asd

asd

Paulo Ricardo Seganfredo Campana

8 de junho de 2023

Capítulo 10

Exemplos

Exemplo 10.1, 10.2, 10.4, 10.5 e 10.7: Está correta a afirmação do fabricante?

Consideremos uma companhia que produz cabos náuticos de determinado tipo. O fabricante garante que a carga de ruptura média de seu produto é de pelo menos 90 kg. Um potencial consumidor, interessado na compra de grande quantidade do produto, decide fazer ensaios de carga de ruptura com 20 espécimes, obtendo uma média amostral de 88.4 kg, supondo que o desvio padrão populacional da carga de ruptura seja conhecido e igual a 10 kg.

```
# H0: μ >= 90
# H1: μ <  90
testar(
    c(m = 88.4, v = 10^2, n = 20),
    tetha0 = 90,
    região_crítica = "menor"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
-0.716	0.237	[-Inf, 92.078]	[86.322, Inf]	menor	z

Supondo sigma desconhecido, S = 10

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
-0.716	0.241	[-Inf, 92.266]	[86.134, Inf]	menor	t

Não rejeita a hipótese nula de que a carga de ruptura média de seu produto é de pelo menos 90 kg, pois o p-valor de 0.237 é maior que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 10.9

Suponha agora que o verdadeiro valor da média populacional é 89 kg, ou seja, um pouco inferior ao valor da carga média de ruptura afirmada pelo fabricante. Para calcularmos o poder do teste.

```
# poder = P(X < xcrit)
pnorm(86.322, 89, sqrt(10^2 / 20))
# [1] 0.115529
```

Exemplo 10.3: Baterias originais ou falsificadas?

A duração da carga das baterias recarregáveis de notebooks de uma certa marca pode ser encarada como uma variável aleatória com distribuição Normal de média 180 minutos e desvio padrão de 40 minutos. Existe uma falsificação do mesmo produto quase perfeita, mas nesse caso a duração da carga da bateria é uma variável aleatória com distribuição Normal de média 150 minutos e o mesmo desvio padrão anterior. Um montador de notebooks recebe 25 baterias dessa marca. Entretanto, ele tem dúvidas quanto à procedência das baterias. Como ele poderia decidir se as baterias são originais ou falsificadas?
Suponha que após submeter à prova as 25 baterias foi encontrada, para a duração de suas cargas, uma média amostral $\bar{X}_{\text{obs}} = 167.4$ min. Qual critério de decisão deve ser adotado se queremos fixar o valor de α em 0.05?

```r
# H0: μ = 180
# H1: μ = 150
testar(
    c(m = 167.4, v = 40^2, n = 25),
    tetha0 = 180,
    região_crítica = "menor"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
-1.575	0.058	[-Inf, 180.559]	[166.841, Inf]	menor	z

Não rejeita a hipótese nula de que a duração da carga das baterias de notebooks tem média de 180 minutos, pois o p-valor de 0.058 é maior que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 10.6: Aumentando a pureza de um produto químico

Um produto químico tem seu teor de pureza distribuído conforme uma normal com média 0.72 e desvio padrão 0.02. A fim de aumentar a pureza, o produto é submetido a um tratamento. 16 unidades amostrais do produto são selecionadas de forma aleatória e submetidas a esse tratamento. Em seguida, a pureza de cada unidade é determinada obtendo-se, para elas, uma média aritmética de 0.73. Podemos dizer que o tratamento contribuiu para o aumento da pureza?

```
# H0: μ = 0.72
# H0: μ > 0.72
testar(
    c(m = 0.73, v = 0.02^2, n = 16),
    tetha0 = 0.72,
    região_crítica = "maior"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
2	0.023	[0.722, Inf]	[-Inf, 0.728]	maior	z

Podemos rejeitar a hipótese nula de que o produto químico tem teor de pureza igual a 0.72, pois o p-valor = 0.023 é menor que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 10.8

Determinemos o poder do teste para rejeitar corretamente a hipótese nula $H_0 : \mu = 0.72$ quando o verdadeiro teor médio de pureza do produto é $\mu = 0.725$. Usaremos o nível de significância $\alpha = 0.01$.

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
2	0.023	[0.718, Inf]	[-Inf, 0.732]	maior	z

```
# H0: μ = 0.72
# H1: μ = 0.725
# poder = P(X > xcrit)
pnorm(0.732, 0.725, sqrt(0.02^2 / 16), lower.tail = FALSE)
# [1] 0.08075666
```

Exemplo 10.10: Pentes de memória: estão corretas as especificações?

As especificações dos pentes de memória RAM para computadores fabricados pela Companhia Boa Memória indicam que a porcentagem de pentes defeituosos não excede 5%. Uma amostra de 100 desses pentes apresentou 7 defeituosos. Com base nesse resultado, podemos afirmar que as especificações estão incorretas?

a) Qual a decisão a ser tomada, ao nível de significância de 0,01?
b) Qual o nível crítico do teste?
c) Qual o poder do teste, se p = 0,12?

```r
# H0: p <= 0.05
# H1: p >  0.05
testar(
    c(m = 0.07, v = 0.05 * 0.95, n = 100),
    tetha0 = 0.05, alpha = 0.01,
    região_crítica = "maior"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
0.918	0.179	[0.019, Inf]	[-Inf, 0.101]	maior	z

Não rejeita a hipótese nula de que a porcentagem de pentes de memória RAM defeituosos não excedem 5%, pois o p-valor = 0.179 é maior que o ponto de corte tomado como $\alpha = 0.01$.

```r
# poder = P(X > xcrit)
pnorm(0.101, 0.12, sqrt(0.12 * 0.88 / 100), lower.tail = FALSE)
# [1] 0.7206201
```

Exercícios

R 10.1: Velocidade média na estrada

Sergio afirma que Raquel dirige seu carro na estrada a uma velocidade média superior a 100 km/h, enquanto Raquel discorda, afirmando dirigir na estrada a uma velocidade média menor ou igual a 100 km/h. Para dirimir essa controvérsia, Sergio resolve cronometrar o tempo (em minutos) que ela gasta ao volante em 10 viagens, sempre pelo mesmo percurso que liga duas cidades 120 km distantes: 73, 68, 73, 61, 70, 78, 63, 64, 74, 62.

a) Quem parece ter razão, ao nível de significância de 5%?
b) Qual o nível crítico?

```r
# H0: μ <= 100 km/h
# H1: μ >  100 km/h
R101 <- c(73, 68, 73, 61, 70, 78, 63, 64, 74, 62)
R101 <- 120 / (R101 / 60)
testar(
    c(m = mean(R101), v = var(R101), n = length(R101)),
    tetha0 = 100,
    região_crítica = "maior", tipo = "t"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
1.974	0.04	[100.403, Inf]	[-Inf, 105.252]	maior	t

Podemos rejeitar a hipótese nula de que Raquel dirige na estrada a uma velocidade média menor ou igual a 100 km/h, pois o p-valor = 0.040 é menor que o ponto de corte tomado como $\alpha = 0.05$.

R 10.2: Teste de hipótese simples com modelo Normal de variância conhecida

Queremos testar a hipótese $H_0 : \mu = 50$ contra a alternativa $H_1 : \mu = 52$, onde μ é a média populacioal de uma distribuição normal com variância conhecida $\sigma^2 = 16$. Temos então n variáveis aleatórias i.i.d. $X, X_2, ..., X_n$, que seguem todas essa lei de probabilidade, e a média amostral $\bar{X}$ será usada como estatística de teste. Como $50 < 52$, é claro que quanto maior for X maiores serão as razões para se rejeitar H_0, em favor de H_1. Então o critério de decisão adequado será da forma: rejeitar H_0, se $X_{\text{obs}} > \bar{x}_c$, e aceitar H_1, se $X_{\text{obs}} < \bar{x}_c$; onde é uma constante a determinar em função de outras condições a serem especificadas.

a) Se $n = 30$ e fixarmos $P(\text{Erro I})$ em $\alpha = 0.01$, quais devem ser o critério de decisão e $\beta = P(\text{Erro II})$?
b) Mostre que se o tamanho da amostra for mantido em $n = 30$, ao fazermos com que o ponto de corte X se mova para a esquerda de forma que β diminua, α necessariamente aumentará.
c) Quais devem ser o critério de decisão e o tamanho n da amostra para que a probabilidade do Erro II se reduza a $\beta = 0.05$, mantendo a probabilidade do Erro I em $\alpha = 0.01$?

```
# H0: μ = 50
# H1: μ = 52
# xcrit
xc <- qnorm(1 - 0.01, 50, sqrt(16 / 30))
xc
# [1] 51.69892
# beta
pnorm(xc, 52, sqrt(16 / 30))
# [1] 0.3400726
```

$$\begin{cases} 0.01 = P\left(Z > \dfrac{\bar{x}_c - 50}{\sqrt{16/n}}\right) \\ 0.05 = P\left(Z < \dfrac{\bar{x}_c - 52}{\sqrt{16/n}}\right) \end{cases} \implies \begin{cases} \dfrac{(\bar{x}_c - 50)\sqrt{n}}{4} = 2.326 \\ \dfrac{(\bar{x}_c - 52)\sqrt{n}}{4} = -1.644 \end{cases} \implies \begin{cases} \bar{x}_c = 50 + \dfrac{9.305}{\sqrt{n}} \\ \bar{x}_c = 52 - \dfrac{6.579}{\sqrt{n}} \end{cases}$$

$$\begin{cases} \dfrac{15.884}{\sqrt{n}} = 2 \\ \bar{x}_c = 52 - \dfrac{6.579}{\sqrt{n}} \end{cases} \implies \begin{cases} \bar{x}_c = 51.17 \\ n = 63.08 \end{cases}$$

R 10.4: Testando hipóteses em uma eleição para governador

Em uma pesquisa eleitoral referente ao primeiro turno de uma eleição para governador foram ouvidos $n = 1000$ eleitores selecionados aleatoriamente e entre eles $m = 510$ declararam-se favoráveis ao candidato A. Deseja-se testar a hipótese H_0, de que a proporção p de eleitores do candidato A é menor ou igual a 0.5 contra a alternativa de que A venceria direto, sem a necessidade do segundo turno.

a) Qual seria a sua decisão ao nível de significância de 5%? Por quê?
b) Qual é o p-valor?
c) Se na realidade $p = 0.55$, qual seria a probabilidade de ser cometido o Erro de tipo II ao ser usada essa mesma regra de decisão? Qual o poder do teste neste caso?

```
# H0: p <= 0.5
# H1: p >  0.5
testar(
    c(m = 0.51, v = 0.5 * 0.5, n = 1000),
    tetha0 = 0.5, região_crítica = "maior"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
0.632	0.264	[0.484, Inf]	[-Inf, 0.526]	maior	z

Não rejeita a hipótese nula de que a proporção de eleitores do candidato A é menor ou igual a 0.5, pois o p-valor $= 0.264$ é maior que o ponto de corte tomado como $\alpha = 0.05$.

```
# beta = P(X < xcrit)
pnorm(0.526, 0.55, sqrt(0.55 * 0.45 / 1000))
# [1] 0.06356221
```

Capítulo 11

Exemplos

Exemplo 11.1: Comparação de dois catalisadores

Um engenheiro químico responsável por um processo produtivo quer verificar se o fato de se empregar um catalisador recentemente lançado no mercado em vez de empregar o catalisador que usa atualmente provoca um aumento significativo no rendimento do processo. Para isso, ele faz 8 ensaios com o catalisador atual (A) obtendo um rendimento médio de 80.5%, e 10 ensaios com o novo catalisador (B), obtendo rendimento médio de 81.3%.
Supondo que: os desvios padrões populacionais são conhecidos e iguais a 1.5% e 3.8%, respectivamente; e admitindo válidas as suposições de normalidade das v.a.'s envolvidas; a que conclusão se chegaria ao aplicar a esse problema um procedimento de teste de hipóteses?

```
# H0: μx >= μy
# H1: μx < μy
testar(
    c(m = 80.5 - 81.3, v = (1.5^2/8 + 3.8^2/10) * (8 + 10), n = 8 + 10),
    região_crítica = "menor"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
-0.609	0.271	[-Inf, 1.36]	[-2.16, Inf]	menor	z

Não rejeita a hipótese nula de que o uso do catalisador recentemente lançado no mercado provoca um aumento significativo no rendimento do processo comparado com o antigo, pois o p-valor $= 0.271$ é maior que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.2: Conteúdo de enxofre no carvão

Uma siderúrgica recebe carvão mineral de duas mineradoras como matéria-prima para a fabricação de aço. São obtidas aleatoriamente 50 unidades amostrais do produto fornecido pela mineradora A, para as quais se mede o conteúdo de enxofre. Com base nessas 50 medições calculam-se para essa variável uma média amostral de 0.61% e um desvio padrão amostral de 0.058%. Enquanto isso, 60 análises do carvão proveniente da mineradora B nos levam a uma média amostral de 0.68% de enxofre e um desvio padrão amostral de 0.065%. Podemos concluir, ao nível de significância de 1%, que o teor médio de enxofre no carvão proveniente da mineradora B é maior do que no carvão fornecido pela mineradora A?

```r
# H0: μx = μy
# H1: μx < μy
testar(
    c(m = 0.61 - 0.68, v = (0.058^2/50 + 0.065^2/60) * (50 + 60), n = 50 + 60),
    alpha = 0.01,
    região_crítica = "menor"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
-5.965	0	[-Inf, -0.043]	[-0.027, Inf]	menor	z

Podemos rejeitar a hipótese nula de que o teor médio de enxofre no carvão proveniente da mineradora B é igual do que no carvão fornecido pela mineradora A, pois o p-valor = 2.44e-09 é menor que o ponto de corte tomado como $\alpha = 0.01$.

Exemplo 11.3: Resistência à tração do concreto

Considere a comparação entre os concretos fabricados com cimentos das duas marcas diferentes. Os dados relativos às cargas de ruptura (em kg/cm^2) resultantes dos ensaios realizados são os seguintes:

Mistura 1	15.4	15.7	14.8	15.7	14.8	15.6	15.4	14.6	15.8	15.3	15.5	15.2
Mistura 2	14.7	14.3	15.4	14.5	14.2	15.0	14.8	15.2	15.3	14.9		

Com base nesses dados, e supondo que ambas as v.a.'s são normais:

a) Teste a hipótese de médias iguais contra médias diferentes relativas às duas misturas, ao nível de significância de 5%.
b) Qual é o p-valor neste caso?

```r
# H0: μx  = μy
# H1: μx != μy
M1 <- c(15.4, 15.7, 14.8, 15.7, 14.8, 15.6, 15.4, 14.6, 15.8, 15.3, 15.5, 15.2)
M2 <- c(14.7, 14.3, 15.4, 14.5, 14.2, 15.0, 14.8, 15.2, 15.3, 14.9)
μ1 <- mean(M1); var1 <- var(M1); n1 <- length(M1)
μ2 <- mean(M2); var2 <- var(M2); n2 <- length(M2)
testar(c(
        m = μ1 - μ2,
        v = ((n1-1) * var1 + (n2-1) * var2) / (n1+n2-2) * (1/n1 + 1/n2) * (n1+n2),
        n = n1 + n2
    ), tipo = "t"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
2.825	0.01	[0.128, 0.845]	[-0.358, 0.358]	bilateral	t

Podemos rejeitar a hipótese nula de que as cargas de ruptura das amostras de concreto são identicas, pois o p-valor = 0.010 é menor que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.4: Conteúdo de oxigênio no cascalho

Consideremos. Seja duas porções idênticas de cascalho coletadas a uma dada profundidade; uma das porções foi analisada pelo Laboratório A, e a outra, pelo Laboratório B, ambos usando o mesmo procedimento. A amostra consiste em 25 pares dessas porções, em que cada par é coletado a uma profundidade diferente. A característica a ser medida é o conteúdo de oxigênio no material coletado. Podemos afirmar que, em média, as medições feitas no Laboratório A tendem a resultar maiores do que as feitas no Laboratório B?
O valor da média amostral da diferença é $\bar{\delta} = 0.27$ e o desvio padrão $s_\Delta = 0.20$

```
# H0: µx <= µy
# H1: µx >  µy
testar(
    c(m = 0.27, v = 0.20^2, n = 25),
    região_crítica = "maior", tipo = "t"
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
6.75	0	[0.202, Inf]	[-Inf, 0.068]	maior	t

Podemos rejeitar a hipótese nula de que as medições feitas no Laboratório A tem resultados iguais aos do Laboratório B, pois o p-valor = 2.78e-07 é menor que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.5: Comparando qualidade de produto

No primeiro dia considerado, uma amostra aleatória de 50 peças foi selecionada da linha de produção, resultando em quatro peças não conformes. Suponha que a máquina que produz as peças sofre uma pane. Teme-se, portanto, que isso afete a qualidade do produto. Assim, uma nova amostra de 50 peças é selecionada no dia seguinte com o objetivo de verificar mais uma vez a proporção de peças não conformes. O número obtido nessa segunda amostra foi de seis peças não con- formes. Com base nessa informação, podemos afirmar que a proporção de peças não conformes produzidas no segundo dia é superior à do primeiro dia? Use $\alpha = 0.05$.
Seja $\hat{p}_1 = 0.08$ e $\hat{p}_2 = 0.12$, portanto $\hat{p} = 0.1$

```
# H0: p1 >= p2
# H1: p1 <  p2
testar(c(
        m = 0.08 - 0.12,
        v = 0.1 * (1 - 0.1) * (1/50 + 1/50) * (50 + 50),
        n = 50 + 50
    ), região_crítica = "menor",
)
```

estatística	p-valor	intervalo	x-crítico	região_crítica	tipo
-0.667	0.252	[-Inf, 0.059]	[-0.099, Inf]	menor	z

Não rejeita a hipótese nula de que a proporção de peças não conformes produzidas no segundo dia é igual ou inferior a do primeiro dia, pois o p-valor $= 0.252$ é maior que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.6: Resistência à tração do concreto

Examinemos mais uma vez a comparação entre os concretos fabricados com cimentos das duas marcas, que foi vista no Exemplo 11.3. As cargas de ruptura obtidas experimentalmente são:

Mistura 1	15.4	15.7	14.8	15.7	14.8	15.6	15.4	14.6	15.8	15.3	15.5	15.2
Mistura 2	14.7	14.3	15.4	14.5	14.2	15.0	14.8	15.2	15.3	14.9		

Com base nesses dados, e supondo que ambas as v.a.'s são normais, teste a hipótese nula de variâncias iguais contra a hipótese alternativa de variâncias diferentes, ao nível de significância de 5%.

```r
# H0: σ1  = σ2
# H1: σ1 != σ2
var.test(M1, M2) |> broom::tidy() |> select(-method)
```

estimate	num.df	den.df	statistic	p.value	conf.low	conf.high	alternative
0.9234355	11	9	0.9234355	0.8856409	0.2360475	3.313193	two.sided

Não rejeita a hipótese nula de que a variância das cargas de ruptura das amostras de concreto são identicas, pois o p-valor $= 0.885$ é maior que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.7: Comparando três catalisadores

Um engenheiro químico deseja comparar três catalisadores analisando o tempo de reação de um determinado processo químico em cada um deles. Para isso, ele realiza 12 ensaios utilizando o catalisador A, 16 ensaios com o catalisador B e 16 com o catalisador C. Os tempos de reação (em minutos) observados nesses ensaios são apresentados a seguir:

A	48	51	45	53	46	62	44	53	36	40	50	40				
B	42	60	47	54	59	53	46	54	36	56	40	54	41	46	47	58
C	47	40	32	36	45	36	39	50	28	52	32	36	41	39	42	28

Teste as hipóteses de médias iguais contra médias diferentes ao nível de significância $\alpha = 5\%$.

```r
# H0: μ1 = μ2 = μ3
# H1: μ1 != μ2 ou μ1 != μ3 ou μ2 != μ3
A <- c(48,51,45,53,46,62,44,53,36,40,50,40)
B <- c(42,60,47,54,59,53,46,54,36,56,40,54,41,46,47,58)
C <- c(47,40,32,36,45,36,39,50,28,52,32,36,41,39,42,28)
E117 <- data.frame(
    tempo = c(A, B, C),
    catalisador = c(rep("A", length(A)), rep("B", length(B)), rep("C", length(C)))
)
lm(tempo ~ catalisador, E117) |> anova() |> broom::tidy()
```

term	df	sumsq	meansq	statistic	p.value
catalisador	2	986.0947	493.04735	9.395561	0.0004375
Residuals	41	2151.5417	52.47663	NA	NA

Podemos rejeitar a hipótese nula de que os tempos de reação dos catalisadores são iguais, pois o p-valor = 4.375e-4 é menor que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.8: Comparando três tipos de catalisador (cont.)

Queremos saber quais combinações dos três catalisadores apresentam diferenças entre sí usando o método LSD de Fisher. Temos $MS_{\text{dentro}} = 52.48$

$$|\bar{Y}_i - \bar{Y}_j| > t_{(1-\alpha/2;N-m)}\sqrt{MS_{\text{dentro}}(1/n_i + 1/n_j)}$$

```r
E118 <- function(i, j, alpha = 0.05) {
    Yi <- E117$tempo[E117$catalisador == i]
    Yj <- E117$tempo[E117$catalisador == j]
    esquerda <- abs(mean(Yi) - mean(Yj))
    direita <- qt(1 - alpha/2, df = 44 - 3) *
      sqrt(52.48 * (1/length(Yi) + 1/length(Yj)))
    esquerda > direita
}
grid <- data.frame(grupos = c("AB", "AC", "BC"))
grid$LSD <- sapply(
    grid$grupos,
    \(g) E118(substr(g, 1,1), substr(g, 2,2))
)
grid
```

grupos	LSD
AB	FALSE
AC	TRUE
BC	TRUE

Ou seja, só ocorre diferença significativa entre os grupo A e C, e B e C

Exemplo 11.9: Comparando marcas de carros

Uma revista especializada decide comparar três marcas de carro (A, B, e C) da mesma categoria e na mesma faixa de preço, analisando diferentes itens. Para isso é realizada uma pesquisa entre usuários dos carros, os quais devem indicar o item que consideraram mais relevante ao comprar o carro. Os resultados são apresentados na tabela a seguir. Ao nível de significância de 5%, podemos afirmar que há alguma relação entre a marca do carro e o item considerado mais relevante pelo usuário?

	A	B	C
Maior capacidade do porta malas	68	31	20
Estética e acabamento	20	30	85
Desempenho	40	81	43
Outros	22	38	22

```r
# H0: marca de carro e item mais relevante independentes
# H1: marca de carro e item mais relevante dependentes
chisq.test(E119) |> broom::tidy()
```

statistic	p.value	parameter	method
110.4041	0	6	Pearson's Chi-squared test

Podemos rejeitar a hipótese nula de que não há alguma relação entre a marca do carro e o item considerado mais relevante pelo usuário, pois o p-valor = 1.68e-21 é menor que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.10: Acidentes de trânsito e dias da semana

A tabela a seguir apresenta a distribuição, por dia da semana, dos acidentes de trânsito ocorridos em 2008 no município de Ribeirão Preto. Os dados foram extraídos do site da Empresa de Trânsito e Transporte Urbano de Ribeirão Preto (Transerp).

segunda	terça	quarta	quinta	sexta	sabado	domingo
2416	2417	2442	2306	2669	2191	1355

Podemos afirmar, ao nível de significância de 5%, que os acidentes ocorrem com a mesma frequência em todos os dias da semana?

```
# H0: acidentes ocorrem com a mesma frequência
# H1: acidentes ocorrem com a diferentes frequências
chisq.test(E1110, p = rep(1/7, 7)) |> broom::tidy()
```

statistic	p.value	parameter	method
476.4794	0	6	Chi-squared test for given probabilities

Podemos rejeitar a hipótese nula de que os acidentes ocorrem com a mesma frequência em todos os dias da semana, pois o p-valor = 9.78e-100 é menor que o ponto de corte tomado como $\alpha = 0.05$.

Exemplo 11.11: Número de defeitos e distribuição de Poisson

Os dados a seguir mostram o número de defeitos encontrados em 350 chapas de aço esmaltado de um metro quadrado cada.

Número de defeitos	0	1	2	3	4	5	6	7
Número de chapas	83	124	73	40	20	6	3	1

Testar, ao nível de significância de 5%, se o número de defeitos por metro quadrado segue uma distribuição de Poisson. Qual é o p-valor do teste?

```r
# H0: número de defeitos segue distribuição Poisson
# H1: número de defeitos não segue distribuição Poisson
lambda <- sum(E1111[[1]] * E1111[[2]]) / sum(E1111[[2]])
probs <- c(dpois(0:7, lambda), 1 - ppois(7, lambda))

chisq.test(c(E1111[[2]], 0), p = probs) |> broom::tidy()
```

statistic	p.value	parameter	method
9.176052	0.3276629	8	Chi-squared test for given probabilities

Não rejeita a hipótese nula de que o número de defeitos por metro quadrado segue uma distribuição de Poisson, pois o p-valor = 0.328 é maior que o ponto de corte tomado como α = 0.05.

Exercícios

asd