

Trabalho 3

Planejamento de Experimentos

Paulo Ricardo Seganfredo Campana

6 de maio de 2024

Experimentos com efeito fixo

Quando queremos testar se um fator com g níveis é significativo na mudança de uma variável resposta y , utilizamos a análise de variância para testar as hipóteses de diferenças de médias entre os níveis.

Seja n o número de observações de cada nível do fator, $N = ng$ o número de observações total e o índice $i = \{1, 2, \dots, n\}$ referente as observações e o índice $j = \{1, 2, \dots, g\}$ referente aos níveis.

A anova se baseia em decompor cada observação da variável medida y_{ij} em uma média global μ uma mudança de média entre os níveis τ_j e um erro normal $\varepsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$, assim temos o seguinte modelo e as hipóteses do teste:

$$y_{ij} = \mu_j + \varepsilon_{ij} \implies y_{ij} = \mu + \tau_j + \varepsilon_{ij}$$

$$\begin{cases} \mathcal{H}_0 : \mu_1 = \mu_2 = \dots = \mu_g \\ \mathcal{H}_1 : \mu_p \neq \mu_q, \text{ para algum } p, q \end{cases} \implies \begin{cases} \mathcal{H}_0 : \tau_1 = \tau_2 = \dots = \tau_g = 0 \\ \mathcal{H}_1 : \tau_j \neq 0, \text{ para algum } j \end{cases}$$

Com os dados em mão, é possível estimar estes valores:

$$\mu = \frac{1}{N} \sum_{j=1}^g \sum_{i=1}^n y_{ij} \qquad \mu_j = \frac{1}{n} \sum_{i=1}^n y_{ij} \qquad \tau_j = \mu_j - \mu$$

Da mesma forma, podemos decompor a soma de quadrados total SS_T na soma de quadrados devido ao fator SS_F e devido aos erros SS_E , estas somas de quadrados irão representar a variância devido a cada componente depois que dividirmos pelos graus de liberdade de cada um.

$$\underbrace{\sum_{j=1}^g \sum_{i=1}^n (y_{ij} - \mu)^2}_{SS_T, N-1 \text{ graus}} = \underbrace{\sum_{j=1}^g \sum_{i=1}^n (\mu_j - \mu)^2}_{SS_F, g-1 \text{ graus}} + \underbrace{\sum_{j=1}^g \sum_{i=1}^n (y_{ij} - \mu_j)^2}_{SS_E, N-g \text{ graus}}$$

Se a maior parte da soma dos quadrados totais é explicada pela soma devido ao fator e menos pela soma devido aos erros, então temos que a diferença entre os níveis é significativa e podemos rejeitar a hipótese nula, assim a estatística do teste é:

$$F = \frac{SS_F/(g-1)}{SS_E/(N-g)} \sim \mathcal{F}_{(g-1, N-g)}$$

Que tem distribuição F -Snedecor, rejeitamos a hipótese nula se a estatística F ultrapassa o quantil $\mathcal{F}_{(g-1, N-g)}(1 - \alpha)$.

Problem 3.6. An article in Bioelectromagnetics (“Electromagnetic Effects on Forearm Disuse Osteopenia: A Randomized, Double-Blind, Sham-Controlled Study”, Vol. 32, 2011, pp. 273-282) described a randomized, double-blind, sham-controlled, feasibility and dosing study to determine if a common pulsing electromagnetic field (PEMF) treatment could moderate the substantial osteopenia that occurs after forearm disuse. Subjects were randomized into four groups after a distal radius fracture, or carpal surgery requiring immobilization in a cast. Active or identical sham PEMF transducers were worn on the distal forearm for 1, 2, or 4h/day for 8 weeks starting after cast removal (“baseline”) when bone density continues to decline. Bone mineral density (BMD) and bone geometry were measured in the distal forearm by dual energy X-ray absorptiometry (DXA) and peripheral quantitative computed tomography (pQCT). The data below are the percent losses in BMD measurements on the radius after 16 weeks for patients wearing the active or sham PEMF transducers for 1, 2, or 4h/day (data were constructed to match the means and standard deviations read from a graph in the paper).

Is there evidence to support a claim that PEMF usage affects BMD loss?

```
data <- data.frame(
  loss = c(
    4.51, 7.95, 4.97, 3    , 7.97, 2.23, 3.95, 5.64, 9.35, 6.52,
    4.96, 6.1  , 7.19, 4.03, 2.72, 9.19, 5.17, 5.7  , 5.85, 6.45,
    5.32, 6    , 5.12, 7.08, 5.48, 6.52, 4.09, 6.28, 7.77, 5.68,
    8.47, 4.58, 4.11, 5.72, 5.91, 6.89, 6.99, 4.98, 9.94, 6.38,
    4.73, 5.81, 5.69, 3.86, 4.06, 6.56, 8.34, 3.01, 6.71, 6.51,
    1.7  , 5.89, 6.55, 5.34, 5.88, 7.5  , 3.28, 5.38, 7.3  , 5.46,
    7.03, 4.65, 6.65, 5.49, 6.98, 4.85, 7.26, 5.92, 5.58, 7.91,
    4.9  , 4.54, 8.18, 5.42, 6.03, 7.04, 5.17, 7.6  , 7.9  , 7.91
  ),
  hours = factor(rep(c(0, 1, 2, 4), each = 20))
)
```

Os dados acima referem a queda em percentual da densidade mineral dos ossos de pacientes com certa doença óssea após usarem um aparelho eletrônico que possivelmente poderia combater os sintomas da doença. 3 grupos de pacientes usaram o aparelho por 1, 2 e 4 horas por dia, e o grupo de controle utilizou um aparelho que não funciona, fazendo papel de placebo. Queremos testar se o aparelho ajuda a diminuir a queda da densidade mineral dos ossos.

Temos 4 níveis com 20 observações cada e iremos realizar uma análise de variância (anova) para testar se há diferença na variável medida em relação ao fator tomando um nível de significância de 5%.

A função `anova` do R nos dá a informação dos graus de liberdade, soma e média de quadrados do fator e dos resíduos, além da estatística do teste e p -valor.

```
model <- lm(loss ~ hours, data)
anova(model) |>
  knitr::kable()
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
hours	3	10.04	3.348	1.298	0.2813
Residuals	76	196.03	2.579	NA	NA

Devido a estatística do teste $F = 1.298$ não é mais extrema que o quantil de 95% da distribuição $\mathcal{F} = 2.725$, não há evidências para rejeitar a hipótese de que os níveis são iguais, ou seja, o aparelho não parece fazer efeito na redução da perda da densidade mineral dos ossos dos pacientes.

Como o teste ANOVA conclui que as diferenças entre os tratamentos e o “placebo” não foram estatisticamente significantes, não é necessário efetuar as comparações múltiplas, pois a diferenças entre quaisquer níveis também não serão significantes, mesmo assim vamos fazer para praticar.

Sendo um caso onde temos um grupo de controle, faz sentido o uso do teste de comparações múltiplas de Dunnett onde testamos a diferença de médias entre cada grupo de tratamento com o grupo de controle, ou seja testar as hipóteses:

$$\begin{cases} \mathcal{H}_0 : \mu_1 = \mu_2 \\ \mathcal{H}_1 : \mu_1 \neq \mu_2 \end{cases} \quad \begin{cases} \mathcal{H}_0 : \mu_1 = \mu_3 \\ \mathcal{H}_1 : \mu_1 \neq \mu_3 \end{cases} \quad \begin{cases} \mathcal{H}_0 : \mu_1 = \mu_4 \\ \mathcal{H}_1 : \mu_1 \neq \mu_4 \end{cases}$$

Neste teste a hipótese nula é rejeitada se acontecer a seguinte desigualdade, onde $d_\alpha(g-1, N-g)$ é o quantil de probabilidade α da distribuição do teste de Dunnett tabelada no apêndice VIII do livro do Montgomery. Para $g-1 = 3$ e $N-g = 76$ com um nível de significância de 5%, este quantil é aproximadamente 2.39.

$$|\bar{y}_j - \bar{y}_1| > d_\alpha(g-1, N-g) \sqrt{\frac{2SS_E}{n(g-1)}}, \quad j = \{2, 3, \dots, g\}$$

Utilizando o pacote **DescTools**, obtemos as informações de diferença de médias entre os grupos, intervalo de confiança para esta diferença e p -valor do teste de comparações múltiplas para as comparações de todos os níveis com o grupo controle.

```
DescTools::DunnettTest(loss ~ hours, data)$`0` |>
  knitr::kable()
```

	diff	lwr.ci	upr.ci	pval
1-0	0.4930	-0.7244	1.710	0.6470
2-0	-0.1945	-1.4119	1.023	0.9636
4-0	0.6780	-0.5394	1.895	0.4048

Assim, pelos p -valores dos testes não podemos rejeitar as hipóteses de que a média da variável resposta é idêntica entre os grupos de tratamento e o grupo de controle. O que pode ser visto também pelo intervalo de confiança das diferenças, que contém o 0.

Experimentos com efeito aleatório

Quando os níveis presentes na amostra do fator que estamos investigando não são todos os possíveis níveis, estamos no caso de ANOVA com efeito aleatório, neste caso o modelo é o mesmo.

Não estamos interessados em saber se os níveis coletados na amostra daquele fator é significativo na mudança da variável resposta y , mas sim se todos os possíveis níveis são significativos. Para isso assumimos que τ_j é uma variável aleatória de variância σ_τ^2 e que τ_j é independente de ε_{ij} , assim a variância total das observações é:

$$\text{Var}[y_{ij}] = \text{Var}[\mu + \tau_j + \varepsilon_{ij}] = \sigma_\tau^2 + \sigma^2$$

Para testar se a população de possíveis níveis do fator tem significância na média da variável resposta, testamos a variância de τ_j :

$$\begin{cases} \mathcal{H}_0 : \sigma_\tau^2 = 0 \\ \mathcal{H}_1 : \sigma_\tau^2 > 0 \end{cases}$$

Para isso usamos a mesma estatística de teste, pois sob a hipótese nula, a variância de todas as observações são as mesmas.

$$F = \frac{\text{SS}_F/(g-1)}{\text{SS}_E/(N-g)} \sim \mathcal{F}_{(g-1, N-g)}$$

É também possível estimar σ^2 e σ_τ^2 e seus intervalos de confiança da seguinte forma:

$$\begin{aligned} \hat{\sigma}^2 &= \text{MS}_E & \hat{\sigma}_\tau^2 &= \frac{\text{MS}_F - \text{MS}_E}{n} \\ \text{IC}(\sigma^2 | \alpha) &= \left[\frac{(N-g)\text{MS}_E}{\chi_{(N-g)}^2(\alpha/2)}, \frac{(N-g)\text{MS}_E}{\chi_{(N-g)}^2(1-\alpha/2)} \right] & \text{IC} \left(\frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma^2} \middle| \alpha \right) &= \left[\frac{L}{1+L}, \frac{U}{1+U} \right] \end{aligned}$$

$$\begin{aligned} L &= \frac{1}{n} \left(\frac{\text{MS}_F}{\text{MS}_{E\mathcal{F}_{(g-1, N-g)}(\alpha/2)}} - 1 \right) \\ U &= \frac{1}{n} \left(\frac{\text{MS}_F}{\text{MS}_{E\mathcal{F}_{(g-1, N-g)}(1-\alpha/2)}} - 1 \right) \end{aligned}$$

3.10. A product developer is investigating the tensile strength of a new synthetic fiber that will be used to make cloth for men's shirts. Strength is usually affected by the percentage of cotton used in the blend of materials for the fiber. The engineer conducts a completely randomized experiment with five levels of cotton content and replicates the experiment five times. The data are shown in the following table.

Is there evidence to support the claim that cotton content affects the mean tensile strength? Use $\alpha = 0.05$.

```
data <- data.frame(
  strength = c(
    7, 7, 15, 11, 9,
    12, 17, 12, 18, 18,
    14, 19, 19, 18, 18,
    19, 25, 22, 19, 23,
    7, 10, 11, 15, 11
  ),
  cotton = rep(c(15, 20, 25, 30, 35), each = 5)
)
```

Como os níveis do fator da porcentagem de algodão são valores numéricos e escolhidos arbitrariamente, temos um caso de ANOVA com efeito aleatório. Para testar se a porcentagem de algodão influencia na média da força de tensão das camisetas, vamos fazer uso da library **ExpDes**.

```
ExpDes::crd(data$cotton, data$strength, quali = FALSE, unfold = 0) |>
  try()
```

```
## -----
```

```
## Analysis of Variance Table
```

```
## -----
```

```
##          DF  SS    MS   Fc   Pr>Fc
```

```
## Treatment  4 476 118.9 14.8 9.13e-06
```

```
## Residuals 20 161   8.1
```

```
## Total      24 637
```

```
## -----
```

```
## CV = 18.88 %
```

```
##
```

```
## -----
```

```
## Shapiro-Wilk normality test
```

```
## p-value: 0.1818
```

```
## According to Shapiro-Wilk normality test at 5% of significance, residuals can be considered normal.
```

```
## -----
```

```
##
```

```
## -----
```

```
## Homogeneity of variances test
```

```
## p-value: 0.9198
```

```
## According to the test of bartlett at 5% of significance, residuals can be considered homocedastic.
```

```
## -----
```

```
##
```

```
## According to the F test, the means can not be considered distinct.
```

```
## -----
```

```
##      Levels Means
## 1         15   9.8
## 2         20  15.4
## 3         25  17.6
## 4         30  21.6
## 5         35  10.8
## -----
## Error in ExpDes::crd(data$cotton, data$strength, quali = FALSE, unfold = 0) :
##   object 'reg' not found
```

Primeiramente, é realizado os testes de normalidade de Shapiro-Wilk e homogeneidade de variâncias entre os grupos de Bartlett, ambos não rejeitam a hipótese nula ao nível de 5% de significância, cumprindo com as suposições do modelo.

Pela tabela da ANOVA temos uma estatística de teste alta de 14.8 e p -valor baixo, inferior a 0.001, assim rejeitamos a hipótese de que a variável aleatória τ_j seja uma constante, portanto há influência entre a porcentagem de algodão nas camisetas e a força de tensão das mesmas.

Também podemos estimar que $\sigma_\tau^2 = \frac{118.94 - 8.06}{5} = 22.176$, bastante diferente de 0.

Experimentos com bloco aleatório

Durante uma análise de variância, alguns fatores externos podem influenciar nos resultados e atrapalhar as conclusões do teste, quando podemos observar e controlar um destes fatores, é possível incluí-lo no modelo de forma a estudar se seu comportamento influencia na medição e separar essa influência do fator do bloco da influência do fator de interesse.

No experimento de bloco aleatório completo, temos que cada observação de cada nível do fator de interesse corresponde a um nível do fator do bloco, assim o índice $i \in \{1, 2, \dots, b\}$ é referente ao bloco, incluindo a influência do bloco, temos o modelo:

$$y_{ij} = \mu + \tau_j + \beta_i + \varepsilon_{ij}$$

Como ainda queremos testar a significância de um fator de interesse na média da variável resposta, as hipóteses são as mesmas. porém agora temos um termo adicional na decomposição da soma de quadrados, a soma de quadrados do bloco, assim podemos também testar a hipótese de que o bloco influencia na variável resposta ou não.

$$\begin{cases} \mathcal{H}_0 : \tau_1 = \tau_2 = \dots = \tau_g = 0 \\ \mathcal{H}_1 : \tau_j \neq 0, \text{ para algum } j \end{cases} \quad \begin{cases} \mathcal{H}_0 : \beta_1 = \beta_2 = \dots = \beta_g = 0 \\ \mathcal{H}_1 : \beta_i \neq 0, \text{ para algum } i \end{cases}$$

$$\underbrace{\sum_{j=1}^g \sum_{i=1}^b (y_{ij} - \mu)^2}_{SS_T, N-1 \text{ graus}} = \underbrace{\sum_{j=1}^g \sum_{i=1}^b (\mu_j - \mu)^2}_{SS_F, g-1 \text{ graus}} + \underbrace{\sum_{j=1}^g \sum_{i=1}^b (\mu_i - \mu)^2}_{SS_B, b-1 \text{ graus}} + \underbrace{\sum_{j=1}^g \sum_{i=1}^b \hat{\varepsilon}_{ij}^2}_{SS_E, (g-1)(b-1) \text{ graus}}$$

E as estatísticas de teste para estas hipóteses são feitas de maneira similar, pela razão de médias de quadrados. Rejeitamos as hipóteses nulas se a estatística do teste ultrapassar o quantil de $(1 - \alpha)\%$ da distribuição sob a hipótese nula.

$$F = \frac{SS_F/(g-1)}{SS_E/(g-1)(b-1)} \sim \mathcal{F}_{(g-1, (g-1)(b-1))}$$

$$F_B = \frac{SS_B/(b-1)}{SS_E/(g-1)(b-1)} \sim \mathcal{F}_{(b-1, (g-1)(b-1))}$$

4.10. An article in the Fire Safety Journal (“The Effect of Nozzle Design on the Stability and Performance of Turbulent Water Jets,”” Vol. 4, August 1981) describes an experiment in which a shape factor was determined for several different nozzle designs at six levels of jet efflux velocity. Interest focused on potential differences between nozzle designs, with velocity considered as a nuisance variable. The data are shown below:

Does nozzle design affect the shape factor? Compare the nozzles with an analysis of variance, using $\alpha = 0.05$.

```
data <- data.frame(
  shape = c(
    0.78, 0.80, 0.81, 0.75, 0.77, 0.78,
    0.85, 0.85, 0.92, 0.86, 0.81, 0.83,
    0.93, 0.92, 0.95, 0.89, 0.89, 0.83,
    1.14, 0.97, 0.98, 0.88, 0.86, 0.83,
    0.97, 0.86, 0.78, 0.76, 0.76, 0.75
  ),
  nozzle = rep(1:5, each = 6),
  velocity = factor(rep(c(11.73, 14.37, 16.59, 20.43, 23.46, 28.74), times = 5))
)
```

Vamos testar se o coeficiente de forma dos jatos d’água é influenciado pelo design do bico, sendo que temos informação adicional das velocidades do fluxo em que esse coeficiente foi calculado. A library **ExpDes** fornece esse experimento.

```
ExpDes::rbd(data$nozzle, data$velocity, data$shape, quali = TRUE)
```

```
## -----
## Analysis of Variance Table
## -----
##          DF      SS      MS    Fc    Pr>Fc
## Treatment  4 0.1022 0.02555 8.92 0.00027
## Block      5 0.0629 0.01257 4.39 0.00736
## Residuals 20 0.0573 0.00287
## Total      29 0.2223
## -----
## CV = 6.23 %
##
## -----
## Shapiro-Wilk normality test
## p-value: 0.5369
## According to Shapiro-Wilk normality test at 5% of significance, residuals can be considered normal.
## -----
##
## -----
## Homogeneity of variances test
## p-value: 0.4645
## According to the test of oneillmathews at 5% of significance, the variances can be considered homoc
## -----
##
## Tukey's test
```



```
## -----
## Groups Treatments Means
## a      4    0.9433
## ab     3    0.9017
## abc    2    0.8533
## bc     5    0.8133
## c      1    0.7817
## -----
```

Primeiramente, os testes de normalidade de Shapiro-Wilk e homogeneidade de variâncias de O'Neill-Matthews não rejeitam as hipóteses nulas, cumprindo com as suposições do modelo.

Em seguida, vemos que com um p -valor de 0.00027 o design do bico é sim significativo no cálculo do coeficiente de forma dos jatos d'água, porém infelizmente os blocos, as velocidades do fluxo, também são significativas, não estamos interessados nessa medida porém ela tem influência na variável resposta e na significância do fator de interesse.

Adicionalmente, vemos pelo teste de Tukey que o design de bico 4 foi o que obteve maior coeficiente de forma, enquanto que o design 1 obteve menor, oque isso significa, eu não sei.

Experimento com quadrado latino

É um experimento similar ao bloco aleatório completo, neste caso teremos mais de um fator de bloco que pode afetar a significância do fator de interesse, para isso adicionamos uma restrição a mais:

Temos a seguinte situação: um conjunto de dados com p^2 observações, sendo elas classificadas em 3 variáveis categóricas cada uma com p categorias distintas, a primeira variável categórica é a variável da “linha” i , a segunda da “coluna” j e a terceira é a variável “tratamento” de interesse k , porém com a restrição de que não temos observações duplicadas em conjuntos de variáveis como por exemplo medições da variável resposta na mesma linha, coluna e tratamento. Sendo assim dividimos a contribuição do valor da variável resposta em 5 componentes:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}, \quad i, j, k \in \{1, \dots, p\}$$

- μ : o componente da média geral
- α_i : o componente relacionado ao fator bloco da linha i
- β_j : o componente relacionado ao fator bloco da coluna j
- τ_k : o componente relacionado ao fator de interesse k
- ε_{ijk} : o componente relacionado ao efeito aleatório

Da mesma maneira, iremos trabalhar em cima da soma de quadrados de cada componente para testar as hipóteses de se as componentes são significantes em relação a variável resposta, as hipóteses e decomposição de somas de quadrados ficam:

$$\begin{cases} \mathcal{H}_0 : \tau_1 = \tau_2 = \dots = \tau_p = 0 \\ \mathcal{H}_1 : \tau_k \neq 0, \text{ para algum } k \end{cases}$$

$$\underbrace{\sum_{i=1}^p \sum_{j=1}^p (y_{ijk} - \mu)^2}_{\text{SS}_T, p^2-1 \text{ graus}} = \underbrace{p \sum_{i=1}^p (\mu_i - \mu)^2}_{\text{SS}_{\text{linha}}, p-1 \text{ graus}} + \underbrace{p \sum_{j=1}^p (\mu_j - \mu)^2}_{\text{SS}_{\text{coluna}}, p-1 \text{ graus}} + \underbrace{p \sum_{k=1}^p (\mu_k - \mu)^2}_{\text{SS}_F, p-1 \text{ graus}} + \underbrace{\sum_{i=1}^p \sum_{j=1}^p \tilde{\varepsilon}_{ijk}^2}_{\text{SS}_E, (p-2)(p-1) \text{ graus}}$$

Padronizando as somas de quadrados pelos seus graus de liberdade temos a média de quadrados, a estatística F que será usada para decisão do teste é a razão entre a média de quadrado dos tratamentos e dos erros, é possível testar também a significância do fator da linha e da coluna caso for de interesse.

$$F = \frac{\text{SS}_F / (p-1)}{\text{SS}_E / ((p-2)(p-1))} \sim \mathcal{F}_{(p-1, (p-2)(p-1))}$$

4.23. An industrial engineer is investigating the effect of four assembly methods (A, B, C, D) on the assembly time for a color television component. Four operators are selected for the study. Furthermore, the engineer knows that each assembly method produces such fatigue that the time required for the last assembly may be greater than the time required for the first, regardless of the method. That is, a trend develops in the required assembly time. To account for this source of variability, the engineer uses the Latin square design shown below.

Analyze the data from this experiment ($\alpha = 0.05$) and draw appropriate conclusions.

```
data <- data.frame(
  order = rep(1:4, times = 4),
  operator = rep(1:4, each = 4),
  method = c(
    "C", "B", "A", "D",
    "D", "C", "B", "A",
    "A", "D", "C", "B",
    "B", "A", "D", "C"
  ),
  time = c(
    10, 7, 5, 10,
    14, 18, 10, 10,
    7, 11, 11, 12,
    8, 8, 9, 14
  )
)
```

Estamos interessados se o método de montagem das televisões (A, B, C, D) é tem impacto estatisticamente significativo no tempo de montagem, porém temos 4 operadores diferentes que fizeram a montagem usando cada um dos 4 métodos em 4 ordens distintas, assim temos um experimento com quadrado latino.

Usando o pacote **ExpDes** temos todos os resultados úteis para a análise de experimentos com quadrados latinos:

```
ExpDes::latsd(
  treat = data$method,
  row = data$order,
  column = data$operator,
  resp = data$time
)
## -----
## Analysis of Variance Table
## -----
##           DF      SS      MS      Fc  Pr>Fc
## Treatment  3  72.5  24.17  13.81 0.0042
## Row        3  18.5   6.17   3.52 0.0885
## Column     3  51.5  17.17   9.81 0.0099
## Residuals  6  10.5   1.75
## Total     15 153.0
## -----
## CV = 12.91 %
##
```

```
## -----
## Shapiro-Wilk normality test
## p-value: 0.3604
## According to Shapiro-Wilk normality test at 5% of significance, residuals can be considered normal.
## -----
##
## Tukey's test
## -----
## Groups Treatments Means
## a      C    13.25
## ab     D    11
## bc     B    9.25
## c      A    7.5
## -----
```

Primeiramente temos testes de normalidade dos resíduos de Shapiro-Wilk que não rejeita a hipótese de normalidade, assim o modelo tem suas suposições validadas.

Pelo p -valor do teste F, temos que os tratamentos (método de montagem) e as colunas (operador) são significantes ao influenciar o tempo de montagem das televisões com p -valores 0.04 e 0.01 respectivamente, porém a variável da linha (ordem dos operários) não é significativa com p -valor 0.09.

Adicionalmente, o pacote proporciona testes de comparações múltiplas, com o padrão de teste de Tukey, isso nos mostra que os métodos de montagem que consomem menor tempo são A, B, D e C respectivamente.

Experimentos com medidas repetidas

Acontece quando os fator de que queremos testar a significância com a análise de variância são medidas de um mesmo indivíduo em tempos diferentes ou espaços diferentes por exemplo. Quando as medidas são de um mesmo indivíduo, temos uma menor variância devido ao componente aleatório e assim o teste é mais poderoso.

Assim, podemos representar esse experimento com o seguinte modelo: onde,

$$y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}, \quad i \in \{1, \dots, n\}, \quad j \in \{1, \dots, g\}$$

- μ : o componente da média geral
- β_i : o componente relacionado ao efeito do indivíduo i
- τ_j : o componente relacionado ao efeito do fator de interesse j
- ε_{ij} : o componente relacionado ao efeito aleatório

É equivalente a um experimento de bloco aleatório completo, onde o fator de bloco são os indivíduos do experimento, a hipótese a ser testada é:

$$\begin{cases} \mathcal{H}_0 : \tau_1 = \tau_2 = \dots = \tau_g = 0 \\ \mathcal{H}_1 : \tau_j \neq 0, \text{ para algum } j \end{cases}$$

A soma de quadrados totais é decomposta na contribuição de cada componente

$$\underbrace{\sum_{i=1}^n \sum_{j=1}^g (y_{ij} - \mu)^2}_{SS_T, N-1 \text{ graus}} = \underbrace{g \sum_{i=1}^n (\mu_i - \mu)^2}_{SS_{\text{indivíduos}}, n-1 \text{ graus}} + \underbrace{n \sum_{j=1}^g (\mu_j - \mu)^2}_{SS_F, g-1 \text{ graus}} + \underbrace{\sum_{i=1}^n \sum_{j=1}^g \hat{\varepsilon}_{ij}^2}_{SS_E, (n-1)(g-1) \text{ graus}}$$

A estatística do teste para verificar se há diferença entre os tratamentos é dado pela razão da soma de quadrados dos tratamentos e dos erros padronizada pelos seus graus de liberdade.

$$F = \frac{SS_F / (g - 1)}{SS_E / ((n - 1)(g - 1))} \sim \mathcal{F}_{(g-1, (n-1)(g-1))}$$

Também é possível testar se há diferenças significantivas entre os indivíduos utilizando a soma de quadrados dos indivíduos e seus graus de liberdade, é esperado que este fator seja significativo pois em geral há sim diferenças entre os indivíduos, a significância desse fator não interfere na significância do fator de interesse pois está incluso no modelo.

Apresenta-se a seguir os resultados obtidos nos testes A, B e C por nove indivíduos. Pode-se afirmar que os testes A, B e C apresentam resultados iguais?

Faça uma análise desse experimento utilizando medidas repetidas.

```
data <- data.frame(
  individuo = factor(rep(1:9, each = 3)),
  teste = factor(rep(c("A", "B", "C"), times = 9)),
  resultado = c(
    98, 95, 77,
    95, 71, 79,
    76, 80, 91,
    95, 81, 84,
    83, 77, 80,
    99, 70, 93,
    82, 80, 87,
    75, 72, 81,
    88, 81, 83
  )
)
```

Primeiramente devemos testar as hipóteses de normalidade e esfericidade do modelo:

```
data |>
  dplyr::group_by(teste) |>
  rstatix::shapiro_test(resultado)
## # A tibble: 3 x 4
##   teste variable statistic    p
##   <fct> <chr>      <dbl> <dbl>
## 1 A      resultado    0.901 0.255
## 2 B      resultado    0.868 0.117
## 3 C      resultado    0.940 0.578

rstatix::anova_test(
  data,
  dv = resultado,
  wid = individuo,
  within = teste
)$Mauchly
##   Effect    W      p p<.05
## 1  teste 0.996 0.987
```

Temos p -valores superiores ao nível de significância adotado de 5% para todos os testes, assim as suposições de normalidade e esfericidade são aceitas. Então construímos o modelo ANOVA de dois fatores sem interação.

```
mod <- lm(resultado ~ individuo + teste, data)
anova(mod) |>
  knitr::kable()
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
indivíduo	8	422.0	52.75	0.8683	0.5616
teste	2	394.7	197.33	3.2483	0.0655
Residuals	16	972.0	60.75	NA	NA

E vemos que as diferenças tanto entre os indivíduos como entre os testes não são significativas pelo p -valor superior a 5%, assim é possível afirmar que os testes A, B e C apresentam resultados que podem ser comparados uns aos outros.

```
TukeyHSD(aov(mod))$teste |>
  knitr::kable()
```

	diff	lwr	upr	p adj
B-A	-9.333	-18.814	0.1474	0.0540
C-A	-4.000	-13.481	5.4807	0.5344
C-B	5.333	-4.147	14.8141	0.3395

E pelas comparações múltiplas de Tukey, os testes que ficaram mais diferentes foram A e B, porém a diferença não é significativa com p -valor de 5.4%.

Experimentos crossover

Este é comumente usado em farmaceuticos para testar se uma droga de teste T é equivalente a droga de referência R , o experimento é organizado de forma que uma amostra aleatória de indivíduos tomam primeiro a droga teste e depois de passado o efeito tomam a droga de referência enquanto que o outro grupo tomam as drogas na ordem contrária.

Para as duas drogas serem bioequivalentes, é preciso que não há diferença significativa nas medições da concentração do efeito ativo no corpo tanto para as duas drogas como para a ordem em que elas foram tomadas (os dois grupos de indivíduos). Para isso montamos o seguinte modelo em que

$$y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk}$$

- μ : o componente da média geral
- α_i : o componente relacionado ao efeito do indivíduo $i \in \{1, 2, \dots, n\}$
- β_j : o componente relacionado ao efeito da ordem de consumo $j \in \{1, 2\}$
- τ_k : o componente relacionado ao efeito das diferentes drogas $k \in \{1, 2\}$
- ε_{ijk} : o componente relacionado ao efeito aleatório

No caso 2×2 onde testamos apenas duas drogas em duas ordens de consumo, para que haja bioequivalência, temos que testar as hipóteses de que o efeito das drogas e da ordem são não significativos com as seguintes hipóteses:

$$\begin{cases} \mathcal{H}_0 : \tau_1 = \tau_2 \\ \mathcal{H}_1 : \tau_1 \neq \tau_2 \end{cases} \quad \begin{cases} \mathcal{H}_0 : \beta_1 = \beta_2 \\ \mathcal{H}_0 : \beta_1 \neq \beta_2 \end{cases} \quad \begin{cases} \mathcal{H}_0 : \alpha_1 = \alpha_2 = \dots = \alpha_n = 0 \\ \mathcal{H}_1 : \alpha_i \neq 0, \text{ para algum } i \end{cases}$$

Utilizando o procedimento o mesmo procedimento da ANOVA, pela soma de quadrados e estatística F , realizamos três testes e precisamos que aceitar a hipótese nula das duas primeiras para as drogas serem consideradas iguais.

$$\underbrace{\sum_{i=1}^n \sum_{j=1}^2 \sum_{k=1}^2 (y_{ijk} - \mu)^2}_{\text{SS}_T, 2n-1 \text{ graus}} = \underbrace{4 \sum_{i=1}^n (\mu_i - \mu)^2}_{\text{SS}_{\text{indivíduos}}, n-1 \text{ graus}} + \underbrace{2n \sum_{j=1}^2 (\mu_j - \mu)^2}_{\text{SS}_{\text{ordem}}, 1 \text{ grau}} + \underbrace{2n \sum_{k=1}^2 (\mu_k - \mu)^2}_{\text{SS}_F, 1 \text{ grau}} + \underbrace{\sum_{i=1}^n \sum_{j=1}^2 \sum_{k=1}^2 \hat{\varepsilon}_{ijk}^2}_{\text{SS}_E, n-2 \text{ graus}}$$

$$F = \frac{\text{SS}_F/1}{\text{SS}_E/(n-2)} \sim \mathcal{F}_{(1, n-2)} \quad F_{\text{ordem}} = \frac{\text{SS}_{\text{ordem}}/1}{\text{SS}_E/(n-2)} \sim \mathcal{F}_{(1, n-2)} \quad F_{\text{indivíduos}} = \frac{\text{SS}_{\text{indivíduos}}/(n-1)}{\text{SS}_E/(n-2)} \sim \mathcal{F}_{(n-1, n-2)}$$

Estes dados são referentes ao Manual de boas práticas em bioequivalência da ANVISA, modulo 3 página 12.

Testar se as duas drogas T e R são de fato bioequivalentes na medição da área sobre a curva farmacocinética.

```
data <- data.frame(
  TRT = rep(c("R", "T"), each = 24),
  SUBJ = rep(1:24, times = 2),
  PRD = c(
    1,2,1,2,1,2,2,1,2,1,2,1,1,2,1,1,2,2,1,2,2,1,1,
    2,1,2,1,2,1,1,2,1,2,1,2,2,1,2,2,1,1,1,2,1,1,2,2
  ),
  GRP = c(
    "RT","TR","RT","TR","RT","TR","TR","RT","TR","RT","TR","RT",
    "RT","TR","RT","RT","TR","TR","TR","RT","TR","TR","RT","RT",
    "RT","TR","RT","TR","RT","TR","TR","RT","TR","RT","TR","RT",
    "RT","TR","RT","RT","TR","TR","TR","RT","TR","TR","RT","RT"
  ),
  asc = c(
    252.92, 785.23, 414.41, 624.60, 367.10, 193.63,
    321.37, 772.56, 530.35, 405.44, 363.72, 697.60,
    356.86, 599.80, 287.29, 562.38, 417.87, 590.51,
    424.91, 334.99, 312.16, 591.43, 798.33, 182.85,

    233.71, 926.39, 534.80, 793.66, 408.92, 246.33,
    518.33, 529.26, 523.34, 376.48, 509.73, 540.35,
    284.84, 640.77, 291.31, 555.81, 546.04, 571.03,
    361.26, 457.09, 470.42, 518.33, 789.24, 257.52
  )
)
```

Iremos usar o pacote **BE** para testar a bioequivalência.

```
BE::be2x2(data, "asc")
## $asc
## $asc$`Analysis of Variance (log scale)`
##           Sum Sq Df Mean Sq F value   Pr(>F)
## SUBJECT      6.61 23   0.287    12.08 8.0e-08 ***
## GROUP         0.33  1   0.331     1.16  0.293
## SUBJECT(GROUP) 6.28 22   0.285    12.00 9.5e-08 ***
## PERIOD        0.07  1   0.073     3.06  0.094 .
## DRUG          0.07  1   0.072     3.02  0.096 .
## ERROR        0.52 22   0.024
## TOTAL        7.28 47
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## $asc$`Between and Within Subject Variability`
##           Between Subject Within Subject
```

```

## Variance Estimate          0.1308      0.02379
## Coefficient of Variation, CV(%) 37.3799    15.51501
##
## $asc$`Least Square Means (geometric mean)`
##           Reference Drug Test Drug
## Geometric Means      430.3      464.9
##
## $asc$`90% Confidence Interval of Geometric Mean Ratio (T/R)`
##           Lower Limit Point Estimate Upper Limit
## 90% CI for Ratio      1.001      1.081      1.166
##
## $asc$`Sample Size`
##           True Ratio=1 True Ratio=Point Estimate
## 80% Power Sample Size      5      8

```

A um nível de significância de 5%, não rejeitamos a hipótese de que a ordem de ingestão e a droga ingerida são significativos na mudança da área da curva farmacocinética, pois estes fatores obtiveram p -valores de 9.4% e 9.6% respectivamente, assim podemos considerar que os dois medicamentos são bioequivalentes.

A influência dos indivíduos da área da curva farmacocinética é sim significativa, o que é esperado pois em geral a diferença entre a reação de medicamentos em indivíduos é verdadeira.

Experimentos fatoriais

Quando temos mais de um fator de interesse, temos um experimento fatorial, é interessantes testar tanto as hipóteses de se cada fator separado é significativo na mudança de média da variável resposta como a hipótese de se a interação entre os fatores é significativa.

O caso de experimento fatorial mais simples é o 2×2 , nesse experimento temos dois fatores, cada um com dois níveis, os níveis serão representados por -1 e 1 e sua interação será o produto dos níveis individuais, assim a interação será positiva quando os dois fatores forem de mesmo nível e negativo quando são de níveis diferentes. O modelo do experimento está a seguir:

$$y_{ijk} = \mu + \beta_i + \tau_j + \gamma_k + \chi_{jk} + \varepsilon_{ijk}$$

- μ : o componente da média geral
- β_i : o componente relacionado ao efeito do bloco $i \in \{1, 2, \dots, n\}$
- τ_j : o componente relacionado ao efeito do primeiro fator de interesse $j \in \{-1, 1\}$
- γ_k : o componente relacionado ao efeito do segundo fator de interesse $k \in \{-1, 1\}$
- χ_{jk} : o componente relacionado ao efeito do da interação entre os fatores de interesse $jk \in \{-1, 1\}$
- ε_{ijk} : o componente relacionado ao efeito aleatório

Temos interesse em três hipóteses principais: se o primeiro e o segundo fator são significativos, se a interação entre eles é significativa e se o bloco ou as medidas repetidas tem influência significativa.

Seja os fatores de interesse A e B , as somas de quadrado e estatística F dos testes são dadas por:

$$\begin{aligned} \underbrace{\sum_{i=1}^n \sum_{j=-1}^1 \sum_{k=-1}^1 (y_{ijk} - \mu)^2}_{SS_T, 4n-1 \text{ graus}} &= \underbrace{4 \sum_{i=1}^n (\mu_i - \mu)^2}_{SS_{\text{bloco}}, n-1 \text{ graus}} + \underbrace{2n \sum_{j=-1}^1 (\mu_j - \mu)^2}_{SS_A, 1 \text{ grau}} + \underbrace{2n \sum_{k=-1}^1 (\mu_k - \mu)^2}_{SS_B, 1 \text{ grau}} \\ &+ \underbrace{n \sum_{jk=-1}^1 (\mu_{jk} - \mu)^2}_{SS_{AB}, 1 \text{ graus}} + \underbrace{\sum_{i=1}^n \sum_{j=-1}^1 \sum_{k=-1}^1 \hat{\varepsilon}_{ijk}^2}_{SS_E, 3n-3 \text{ graus}} \end{aligned}$$

$$F_A = \frac{SS_A/1}{SS_E/(3n-3)} \sim \mathcal{F}_{(1, 3n-3)}$$

$$F_{AB} = \frac{SS_{AB}/3}{SS_E/(3n-3)} \sim \mathcal{F}_{(3, 3n-3)}$$

$$F_B = \frac{SS_B/1}{SS_E/(3n-3)} \sim \mathcal{F}_{(1, 3n-3)}$$

$$F_{\text{bloco}} = \frac{SS_{\text{bloco}}/(n-1)}{SS_E/(3n-3)} \sim \mathcal{F}_{(n-1, 3n-3)}$$

Problem 5.31. An article in Quality Progress (May 2011, pp. 42–48) describes the use of factorial experiments to improve a silver powder production process. This product is used in conductive pastes to manufacture a wide variety of products ranging from silicon wafers to elastic membrane switches. Powder density (g/cm^2) and surface area (cm^2/g) are the two critical characteristics of this product. The experiments involved three factors - reaction temperature, ammonium percent, and stirring rate. Each of these factors had two levels and the design was replicated twice. The design is shown below.

Analyze the density response. Are any interactions significant? Draw appropriate conclusions about the effects of the significant factors on the response.

```
data <- data.frame(
  stirrate = factor(rep(rep(c(100, 150), each = 4), times = 2)),
  temperature = factor(rep(c(8, 40), each = 16)),
  replicate = factor(rep(1:4, times = 8)),
  density = c(
    14.68, 15.18, 15.12, 17.48,
    7.54, 6.66, 12.46, 12.62,
    10.95, 17.68, 12.65, 15.96,
    8.03, 8.84, 14.96, 14.96
  )
)
```

Usamos a função `aov` do R base, adicionando o bloco, os fatores de interesse e sua interação utilizando o sinal `:`.

```
mod <- aov(
  density ~ replicate + stirrate + temperature + stirrate:temperature,
  data
)
summary(mod)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
## replicate	3	110	36.7	7.49	0.00097	***
## stirrate	1	141	141.4	28.87	1.4e-05	***
## temperature	1	0	0.0	0.00	1.00000	
## stirrate:temperature	1	0	0.0	0.00	1.00000	
## Residuals	25	122	4.9			
## ---						
## Signif. codes:	0	'***'	0.001	'**'	0.01	'*' 0.05
		'.'	0.1	' ' 1		

Pelos p -valores dos testes, vemos que a velocidade de mistura tem impacto significativo na densidade do pó de prata produzido pelo processo, também houve diferença significativa nas réplicas do experimento.

A média da densidade nos dois valores de temperaturas usadas no experimento foi exatamente a mesma pela acurácia de medição de duas casas decimais, assim tanto a temperatura como a interação dela com a velocidade de mistura não são significantes na mudança da média da densidade.