressão intra-ocular (PIO) elevada irão participar do estudo.
- A pressão será medida após dois meses de uso do colírio.
- O objetivo é comparar a redução média de PIO dos dois colírios.

Denotando os grupos por A e B, queremos o seguinte:

$$\delta = \mu_A - \mu_B.$$

O interesse é então testar a hipótese:

$$H_0 : \delta = 0.$$

Comparação de Médias

- Existem duas formas básicas de conduzir o estudo:
  - 50 pacientes são submetidos ao colírio A e ao colírio B (medidas repetidas). Considera-se uma período de descanso de dois meses entre a aplicação dos colírios. É indicado aleatorizar a ordem de aplicação de A e B.
  - 100 pacientes são selecionados e 50 são sorteados para receber o colírio A e os demais recebem o B.
- Ambos estudos são experimentais
  - Pareado: Estudo Cross-over.
  - Amostras Independentes: Estudo Clínico Aleatorizado.
- Qual forma você utilizaria?

Amostra Pareada ou Independente?

1. Vantagens de Parear as Amostras
   - Controlar por possíveis fatores de confusão.
   - Menos pacientes/unidades na amostra.
   - Teste mais preciso com menos suposições.
   - Controla pelo efeito de coorte.
2. Vantagens de Amostras Independentes
   - Dados são obtidos de forma mais rápida.

Teste-t pareado

O objetivo é comparar duas medidas pareadas.

$$\delta = \mu_A - \mu_B.$$

Uma estimativa natural para δ é a diferença das médias. Ou seja

$$\widehat{\delta} = \widehat{\mu}_A - \widehat{\mu}_B.$$

A variância de $\widehat{\delta}$ é

$$Var(\widehat{\delta}) = \frac{1}{N}(\sigma_A^2 + \sigma_B^2 - 2\sigma_{AB}).$$

Teste-t pareado

Usualmente dados longitudinais têm correlação positiva. Ou seja $\sigma_{AB} > 0$.

Isto significa que a estatística a ser utilizada tem menor variância do que aquela obtida com dados independentes.

Considere as diferenças:

$$d_i = y_{i1} - y_{i2} \qquad i = 1, \ldots, n.$$

A estatística de teste é:

$$t = \frac{\overline{d}}{s_d/\sqrt{n}},$$

que, sob H_0, tem uma distribuição t com $n-1$ graus de liberdade.
Suposição: d_i vem de uma distribuição normal.

Teste-t pareado: exemplo colírios

```r
file <- "http://www.est.ufmg.br/~enricoc/pdf/longitudinais/scr
colirio <- read.table(file, h=TRUE, dec=',')
head(colirio)
```

```
  PRONT. IDADE SEXO IOP_sem IOP_betop IOP_timo Ordem
1   6649    75    F    19.1      17.7     14.4     0
2   3106    61    M    30.5      21.7     21.4     1
3  15231    57    F    19.1      18.4     17.8     0
4    799    42    F    20.0      14.8     18.1     1
5   9371    59    M    24.0      21.8     14.7     0
6    757    65    M    19.3      20.3     20.8     1
```

Teste-t pareado: exemplo colírios

```r
colMeans(cbind(colirio$IOP_beto,colirio$IOP_timo))
```

```
[1] 18.2375 16.1375
```

```r
var(cbind(colirio$IOP_beto,colirio$IOP_timo))
```

```
         [,1]     [,2]
[1,] 16.20242  5.65629
[2,]  5.65629 10.85919
```

```r
cor(colirio$IOP_betop, colirio$IOP_timo)
```

```
[1] 0.4264253
```

Teste-t pareado[1]: exemplo colírios

```
t.test(colirio$IOP_betop, colirio$IOP_timo,paired=TRUE)
```

```
    Paired t-test

data:  colirio$IOP_betop and colirio$IOP_timo
t = 2.9934, df = 31, p-value = 0.005378
alternative hypothesis: true difference in means is not equal
95 percent confidence interval:
 0.6692014 3.5307986
sample estimates:
mean of the differences
                    2.1
```

[1] o teste somente é válido se não houver efeito de período e de "carry-over"

Teste-t pareado: exemplo colírios

```
t.test(colirio$IOP_betop, colirio$IOP_timo)
```

```
    Welch Two Sample t-test

data:  colirio$IOP_betop and colirio$IOP_timo
t = 2.2836, df = 59.674, p-value = 0.02597
alternative hypothesis: true difference in means is not equal
95 percent confidence interval:
 0.2603069 3.9396931
sample estimates:
mean of x mean of y
  18.2375   16.1375
```

Teste-t pareado: exemplo colírios

```
x=colirio$IOP_betop; y=colirio$IOP_timo; n=length(x)
#delta
delta=mean(x)-mean(y);delta
```

```
[1] 2.1
```

```
#variância de delta, independente
delta.var.ind=(var(x)+var(y))/n;delta.var.ind
```

```
[1] 0.8456754
```

```
#variância de delta, pareado
delta.var.par=(var(x)+var(y)-2*cov(x,y))/n;delta.var.par
```

```
[1] 0.4921573
```

```
var(x-y)/n
```

```
[1] 0.4921573
```

Teste-t pareado: exemplo colírios

```r
#t independente
t.ind=delta/sqrt(delta.var.ind);t.ind
```

```
[1] 2.283586
```

```r
#t pareado
t.par=delta/sqrt(delta.var.par);t.par
```

```
[1] 2.993418
```

Comparação de mais de duas médias

Comparação dos colírios A, B e C:

- Pacientes com pressão intra-ocular elevada irão participar do estudo.
- A pressão será medida após dois meses de uso do colírio.
- O objetivo é comparar a redução média dos três colírios.

Então, queremos testar a seguinte hipótese:

$$H_0 : \mu_A = \mu_B = \mu_C.$$

A ANOVA é válida neste caso?

Caracterização dos Dados Longitudinais

Análise de Dados Longitudinais

- Características:
  - As respostas de diferentes unidades são independentes.
  - As respostas para a mesma unidade são correlacionadas.
  - De uma forma geral, as respostas próximas no tempo devem ser mais correlacionadas.
- Medida Temporal
  - Idade.
  - Calendário medido a partir de um certo evento.

Análise de Dados Longitudinais

- Objetivos do Estudo:
  - Avaliar o comportamento temporal.
  - Avaliar o efeito de covariáveis sobre a resposta.
  - Predição.
- Modelos de Regressão
  - Modelos marginais (modelar a média e a estrutura de covariância).
  - Modelo de efeitos aleatórios.
  - Modelo de transição.

Características da Correlação dos Dados

- As correlações usualmente são positivas.
- As correlações usualmente diminuem à medida que aumenta a separação no tempo.
- As correlações entre medidas repetidas raramente aproximam do zero.
- Medidas muito próximas tendem a ter correlação um.