

Relatório Estatística Multivarida I

Thiago Tavares Lopes

12 dezembro 2024

Atividade

- 1) Um novo método de aprendizado foi aplicado a 8 alunos de uma escola. Foram obtidas notas antes e depois da aplicação do novo método nas disciplinas de matemática e biologia. Podemos afirmar que houve diferença entre antes e depois?

```
dados <- data.frame(
  Indivíduo = 1:8,
  Matematica_Antes = c(8.9, 7.5, 5.6, 5.0, 6.5, 6.0, 6.9, 6.3),
  Biologia_Antes = c(4.4, 4.3, 7.4, 4.8, 4.8, 5.6, 4.3, 7.0),
  Matematica_Depois = c(8.3, 7.0, 8.4, 8.7, 7.7, 6.3, 7.1, 9.2),
  Biologia_Depois = c(7.8, 8.7, 8.7, 8.1, 6.8, 6.0, 6.6, 7.8)
)

# Diferenças (Depois - Antes)
dif <- dados[, 4:5] - dados[, 2:3]

# Teste  $T^2$  de Hotelling
# library(ICSNP)
# resultado <- HotellingsT2(dif)
# resultado

library(rrcov)

## Carregando pacotes exigidos: robustbase
## Scalable Robust Estimators with High Breakdown Point (version 1.7-6)
T2.test(dif)
```

```
##
## One-sample Hotelling test
##
## data: dif
## T2 = 35,6, F = 15,3, df1 = 2, df2 = 6, p-value = 0,0044
## alternative hypothesis: true mean vector is not equal to (0, 0)'
##
## sample estimates:
##               Matematica_Depois Biologia_Depois
## mean x-vector           1,25           2,2375
```

Resposta: O valor-p associado ao teste é 0,0044, indicando que rejeitamos a hipótese nula.

- 2) Considere os dados abaixo, coletados para dois grupos independentes (A e B) com relação às variáveis comprimento, peso e volume.

```

# Dados do Grupo A
grupo_A <- data.frame(
  Comprimento = c(13.53, 18.92, 10.91, 13.92, 13.38, 12.35, 13.97),
  Peso = c(19.22, 17.77, 15.99, 11.38, 18.7, 11.7, 18.46),
  Volume = c(17.7, 14.72, 11.5, 11.32, 12.37, 19.23, 16.23)
)

# Dados do Grupo B
grupo_B <- data.frame(
  Comprimento = c(16.3, 14.65, 14.18, 16.86, 16.14, 15.77, 18.2),
  Peso = c(13.2, 12.81, 15.29, 11.77, 13.65, 11.23, 13.56),
  Volume = c(11.5, 17.96, 16.91, 11.5, 14.94, 18.17, 15.46)
)

# Cálculo das médias
media_A <- colMeans(grupo_A)
media_B <- colMeans(grupo_B)
media_A

## Comprimento      Peso      Volume
##      13,854      16,174      14,724

media_B

## Comprimento      Peso      Volume
##      16,014      13,073      15,206

# Matriz de covariância combinada
n_A <- nrow(grupo_A)
n_B <- nrow(grupo_B)
Sp <- ((n_A - 1) * cov(grupo_A) + (n_B - 1) * cov(grupo_B)) / (n_A + n_B - 2)
Sp

##           Comprimento      Peso      Volume
## Comprimento      3,98471 0,82785 -0,70688
## Peso              0,82785 6,42748  0,43554
## Volume            -0,70688 0,43554  8,81908

# Teste T² de Hotelling

resultado_grupos <- T2.test(grupo_A, grupo_B)
resultado_grupos

##
## Two-sample Hotelling test
##
## data: grupo_A and grupo_B
## T2 = 11,73, F = 3,26, df1 = 3, df2 = 10, p-value = 0,068
## alternative hypothesis: true difference in mean vectors is not equal to (0,0,0)
## sample estimates:
##           Comprimento      Peso      Volume
## mean x-vector      13,854 16,174 14,724
## mean y-vector      16,014 13,073 15,206

```

Resposta: A média do grupo A: comprimento(13.854), Peso(16,174), Volume(14,724)

Media do grupo B: comprimento(16,014), Peso(13,073), Volume(15,206)

como o p-valor foi de 0,068 não rejeitamos H0.

- 3) Um teste foi realizado com 15 pessoas, em que foram avaliados o nível de Ansiedade, de Cortisol e também o escore da qualidade do sono.

```
# Dados
dados_q3 <- data.frame(
  Ansiedade = c(42, 38, 44, 40, 37, 41, 43, 39, 42, 38, 41, 40, 39, 43, 37),
  Cortisol = c(26, 24, 27, 25, 24, 26, 28, 25, 27, 24, 26, 25, 23, 28, 24),
  Qualidade_Sono = c(6.5, 7.2, 6.8, 7.0, 7.5, 6.9, 6.4, 7.3, 6.6, 7.1, 6.7, 7.4, 7.0, 6.8, 7.6)
)

# Matriz de correlação
correlacao <- cor(dados_q3)
print(correlacao)

##               Ansiedade Cortisol Qualidade_Sono
## Ansiedade      1,00000  0,90222      -0,83173
## Cortisol        0,90222  1,00000      -0,73359
## Qualidade_Sono -0,83173 -0,73359       1,00000

# Teste de significância das correlações
library(psych)
cortest.bartlett(dados_q3)

## R was not square, finding R from data
## $chisq
## [1] 34,844
##
## $p.value
## [1] 0,00000013145
##
## $df
## [1] 3

# Hipótese nula: vetor de médias igual a (39, 25, 7.0)
mu_hipotetico <- c(39, 25, 7.0)

# Teste T² de Hotelling

resultado_q3 <- T2.test(dados_q3, mu = mu_hipotetico)
print(resultado_q3)

##
## One-sample Hotelling test
##
## data: dados_q3
## T2 = 18,12, F = 5,18, df1 = 3, df2 = 12, p-value = 0,016
## alternative hypothesis: true mean vector is not equal to (39, 25, 7)'
##
## sample estimates:
##               Ansiedade Cortisol Qualidade_Sono
## mean x-vector    40,267    25,467         6,9867
```

- A) Pelo teste de Bartlett o p-valor é menor que 0,05, logo rejeitamos H0 e matriz de correlação é significativamente diferente de uma matriz identidade. Ou seja, há correlação entre as variáveis.
- B) Como o p-valor foi menor que 0,05, rejeitamos H0 e podemos dizer que o vetor de amostras é diferente do valor da população.
- 4) A tabela abaixo apresenta os dados fornecidos, organizados por tratamentos (variedades), altura das plantas e número de frutos. Execute uma análise de variância multivariada para comparar as variedades.

```

dados_q4 <- data.frame(
  Variedade = c("A", "B", "C", "D", "A", "B", "C", "D", "A", "B", "C", "D"),
  Altura_Plantas = c(45.2, 46.5, 47.0, 42.1, 43.0, 42.5, 50.1, 49.8, 51.2, 40.5, 41.0, 40.8),
  Numero_Frutos = c(15, 14, 16, 12, 13, 11, 17, 18, 16, 10, 9, 8)
)

# MANOVA
manova_result <- manova(cbind(Altura_Plantas, Numero_Frutos) ~ Variedade, data = dados_q4)
summary(manova_result)

##              Df Pillai approx F num Df den Df Pr(>F)
## Variedade    3  0,147    0,211      6     16  0,97
## Residuals    8

#summary.aov(manova_result) # Análise de variância univariada para cada variável

```

O teste da MANOVA indicou que as variedades não tem efeito significativos nas variáveis analisadas