

Trabalho 2

Experimentos com quadrados latinos e ANOVA de medidas repetidas

Paulo Ricardo Seganfredo Campana

11 de abril de 2024

Em experimentos com quadrados latinos 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 efeito da linha i
- β_j : o componente relacionado ao efeito da coluna j
- τ_k : o componente relacionado ao tratamento k
- ε_{ijk} : o componente relacionado ao efeito aleatório

Da mesma maneira com os modelos ANOVA, iremos trabalhar em cima da soma de quadrados de cada componente para medir se as componentes são significantes em relação ao erro, a decomposição fica:

$$SS_{\text{total}} = SS_{\text{linha}} + SS_{\text{coluna}} + SS_{\text{tratamento}} + SS_{\text{erro}}$$

$$\underbrace{\sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p (y_{ijk} - \bar{y})^2}_{SS_{\text{total}}, p^2-1 \text{ graus}} = \underbrace{p \sum_{i=1}^p (\bar{y}_i - \bar{y})^2}_{SS_{\text{linha}}, p-1 \text{ graus}} + \underbrace{p \sum_{j=1}^p (\bar{y}_j - \bar{y})^2}_{SS_{\text{coluna}}, p-1 \text{ graus}} + \underbrace{p \sum_{k=1}^p (\bar{y}_k - \bar{y})^2}_{SS_{\text{tratamento}}, p-1 \text{ graus}} + \underbrace{\sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p \hat{\varepsilon}_{ijk}^2}_{SS_{\text{erro}}, (p-2)(p-1) \text{ graus}}$$

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 da variável da linha e da coluna.

$$F = \frac{SS_{\text{tratamento}}/(p-1)}{SS_{\text{erro}}/(p-2)(p-1)} \mathcal{F}_{p-1, (p-2)(p-1)}$$

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
  )
)
```

```
head(data, 8)
##   order operator method time
## 1     1         1      C    10
## 2     2         1      B     7
## 3     3         1      A     5
## 4     4         1      D    10
## 5     1         2      D    14
## 6     2         2      C    18
## 7     3         2      B    10
## 8     4         2      A    10
```

ExpDes

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
## -----
```

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 operarios) não é significativa com p-valor 0.09.

Adicionalmente, temos testes de normalidade dos resíduos que não rejeitam a hipótese de normalidade, assim o modelo tem suas suposições validadas, além disso o pacote proporciona testes de comparações múltiplas, com o padrão de teste de Tukey, isso nos mostra que a ordem dos métodos de montagem com menor tempo gasto é A, B, D e C.

Na mão

Exemplificando os cálculos no R, obtemos os mesmos resultados da tabela ANOVA:

```
p <- 4
ybar <- mean(data$time) # 10.25
ybar_i <- tapply(data$time, data$order, mean) # 9.75 11.00 8.75 11.50
ybar_j <- tapply(data$time, data$operator, mean) # 8.00 13.00 10.25 9.75
ybar_k <- tapply(data$time, data$method, mean) # 7.50 9.25 13.25 11.00

SS_tot <- sum((data$time - ybar)^2) # 153
SS_linha <- p * sum((ybar_i - ybar)^2) # 18.5
SS_coluna <- p * sum((ybar_j - ybar)^2) # 51.5
SS_tratamento <- p * sum((ybar_k - ybar)^2) # 72.5
SS_erro <- SS_tot - SS_linha - SS_coluna - SS_tratamento # 10.5

df_erro <- (p - 2) * (p - 1)
MSE_erro <- SS_erro / df_erro

F_linha <- (SS_linha / (p - 1)) / MSE_erro # 3.52
F_coluna <- (SS_coluna / (p - 1)) / MSE_erro # 9.80
F_tratamento <- (SS_tratamento / (p - 1)) / MSE_erro # 13.81

pval_linha <- pf(F_linha, p - 1, df_erro, lower.tail = FALSE) # 0.0885
pval_coluna <- pf(F_coluna, p - 1, df_erro, lower.tail = FALSE) # 0.0099
pval_tratamento <- pf(F_tratamento, p - 1, df_erro, lower.tail = FALSE) # 0.0042
```

ANOVA de medidas repetidas

Acontece quando os “tratamentos” que queremos testar a significância com ANOVA 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 tratamento j
- ε_{ij} : o componente relacionado ao efeito aleatório

E a hipótese a ser testada é

$$\begin{cases} \mathcal{H}_0 : \tau_1 = \tau_2 = \dots = \tau_g \\ \mathcal{H}_1 : \exists a, b ; \tau_a \neq \tau_b \end{cases}$$

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

$$SS_{\text{total}} = SS_{\text{indivíduos}} + SS_{\text{tratamento}} + SS_{\text{erro}}$$

$$\underbrace{\sum_{i=1}^n \sum_{j=1}^g (y_{ij} - \bar{y})^2}_{SS_{\text{total}}, rk-1 \text{ graus}} = \underbrace{g \sum_{i=1}^n (\bar{y}_i - \bar{y})^2}_{SS_{\text{indivíduos}}, n-1 \text{ graus}} + \underbrace{n \sum_{j=1}^g (\bar{y}_j - \bar{y})^2}_{SS_{\text{tratamento}}, g-1 \text{ graus}} + \underbrace{\sum_{i=1}^n \sum_{j=1}^g \hat{\varepsilon}_{ij}^2}_{SS_{\text{erro}}, (n-1)(g-1) \text{ graus}}$$

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.

$$F = \frac{SS_{\text{tratamento}}/(g-1)}{SS_{\text{erro}}/(n-1)(g-1)} \mathcal{F}_{g-1, (n-1)(g-1)}$$

Também é possível testar se há diferenças significantivas entre os indivíduos utilizando a soma de quadrados dos indivíduos e seus graus de liberdade ao invés dos tratamentos.

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 = rep(1:9, each = 3),
  teste = 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
  )
)
```

```
head(data, 9)
```

##	individuo	teste	resultado
## 1	1	A	98
## 2	1	B	95
## 3	1	C	77
## 4	2	A	95
## 5	2	B	71
## 6	2	C	79
## 7	3	A	76
## 8	3	B	80
## 9	3	C	91

Com ANOVA

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

```
library(rstatix)

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

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 temos as suposições de normalidade e esfericidade. Então construímos o modelo anova de dois fatores sem interação.

```
mod <- lm(resultado ~ as.factor(individuo) + as.factor(teste), data)
anova(mod)
## Analysis of Variance Table
##
## Response: resultado
##              Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(individuo)  8    422    52.8    0.87  0.562
## as.factor(teste)      2    395   197.3    3.25  0.065 .
## Residuals           16    972    60.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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 iguais.

Na mão

Exemplificando os cálculos no R, obtemos os mesmos resultados da tabela ANOVA:

```
n <- 9
g <- 3
ybar <- mean(data$resultado) # 83.44
ybar_i <- tapply(data$resultado, data$individuo, mean) # 90.00 81.67 82.33 ...
ybar_j <- tapply(data$resultado, data$teste, mean) # 87.89 78.56 83.89

SS_tot <- sum((data$resultado - ybar)^2) # 1789
SS_individuo <- g * sum((ybar_i - ybar)^2) # 422
SS_tratamento <- n * sum((ybar_j - ybar)^2) # 394.7
SS_erro <- SS_tot - SS_individuo - SS_tratamento # 972

df_erro <- (n - 1) * (g - 1)
MSE_erro <- SS_erro / df_erro

F_individuo <- (SS_individuo / (n - 1)) / MSE_erro # 0.8683
F_tratamento <- (SS_tratamento / (g - 1)) / MSE_erro # 3.248

pval_individuo <- pf(F_individuo, n - 1, df_erro, lower.tail = FALSE) # 0.5616
pval_tratamento <- pf(F_tratamento, g - 1, df_erro, lower.tail = FALSE) # 0.06547
```