

Laboratorio(participacao)

Comparacao entre proporção de células mortas entre os dias de experimentos

AUTHOR

Leo Alec

1 Contexto

Em estudos com células vivas, a variabilidade intrínseca do material biológico e as condições experimentais podem levar a flutuações nos resultados. Para garantir a confiabilidade dos dados, realizamos **três experimentos independentes em duplicata** (totalizando 6 réplicas técnicas por tratamento)

Nosso objetivo é **comparar as proporções de células mortas entre os experimentos**

2 X^2

O teste X^2 serve para avaliarmos se existe ou não uma associação significativa entre variáveis categóricas ou **proporções observadas**(Problema atual).

2.0.1 O que é o X^2 ?

- Hipoteses:
 - H_0 : A proporção é a mesma entre as amostras
 - H_1 : A proporção é diferente em pelo menos uma das amostras
- X^2 : Mede o **quão distantes** as contagens observadas estão das esperadas se H_0 fosse verdadeira.

$$X^2 = \sum \frac{O_i - E_i^2}{E_i}, \text{ onde } O_i = \text{valor observado do índice } i; E_i = \text{valor esperado do índice } i$$

- p : índice de rejeição ou aceitação da nossa hipótese
 - Se $p < 0.05$ (valor comumente utilizado), rejeitamos H_0 , ou seja, há diferenças significativas entre as amostras

2.0.1.1 Requerimentos

- Tamanho amostral ≥ 5
- Não Exige Normalidade
- Independência entre as amostras

Que ótimo! A situação se enquadra muito bem no nosso teste previsto.

2.0.2 Resolução do problema

- Hipoteses
 - H_0 : A proporção de células mortas é a mesma nos três dias

- H_1 : A proporcao das celulas mortas e difente em pelo menos 1 dia

▼ Code

```
df1 <- data.frame(
  Tratamento = c('Controle', 'Veículo', '1,25%', '10%', '25%'),
  V_Dupl1 = c(188, 110, 91, 134, 45),
  M_Dupl1 = c(38, 38, 27, 42, 55),
  Total_Dupl1 = c(226, 148, 118, 176, 100),
  V_Dupl2 = c(108, 122, 91, 95, 62),
  M_Dupl2 = c(22, 35, 38, 46, 68),
  Total_Dupl2 = c(130, 157, 129, 141, 130)
)

df2 <- data.frame(
  Tratamento = c('Controle', 'Veículo', '1,25%', '10%', '25%'),
  V_Dupl1 = c(137, 59, 85, 74, 73),
  M_Dupl1 = c(45, 23, 52, 53, 59),
  Total_Dupl1 = c(182, 82, 137, 127, 132),
  V_Dupl2 = c(63, 59, 97, 102, 79),
  M_Dupl2 = c(31, 23, 44, 55, 41),
  Total_Dupl2 = c(94, 82, 141, 157, 120)
)

df3 <- data.frame(
  Tratamento = c('Controle', 'Veículo', '1,25%', '10%', '25%'),
  V_Dupl1 = c(58, 46, 69, 33, 9),
  M_Dupl1 = c(16, 17, 23, 19, 30),
  Total_Dupl1 = c(74, 63, 92, 52, 39),
  V_Dupl2 = c(5, 66, 71, 30, 6),
  M_Dupl2 = c(3, 16, 27, 21, 36),
  Total_Dupl2 = c(8, 82, 98, 51, 42)
)
```

▼ Code

```
calcular_proporcao_mortas <- function(df) {
  total_mortas <- sum(df$M_Dupl1, df$M_Dupl2)
  total_celulas <- sum(df$Total_Dupl1, df$Total_Dupl2)
  return(total_mortas / total_celulas)
}

# Proporcao de celularsmortas
prop_2404 <- calcular_proporcao_mortas(df1)
prop_3004 <- calcular_proporcao_mortas(df2)
prop_0805 <- calcular_proporcao_mortas(df3)

cat("Proporções recalculadas CORRETAMENTE:\n")
```

Proporções recalculadas CORRETAMENTE:

▼ Code

```
cat(sprintf("24/04: %.4f\n", prop_2404))
```

24/04: 0.2811

▼ Code

```
cat(sprintf("30/04: %.4f\n", prop_3004))
```

30/04: 0.3397

▼ Code

```
cat(sprintf("08/05: %.4f\n\n", prop_0805))
```

08/05: 0.3461

▼ Code

```
tabela <- matrix(c(
  sum(df1$M_Dupl1, df1$M_Dupl2), sum(df1$V_Dupl1, df1$V_Dupl2),
  sum(df2$M_Dupl1, df2$M_Dupl2), sum(df2$V_Dupl1, df2$V_Dupl2),
  sum(df3$M_Dupl1, df3$M_Dupl2), sum(df3$V_Dupl1, df3$V_Dupl2)
), nrow = 3, byrow = TRUE)

# Teste Qui-Quadrado
teste_chi2 <- chisq.test(tabela)
cat(sprintf("Teste Qui-Quadrado: p = %.4f\n\n", teste_chi2$p.value))
```

Teste Qui-Quadrado: p = 0.0009

2.0.3 Interpretação

Nosso o valor p está muito abaixo do nosso limite pré-determinado.

- $p = 0.0009 < 0.05$, ou seja, existe uma diferença significativa na proporção de células mortas entre os dias amostrados

2.0.4 Mas quais dias tiveram mais discrepância?

Para responder a essa pergunta vamos aplicar:

3 Post Hoc

Teste Z com Bonferroni:

- *Por que Z?* Porque estamos comparando proporções (mortas/total).
- *Por que Bonferroni?* Para controlar o erro acumulado ao fazer múltiplas comparações.

3.0.1 Como ele se enquadra?

- Vai comparar cada par de dias (24/04 vs 30/04, 24/04 vs 08/05, 30/04 vs 08/05)

- Usa correção de Bonferroni ($\alpha = 0.05/3$ (cada dia) = 0.0167)
- Mostra o aumento percentual na proporção de células mortas

▼ Code

```
tabela_2404_vs_3004 <- matrix(c(
  sum(df1$M_Dupl1, df1$M_Dupl2), sum(df1$V_Dupl1, df1$V_Dupl2),
  sum(df2$M_Dupl1, df2$M_Dupl2), sum(df2$V_Dupl1, df2$V_Dupl2)
), nrow = 2, byrow = TRUE)

tabela_2404_vs_0805 <- matrix(c(
  sum(df1$M_Dupl1, df1$M_Dupl2), sum(df1$V_Dupl1, df1$V_Dupl2),
  sum(df3$M_Dupl1, df3$M_Dupl2), sum(df3$V_Dupl1, df3$V_Dupl2)
), nrow = 2, byrow = TRUE)

tabela_3004_vs_0805 <- matrix(c(
  sum(df2$M_Dupl1, df2$M_Dupl2), sum(df2$V_Dupl1, df2$V_Dupl2),
  sum(df3$M_Dupl1, df3$M_Dupl2), sum(df3$V_Dupl1, df3$V_Dupl2)
), nrow = 2, byrow = TRUE)
```

▼ Code

```
# Realizar os testes
teste_2404_vs_3004 <- chisq.test(tabela_2404_vs_3004)
teste_2404_vs_0805 <- chisq.test(tabela_2404_vs_0805)
teste_3004_vs_0805 <- chisq.test(tabela_3004_vs_0805)

# Coletar os p-valores
p_values <- c(
  teste_2404_vs_3004$p.value,
  teste_2404_vs_0805$p.value,
  teste_3004_vs_0805$p.value
)
```

▼ Code

```
p_ajustados <- p.adjust(p_values, method = "bonferroni")

# Nomes das comparações
comparacoes <- c("24/04 vs 30/04", "24/04 vs 08/05", "30/04 vs 08/05")

# Mostrar resultados
resultados <- data.frame(
  Comparacao = comparacoes,
  p_value = p_values,
  p_ajustado = p_ajustados,
  Significativo = ifelse(p_ajustados < 0.05, "Sim", "Não")
)

print(resultados)
```

	Comparacao	p_value	p_ajustado	Significativo
1	24/04 vs 30/04	0.001143957	0.003431871	Sim
2	24/04 vs 08/05	0.004083701	0.012251103	Sim
3	30/04 vs 08/05	0.826884336	1.000000000	Não

3.0.2 Interpretação

- Há um aumento significativo e progressivo na proporção de células mortas ao longo do tempo
- Maior salto de 24/04 - 30/04