MAD211 - Estatística para Administração

Testes para proporções

Prof. Carlos Trucíos
carlos.trucios@facc.ufrj.br
ctruciosm.github.io

Faculdade de Administração e Ciências Contábeis,
Universidade Federal do Rio de Janeiro

Aula 20

Universidade Federal do Rio de Janeiro

Motivação

Teste para a proporção

Diferença de proporções

Motivação

Motivação

- Até agora, temos focado em testes de hipóteses a respeito da:
  - média,
  - diferença de médias,
  - variância.

Motivação

- Até agora, temos focado em testes de hipóteses a respeito da:
  - média,
  - diferença de médias,
  - variância.
- Contudo, existe mais um parâmetro populacional que é de bastante interesse: **a proporção**.

Motivação

- Até agora, temos focado em testes de hipóteses a respeito da:
  - média,
  - diferença de médias,
  - variância.
- Contudo, existe mais um parâmetro populacional que é de bastante interesse: **a proporção**.
- Hoje, focaremos em testes de hipóteses para a proporção populacional e para a diferença de proporções.

Motivação

- Até agora, temos focado em testes de hipóteses a respeito da:
  - média,
  - diferença de médias,
  - variância.
- Contudo, existe mais um parâmetro populacional que é de bastante interesse: **a proporção**.
- Hoje, focaremos em testes de hipóteses para a proporção populacional e para a diferença de proporções.
- Estes testes são úteis, por exemplo, para:
  - fazer inferência sobre a intenção de voto,
  - fazer inferência sobre a occorencia de determinadas características na população (proporção de pessoas a favor ou contra uma determinada politica de governo, proporção de vacinados, etc).

Teste para a proporção

Teste para a proporção

Seja p_0 o valor hipotético da proporção populacional. Podemos estar interessados em algum dos seguintes testes:

$$H_0 : p = p_0 \quad vs. \quad H_1 : p \neq p_0,$$

$$H_0 : p \leq p_0 \quad vs. \quad H_1 : p > p_0,$$

$$H_0 : p \geq p_0 \quad vs. \quad H_1 : p < p_0,$$

Teste para a proporção

Seja p_0 o valor hipotético da proporção populacional. Podemos estar interessados em algum dos seguintes testes:

$$H_0 : p = p_0 \quad vs. \quad H_1 : p \neq p_0,$$
$$H_0 : p \leq p_0 \quad vs. \quad H_1 : p > p_0,$$
$$H_0 : p \geq p_0 \quad vs. \quad H_1 : p < p_0,$$

Assim como no caso da média populacional μ, precisaremos também uma estatística de teste e uma regra de decisão para sabermos se rejeitar ou não rejeitar H_0.

Teste para a proporção

A forma de construir a estatística de teste é muito semelhante ao caso da média, mas, substituiremos $\bar{x}$ pela proporção amostral e subtituiremos o desvio padrão da média pelo desvio padrão da proporção.

Teste para a proporção

A forma de construir a estatística de teste é muito semelhante ao caso da média, mas, substituiremos $\bar{x}$ pela proporção amostral e subtituiremos o desvio padrão da média pelo desvio padrão da proporção.

Sejam $X_1, \ldots, X_n$ variáveis aleatórias Bernoulli, com probabilidade de sucesso p. Pelo Teorema Central do Limite, sabemos que, quando n for grande:

$$\frac{\bar{X} - p}{\sqrt{p(1-p)/n}} \sim N(0,1).$$

Teste para a proporção

A forma de construir a estatística de teste é muito semelhante ao caso da média, mas, substituiremos $\bar{x}$ pela proporção amostral e subtituiremos o desvio padrão da média pelo desvio padrão da proporção.

Sejam $X_1, \ldots, X_n$ variáveis aleatórias Bernoulli, com probabilidade de sucesso p. Pelo Teorema Central do Limite, sabemos que, quando n for grande:

$$\frac{\bar{X} - p}{\sqrt{p(1-p)/n}} \sim N(0, 1).$$

No caso da Bernoulli, $\bar{X}$ tranforma-se em $\hat{p}$ (a proporção estimada), pois os Xs são apenas zeros e uns. Alguns livros denotam $\hat{p}$ por $\bar{p}$.

Teste para a proporção

A forma de construir a estatística de teste é muito semelhante ao caso da média, mas, substituiremos $\bar{x}$ pela proporção amostral e subtituiremos o desvio padrão da média pelo desvio padrão da proporção.

Sejam $X_1, \ldots, X_n$ variáveis aleatórias Bernoulli, com probabilidade de sucesso p. Pelo Teorema Central do Limite, sabemos que, quando n for grande:

$$\frac{\bar{X} - p}{\sqrt{p(1-p)/n}} \sim N(0,1).$$

No caso da Bernoulli, $\bar{X}$ tranforma-se em $\hat{p}$ (a proporção estimada), pois os Xs são apenas zeros e uns. Alguns livros denotam $\hat{p}$ por $\bar{p}$.

Sob H_0, ou seja assumindo que H_0 é verdade, temos que

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1)$$

Teste para a proporção

Com nossa estatística de teste definida

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1),$$

nossa regra de decisão é dada por:

- Se $H_0 : p = p_0 \quad vs \quad H_1 : p \neq p_0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$

Teste para a proporção

Com nossa estatística de teste definida

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} \sim N(0, 1),$$

nossa regra de decisão é dada por:

- Se $H_0 : p = p_0 \quad vs \quad H_1 : p \neq p_0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$
- Se $H_0 : p \leq p_0 \quad vs \quad H_1 : p > p_0$, rejeitamos H_0 se $z > z_{1-\alpha}$

Teste para a proporção

Com nossa estatística de teste definida

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1),$$

nossa regra de decisão é dada por:

- Se $H_0 : p = p_0 \quad vs \quad H_1 : p \neq p_0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$
- Se $H_0 : p \leq p_0 \quad vs \quad H_1 : p > p_0$, rejeitamos H_0 se $z > z_{1-\alpha}$
- Se $H_0 : p \geq p_0 \quad vs \quad H_1 : p < p_0$, rejeitamos H_0 se $z < z_{\alpha}$

Teste para a proporção

Com nossa estatística de teste definida

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1),$$

nossa regra de decisão é dada por:

- Se $H_0 : p = p_0 \quad vs \quad H_1 : p \neq p_0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$
- Se $H_0 : p \leq p_0 \quad vs \quad H_1 : p > p_0$, rejeitamos H_0 se $z > z_{1-\alpha}$
- Se $H_0 : p \geq p_0 \quad vs \quad H_1 : p < p_0$, rejeitamos H_0 se $z < z_{\alpha}$

Teste para a proporção

Com nossa estatística de teste definida

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} \sim N(0,1),$$

nossa regra de decisão é dada por:

- Se $H_0 : p = p_0 \quad vs \quad H_1 : p \neq p_0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$
- Se $H_0 : p \leq p_0 \quad vs \quad H_1 : p > p_0$, rejeitamos H_0 se $z > z_{1-\alpha}$
- Se $H_0 : p \geq p_0 \quad vs \quad H_1 : p < p_0$, rejeitamos H_0 se $z < z_{\alpha}$

Se utilizarmos o computador e tivermos o $p - valor$, rejeitamos H_0 se $p - valor < alpha$.

Teste para a proporção: Exemplos

1. Sejam as hipóteses $H_0 : p = 0.2 \quad vs. \quad H_1 : p \neq 0.2$. Sabendo que uma amostra de tamanho 400 produziu uma proporção amostral $\hat{p} = 0.175$, rejeitamos ou não rejeitamos H_0?

Teste para a proporção: Exemplos

1. Sejam as hipóteses $H_0: p = 0.2 \quad vs. \quad H_1: p \neq 0.2$. Sabendo que uma amostra de tamanho 400 produziu uma proporção amostral $\hat{p} = 0.175$, rejeitamos ou não rejeitamos H_0?

Solução:

- Calculamos a estatística de teste,

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.175 - 0.2}{\sqrt{0.2(1-0.2)/400}} = -1.25$$

```
alpha = 0.05
qnorm(1 - alpha/2)
```

```
## [1] 1.959964
```

Teste para a proporção: Exemplos

1. Sejam as hipóteses $H_0 : p = 0.2 \quad vs. \quad H_1 : p \neq 0.2$. Sabendo que uma amostra de tamanho 400 produziu uma proporção amostral $\hat{p} = 0.175$, rejeitamos ou não rejeitamos H_0?

Solução:

- Calculamos a estatística de teste,

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{0.175 - 0.2}{\sqrt{0.2(1 - 0.2)/400}} = -1.25$$

- Calculamos o quantil apropriado, $z_{1-\alpha/2}$

```
alpha = 0.05
qnorm(1 - alpha/2)
```

```
## [1] 1.959964
```

Teste para a proporção: Exemplos

1. Sejam as hipóteses $H_0 : p = 0.2 \quad vs. \quad H_1 : p \neq 0.2$. Sabendo que uma amostra de tamanho 400 produziu uma proporção amostral $\hat{p} = 0.175$, rejeitamos ou não rejeitamos H_0?

Solução:

- Calculamos a estatística de teste,

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.175 - 0.2}{\sqrt{0.2(1-0.2)/400}} = -1.25$$

- Calculamos o quantil apropriado, $z_{1-\alpha/2}$

```
alpha = 0.05
qnorm(1 - alpha/2)
```

```
## [1] 1.959964
```

- $|z| = 1.25 > 1.959964$? **Não**, então não rejeitamos H_0

Teste para a proporção: R

A função `prop.test()` nos ajuda a fazer o teste para a proporção.

```
# x: número de sucessos
# n: número de observações na amostra
# p: nosso "p0"
# correct = FALSE
# conf.level = 1 - alpha
# alternative: "two.sided", "less" ou "greater"
prop.test(x, n, p, correct = FALSE, conf.level, alternative)
```

Teste para a proporção: R

A função `prop.test()` nos ajuda a fazer o teste para a proporção.

```
# x: número de sucessos
# n: número de observações na amostra
# p: nosso "p0"
# correct = FALSE
# conf.level = 1 - alpha
# alternative: "two.sided", "less" ou "greater"
prop.test(x, n, p, correct = FALSE, conf.level, alternative)
```

No exemplo anterior temos:

- $H_0 : p = 0.2 \quad vs. \quad H_1 : p \neq 0.2.$
- $\alpha = 0.05$, $\hat{p} = 0.175$, $n = 400$, ou seja $x = \hat{p} \times n = 50$
- `alternative = "two.sided"`

Teste para a proporção: Exemplos

```
prop.test(x = 0.175*400, n = 400, p = 0.2, correct = FALSE,
          conf.level = 1 - 0.05, alternative = "two.sided")
```

```
##
##  1-sample proportions test without continuity correction
##
## data:  0.175 * 400 out of 400, null probability 0.2
## X-squared = 1.5625, df = 1, p-value = 0.2113
## alternative hypothesis: true p is not equal to 0.2
## 95 percent confidence interval:
##  0.1409042 0.2152788
## sample estimates:
##     p
## 0.175
```

Como `p-value` $> \alpha$, não rejeitamos H_0. (**Obs:** `X-squared` $= z^2$).

Teste para a proporção: Exemplos

2. O Departamento de Administração da FACC/UFRJ, gostaria saber se a maioria dos alunos está satisfeito com a forma como o ensino remoto está acontecendo. Uma amostra de 200 alunos é escolhida de forma aleatória e 120 dos alunos selecionados mencionaram estarem satisfeitos. Defina a hipótese apropriada e faça o respectivo teste de hipóteses.

Teste para a proporção: Exemplos

2. O Departamento de Administração da FACC/UFRJ, gostaria saber se a maioria dos alunos está satisfeito com a forma como o ensino remoto está acontecendo. Uma amostra de 200 alunos é escolhida de forma aleatória e 120 dos alunos selecionados mencionaram estarem satisfeitos. Defina a hipótese apropriada e faça o respectivo teste de hipóteses.

Solução:

- Definimos as hipóteses: $H_0 : p \leq 0.5 \quad vs \quad H_1 : p > 0.5$

Teste para a proporção: Exemplos

2. O Departamento de Administração da FACC/UFRJ, gostaria saber se a maioria dos alunos está satisfeito com a forma como o ensino remoto está acontecendo. Uma amostra de 200 alunos é escolhida de forma aleatória e 120 dos alunos selecionados mencionaram estarem satisfeitos. Defina a hipótese apropriada e faça o respectivo teste de hipóteses.

Solução:

- Definimos as hipóteses: $H_0 : p \leq 0.5 \quad vs \quad H_1 : p > 0.5$
- Calculamos a estatística de teste:

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{0.6 - 0.5}{\sqrt{0.5(1 - 0.5)/200}} = 2.828427$$

Teste para a proporção: Exemplos

2. O Departamento de Administração da FACC/UFRJ, gostaria saber se a maioria dos alunos está satisfeito com a forma como o ensino remoto está acontecendo. Uma amostra de 200 alunos é escolhida de forma aleatória e 120 dos alunos selecionados mencionaram estarem satisfeitos. Defina a hipótese apropriada e faça o respectivo teste de hipóteses.

Solução:

- Definimos as hipóteses: $H_0 : p \leq 0.5 \quad vs \quad H_1 : p > 0.5$
- Calculamos a estatística de teste:
$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}} = \frac{0.6 - 0.5}{\sqrt{0.5(1 - 0.5)/200}} = 2.828427$$
- Calculamos o Quantil apropriado, $z_{1-\alpha} = \texttt{qnorm(1 - 0.05)} = 1.64$

Teste para a proporção: Exemplos

2. O Departamento de Administração da FACC/UFRJ, gostaria saber se a maioria dos alunos está satisfeito com a forma como o ensino remoto está acontecendo. Uma amostra de 200 alunos é escolhida de forma aleatória e 120 dos alunos selecionados mencionaram estarem satisfeitos. Defina a hipótese apropriada e faça o respectivo teste de hipóteses.

Solução:

- Definimos as hipóteses: $H_0 : p \leq 0.5 \quad vs \quad H_1 : p > 0.5$
- Calculamos a estatística de teste:
$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}} = \frac{0.6 - 0.5}{\sqrt{0.5(1-0.5)/200}} = 2.828427$$
- Calculamos o Quantil apropriado, $z_{1-\alpha} =$ `qnorm(1 - 0.05)` $= 1.64$
- $2.828427 > 1.64$? **Sim**, então rejeitamos H_0.

Teste para a proporção: Exemplos

```r
prop.test(x = 120, n = 200, p = 0.5, correct = FALSE,
          conf.level = 1 - 0.05, alternative = "greater")
```

```
##
##  1-sample proportions test without continuity correction
##
## data:  120 out of 200, null probability 0.5
## X-squared = 8, df = 1, p-value = 0.002339
## alternative hypothesis: true p is greater than 0.5
## 95 percent confidence interval:
##  0.5420517 1.0000000
## sample estimates:
##   p
## 0.6
```

Como `p-value` $< \alpha$, rejeitamos H_0. (**Obs:** `X-squared` $= z^2$).

Diferença de proporções

Diferença de proporções

- Assim como existem casos onde estamos interessados em fazer inferência para a diferença de médias, também podemos estar interessados em fazer inferência para a diferença de proporções.

Diferença de proporções

- Assim como existem casos onde estamos interessados em fazer inferência para a diferença de médias, também podemos estar interessados em fazer inferência para a diferença de proporções.
- Com isto, podemos saber se, por exemplo:
  - a intenção de voto, por um determinado candidato, em dois estados diferentes é a mesma o não,
  - a proporção de desempregados,
  - em um teste de seleção, ver se a proporão de erros de digitação do candidato A é menor do que a proporção de erros de digitação do candidato B,
  - etc

Diferença de proporções

- Assim como existem casos onde estamos interessados em fazer inferência para a diferença de médias, também podemos estar interessados em fazer inferência para a diferença de proporções.
- Com isto, podemos saber se, por exemplo:
  - a intenção de voto, por um determinado candidato, em dois estados diferentes é a mesma o não,
  - a proporção de desempregados,
  - em um teste de seleção, ver se a proporão de erros de digitação do candidato A é menor do que a proporção de erros de digitação do candidato B,
  - etc
- O teste funciona de forma semelhante ao caso da comparação de médias, mas agora precisaremos das proporções estimadas e do desvio padrão da diferença de proporções.

Diferença de proporções

Podemos estar interessados em algum dos seguintes testes:

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0,$$
$$H_0 : p_1 - p_2 \leq 0 \quad vs. \quad H_1 : p_1 - p_2 > 0,$$
$$H_0 : p_1 - p_2 \geq 0 \quad vs. \quad H_1 : p_1 - p_2 < 0,$$

Diferença de proporções

Podemos estar interessados em algum dos seguintes testes:

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0,$$
$$H_0 : p_1 - p_2 \leq 0 \quad vs. \quad H_1 : p_1 - p_2 > 0,$$
$$H_0 : p_1 - p_2 \geq 0 \quad vs. \quad H_1 : p_1 - p_2 < 0,$$

Precisaremos de uma estatística de teste e uma regra de decisão para sabermos se rejeitar ou não rejeitar H_0.

Diferença de proporções

Do teste para proporções sabemos que, se $X_1, \ldots, X_{n_1}$ são variáveis aleatórias Bernoulli com probabilidade de sucesso p_1,

$$\frac{\hat{p}_1 - p_1}{\sqrt{p_1(1-p_1)/n_1}} \sim N(0,1).$$

Diferença de proporções

Do teste para proporções sabemos que, se $X_1, \dots, X_{n_1}$ são variáveis aleatórias Bernoulli com probabilidade de sucesso p_1,

$$\frac{\hat{p}_1 - p_1}{\sqrt{p_1(1-p_1)/n_1}} \sim N(0,1).$$

Por outro lado, se $Y_1, \dots, Y_{n_2}$ são variáveis aleatórias Bernoulli com probabilidade de sucesso p_2,

$$\frac{\hat{p}_2 - p_2}{\sqrt{p_2(1-p_2)/n_2}} \sim N(0,1).$$

Diferença de proporções

Utilizando uma propriedade da distribuição normal, temos que

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}} \sim N(0,1).$$

Diferença de proporções

Utilizando uma propriedade da distribuição normal, temos que

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}} \sim N(0,1).$$

Sob H_0, ou seja assumindo que H_0 é verdade, temos que $p_1 = p_2$ e podemos fazer $p_1 = p_2 = p$. Assim, temos que

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - \overbrace{(p-p)}^{0}}{\sqrt{p(1-p)/n_1 + p(1-p)/n_2}} \sim N(0,1).$$

Diferença de proporções

Utilizando uma propriedade da distribuição normal, temos que

$$\frac{(\hat{p}_1 - \hat{p}_2) - (p_1 - p_2)}{\sqrt{p_1(1-p_1)/n_1 + p_2(1-p_2)/n_2}} \sim N(0,1).$$

Sob H_0, ou seja assumindo que H_0 é verdade, temos que $p_1 = p_2$ e podemos fazer $p_1 = p_2 = p$. Assim, temos que

$$z = \frac{(\hat{p}_1 - \hat{p}_2) - \overbrace{(p - p)}^{0}}{\sqrt{p(1-p)/n_1 + p(1-p)/n_2}} \sim N(0,1).$$

Como p é desconhecido, podemos substituir p por $\hat{p} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$

Diferença de proporções

Em amostras grandes, temos que

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}} \sim N(0,1).$$

n ser grande faz com que substituir p por $\hat{p}$ não mude a distribuição da estatística de teste.

Diferença de proporções

Em amostras grandes, temos que

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}} \sim N(0,1).$$

n ser grande faz com que substituir p por $\hat{p}$ não mude a distribuição da estatística de teste.

Assim, com a estatística de teste definida, a regrade decisão é dada por:

- Se $H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$

Diferença de proporções

Em amostras grandes, temos que

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}} \sim N(0,1).$$

n ser grande faz com que substituir p por $\hat{p}$ não mude a distribuição da estatística de teste.

Assim, com a estatística de teste definida, a regrade decisão é dada por:

- Se $H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$
- Se $H_0 : p_1 - p_2 \leq 0 \quad vs. \quad H_1 : p_1 - p_2 > 0$, rejeitamos H_0 se $z > z_{1-\alpha}$

Diferença de proporções

Em amostras grandes, temos que

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}} \sim N(0,1).$$

n ser grande faz com que substituir p por $\hat{p}$ não mude a distribuição da estatística de teste.

Assim, com a estatística de teste definida, a regrade decisão é dada por:

- Se $H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$, rejeitamos H_0 se $|z| > z_{1-\alpha/2}$
- Se $H_0 : p_1 - p_2 \leq 0 \quad vs. \quad H_1 : p_1 - p_2 > 0$, rejeitamos H_0 se $z > z_{1-\alpha}$
- Se $H_0 : p_1 - p_2 \geq 0 \quad vs. \quad H_1 : p_1 - p_2 < 0$, rejeitamos H_0 se $z < z_\alpha$

Diferença de proporções: Exemplos

1. Com o objetivo de saber a opinião dos alunos dos cursos de Administração e Contabilidade da FACC, uma amostra de 400 alunos do curso de Administração e outra amostra de 300 alunos do curso de Contabilidade são escolhidas de forma aleatória. Em cada uma das amostras, a proporção de alunos que está satisfeito com o ensino remoto é a seguinte: $\hat{p}_{adm} = 0.48$ e $\hat{p}_{cont} = 0.5$. Com os ados fornecidos, podemos dizer que a proporção de alunos que está satisfeito com o ensino remoto é o mesmo em ambos os cursos?

Diferença de proporções: Exemplos

1. Com o objetivo de saber a opinião dos alunos dos cursos de Administração e Contabilidade da FACC, uma amostra de 400 alunos do curso de Administração e outra amostra de 300 alunos do curso de Contabilidade são escolhidas de forma aleatória. Em cada uma das amostras, a proporção de alunos que está satisfeito com o ensino remoto é a seguinte: $\hat{p}_{adm} = 0.48$ e $\hat{p}_{cont} = 0.5$. Com os ados fornecidos, podemos dizer que a proporção de alunos que está satisfeito com o ensino remoto é o mesmo em ambos os cursos?

Solução:

- Definimos as hipóteses:

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

Diferença de proporções: Exemplos

1. Com o objetivo de saber a opinião dos alunos dos cursos de Administração e Contabilidade da FACC, uma amostra de 400 alunos do curso de Administração e outra amostra de 300 alunos do curso de Contabilidade são escolhidas de forma aleatória. Em cada uma das amostras, a proporção de alunos que está satisfeito com o ensino remoto é a seguinte: $\hat{p}_{adm} = 0.48$ e $\hat{p}_{cont} = 0.5$. Com os ados fornecidos, podemos dizer que a proporção de alunos que está satisfeito com o ensino remoto é o mesmo em ambos os cursos?

Solução:

- Definimos as hipóteses:

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

- Anotamos as informações: $n_1 = 400$, $\hat{p}_1 = 0.48$, $n_2 = 300$, $\hat{p}_2 = 0.5$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

- $n_1 = 400$, $\hat{p}_1 = 0.48$, $n_2 = 300$, $\hat{p}_2 = 0.5$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

- $n_1 = 400$, $\hat{p}_1 = 0.48$, $n_2 = 300$, $\hat{p}_2 = 0.5$
- Calculamos a estatística de teste

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}},$$

em que $\hat{p} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \dfrac{400 \times 0.48 + 300 \times 0.5}{400 + 300} \approx 0.49$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

- $n_1 = 400,\ \hat{p}_1 = 0.48,\ n_2 = 300,\ \hat{p}_2 = 0.5$
- Calculamos a estatística de teste

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}},$$

em que $\hat{p} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \dfrac{400 \times 0.48 + 300 \times 0.5}{400 + 300} \approx 0.49$

$$z = \frac{0.48 - 0.50}{\sqrt{0.49(1-0.49)/400 + 0.49(1-0.49)/300}} \approx -0.52$$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}} \approx -0.52$$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}} \approx -0.52$$

Calculamos o quantil apropriado, $z_{1-\alpha/2} =$ `qnorm(1 - 0.05/2)` $= 1.96$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 = 0 \quad vs. \quad H_1 : p_1 - p_2 \neq 0$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}} \approx -0.52$$

Calculamos o quantil apropriado, $z_{1-\alpha/2} = \texttt{qnorm(1 - 0.05/2)} = 1.96$

$|\mathrm{z}| = 0.52 > 1.96$? **Não**, então não rejeitamos H_0

Diferença de proporções: Exemplos

```
# x: vetor bidimensional com o número de sucessos
# n: vetor bidimensional com o número de observações nas amostra
# p: nao colocar nada
# correct = FALSE
# conf.level = 1 - alpha
# alternative: "two.sided", "less" ou "greater"
prop.test(x, n, p, correct = FALSE, conf.level, alternative)
```

Diferença de proporções: Exemplos

```
prop.test(correct = FALSE, conf.level = 1 - 0.05,
x = c(192, 150), n = c(400, 300), alternative = "two.sided")
```

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(192, 150) out of c(400, 300)
## X-squared = 0.27443, df = 1, p-value = 0.6004
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.09482169  0.05482169
## sample estimates:
## prop 1 prop 2
##   0.48   0.50
```

Como p-value $> \alpha$, não rejeitamos H_0. (**Obs:** X-squared $= z^2$).

Diferença de proporções: Exemplos

2. A corretora `XP` está definindo uma nova estratégia de captação de clientes. O CEO suspeita que a mesma estratégia funciona pior para pessoas com idades `abaixo dos 35 anos` (Grupo 1) do que em pessoas com idades de `35 anos a mais` (Grupo 2). Verificou-se que, das 100 pessoas contatadas com idades abaixo dos 35 anos, 49 delas viraram clientes da corretora. Por outro lado, das 150 pessoas contatadas com idade de 35 ou mais anos, 60 delas viraram clientes da corretora. Assumindo que as pessoas contatadas foram selecionadas de forma aleatória, poderiamos dizer que as suspeitas do CEO são verificadas?

Diferença de proporções: Exemplos

2. A corretora `XP` está definindo uma nova estratégia de captação de clientes. O CEO suspeita que a mesma estratégia funciona pior para pessoas com idades `abaixo dos 35 anos` (Grupo 1) do que em pessoas com idades de `35 anos a mais` (Grupo 2). Verificou-se que, das 100 pessoas contatadas com idades abaixo dos 35 anos, 49 delas viraram clientes da corretora. Por outro lado, das 150 pessoas contatadas com idade de 35 ou mais anos, 60 delas viraram clientes da corretora. Assumindo que as pessoas contatadas foram selecionadas de forma aleatória, poderiamos dizer que as suspeitas do CEO são verificadas?

Solução:

- $H_0 : p_1 \geq p_2 \quad vs \quad H_1 : p_1 < p_2$

Diferença de proporções: Exemplos

2. A corretora `XP` está definindo uma nova estratégia de captação de clientes. O CEO suspeita que a mesma estratégia funciona pior para pessoas com idades `abaixo dos 35 anos` (Grupo 1) do que em pessoas com idades de `35 anos a mais` (Grupo 2). Verificou-se que, das 100 pessoas contatadas com idades abaixo dos 35 anos, 49 delas viraram clientes da corretora. Por outro lado, das 150 pessoas contatadas com idade de 35 ou mais anos, 60 delas viraram clientes da corretora. Assumindo que as pessoas contatadas foram selecionadas de forma aleatória, poderiamos dizer que as suspeitas do CEO são verificadas?

Solução:

- $H_0 : p_1 \geq p_2 \quad vs \quad H_1 : p_1 < p_2$
- $n_1 = 100$, $\hat{p}_1 = 49/100 = 0.49$, $n_2 = 150$, $\hat{p}_2 = 60/150 = 0.4$

Diferença de proporções: Exemplos

$$H_0 : p_1 \geq p_2 \quad vs \quad H_1 : p_1 < p_2$$

Diferença de proporções: Exemplos

$$H_0 : p_1 \geq p_2 \quad vs \quad H_1 : p_1 < p_2$$

é equivalente a

$$H_0 : p_1 - p_2 \geq 0 \quad vs \quad H_1 : p_1 - p_2 < 0$$

Diferença de proporções: Exemplos

$$H_0 : p_1 \geq p_2 \quad vs \quad H_1 : p_1 < p_2$$

é equivalente a

$$H_0 : p_1 - p_2 \geq 0 \quad vs \quad H_1 : p_1 - p_2 < 0$$

Calculando a estatística de teste:

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})/n_1 + \hat{p}(1-\hat{p})/n_2}},$$

em que $\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{100 \times 0.49 + 150 \times 0.4}{100 + 150} \approx 0.49 = 0.436$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 \geq 0 \quad vs \quad H_1 : p_1 - p_2 < 0$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} = \frac{0.49 - 0.4}{\sqrt{\frac{0.436(1-0.436)}{100} + \frac{0.436(1-0.436)}{150}}} \approx 1.4$$

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 \geq 0 \quad vs \quad H_1 : p_1 - p_2 < 0$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} = \frac{0.49 - 0.4}{\sqrt{\frac{0.436(1-0.436)}{100} + \frac{0.436(1-0.436)}{150}}} \approx 1.4$$

Calculamos o Quantil apropriado, $z_\alpha = \texttt{qnorm(0.05)} = -1.64$.

Diferença de proporções: Exemplos

$$H_0 : p_1 - p_2 \geq 0 \quad vs \quad H_1 : p_1 - p_2 < 0$$

$$z = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2}}} = \frac{0.49 - 0.4}{\sqrt{\frac{0.436(1-0.436)}{100} + \frac{0.436(1-0.436)}{150}}} \approx 1.4$$

Calculamos o Quantil apropriado, $z_\alpha = \texttt{qnorm(0.05)} = -1.64$.

z $= 1.4 <$ -1.64? **Não**, então não rejeitamos H_0.

Diferença de proporções: Exemplos

```r
prop.test(correct = FALSE, conf.level = 1 - 0.05,
x = c(49, 60), n = c(100, 150), alternative = "less")
```

```
##
##  2-sample test for equality of proportions without continuity
##  correction
##
## data:  c(49, 60) out of c(100, 150)
## X-squared = 1.9764, df = 1, p-value = 0.9201
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.0000000  0.1953092
## sample estimates:
## prop 1 prop 2
##   0.49   0.40
```

Como `p-value` $> \alpha$, não rejeitamos H_0. (**Obs:** `X-squared` $= z^2$).

Leituras recomendadas

- Anderson, D. R; Sweeney, D. J.; e Williams, T. A. (2008). *Estatística Aplicada à Administração e Economia*. 2ed. Cengage Learning. **Cap 9.5**
- Anderson, D. R; Sweeney, D. J.; e Williams, T. A. (2008). *Estatística Aplicada à Administração e Economia*. 2ed. Cengage Learning. **Cap 11.1**