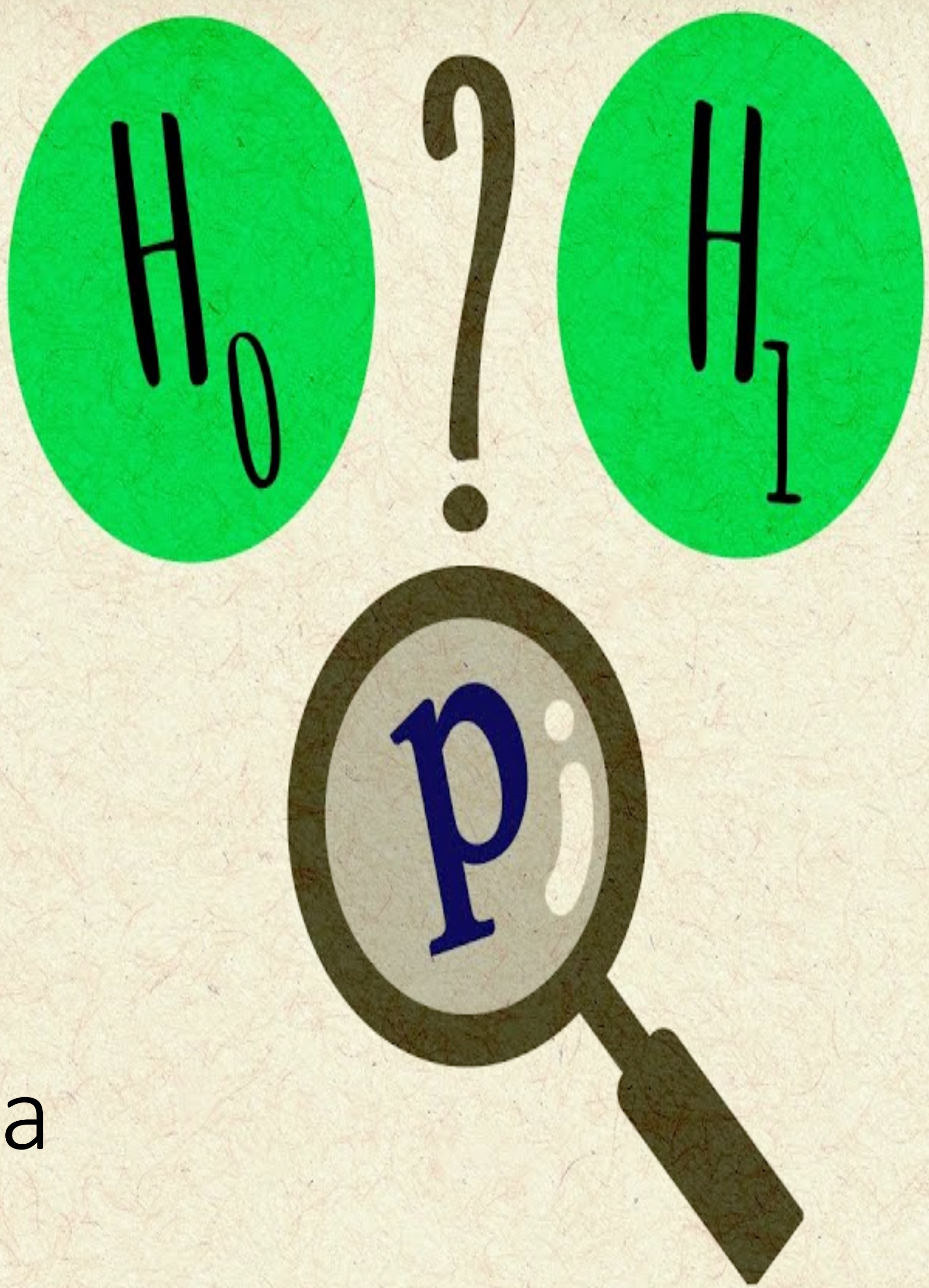


Estatística Aplicada

Cap. 11

Conceito

- Objetivo: decidir se uma afirmação sobre um parâmetro **populacional** é verdadeira a partir de informações obtidas de uma **amostra.**

- A média de uma amostra nem sempre coincide com a média suposta da população

- Se encontrarmos em uma amostra uma média diferente da esperada para a população, podemos supor que a média populacional é diferente da esperada?

Termos usados

- Hipótese
  - Hipótese: uma afirmação feita sobre a população que será testada
  - $H_0 \Rightarrow$ hipótese nula: afirmação considerada verdadeira
  - $H_1 \Rightarrow$ hipótese alternativa: a que será aceita se a nula não for verdadeira
- Nível de significância
  - Probabilidade de a hipótese nula ser rejeitada, sendo ela verdadeira
  - Isto é, probabilidade de errar na escolha da hipótese
  - Em geral usa-se 5% ou 1%
- Tipos
  - Podem ser referentes à **média** ou à **proporção**
  - Podem ser **bicaudais** ou **unicaudais** (esquerda ou direita)

Testes para médias, bicaudal

- O teste é chamado **bicaudal** quando a diferença entre a média amostral e a média populacional pode estar tanto no extremo esquerdo como no direito da curva de distribuição
- É usado para testes em que queremos verificar se o valor da média populacional é **igual** ou **diferente** da média hipotética

$$H_0 \Rightarrow M = M_0$$
$$H_1 \Rightarrow M \neq M_0$$

Onde:

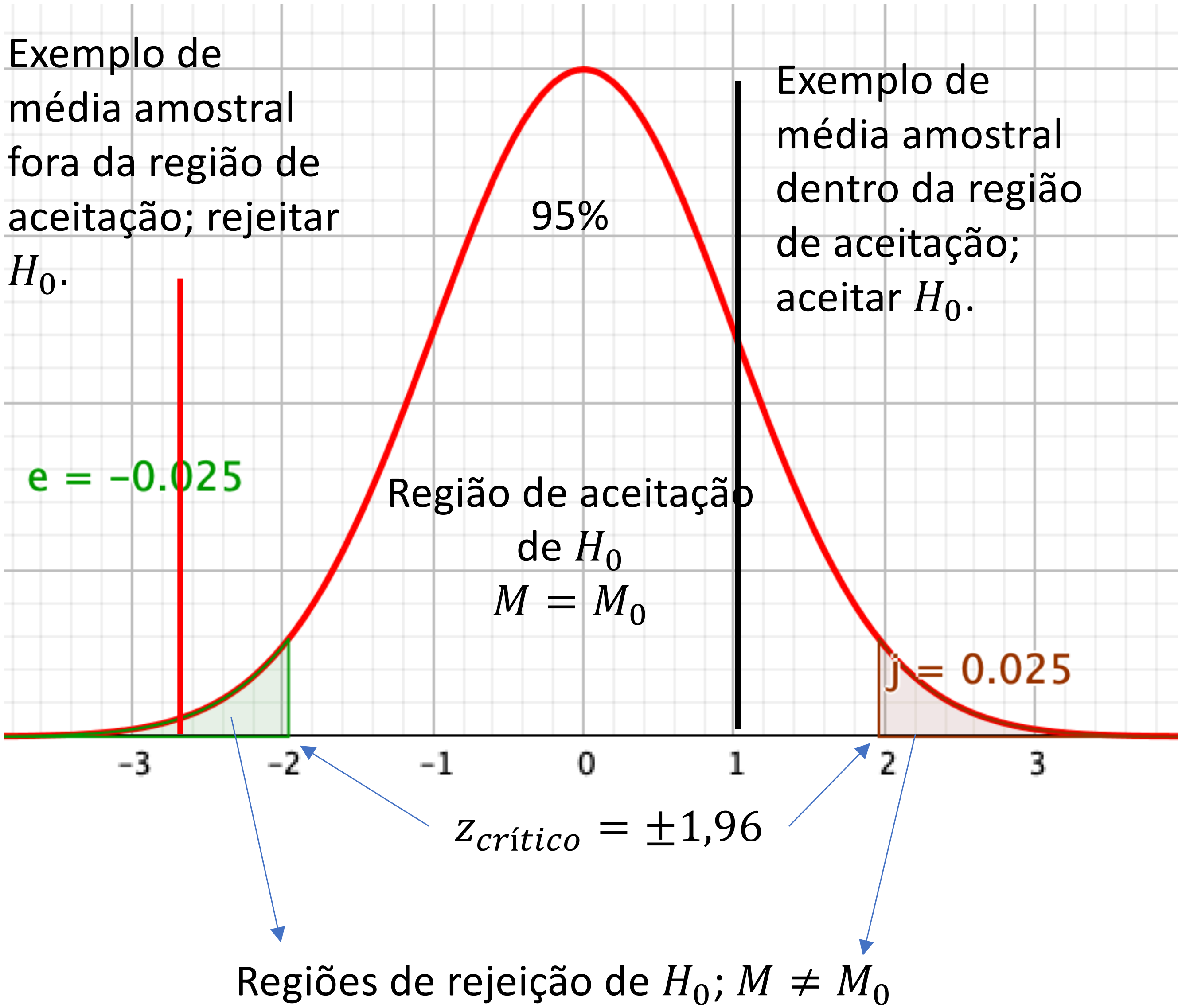
$M \Rightarrow média\ real\ da\ população$

$M_0 \Rightarrow média\ hipotética\ da\ população$

- Exemplos:
  - O diâmetro dos eixos está dentro da tolerância de fabricação?
  - O teor de açúcar de um refrigerante está entre os valores especificados?
  - A quantidade de gordura no leite está entre os valores máximo e mínimo determinados?

Testes para médias, bicaudal

- Exemplo de um teste de hipótese com 5% de significância (= 95% de confiança)



Passos:

1) Estabelecer as **hipóteses**

2) Com base na significância, calcular os **valores críticos** (máximo e mínimo aceitáveis)

3)Calcular o z da amostra, **valor de teste**

4)Comparar e **concluir**

Exemplo

Um engenheiro afirma que o tempo de montagem de um aparelho é de 82,8 minutos. O gerente deseja confirmar o dado, e seleciona uma amostra de 36 aparelhos, e constata um tempo médio de montagem de 78,3 minutos com desvio padrão de 15,3 minutos. Testar com significância de 5% a afirmação do engenheiro (o tempo de montagem é 82,8 minutos), contra a afirmação de que o tempo de montagem seja diferente (maior ou menor que 82,8 minutos).

Dados:

$$M_0 = 82{,}8\ min. \Rightarrow média\ populacional\ hipotética$$

$$m = 78{,}3\ min. \Rightarrow média\ amostral$$

$$s = 15{,}4\ min. \Rightarrow desvio\ padrão\ amostral$$

$$n = 36\ aparelhos \Rightarrow tamanho\ da\ amostra$$

$$\alpha = 0{,}05\ (5\%) \Rightarrow nível\ de\ significância$$

Resolução

- Passo 1: estabelecer as **hipóteses**

$H_0 \Rightarrow M = 82{,}8\ min.$ (hipótese nula; afirmação do engenheiro)

$H_1 \Rightarrow M \neq 82{,}8\ min.$ (hipótese alternativa; afirmação do gerente)

$$M \Rightarrow média\ real\ da\ população$$

- Passo 2: estabelecer os **valores críticos** de z para significância de 5%
  - Como o teste é **bilateral***, fazemos:

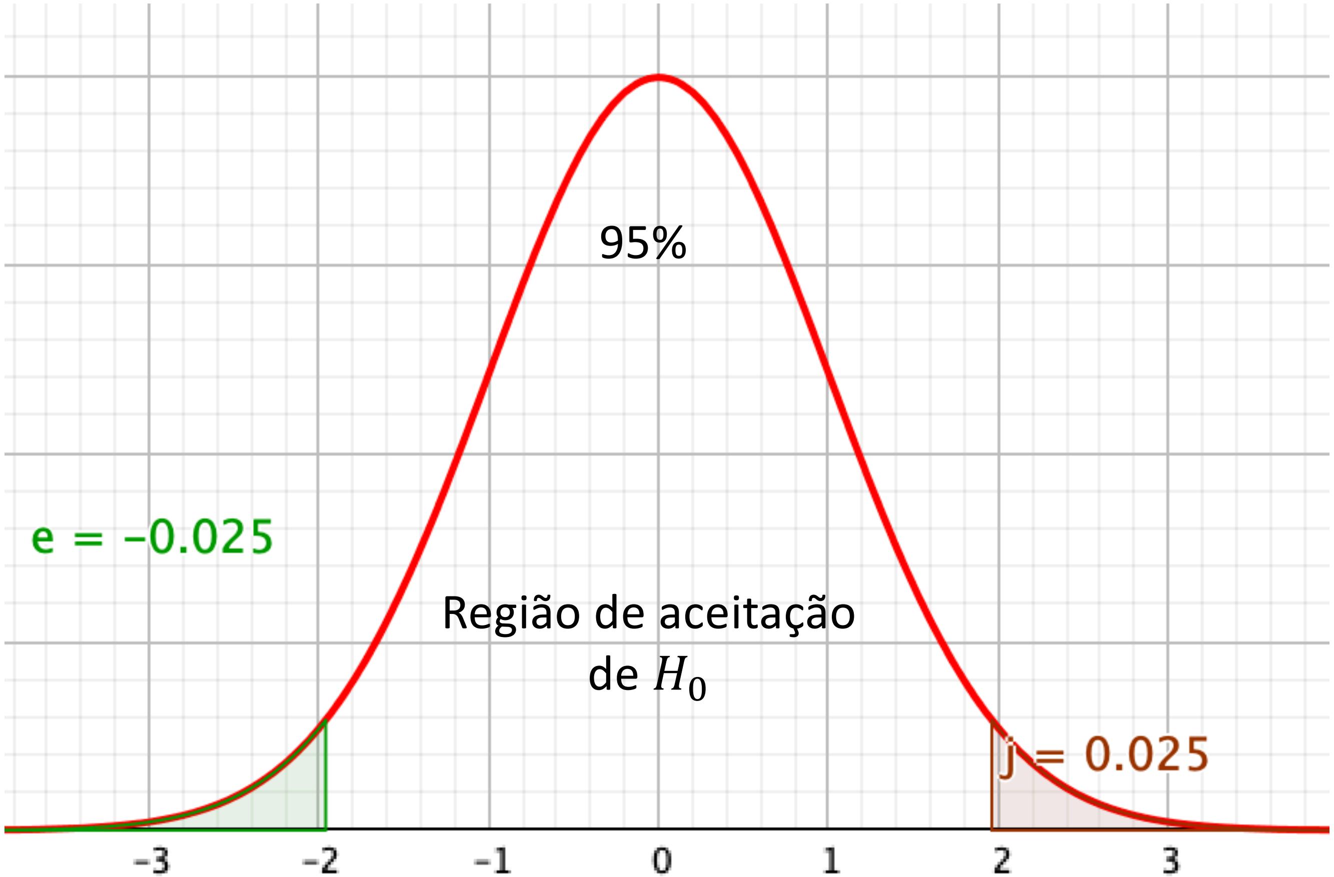
$$\frac{0{,}05}{2} = 0{,}025 \Rightarrow \text{corresponde a } 0{,}475 \text{ na tabela do z}$$

  - Da tabela temos que $z_{crítico} = \pm 1{,}96$

*ou seja, a significância deve ser distribuída metade para cada cauda.

Resolução, cont.

- Passo 2, cont.: interpretação do $z_{crítico}$



Os limites do $z_{crítico}$ representam os limites dentro dos quais considera-se que a diferença da amostra está dentro da variação amostral estatística prevista, e não representa necessariamente que a população tenha uma outra média.

$$z_{crítico} = \pm 1{,}96$$

Resolução, cont.

- Passo 3: **valor de teste**. Calcular o z correspondente à média amostral

$$z_{teste} = \frac{m - M_0}{\frac{s}{\sqrt{n}}}$$

Onde:

$m \Rightarrow média\ amostral$

$M_0 \Rightarrow média\ populacional\ hipotética$

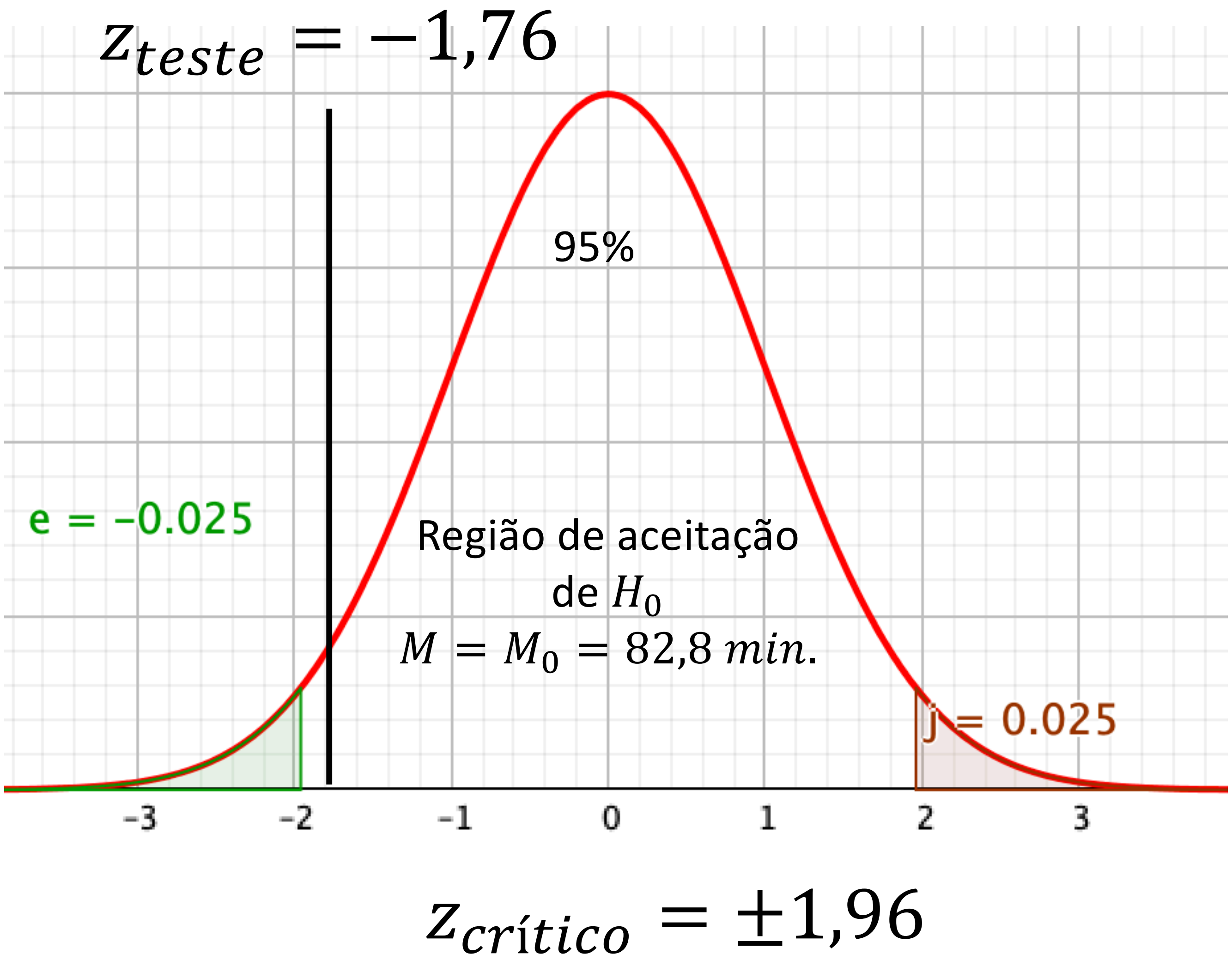
$s \Rightarrow desvio\ padrão\ amostral$

$n \Rightarrow número\ de\ elementos\ da\ amostra$

$$z_{teste} = \frac{78{,}3 - 82{,}8}{\frac{15{,}3}{\sqrt{36}}} \Rightarrow z_{teste} = -1{,}76$$

Resolução, cont.

- Passo 4: comparação e **conclusão**



Conclusão: como a média amostral (z_{teste}) está na região de aceitação, podemos aceitar a hipótese nula, isso é, o gerente deve aceitar a afirmação do engenheiro: o tempo médio de montagem pode ser 82,8 minutos, com uma significância de 5%. A diferença pode ser atribuída ao acaso.

Possíveis variações na fórmula do $\boldsymbol{z_{teste}}$

- Com desvio padrão **populacional**

$$z_{teste} = \frac{m - M_0}{\frac{\sigma}{\sqrt{n}}}$$

- Com desvio padrão **amostral**

$$n > 30$$

$$z_{teste} = \frac{m - M_0}{\frac{s}{\sqrt{n}}}$$

$$n \leq 30$$

$$t_{teste} = \frac{m - M_0}{\frac{s}{\sqrt{n}}}$$

Exemplo

Supor no caso anterior que a amostra selecionada tenha sido de 25 aparelhos, e testar novamente as hipóteses

- Passo 1: estabelecer as hipóteses

$H_0 \Rightarrow M = 82{,}8\ min.$ (hipótese nula; afirmação do engenheiro)
$H_1 \Rightarrow M \neq 82{,}8\ min.$ (hipótese alternativa; engenheiro está errado)

- Passo 2: como $(n \leq 30)$ estabelecer os valores críticos de t para significância de 5%

$$G_l = 25 - 1 = 24$$

$$\frac{0{,}05}{2} = 0{,}025 \Rightarrow t_{crítico} = \pm 2{,}064$$

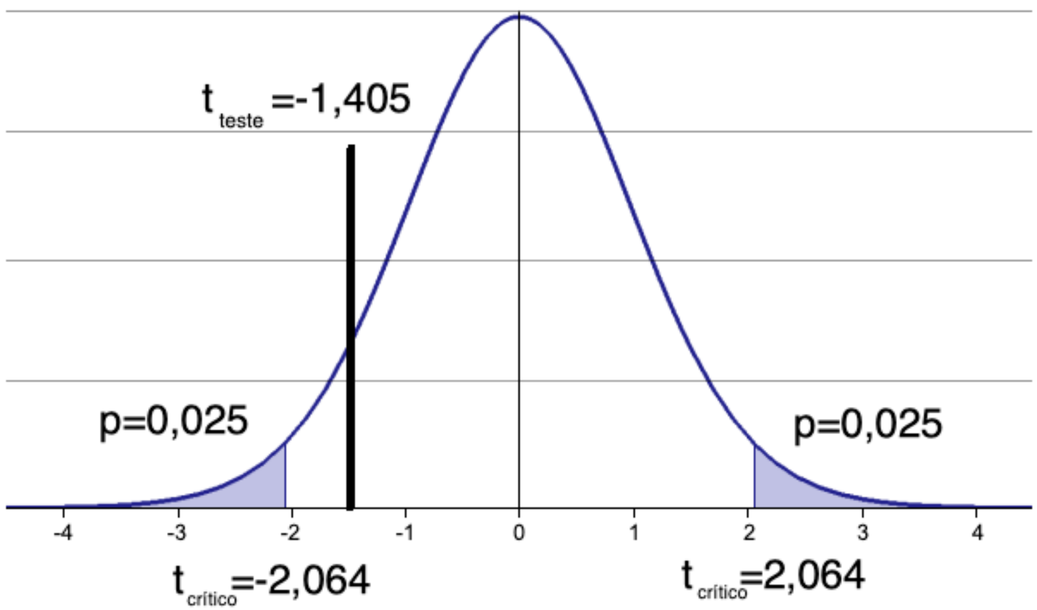
	Área na cau				
gl	0,25	0,10	0,05	0,025	0,0
20	0,687	1,325	1,725	2,086	2,52
21	0,686	1,323	1,721	2,080	2,51
22	0,686	1,321	1,717	2,074	2,50
23	0,685	1,319	1,714	2,069	2,50
24	0,685	1,318	1,711	2,064	2,49
25	0,684	1,316	1,708	2,060	2,48

Exemplo, cont.

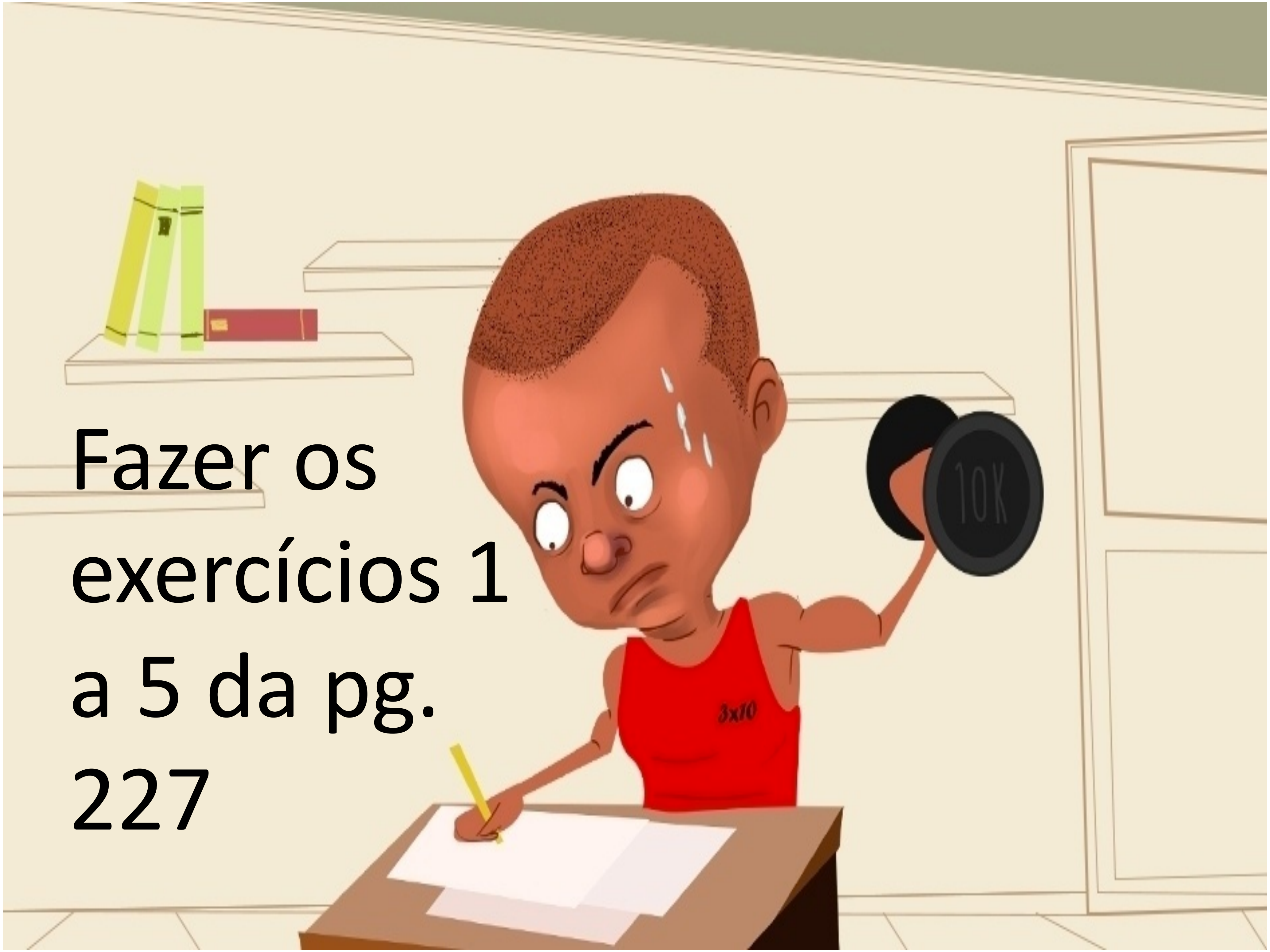
- Passo 3: calcular o t_{teste}

$$t_{teste} = \frac{78{,}3 - 82{,}8}{\frac{15{,}3}{\sqrt{25}}} \Rightarrow t_{teste} = -1{,}405$$

- Passo 4: comparação e conclusão



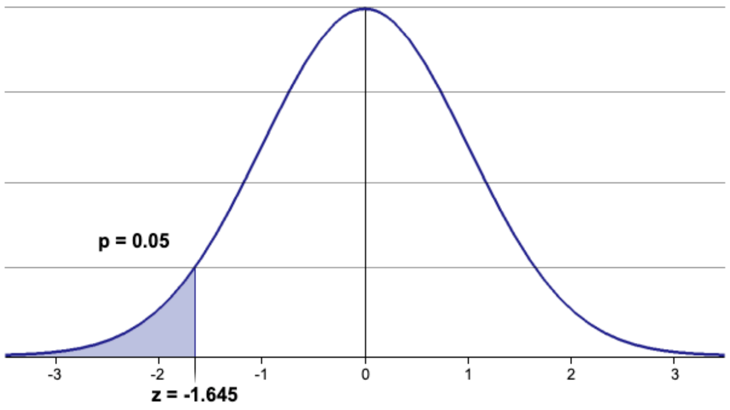
Conclusão: como a média amostral está na região de aceitação, podemos aceitar a hipótese nula, isso é, o gerente deve aceitar a afirmação do engenheiro. A diferença pode ser atribuída ao acaso.



Fazer os exercícios 1 a 5 da pg. 227

Teste de médias, unilateral ou unicaudal

- Aplica-se quando a região crítica se concentra em apenas uma das extremidades da distribuição, esquerda ou direita
- Unicaudal à **esquerda**: a região crítica (de rejeição) está apenas à esquerda.



$$H_0: M \geq M_0$$

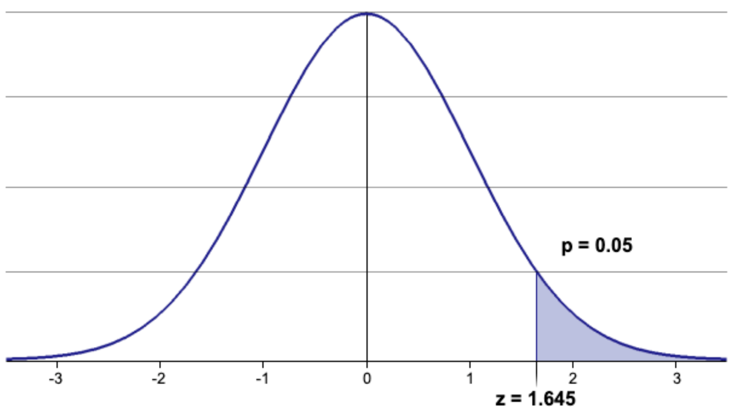
$$H_1: M < M_0$$

Exemplos:

- O fabricante afirma que a resistência mínima à tração de um componente é 200 N
- A vida útil de um produto é no mínimo 36 meses
- O peso mínimo de um produto embalado é pelo menos 250 g

Teste de médias, unilateral ou unicaudal

- Unicaudal a **direita**: a região crítica (de rejeição) está apenas à esquerda.



$$H_0: M \leq M_0$$

$$H_1: M > M_0$$

Exemplos:

- O máximo teor de agrotóxico em um alimento é de 0,1 mg/kg
- O consumo máximo dos aparelhos é de 230 kWh
- A quantidade máxima de peças defeituosas de um fornecedor é de 30 por lote de 1000

Exemplo, unicaudal esquerda

Um fabricante afirma que seus cabos suportam um peso mínimo de 600 kg, com desvio padrão de 50 kg. A análise de uma amostra de 12 cabos apresentou uma média de 575 kg. Determine, a um nível de 5% de significância se a afirmação do fabricante é válida.

Resolução

- Passo 1: estabelecer as hipóteses

$H_0 \Rightarrow M \geq 600\ kg.$ (hipótese nula; afirmação do fabricante)

$H_1 \Rightarrow M < 600\ kg.$ (hipótese alternativa; a resistência dos cabos é inferior à afirmada pelo fabricante)

- Passo 2: estabelecer os valores críticos de z para significância de 5%
  - Utilizaremos o valor z mesmo com $n < 30$, pois o desvio padrão é o **populacional**
  - Como o teste é unilateral, não dividimos a região de significância por 2

$$5\% = 0{,}05 \Rightarrow \text{corresponde a } 0{,}45 \text{ na tabela do z} \Rightarrow \mathrm{z}_{crítico} = -1{,}65$$

Resolução

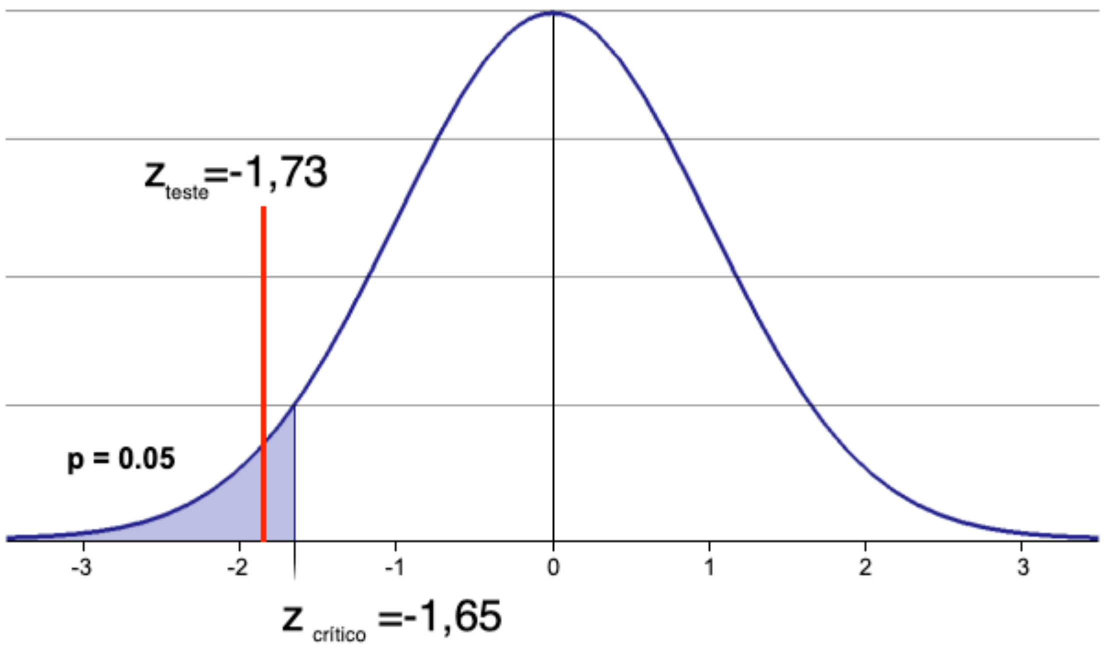
- Passo 3: cálculo do z_{teste}

$$z_{teste} = \frac{m - M_0}{\frac{\sigma}{\sqrt{n}}} = \frac{575 - 600}{\frac{50}{\sqrt{12}}} \Rightarrow z_{teste} = -1{,}73$$

Obs.: usamos o valor **negativo** de z, pois a região crítica é a **inferior** a esse valor

REsolução

- Passo 4: comparar e concluir



Conclusão: a média amostral encontra-se na **região de rejeição**, o que permite afirmar que a diferença entre as médias é relevante a um nível de significância de 5%. Portanto, a afirmação do fabricante (hipótese nula) deve ser **rejeitada**, pois os cabos suportam, em média, **menos** que 600 kg.

Exemplo, unicaudal direita

Um engenheiro de produção fabricante afirma que uma máquina automática de refrigerante está regulada para fornecer 340 ml por copo. O gerente da lanchonete, também um engenheiro de produção, afirma que a média fornecida por copo é maior, e solicita recalibragem da máquina. Ambos decidem testar, com um nível de significância de 5%, as hipóteses abaixo:

- Passo 1: hipóteses

$H_0 \Rightarrow M \leq 340\ kg$. (hipótese nula; o fabricante está correto)

$H_1 \Rightarrow M > 340\ kg$. (hipótese alternativa; o gerente está correto)

- Os testes feitos são
  - Fabricante: Em uma amostra de 10 copos, ele obtém uma média de 350 ml por copo, com desvio padrão de 20 ml.
  - Gerente da lanchonete: O gerente, não conformado com a conclusão, realiza um novo teste, usando porém uma amostra de 50 copos. Ele obtém a mesma média (350 ml) e o mesmo desvio padrão (20 ml)

Exemplo, resolução

1) Teste do engenheiro fabricante da máquina

Em uma amostra de 10 copos, ele obtém uma média de 350 ml por copo, com desvio padrão de 20 ml.

- Passo 2: valor crítico. Como a amostra é menor que 30 e o desvio padrão conhecido é o da amostra, deve-se usar a tabela $t - student$

$$G_l = 10 - 1 = 9$$

$$5\% \Rightarrow 0{,}05 \Rightarrow t_{crítico} = +1{,}833$$

Obs.: usamos o valor positivo de t, pois a região crítica é a superior a esse valor

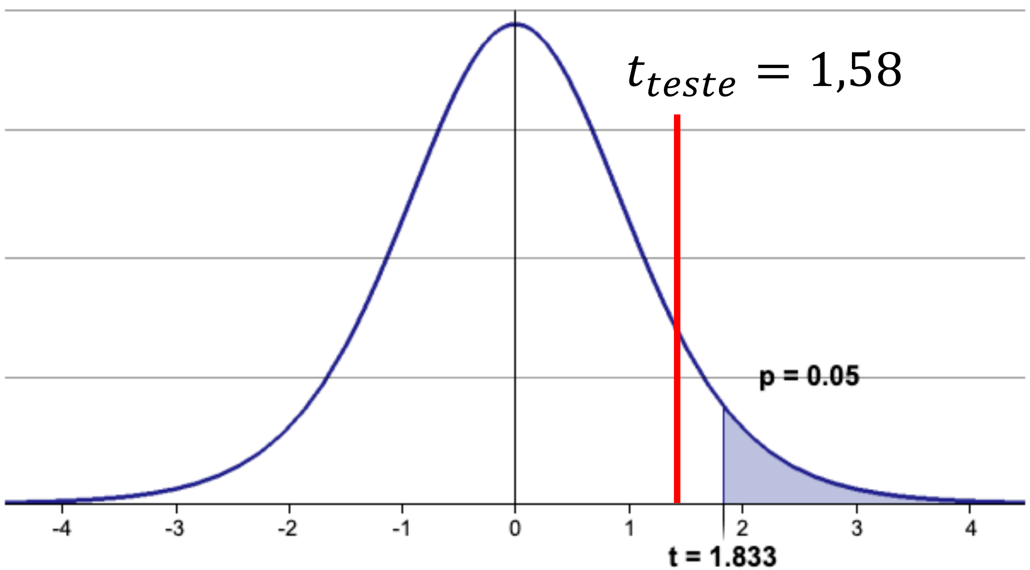
gl	0,25	0,10	0,05
1	1,000	3,078	6,314
2	0,816	1,886	2,920
3	0,765	1,638	2,353
4	0,741	1,533	2,132
5	0,727	1,476	2,015
6	0,718	1,440	1,943
7	0,711	1,415	1,895
8	0,706	1,397	1,860
9	0,703	1,383	1,833
10	0,700	1,372	1,812

Exemplo, resolução

- Passo 3: cálculo do t_{teste}

$$t_{teste} = \frac{350 - 340}{\frac{20}{\sqrt{10}}} \Rightarrow t_{teste} = 1{,}58$$

- Passo 4: comparar e concluir



Conclusão do fabricante: o valor de teste caiu na região de aceitação, o que leva o fabricante a afirmar que a máquina está calibrada corretamente (a hipótese nula deve ser aceita).

Exemplo, resolução

2) Teste do gerente da lanchonete

O gerente, não conformado com a conclusão, realiza um novo teste, usando porém uma amostra de 50 copos. Ele obtém a mesma média (350 ml) e o mesmo desvio padrão (20 ml)

- Passo 2: valor crítico. Agora, como a amostra é superior a 30, usamos a tabela z.

$$5\% = 0{,}05 \Rightarrow corresponde\ a\ 0{,}45\ na\ tabela \Rightarrow z_{crítico} = +1{,}65$$

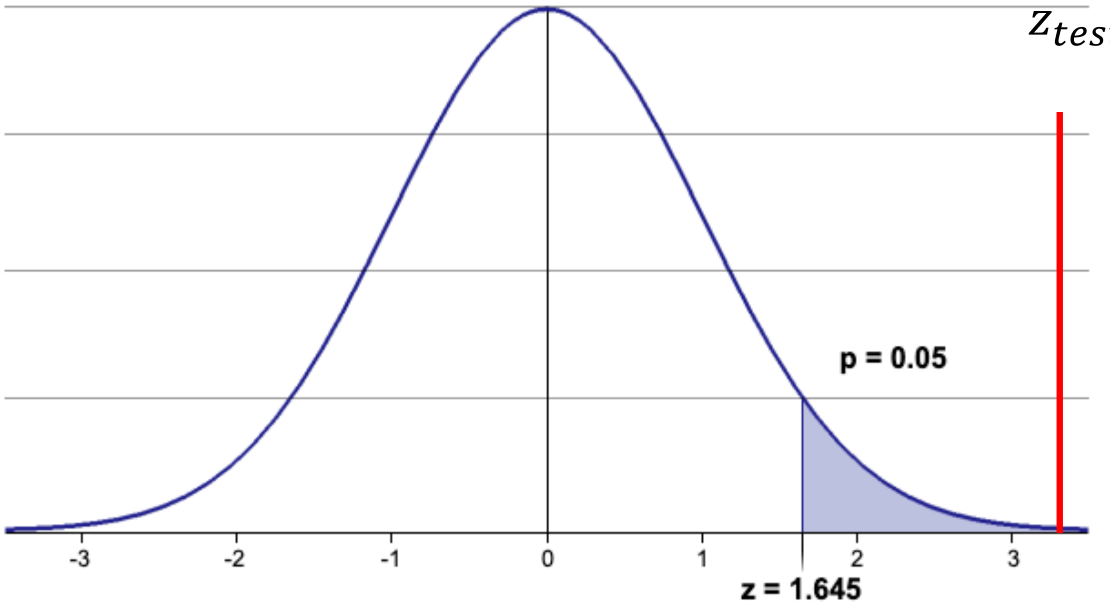
- Passo 3: cálculo do z_{teste}

$$z_{teste} = \frac{350 - 340}{\frac{20}{\sqrt{50}}} \Rightarrow z_{teste} = 3{,}54$$

Exemplo, resolução

- Passo 4: comparar e concluir

$$z_{teste} = 3{,}54$$



Conclusão: o gerente (um engenheiro que não usava WhatsApp durante a aula de Estatística Aplicada, mesmo online) demonstrou que o teste caiu na região de rejeição, e que, portanto, a afirmação do fabricante deve ser rejeitada, e a máquina precisa sim ser calibrada.

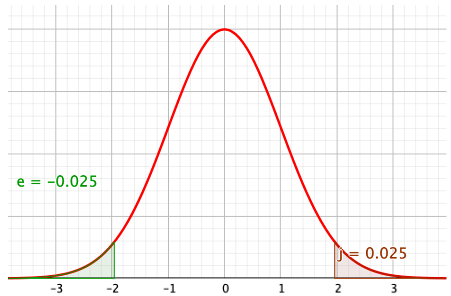
Fazer os exercícios 1 e 2 da pg. 235 e 1 a 3 da pg. 241

Teste de proporções

- Objetivo: testar se uma afirmação de **proporção** feita para uma população é ou não aceitável, a partir de uma proporção amostral.
- Quanto à lateralidade, valem as mesmas considerações

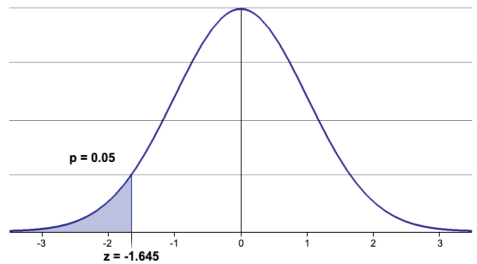
Bilateral

$$H_0: P = P_0$$
$$H_1: P \neq P_0$$



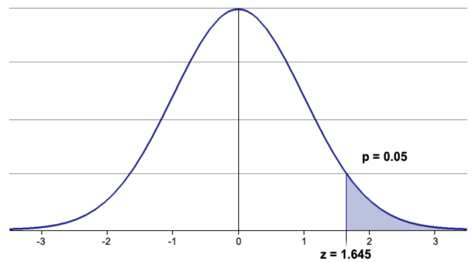
Unilateral à esquerda

$$H_0: P \geq P_0$$
$$H_1: P < P_0$$



Unilateral à direita

$$H_0: P \leq P_0$$
$$H_1: P > P_0$$



Teste de proporções

- Sempre usamos o valor z; não usamos a tabela $t - student$
- O valor de teste é calculado com a fórmula:

$$z_{teste} = \frac{p_a - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}}$$

Onde:

$p_a \Rightarrow porcentagem\ amostral$

$p_0 \Rightarrow porcentagem\ populacional\ suposta$

$q_0 \Rightarrow 1 - p_0$

Exemplo

Um fabricante de pendrives afirma que no mínimo 90% de sua produção não apresenta problemas. Uma amostra de 200 peças foi testada, e 30 pendrives apresentaram defeito. Testar, ao nível de significância de 2%, a afirmação do fabricante contra a alternativa de que a proporção de pendrives sem defeito é inferior a 90%

Resolução

- Passo 1: hipóteses

$H_0 \Rightarrow p \geq 90\%$. (hipótese nula; pelo menos 90% perfeitas)

$H_1 \Rightarrow p < 90\%$. (hipótese alternativa; menos que 90% perfeitas)

- Passo 2: valor crítico. Teste unilateral à esquerda

$$2\% = 0{,}02\ corresponde\ a\ 0{,}48\ na\ tabela \Rightarrow z_{crítico} = -2{,}06$$

- Passo 3: valor de teste z_{teste}

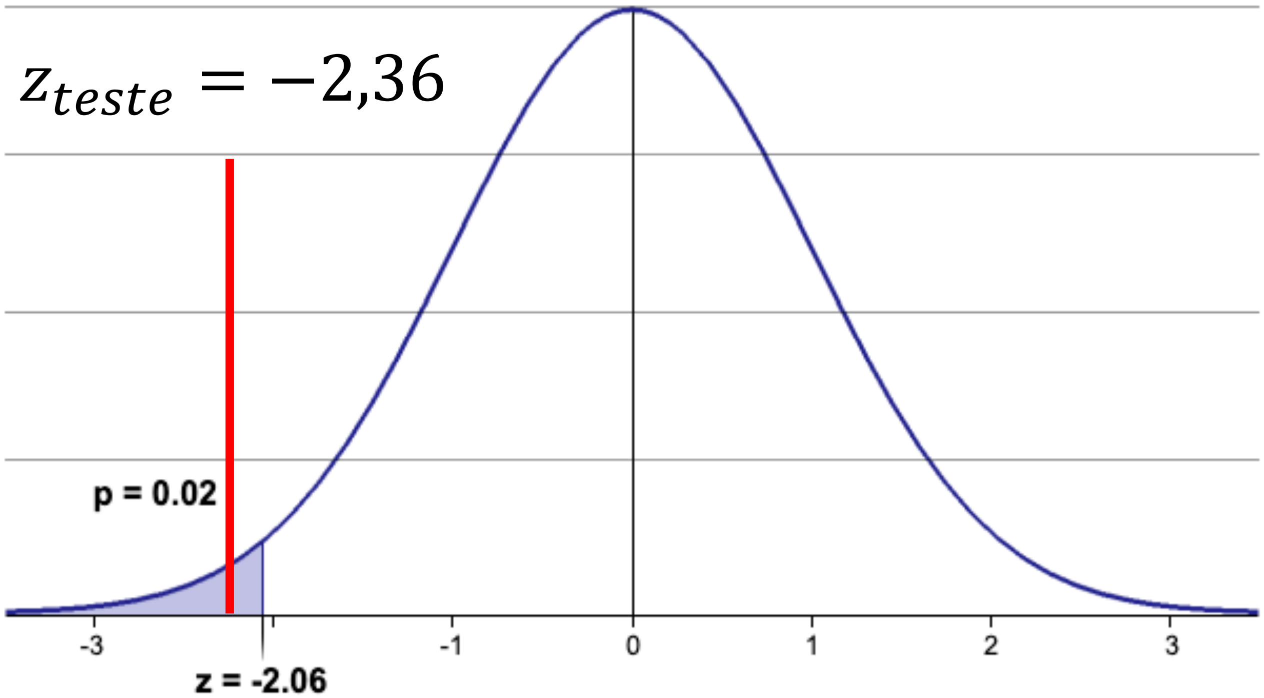
$$p_a = \frac{200 - 30}{200} = \frac{170}{200}$$

$$p_a = 0{,}85\ (porcentagem\ de\ peças\ boas\ na\ amostra)$$

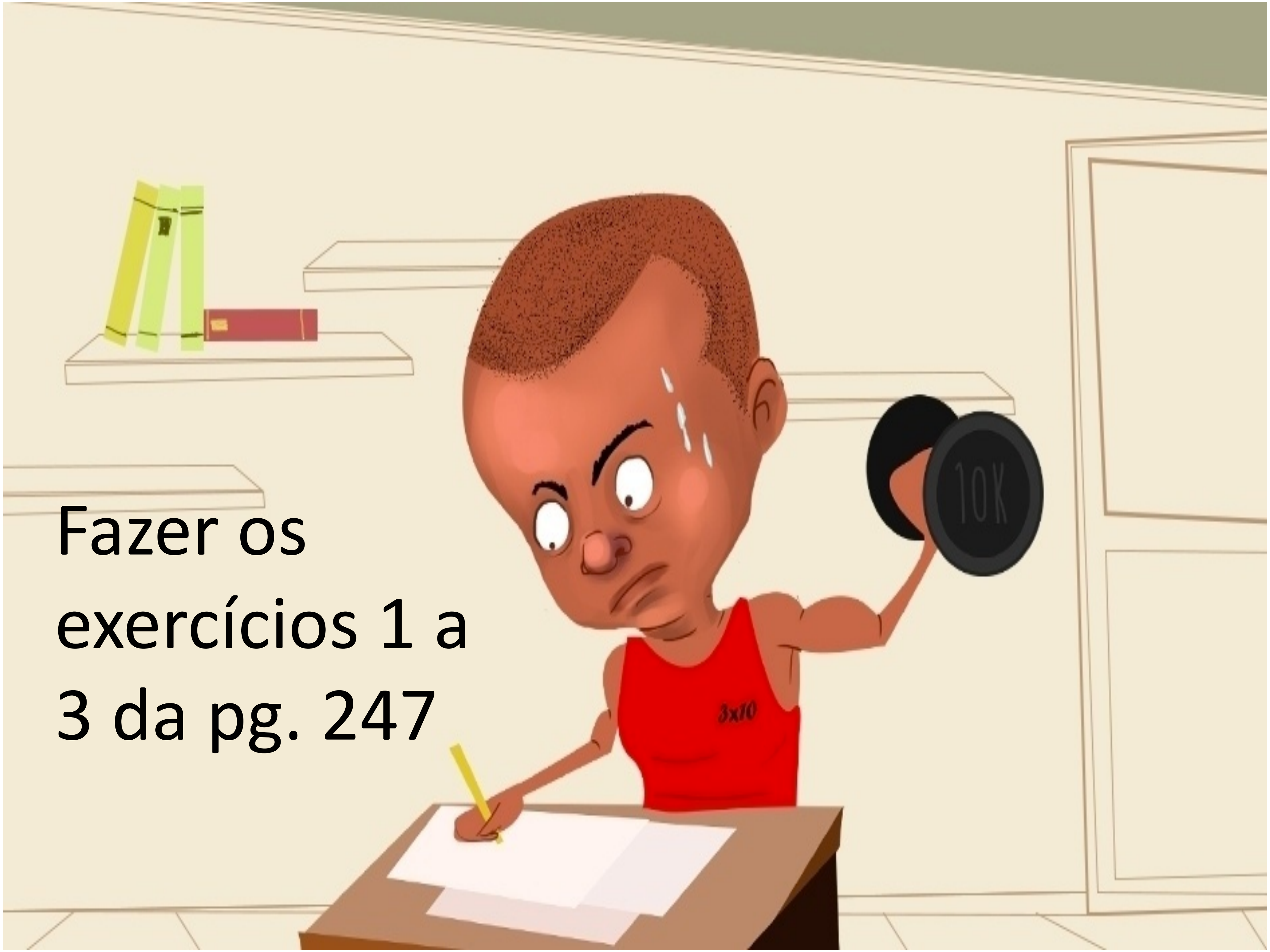
$$z_{teste} = \frac{p_a - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}} = \frac{0{,}85 - 0{,}90}{\sqrt{\frac{0{,}90 \cdot 0{,}10}{200}}} \Rightarrow z_{teste} = -2{,}36$$

Resolução, cont.

- Passo 4: comparar e concluir



Conclusão: o valor de teste está na região de rejeição. Assim, a afirmação do fabricante é inaceitável; a proporção de pendrives sem defeito é inferior a 90%


Fazer os exercícios 1 a 3 da pg. 247
10K
3x10