Testes de hipóteses

Prof. Wagner Hugo Bonat

Departamento de Estatística
Universidade Federal do Paraná

Motivação

Inferência estatística

Falar sobre a **população** a partir da observação da **amostra**.

Na **inferência estatística** os dois principais objetivos são:

1. **Estimar** um parâmetro populacional.
    - Estimativa pontual.
    - Estimativa intervalar.
2. **Testar** uma hipótese ou afirmativa sobre um parâmetro populacional.

Figura 1. Analogia ao processo de estimação. Extraído de bestbowreviews.com.

Testes de hipótese

- **Hipótese**
  - É uma afirmativa sobre uma **propriedade** da população.
- **Teste de hipótese**
  - É um procedimento para se testar uma **afirmativa** sobre uma propriedade da população.
  - Permite tomar **decisões** sobre a população com base em informações de dados amostrais.

Exemplo: proporção sexual em peixes

- Deseja-se estudar a proporção de peixes machos e fêmeas de uma mesma espécie em uma lagoa.
- Sem nenhuma informação prévia, supõe-se que a proporção sexual é de 50% ($p = 0.5$).
- Se, em uma amostra de 100 peixes:
  - 54 forem fêmeas.
  - 65 forem fêmeas.
  - 92 forem fêmeas.
- Qual a evidência necessária para concluir que a proporção de fêmeas é maior que a de machos nessa população?

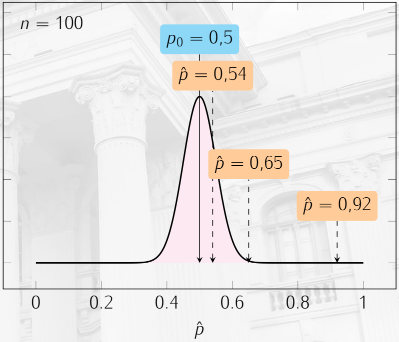

Figura 2. Proporções amostrais supondo $p = 0.5$ na população.

Exemplo: cardápio vegano

- **Questão:** Há interesse por opções veganas? ($\hat{p} = 0.12$ em uma amostra).

- **Intervalo de confiança**
  - **Pergunta**: Qual a proporção que prefere pratos veganos?
  - 0.12 ± 0.035 ou $(0.085, 0.155)$

- **Teste de hipótese**
  - **Pergunta**: A proporção de clientes que prefere pratos veganos supera 10%?
  - $\hat{p} = 0.12$ é **significativamente maior** do que $p = 0.10$?
  - Em outras palavras: é muito ou pouco provável que $\hat{p} = 0.12$ tenha sido gerado por uma população com $p = 0.10$?
  - A questão é: $\hat{p} = 0.12$ é significativamente diferente de $p = 0.10$?

Exemplo: caracterização dos clientes

Questão: Qual será a *idade média* dos clientes? ($\overline{y} = 32$ em uma amostra)

- **Intervalo de confiança**
  - **Pergunta**: Qual a idade média dos clientes?
  - 32 ± 2.5 ou $(29.5, 34.5)$
- **Teste de hipótese**
  - **Pergunta**: A idade média dos clientes é igual a 35 anos?
  - $\overline{y} = 32$ é **significativamente diferente** de $\mu = 35$?

Exemplo: caracterização dos clientes por sexo

Questão: Qual será a *idade média* dos clientes **por sexo**?
($\overline{y}_h = 34$ para homens e $\overline{y}_m = 31$ para mulheres em uma amostra)

- **Intervalo de confiança**
  - **Pergunta**: Qual a idade média dos clientes homens e mulheres?
  - 34 ± 2.3 ou $(31.7, 36.3)$ para homens
  - 31 ± 2.8 ou $(28.2, 33.8)$ para mulheres
  - ou ainda a diferença de idade
  - 3 ± 2.5 ou $(0.5, 5.5)$
- **Teste de hipótese**
  - **Pergunta**: Existe diferença (**significativa**) entre a idade média dos clientes homens e mulheres?
  - $\overline{y}_h - \overline{y}_m = 34 - 31 = 3$ é **significativamente diferente** de $\mu_h - \mu_m = 0$?

Fundamento lógico do teste de hipótese

- Testamos uma afirmativa na tentativa de distinguir entre resultados que:
  - Podem **facilmente** ocorrer por *acaso* na amostra.
  - São **altamente improváveis** de ocorrer por *acaso* na amostra.
- A ocorrência de **resultados altamente improváveis** pode ser explicada de uma das duas formas:
  - Ou um evento raro realmente **ocorreu**.
  - Ou a **suposição** subjacente não é verdadeira.
- Regra do evento raro Se, sob uma dada **suposição**, a probabilidade de um evento observado particular é **extremamente pequena**, concluímos que a suposição **provavelmente** não é verdadeira.

Procedimentos gerais para um teste de hipótese

1. **Definir a hipótese nula (H_0) e a alternativa (H_a).**
2. Definir um nível de significância α, que irá determinar o nível de confiança $100(1-\alpha)\%$ do teste.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste, com base na distribuição amostral do estimador do parâmetro sob teste $\rightarrow$ valor calculado.
5. Determinar a região crítica (região de rejeição), com base no nível de significância α $\rightarrow$ valor crítico.
6. Concluir o teste.

1. Definição das hipóteses · tipos de hipóteses

1. Definição das hipóteses · tipos de hipóteses

- **Hipótese nula** H_0
  - É uma afirmativa de que o valor de um parâmetro populacional é **igual** a algum valor especificado.
  - O termo *nula* é usado para indicar nenhuma mudança ou nenhum efeito.
  - Exemplos:

$$\mu = 10$$
$$p = 0.5$$
$$\sigma^2 = 4.$$

- **Hipótese alternativa**
  - É uma afirmativa de que o parâmetro tem um valor que, de alguma forma, **difere** da hipótese nula.
  - Exemplos:

$$\mu \neq 10$$
$$p > 0.5$$
$$\sigma^2 < 4.$$

1. Definição de hipóteses · decisões sobre a hipótese

Quando fazemos um teste de hipótese, chegamos a um dos dois possíveis resultados:

- **Rejeitar** H_0: em favor da hipótese alternativa H_a.
- **Não rejeitar** H_0: e conclui-se que não existem diferenças.

Atenção!

- O termo **aceitar** a hipótese nula é filosoficamente incorreto, pois não se pode aceitar uma hipótese baseada apenas em evidências amostrais.
- E ainda existe um **erro** associado a todo teste de hipótese.

Teoria do falsificacionismo de K. Popper

- Uma hipótese **não pode ser provada**, apenas **desprovada**.
- Se a hipótese permanece válida então ela **não é validada**, mas adquire um certo "**grau de confiança**".
- Se você está fazendo um estudo e deseja usar um teste de hipótese para **apoiar** sua afirmativa, esta deve ser escrita de modo a se tornar a **hipótese alternativa**.
- Você nunca pode apoiar uma afirmativa de que um parâmetro **seja igual** a algum valor específico.
- Nesse contexto de se tentar apoiar o resultado de pesquisa, a hipótese alternativa é, algumas vezes, chamada de **hipótese de pesquisa**.

1. Definição de hipóteses · exemplo

Em um estudo sobre a proporção sexual de peixes de uma mesma espécie em uma lagoa, deseja-se testar a hipótese de que a proporção de fêmeas é maior do que a proporção de machos.

- Supondo inicialmente que a proporção de fêmeas é de 50% ($p = 0.5$), então

$$H_0 : p = 0.5 \quad \text{vs} \quad H_a : p > 0.5$$

  - Com isso, deseja-se que a **hipótese nula** $p = 0.5$ seja rejeitada, de modo que a **hipótese alternativa** $p > 0.5$ seja apoiada.
- Apoiar a hipótese alternativa de que $p > 0.5$ é o mesmo que apoiar a afirmativa de a proporção de fêmeas na população é maior do que a de machos.

Procedimentos gerais para um teste de hipótese

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. **Definir um nível de significância** α**, que irá determinar o nível de confiança** $100(1-\alpha)$**% do teste**.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste, com base na distribuição amostral do estimador do parâmetro sob teste $\rightarrow$ valor calculado.
5. Determinar a região crítica (região de rejeição), com base no nível. de significância α $\rightarrow$ valor crítico.
6. Concluir o teste.

2. Nível de significância · ilustração dos erros

2. Nível de significância · erros de decisão

Para entendermos o que é o nível de significância (α), precisamos saber que, ao realizar um teste de hipótese, estamos sujeitos a dois tipos de erros.

- **Erro Tipo I**: rejeitar H_0, quando H_0 é verdadeira (falso negativo).
- **Erro Tipo II**: não rejeitar H_0, quando H_0 é falsa (falso positivo).

	H_o **verdadeira**	H_o **falsa**
Não rejeitar H_0	Decisão correta	Erro tipo II
Rejeitar H_0	Erro tipo I	Decisão correta

2. Nível de significância · ilustração dos erros

Figura 3. Erros de decisão em testes de hipótese. Modificado de www.irishmirror.ie.

2. Nível de significância · definição dos erros

Definimos por α e β as probabilidades de cometer os erros do tipo I e II:

- $\alpha =$ P(erro tipo I) $=$ P(rejeitar $H_0 \,|\, H_0$ verdadeira).
- $\beta =$ P(erro tipo II) $=$ P(não rejeitar $H_0 \,|\, H_0$ falsa).
- α é o **nível de significância** do teste.
- $1 - \alpha$ é o **nível de confiança** do teste.

No exemplo anterior, se $H_0 : p = 0.5$ e $H_a : p > 0.5$, então:

- $\alpha =$ P(concluir que a proporção de fêmeas é maior quando na verdade não é).
- $\beta =$ P(concluir que a proporção sexual é igual quando na verdade não é).

2. Nível de significância · balanço entre os erros

- A situação ideal é aquela em que ambas as probabilidades, α e β, são próximas de zero.
- No entanto, à medida que diminuimos α, a probabilidade β tende a aumentar.
- Levando isso em conta, ao formular as hipóteses, **devemos cuidar para que o erro (usualmente) mais importante a ser evitado seja o erro do tipo I**.
- Por isso, a probabilidade α recebe o nome de **nível de significância** do teste, e é esse erro que devemos controlar.

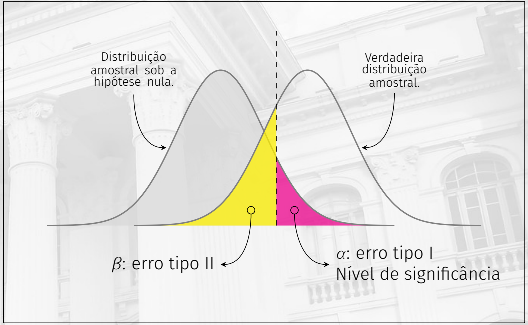

Figura 4. Probabilidade dos tipos de erros em testes de hipótese.

Procedimentos gerais para um teste de hipótese

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir um nível de significância α, que irá determinar o nível de confiança $100(1-\alpha)\%$ do teste.
3. **Definir o tipo de teste, com base na hipótese alternativa**.
4. Calcular a estatística de teste, com base na distribuição amostral do estimador do parâmetro sob teste $\rightarrow$ valor calculado.
5. Determinar a região crítica (região de rejeição), com base no nível de significância α $\rightarrow$ valor crítico.
6. Concluir o teste.

3. Tipos de testes

3. Tipos de testes

A hipótese alternativa determinará o **sentido** do teste de hipótese, que pode ser:

- Bilateral:

$$H_a : \theta \neq \theta_0.$$

- Unilateral à esquerda:

$$H_a : \theta < \theta_0.$$

- Unilateral à direita:

$$H_a : \theta > \theta_0.$$

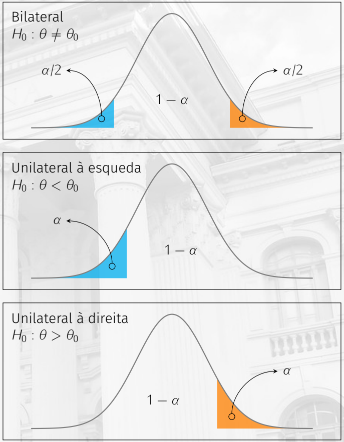

Figura 5. A região de rejeição de H_0 conforme o tipos de hipótese alternativa.

3. Tipos de testes · bilateral

Uma hipótese do tipo

$$H_0 : \theta = \theta_0$$
$$H_a : \theta \neq \theta_0$$

é **bilateral**.

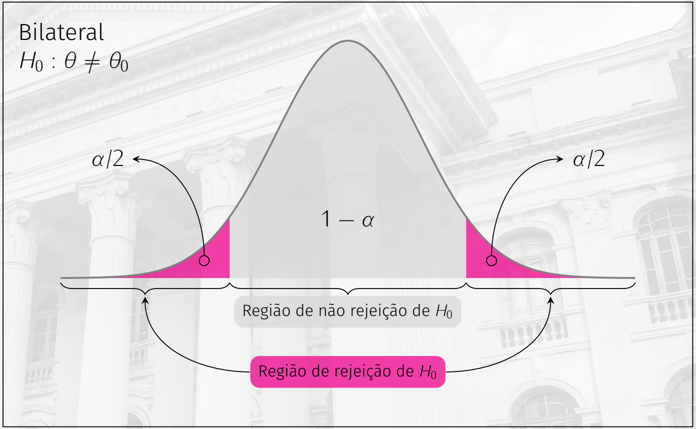

Figura 6. A região de rejeição de H_0 para uma hipótese alternativa bilateral.

3. Tipos de testes · unilateral à esquerda

Uma hipótese do tipo

$$H_0 : \theta = \theta_0$$
$$H_a : \theta < \theta_0$$

é **unilateral à esquerda**.

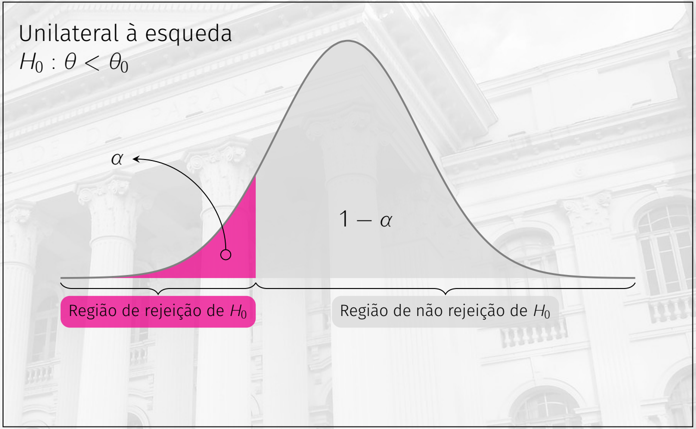

Figura 7. A região de rejeição de H_0 para uma hipótese alternativa unilateral à esquerda.

3. Tipos de testes · unilateral à direita

Uma hipótese do tipo

$$H_0: \theta = \theta_0$$
$$H_a: \theta > \theta_0$$

é **unilateral à direita**.

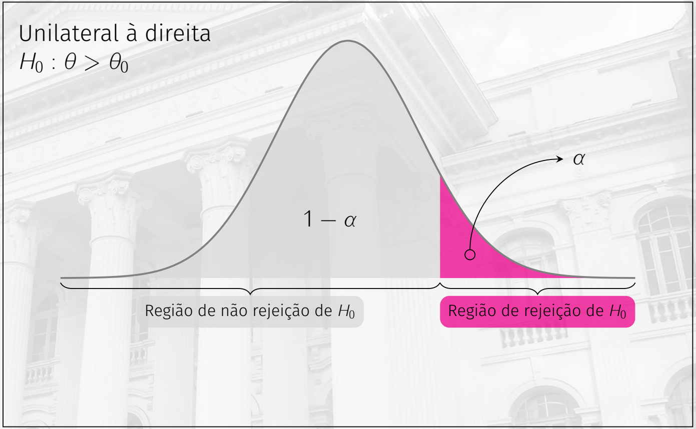

Figura 8. A região de rejeição de H_0 para uma hipótese alternativa unilateral à direita.

Procedimentos gerais para um teste de hipótese

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir um nível de significância α, que irá determinar o nível de confiança $100(1 - \alpha)\%$ do teste.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. **Calcular a estatística de teste, com base na distribuição amostral do estimador do parâmetro sob teste** $\rightarrow$ **valor calculado**.
5. Determinar a região crítica (região de rejeição), com base no nível de significância α $\rightarrow$ valor crítico.
6. Concluir o teste.

4. Estatística de teste

A **estatística de teste** é um valor usado para tomar a decisão sobre a hipótese nula, supondo que ela seja verdadeira.

Considera a distribuição amostral do estimador **sob a hipótese nula**.

- Estatística de teste para a **média** (μ)

$$z = \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}} \quad \text{ou} \quad t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}.$$

- Estatística de teste para a **proporção** (p)

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}}.$$

- Estatística de teste para a **variância** (σ^2)

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}.$$

Procedimentos gerais para um teste de hipótese

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir um nível de significância α, que irá determinar o nível de confiança $100(1-\alpha)\%$ do teste.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste, com base na distribuição amostral do estimador do parâmetro sob teste $\rightarrow$ valor calculado.
5. **Determinar a região crítica (região de rejeição), com base no nível de significância** $\alpha \rightarrow$ **valor crítico**.
6. Concluir o teste.

5. Região crítica · definição

- A estatística de teste **sozinha** não nos dá informação suficiente para a tomada de decisão sobre a afirmativa em um teste.
- É necessário comparar esta estatística com algum **valor de referência**, que nos informe o quão extrema é a estatística de teste para rejeição de H_0.
- Este valor de referência é chamado de **valor crítico**, que divide a região de rejeição da região de não rejeição da hipótese nula. Depende:
  - Da distribuição amostral da estatística de teste sob H_0.
  - Do nível de significância α.
- A **região crítica** de um teste de hipótese é a **região de rejeição** da hipótese nula.

5. Região crítica · exemplos

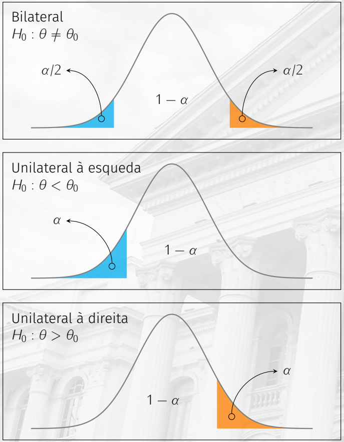

Figura 9. A região de rejeição de H_0 conforme o tipos de hipótese alternativa.

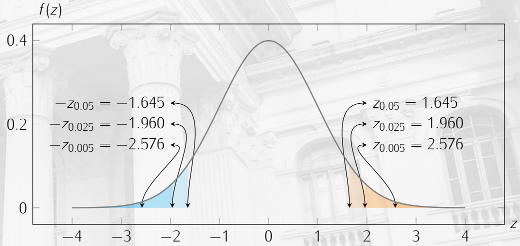

Figura 10. Valores críticos para determinar a região crítica.

Procedimentos gerais para um teste de hipótese

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir um nível de significância α, que irá determinar o nível de confiança $100(1 - \alpha\%)$ do teste.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste, com base na distribuição amostral do estimador do parâmetro sob teste $\rightarrow$ valor calculado.
5. Determinar a região crítica (região de rejeição), com base no nível de significância α $\rightarrow$ valor crítico.
6. **Concluir o teste**.

6. Conclusão do teste · valor da estatística

Com base na estatística de teste e valor crítico

- Se a estatística de estiver **dentro** da região crítica → **rejeita** H_0.
- Se a estatística de teste estiver **fora** da região crítica → **não rejeita** H_0.

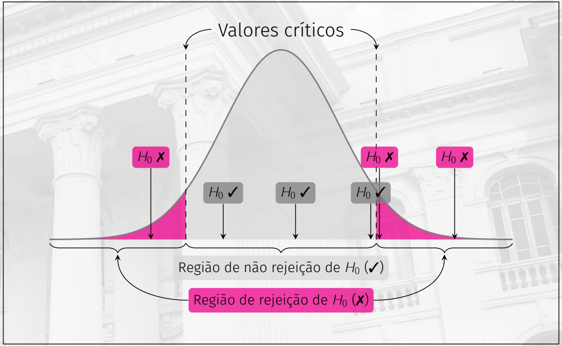

Figura 11. Decisões conforme o valor da estatística de teste.

6. Conclusão do teste · nível descritivo

Com base no nível descritivo ou p-valor

- Em geral, α é pré-fixado para construir a regra de decisão.
- Uma alternativa é deixar em aberto a escolha de α para quem for tomar a decisão.
- A ideia é calcular, **supondo que a hipótese nula é verdadeira**, a probabilidade de se obter **estatísticas mais extremas** do que aquela fornecida pela amostra.
- Essa probabilidade é chamada de **nível descritivo**, denotada por α^* (ou p**-valor**).
- Valores pequenos de α^* **evidenciam** que a hipótese nula é falsa.
- O conceito de "pequeno" **fica para quem decide** qual α deve usar para comparar com α^*.

6. Conclusão do teste · ilustração do caso unilateral

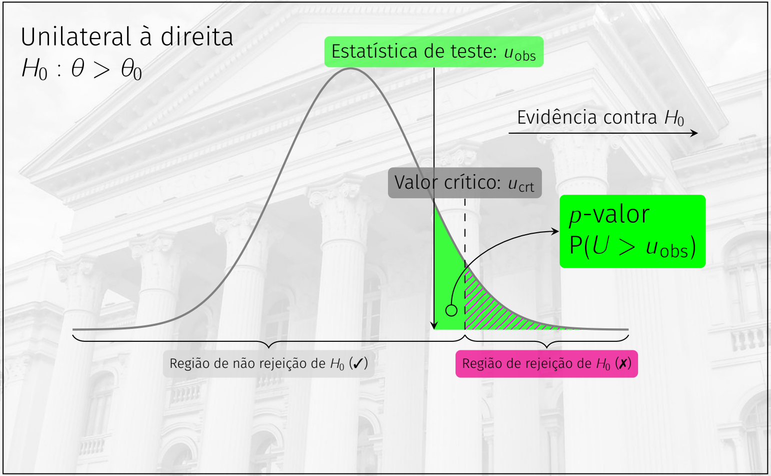

Figura 12. Nível descritivo para um teste com hipótese unilateral à direita.

6. Conclusão do teste · ilustração do caso bilateral

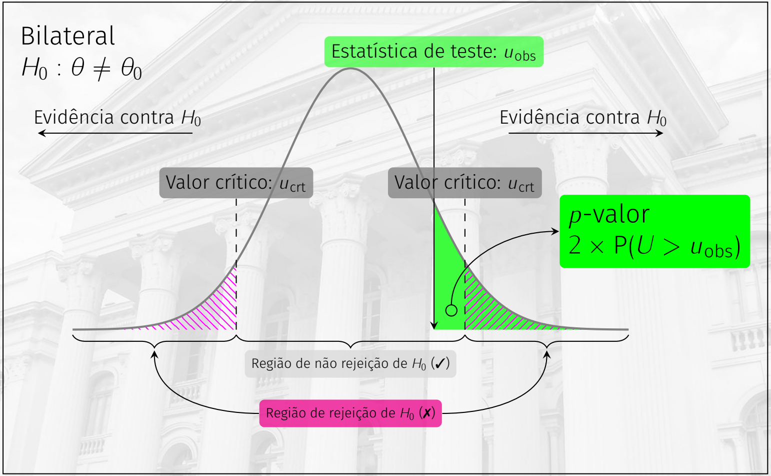

Figura 13. Nível descritivo para um teste com hipótese bilateral.

6. Conclusão do teste · formalização do p-valor

Com base no nível descritivo ou p-valor

Para **testes unilaterais**, sendo $H_0 : \theta = \theta_0$, a expressão de α^* depende da hipótese alternativa:

$$\alpha^* = \text{P}(U < u_{\text{obs}} \,|\, H_0 \text{ verdadeira}) \quad \text{para } H_a : \theta < \theta_0$$
$$\alpha^* = \text{P}(U > u_{\text{obs}} \,|\, H_0 \text{ verdadeira}) \quad \text{para } H_a : \theta > \theta_0,$$

em que U é a estatística de teste, u_{obs} o seu valor observado.

Para **testes bilaterais**, temos $H_0 : \theta = \theta_0$ contra $H_0 : \theta \neq \theta_0$, a definição do nível descritivo depende da relação entre u_{obs} e θ_0:

$$\alpha^* = 2 \times \text{P}(U < u_{\text{obs}} \,|\, H_0 \text{ verdadeira}) \quad \text{se } u_{\text{obs}} < \theta_0$$
$$\alpha^* = 2 \times \text{P}(U > u_{\text{obs}} \,|\, H_0 \text{ verdadeira}) \quad \text{se } u_{\text{obs}} > \theta_0.$$

Como estamos calculando a probabilidade para apenas uma das caudas, então esse valor é multiplicado por 2.

Relação de teste de hipótese com intervalo de confiança

Seja $\text{IC}_{1-\alpha}(\theta)$ o **intervalo de confiança** de $100(1-\alpha)\%$ para o parâmetro θ.

O **teste de hipótese** com nível de significância α para

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_a : \theta \neq \theta_0$$

conduzirá a rejeição de H_0, se e somente se, θ_0 **não** estiver contido no $\text{IC}_{1-\alpha}(\theta)$.

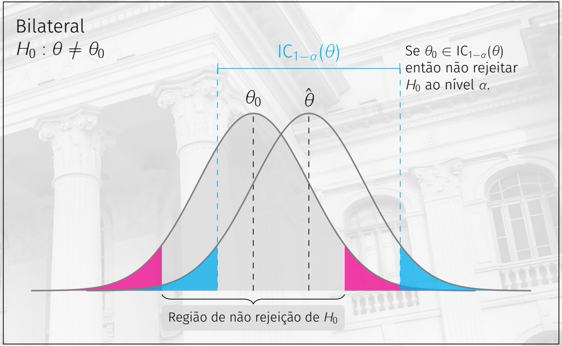

Figura 14. Relação entre teste de hipótese e intervalo de confiança.

Testes de hipótese para médias, proporções e variâncias

Testes de hipótese para uma população

Quando queremos fazer inferência para um parâmetro de uma única população.

- Testar se a média de altura dos estudantes da UFPR é igual a 170 cm ($\mu = 170$ de uma distribuição normal).
  - Testes para a **média** de uma população.
    - σ^2 conhecido.
    - σ^2 desconhecido.
- Testar se a proporção de peixes fêmeas em uma lagoa é de 50% ($p = 0.5$ de uma distribuição Bernoulli).
  - Teste para a **proporção** de uma população.
- Testar se a variabilidade do diâmetro de um lote de parafusos se mantém em torno de 0.02 mm ($\sigma = 0.02$ de uma distribuição normal).
  - Teste para a **variância** de uma população.

Testes de hipótese para duas populações

Quando queremos comparar parâmetros de duas populações.

- Testar se IRA dos estudantes da UFPR difere entre alunos e alunas ($\mu_M = \mu_F$ de distribuições normais).
  - Testes para comparar **médias** de duas populações.
    - σ^2 conhecidos.
    - σ^2 desconhecido(s) (pareadas/independentes; iguais/diferentes).
- Testar se as proporções de pacientes recuperados são distintas entre dois tratamentos ($p_1 = p_2$ de distribuições Bernoulli).
  - Teste para comparar duas **proporções**.
- Testar se a variabilidade do diâmetro parafusos difere entre dois fornecedores ($\sigma_1^2 = \sigma_2^2$ de duas distribuições normais).
  - Teste para comparar **variâncias** de uma população.

Procedimentos gerais para um teste de hipótese

Em qualquer tipo de teste, os passos serão sempre os mesmos

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir um nível de significância α, que irá determinar o nível de confiança $100(1 - \alpha)\%$ do teste.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a **estatística de teste**, com base na distribuição amostral da mesma sob a hipótese nula $\rightarrow$ valor calculado.
5. Determinar a região crítica (região de rejeição), sob hipótese nula, com base no nível de significância $\alpha \rightarrow$ valor crítico.
6. Conclusão.

Testes de hipótese para a média μ: (σ^2 conhecido)

Condições para o teste

Quando temos os seguintes requisitos:

- Temos uma AAS.
- σ^2 é conhecido.
- A população tem distribuição Normal ou $n > 30$.

Podemos usar o Teorema do Limite Central para afirmar que a média segue distribuição Normal, e a **estatística de teste** é dada por

$$z = \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}},$$

em que μ_0 é o valor de teste sob a hipótese nula.

Etapas do teste

Procedimentos gerais para um teste de hipótese para a média μ com σ^2 conhecido:

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste
$$z = \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}}.$$
5. Determinar a região crítica (região de rejeição), com base no nível de significância α.
6. Conclusões.

Exercício: enchimento de embalagens

- Uma máquina de encher embalagens de café está funcionando adequadamente se colocar 700 g em cada embalagem.
- Para verificar a calibração da máquina, foi coletada uma amostra de 40 embalagens, que resultou em uma média de 698 g.
- Sabe-se que o desvio-padrão do enchimento da máquina é de 10 g.
- Teste a hipótese de o peso médio das embalagens na população ser 700 g, com um nível de significância de 5%.

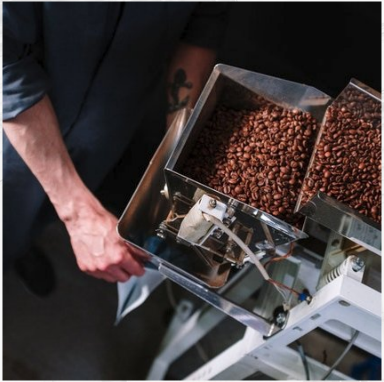

Figura 15. Foto de cottonbro no Pexels.

Solução

1. Definir a hipótese nula (H_0) e a alternativa (H_a): $H_0 : \mu = 700 \quad \text{vs} \quad H_a : \mu \neq 700$.
2. Definir o nível de significância: $\alpha = 0.05$.
3. Definir o tipo de teste, com base na hipótese alternativa: *teste bilateral*.
4. Calcular a estatística de teste

$$z = \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}} = \frac{698 - 700}{10/\sqrt{40}} = -1.265.$$

5. Determinar a região crítica (região de rejeição), com base no nível de significância α.

$$\text{RC} = \{z < -1.96 \text{ ou } z > 1.96\}.$$

6. Conclusão: $z \notin \text{RC}$, portanto **não rejeita** H_0.

Nível descritivo do teste

p-valor do teste:

$$2 \times P(Z < -1.265) = 2 \times 0.103 = 0.206.$$

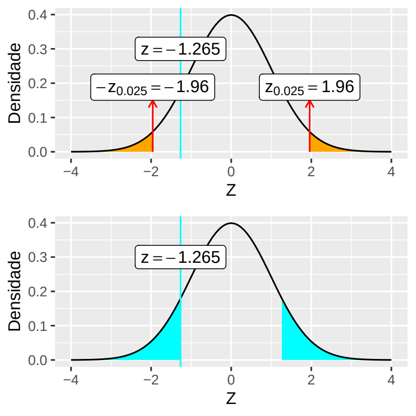

Figura 16. Região de rejeição da hipótese nula e nível descritivo.

Exercício: resistência das lajotas

- Um fabricante de lajotas introduz um novo material em sua fabricação e acredita que aumentará a resistência média, que é de 206 kg.
- A resistência das lajotas tem distribuição normal com desvio-padrão de 12 kg.
- Retirou-se uma amostra de 30 lajotas, e obteve-se uma média amostral de 210 kg.
- Ao nível de 10%, pode o fabricante afirmar que a resistência média de suas lajotas tenha aumentado?

Figura 17. Foto de Rodolfo Quirós no Pexels.

Solução

1. Hipóteses: $H_0 : \mu = 206 \quad \text{vs} \quad H_a : \mu > 206$ (teste unilateral à direita).
2. Estatística de teste

$$z = \frac{\overline{y} - \mu_0}{\sigma/\sqrt{n}} = \frac{210 - 206}{12/\sqrt{30}} = 1.826.$$

3. Nível de significância $\alpha = 0.1 \rightarrow \text{RC} = \{z > 1.282\}$.
4. Conclusão do teste:
   - $z \in \text{RC}$, portanto **rejeita** H_0.
   - **p-valor** $= \text{P}(Z > 1.826) = 0.034$: probabilidade muita baixa de valor tão ou mais extremo que a média amostral ocorrer por acaso.
   Portanto, existem evidências de que a resistência média das lajotas tenha aumentado.

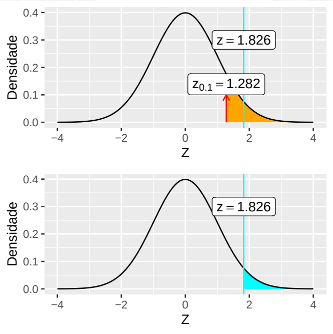

Figura 18. Região de rejeição da hipótese nula e nível descritivo.

Testes de hipótese para a média μ: (σ^2 desconhecido)

Condições para o teste

Quando temos os seguintes requisitos:

- Temos uma AAS.
- σ^2 é **desconhecido**.
- A população tem distribuição Normal ou $n > 30$.

Podemos usar o Teorema do Limite Central para afirmar que a média segue uma distribuição Normal, e a **estatística de teste** é dada por

$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}} \sim t_\nu,$$

com $\nu = n - 1$ graus de liberdade, e onde μ_0 é o valor de teste na hipótese nula.

Etapas do teste

Procedimentos gerais para um teste de hipótese para a média μ com σ desconhecido:

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste
$$t = \frac{\overline{y} - \mu_0}{s/\sqrt{n}}.$$
5. Determinar a região crítica (região de rejeição), com base no nível de significância α.
6. Conclusões.

Exercício: uso do cartão de crédito

Um estudo foi idealizado para estimar a média anual dos débitos de cartão de crédito da população de famílias brasileiras. Uma amostra de $n = 15$ famílias forneceu média de saldos de R\$ 5200.00 e o desvio padrão foi de R\$ 3058.00.

1. Obtenha um intervalo com 95% de confiança.
2. Teste a hipótese de que a média anual de débitos é de R\$ 6000.00, com o mesmo nível de confiança.

Figura 19. Foto de Norma Mortenson no Pexels.

Solução

O intervalo de confiança é obtido por

$$IC_{1-0.95}(\mu) = \left(5200 \pm 2.145 \cdot \frac{3058}{\sqrt{15}}\right) \approx (3506.4, 6893.6).$$

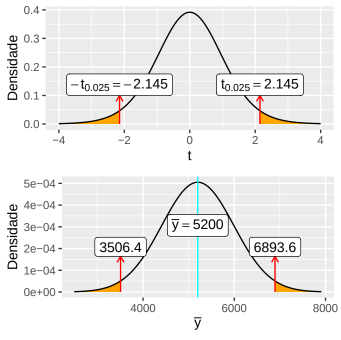

Figura 20. Quantis da distribuição t-Student e intervalo de confiança para a média.

Solução

1. Hipóteses: $H_0 : \mu = 6000 \quad \text{vs} \quad H_a : \mu \neq 6000$ (teste bilateral).
2. Estatística de teste

$$t = \frac{5200 - 6000}{3058/\sqrt{15}} = 1.013.$$

3. Nível de significância $\alpha = 0.05$

$$\text{RC} = \{t < -2.145 \text{ ou } t > 2.145\}.$$

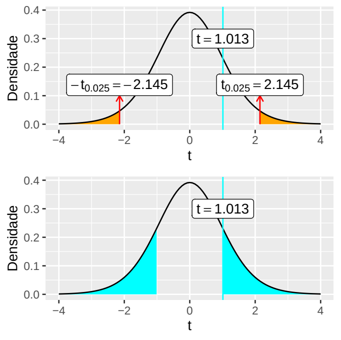

Figura 21. Resultado do teste de hipótese.

Solução

4. Conclusão do teste:
   - $t \notin$ RC, portanto **não rejeita** H_0.
   - **p-valor** $= 2 \times P(T > 1.013) = 2 \times 0.164 = 0.328$: probabilidade alta de ocorrência de um valor médio tão ou mais extremo do que o obtido nesta amostra (assumindo que o valor populacional é de R$ 6000.00).
     Portanto, não existe evidência de que a média de débitos seja diferente de R$ 6000.00.

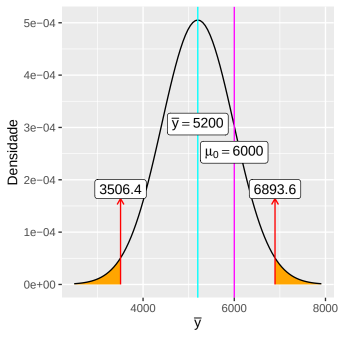

Figura 22. Intervalo de confiança para a média.

Note que o $IC_{1-\alpha}(\mu)$ contém o valor sob hipótese nula, ou seja, $H_0 : \mu = 6000$.

Testes de hipótese para a proporção p

A distribuição amostral

Se $Y \sim \text{Ber}(p)$, então a proporção amostral

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

é a **melhor estimativa** para a proporção populacional p.

Já vimos que, quando ambas condições são satisfeitas,

- $np \geq 5$
- $n(1-p) \geq 5$,

a distribuição amostral de $\hat{p}$ pode ser aproximada (pelo TLC) por

$$\hat{p} \overset{\text{aprox}}{\sim} \text{N}\left(p, \frac{p(1-p)}{n}\right).$$

Condições para o teste

Quando temos os seguintes requisitos:

- Temos uma AAS.
- As condições para a distribuição Binomial são satisfeitas:
  - As tentativas são independentes.
  - Há duas categorias de resultado ("sucesso", "fracasso").
  - A probabilidade de sucesso p permanece constante.
- $np_0 \geq 5$ e $n(1 - p_0) \geq 5$.

Podemos usar a distribuição Normal como aproximação da Binomial e, portanto, usamos a **estatística de teste**

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

em que p_0 é o valor de proporção de teste na hipótese nula.

Etapas do teste

Procedimentos gerais para um teste de hipótese para a proporção p

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}.$$
5. Determinar a região crítica (região de rejeição), com base no nível de significância α.
6. Conclusão.

Exemplo

- Uma empresa desenvolveu uma nova vacina para uma doença, e afirma que a proporção de imunizados é maior do que 50%.
- Em uma amostra de 726 pessoas que tomaram a vacina, 668 estavam imunizadas.
- Use este resultado, com um nível de significância de 5%, para testar a afirmativa de que a proporção de imunizados é maior do que 50%.

Figura 23. Foto de cottonbro no Pexels.

Solução

1. Hipóteses: $H_0 : p = 0.5 \quad \text{vs} \quad H_a : p > 0.5$ (teste unilateral à direita).
2. Estatística de teste

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.92 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{726}}} = 22.633.$$

3. Nível de significância $\alpha = 0.05 \rightarrow \text{RC} = \{z > 1.645\}$.
4. Conclusão do teste:
   - $z \in \text{RC}$, portanto **rejeita** H_0.
   - **p-valor** $= \text{P}(Z > 22.633) \approx 0$: a probabilidade de ocorrência **ao acaso** de uma proporção tão ou mais extrema do que essa é praticamente nula. Portanto, existem fortes evidências de que a vacina imuniza mais que 50%.

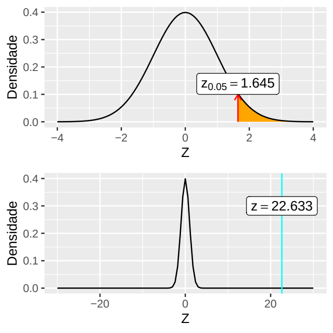

Figura 24. Região de rejeição da hipótese nula e nível descritivo.

Exemplo (cont.)

- Ainda no contexto da vacina, acredita-se que, devido ao seu alto custo, seu uso só seria viável se pelo menos 90% das pessoas forem imunizadas.
- Neste caso, qual seria a decisão sobre a adoção da vacina, com um nível de significância de 5%.

Figura 25. Foto de cottonbro no Pexels.

Solução

1. Hipóteses: $H_0 : p = 0.9 \quad \text{vs} \quad H_a : p > 0.9$ (teste unilateral à direita).
2. Estatística de teste

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.92 - 0.9}{\sqrt{\frac{0.9(1-0.9)}{726}}} = 1.986.$$

3. Nível de significância $\alpha = 0.05 \rightarrow \text{RC} = \{z > 1.645\}$.
4. Conclusão do teste:
   - $z \in \text{RC}$, portanto **rejeita** H_0.
   - **p-valor** $= \text{P}(Z > 1.986) = 0.024$: portanto, existe evidência, a este nível de significância, de que a adoção da vacina seria viável.

OBS: E se fosse adotado $\alpha = 0.01$ ($\text{RC} = \{z > 2.326\}$)?

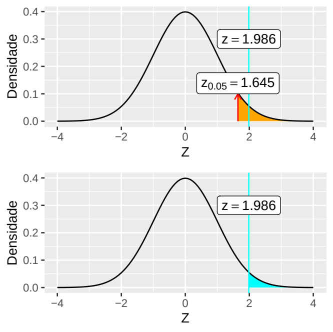

Figura 26. Região de rejeição da hipótese nula e nível descritivo.

Teste de hipótese para a variância σ^2

Condições para o teste

Quando temos os seguintes requisitos:

- Temos uma AAS.
- A população tem distribuição Normal (essa é uma exigência mais estrita).

Então, usamos a **estatística de teste**

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2},$$

em que σ_0^2 é o valor de variância na hipótese nula.

Etapas do teste

Procedimentos gerais para um teste de hipótese para a variância σ^2

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste
$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}.$$
5. Determinar a região crítica (região de rejeição), com base no nível de significância α.
6. Conclusão.

Exemplo: concentração de princípio ativo

Na indústria, baixa variabilidade é sinônimo de qualidade. Para isso, constantemente monitora-se e corrige-se a produção para manter níveis aceitáveis de qualidade.

Uma amostra de frascos de medicamento foi inspecionada em relação a concentração (m/m) de princípio ativo na solução. O lote é rejeitado se claramente ultrapassar o limite de $\sigma^2 = 0.0009$. Os dados estão abaixo.

0.15	0.18	0.18	0.20	0.21	0.22	0.25	0.27
0.17	0.18	0.19	0.20	0.21	0.22	0.26	0.27

Faça um teste para verificar se a variância é maior do que 0.0009, com $\alpha = 5\%$.

Figura 27. Foto de Karolina Grabowska no Pexels.

Solução

1. Hipóteses: $H_0 : \sigma^2 = 0.0009 \quad \text{vs} \quad H_a : \sigma^2 > 0.0009$ (teste unilateral à direita).
2. Estatística de teste

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(16-1)0.0013}{0.0009} = 21.667.$$

3. Nível de significância $\alpha = 0.05 \rightarrow \text{RC} = \{\chi^2 > 24.996\}$.
4. Conclusão do teste:
   - $\chi^2 \notin \text{RC}$, portanto **não rejeita** H_0.
   - **p-valor** $= \text{P}(\chi^2 > 21.667) = 0.117$: portanto, ao nível de 5% de significância, **não rejeita-se** a hipótese de que a variância seja igual a 0.0009.

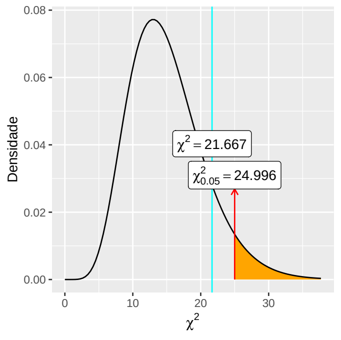

Figura 28. Região de rejeição da hipótese nula.

Testes para comparar médias de duas populações

Testes para comparar duas médias

Exemplo: comparação de IRA entre alunos e alunas de uma universidade

Diferentes possibilidades de testes:

- **Testes de hipótese para a diferença de médias de duas populações: σ^2's conhecidos**.
- **Testes de hipótese para a diferença de médias de duas populações: σ^2's desconhecidos**.
  - **Testes de hipótese para amostras emparelhadas**.
  - **Testes de hipótese para amostras independentes**.
    - **Variâncias iguais** $\sigma_1^2 = \sigma_2^2$.
    - **Variâncias diferentes** $\sigma_1^2 \neq \sigma_2^2$.

Testes de hipótese para duas populações

Ao testar uma hipótese para duas populações, devem ser consideradas

- **Amostras independentes**: quando os valores amostrados de uma população **não estão relacionados** ou emparelhados com os da outra população.
  - Exemplo: teste para pressão sanguínea do grupo controle vs grupo medicado.
- **Amostras dependentes ou emparelhadas**: quando cada elemento de uma amostra corresponde ao mesmo elemento da outra amostra (geralmente o mesmo indivíduo analisado antes e depois de um experimento).
  - Exemplo: teste para a diferença de peso **de uma mesma pessoa** antes e depois de uma dieta.

Testes de hipótese para médias de duas populações

Até agora testamos hipóteses para uma único parâmetro populacional

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0 \quad \text{ou} \quad H_a : \mu < \mu_0 \quad \text{ou} \quad H_a : \mu > \mu_0.$$

Podemos estender o teste de hipótese quando queremos comparar o mesmo parâmetro para **duas populações diferentes**.

Em geral, faremos testes para verificar se a diferença entre estes dois parâmetros é igual a zero

$$H_0 : \mu_1 - \mu_2 = 0 \quad \Rightarrow \quad H_0 : \mu_1 = \mu_2 \qquad \text{vs} \qquad \begin{aligned} H_a &: \mu_1 - \mu_2 \neq 0 & \Rightarrow \quad & H_a : \mu_1 \neq \mu_2 \quad \text{ou} \\ H_a &: \mu_1 - \mu_2 < 0 & \Rightarrow \quad & H_a : \mu_1 < \mu_2 \quad \text{ou} \\ H_a &: \mu_1 - \mu_2 > 0 & \Rightarrow \quad & H_a : \mu_1 > \mu_2. \end{aligned}$$

Testes de hipótese para a diferença de médias de duas populações: σ^2 conhecido

Distribuição amostral da diferença

Considere duas populações Y_1 e Y_2 com médias μ_1 e μ_2, e desvios-padrão σ_1 e σ_2, ou seja

$$Y_1 \sim \mathrm{N}(\mu_1, \sigma_1^2) \qquad \text{e} \qquad Y_2 \sim \mathrm{N}(\mu_2, \sigma_2^2).$$

A “**nova**” variável $\overline{Y}_d = (\overline{Y}_1 - \overline{Y}_2)$ também possui distribuição normal com

$$\begin{aligned} \mathrm{E}(\overline{Y}_1 - \overline{Y}_2) &= \mu_1 - \mu_2 \\ \mathrm{V}(\overline{Y}_1 - \overline{Y}_2) &= \mathrm{V}(\overline{Y}_1) + \mathrm{V}(\overline{Y}_2) \\ &= \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}, \end{aligned}$$

ou seja, a **distribuição amostral da diferença de médias é**

$$\overline{Y}_d = (\overline{Y}_1 - \overline{Y}_2) \sim \mathrm{N}\left(\mu_1 - \mu_2, \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right).$$

Condições para o teste

Quando temos os seguintes requisitos:

- Ambas amostras são AAS.
- Ambas amostras são independentes.
- Ambas populações tem distribuição normal ou $n_1 > 30$ e $n_2 > 30$.

Podemos usar o Teorema do Limite Central para afirmar que a diferença entre as duas médias segue uma distribuição normal, e a **estatística de teste** é dada por

$$z = \frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

Etapas do teste

Procedimentos gerais

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste

$$z = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

5. Determinar a região crítica (região de rejeição), com base no nível de significância α.
6. Conclusão do teste.

Exercício: tempos de entrega

Uma transportadora de mercadorias tem duas possibilidades de trajeto para realizar entregas. O gerente de logística *desconfia não haver diferença significativa* entre o tempo de cada trajeto.

Foram selecionadas aleatoriamente 45 entregas realizadas no **primeiro** trajeto, resultando em uma média amostral de 57 minutos. No **segundo** trajeto, foram selecionadas aleatoriamente 30 entregas, e o tempo médio foi de 54 minutos.

O desvio-padrão populacional do primeiro trajeto é de $\sigma_1 = 8$ minutos, e o do segundo trajeto é de $\sigma_2 = 6$ minutos. Teste a hipótese de que não existe diferença significativa entre o tempo médio dos dois trajetos, ao nível de 1% de significância.

Figura 29. Foto de Norma Mortenson no Pexels.

Solução

1. Hipóteses: $H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_a : \mu_1 \neq \mu_2$ (teste bilateral).
2. Estatística de teste

$$z = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{57 - 54}{\sqrt{\frac{8^2}{45} + \frac{6^2}{30}}} = 1.311.$$

3. Nível de significância
   $\alpha = 0.01 \rightarrow \text{RC} = \{z < -2.576 \text{ ou } z > 2.576\}$.
4. Conclusão do teste:
   - $z \notin \text{RC}$, portanto **não rejeita** H_0.
   - **p-valor** $= 2 \times \text{P}(Z > 1.311) = 0.19$: não existem evidências para rejeitar a hipótese de que os tempos dos trajetos sejam iguais.

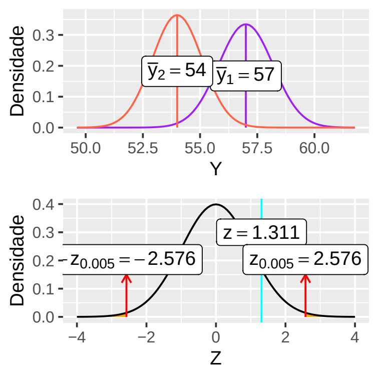

Figura 30. Distribuição amostral das médias e região de rejeição da hipótese nula.

Testes de hipótese para a diferença de médias de duas populações: σ^2 desconhecido

Suposições sobre as variâncias

- Quando não conhecemos σ^2, usamos a **estimativa** amostral s^2.
- Nesse caso, já vimos que usamos a **distribuição** t **no lugar da distribuição** z.
- No entanto, quando temos duas amostras, devem ser considerados dois casos distintos:
  - **Variâncias iguais:** quando é razoável supor que as variâncias populacionais são iguais, ou seja $\sigma_1^2 = \sigma_2^2$.
  - **Variâncias diferentes:** quando não se pode fazer nenhuma suposição sobre a igualdade das variâncias populacionais, ou seja $\sigma_1^2 \neq \sigma_2^2$.

Estatística do teste para o caso de variâncias iguais $\sigma_1^2 = \sigma_2^2$

Neste caso, calculamos a **média ponderada das variâncias amostrais** s_1^2 e s_2^2 para obter uma estimativa da variância populacional comum

$$\hat{s}^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2}.$$

A estatística de teste fica

$$t = \frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\hat{s}^2}{n_1} + \frac{\hat{s}^2}{n_2}}} = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{\hat{s}^2}{n_1} + \frac{\hat{s}^2}{n_2}}} \sim t_\nu,$$

em que

$$\nu = n_1 + n_2 - 2$$

são os graus de liberdade.

Etapas do teste

Procedimentos gerais

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste

$$t = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{\hat{s}^2}{n_1} + \frac{\hat{s}^2}{n_2}}} = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\hat{s}^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}},$$

   em que $\hat{s}^2$ é a variância combinada das amostras.
5. Determinar a região crítica (região de rejeição), com base no nível de significância α (**Obs:** use $\nu = n_1 + n_2 - 2$).
6. Conclusão do teste.

Exercício: rendimento das turmas

- Em uma avaliação de estatística, foi selecionada uma amostra de 12 alunos da turma A, resultando em uma média de 7.9 com desvio-padrão 0.6.
- Na turma B, foram selecionados 15 alunos, os quais tiraram nota média 6.7 com desvio-padrão 0.8.
- As notas possuem distribuição normal e assume-se que $\sigma_1^2 = \sigma_2^2$.
- Teste a hipótese de que a turma A tem média maior do que a turma B, com um nível de significância de 1%.

Figura 31. Foto de Kaboompics.com do Pexels.

Solução

1. Hipóteses: $H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_a : \mu_1 > \mu_2 = \mu_1 - \mu_2 > 0$ (unilateral à direita).
2. Variância combinada

$$\hat{s}^2 = \frac{(n_1 - 1) \cdot s_1^2 + (n_2 - 1) \cdot s_2^2}{n_1 + n_2 - 2} = \frac{(12 - 1) \cdot 0.6^2 + (15 - 1) \cdot 0.8^2}{12 + 15 - 2} = 0.517.$$

3. Estatística de teste

$$t = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\hat{s}^2 \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{7.9 - 6.7}{\sqrt{0.517 \left(\frac{1}{12} + \frac{1}{15}\right)}} = 4.309.$$

4. Nível de significância $\alpha = 0.01 \rightarrow \text{RC} = \{t > 2.485\}$.

Solução

4. Conclusão do teste:
   - $t \in$ RC, portanto **rejeita** H_0.
   - **p-valor** $= \text{P}(T > 4.309) = 0.0001$.

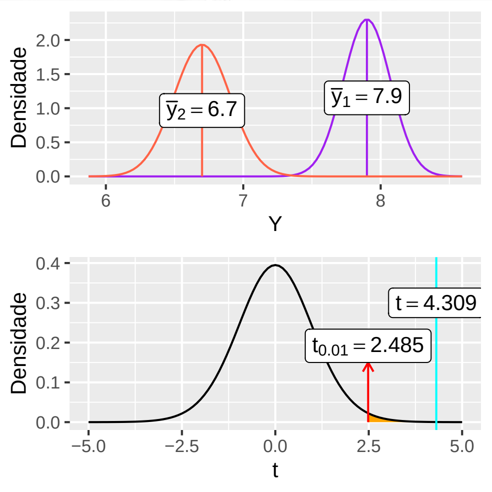

Figura 32. Distribuição amostral das médias e região de rejeição da hipótese nula.

Estatística do teste para o caso de variâncias diferentes $\sigma_1^2 \neq \sigma_2^2$

Neste caso, ainda usamos as variâncias amostrais s_1^2 e s_2^2 para determinar o **erro-padrão da diferença** entre as duas médias.

A estatística de teste fica

$$t = \frac{(\overline{y}_1 - \overline{y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_{\nu}.$$

Porém, como as variâncias são diferentes, os **graus de liberdade** devem ser ajustados

$$\nu = \frac{(w_1 + w_2)^2}{\frac{w_1^2}{n_1 - 1} + \frac{w_2^2}{n_2 - 1}},$$

em que

$$w_1 = \frac{s_1^2}{n_1} \quad \text{e} \quad w_2 = \frac{s_2^2}{n_2}.$$

Etapas para o teste

Procedimentos gerais

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste

$$t = \frac{\overline{y}_1 - \overline{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t_\nu .$$

5. Determinar a região crítica (região de rejeição), com base no nível de significância α (**Obs**: o valor de graus de liberdade ν deve ser calculado conforme equação anterior).
6. Conclusão do teste.

Exemplo: tempo de uma tarefa doméstica

- Uma pesquisa avaliou a eficácia de dois tipos de treinamento, com a finalidade de reduzir o tempo médio de determinada tarefa doméstica.
- Foram selecionadas duas amostras aleatórias de populações com distribuição Normal, onde assume-se que $\sigma_1^2 \neq \sigma_2^2$.
- Os dados da pesquisa estão no quadro abaixo. Teste a hipótese de que o tempo médio para a realização da tarefa é igual para os dois treinamentos, ao nível de 5% de significância.

Treinamento 1	$n_1 = 15$	$\overline{y}_1 = 24.2$ min	$s_1 = 3.16$ min
Treinamento 2	$n_2 = 10$	$\overline{y}_2 = 23.9$ min	$s_2 = 4.47$ min

Figura 33. Foto de cottonbro no Pexels.

Solução

1. Hipóteses: $H_0 : \mu_1 = \mu_2 \quad \text{vs} \quad H_a : \mu_1 \neq \mu_2$ (teste bilateral).
2. Estatística de teste

$$t = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{24.2 - 23.9}{\sqrt{\frac{3.16^2}{15} + \frac{4.47^2}{10}}} = 0.184.$$

3. Nível de significância $\alpha = 0.05$.
4. Graus de liberdade:

$$w_1 = \frac{s_1^2}{n_1} = \frac{3.16^2}{15} = 0.666 \quad \text{e} \quad w_2 = \frac{s_2^2}{n_2} = \frac{4.47^2}{10} = 1.998.$$

$$\nu = \frac{(w_1 + w_2)^2}{\frac{w_1^2}{n_1 - 1} + \frac{w_2^2}{n_2 - 1}} = \frac{(0.666 + 1.998)^2}{\frac{0.666^2}{15 - 1} + \frac{1.998^2}{10 - 1}} = 14.933.$$

Solução

3. Nível de significância $\alpha = 0.05 \rightarrow$ RC $= \{t < -2.132$ ou $t > 2.132\}$.
4. Conclusão do teste:
   - $t \notin$ RC, portanto **não rejeita** H_0.
   - **p-valor** $= 2 \times P(T > 0.184) = 0.856$.

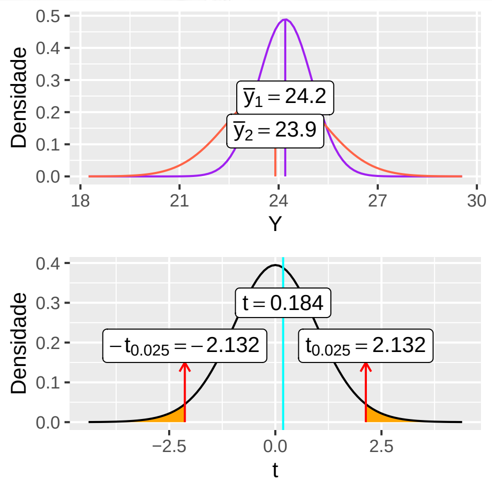

Figura 34. Distribuição amostral das médias e região de rejeição da hipótese nula.

Testes de hipótese para a diferença de médias de duas populações: σ desconhecido e amostras emparelhadas

Amostras emparelhadas

Fazemos testes de comparação de médias para **dados emparelhados** quando os resultados das duas amostras são **relacionados** de acordo com algum critério.

Para cada **par** (Y_{1i}, Y_{2i}), o valor da primeira amostra deve estar claramente associado ao valor da segunda amostra (estudos do tipo "antes" e "depois").

Este teste verifica se o processo ao qual os indivíduos em estudo foram submetidos **produziu alguma alteração**.

Exemplos:

- Influência de uma nova dieta sobre os mesmos indivíduos.
- Influência de uma campanha publicitária sobre a intenção de compra do consumidor.
- Influência de hábitos de saúde acompanhando pares de gêmeos.

Distribuição amostral da diferença

Ao invés de analisarmos cada grupo separadamente, observamos somente a **diferença** D_i entre as duas amostras Y_{1i} e Y_{2i},

$$D_i = Y_{1i} - Y_{2i},$$

e calculamos a média destas diferenças

$$\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i,$$

que terá distribuição

$$\overline{D} \sim \mathrm{N}\left(\mu_D, \frac{\sigma_D^2}{n}\right).$$

O parâmetro μ_D é estimado pela média amostral $\overline{D}$, e como usualmente não temos informações sobre σ_D^2, estimamos seu valor por s_D^2.

Média e variância

Além da média das diferenças,

$$\overline{d} = \frac{1}{n}\sum_{i=1}^{n} d_i,$$

precisamos calcular também a **variância das diferenças** entre os pares, dado por

$$s_d^2 = \frac{1}{n-1}\sum_{i=1}^{n}(d_i - \overline{d})^2$$
$$= \frac{1}{n-1}\left[\sum_{i=1}^{n} d_i^2 - n\overline{d}^2\right].$$

i	y_1	y_2	d
1	y_{11}	y_{21}	$d_1 = y_{11} - y_{21}$
2	y_{12}	y_{22}	$d_2 = y_{12} - y_{22}$
3	y_{13}	y_{23}	$d_3 = y_{13} - y_{23}$
4	y_{14}	y_{24}	$d_4 = y_{14} - y_{24}$
5	y_{15}	y_{25}	$d_5 = y_{15} - y_{25}$
6	y_{16}	y_{26}	$d_6 = y_{16} - y_{26}$
7	y_{17}	y_{27}	$d_7 = y_{17} - y_{27}$
8	y_{18}	y_{28}	$d_8 = y_{18} - y_{28}$
9	y_{19}	y_{29}	$d_9 = y_{19} - y_{29}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	y_{1n}	y_{2n}	$d_n = y_{1n} - y_{2n}$

Estatística do teste

Uma vez que a diferença média é calculada com base nas diferenças entre amostras **emparelhadas** (isto é, σ^2 é **desconhecido**), e que os valores de D_i geralmente tem distribuição normal, **usamos a distribuição** t, com **estatística de teste** dada por

$$t = \frac{\overline{d} - \mu_d}{s_d/\sqrt{n}} = \frac{\overline{d}}{s_d/\sqrt{n}} \sim t_{\nu},$$

em que

- $\nu = n - 1$, com n sendo o **número de pares** observados.
- μ_d é a média das diferenças na população (normalmente $\mu_d = 0$).

Observação para H_a

Normalmente $H_0 : \mu_d = 0$ e

- $\mu_d < 0$ significa que houve **aumento** “depois”.
- $\mu_d > 0$ significa que houve **diminuição** “depois”.

Etapas do teste

Procedimentos gerais

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste
$$t = \frac{\overline{d} - \mu_d}{s_d/\sqrt{n}}.$$
5. Determinar a região crítica (região de rejeição), com base no nível de significância α (**Obs:** $\nu = n - 1$, sendo n o **número de pares** observados).
6. Conclusão do teste.

Exemplo: manutenção preventiva

- Em uma fábrica, sete máquinas foram selecionadas aleatoriamente a fim de determinar o **efeito da manutenção preventiva** na produção.
- Inicialmente, as máquinas trabalharam por um período na forma habitual, e **depois** trabalham o mesmo período recebendo manutenções preventivas.
- O total de trabalho produzido, **antes** e **depois** da adoção das manutenções, está na tabela ao lado.
- Ao nível de 5%, podemos concluir que o trabalho médio produzido **é maior depois** da adoção das manutenções preventivas?

i	Antes	Depois	Diferença
1	12.10	12.50	−0.40
2	12.30	16.00	−3.70
3	11.10	12.90	−1.80
4	12.80	14.00	−1.20
5	14.10	12.90	1.20
6	8.40	12.50	−4.10
7	13.30	13.50	−0.20

Solução

1. Hipóteses: $H_0 : \mu_d = 0 \quad \text{vs} \quad H_a : \mu_d < 0$ (unilateral à esquerda).
2. Estatística de teste

$$t = \frac{\overline{d} - \mu_d}{s_d/\sqrt{n}} = \frac{-1.457 - 0}{1.913/\sqrt{7}} = -2.015.$$

3. Nível de significância
$\alpha = 0.05 \rightarrow \text{RC} = \{t < -1.943\}$.
4. Conclusão do teste:
   - $t \in \text{RC}$, portanto **rejeita** H_0.
   - **p-valor** $= \text{P}(T < -2.015) = 0.045$: existem evidências de que o tempo médio de funcionamento das máquinas é maior quando recebem manutenções preventivas.

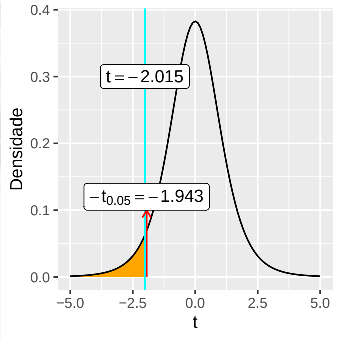

Figura 35. Resultado do teste de hipótese.

Teste de hipótese para a diferença de proporção de duas populações

Hipóteses

Se a amostra for suficientemente grande, sabemos, pelo **Teorema do Limite Central**, que a distribuição de probabilidade da proporção amostral tem um comportamento aproximadamente **Normal**.

Na **comparação** de duas proporções populacionais p_1 e p_2, usaremos como estimador a diferença entre as respectivas proporções amostrais, $\hat{p}_1$ e $\hat{p}_2$.

Supondo que duas amostras foram retiradas de duas populações **independentes**, teremos duas proporções amostrais independentes, e a **diferença entre elas** também terá distribuição aproximadamente **Normal**.

Assim, o interesse será em testar

$$
\begin{array}{lcll}
H_0 : p_1 - p_2 = 0 & \Rightarrow & H_0 : p_1 = p_2 & \\
 & \text{vs} & & \\
H_a : p_1 - p_2 \neq 0 & \Rightarrow & H_a : p_1 \neq p_2 & \text{ou} \\
H_a : p_1 - p_2 < 0 & \Rightarrow & H_a : p_1 < p_2 & \text{ou} \\
H_a : p_1 - p_2 > 0 & \Rightarrow & H_a : p_1 > p_2. &
\end{array}
$$

Distribuição amostral

Desse modo, o estimador a ser utilizado será $\hat{p}_1 - \hat{p}_2$, cuja distribuição será aproximada pela **Normal**, com parâmetros

$$
\begin{aligned}
E(\hat{p}_1 - \hat{p}_2) &= p_1 - p_2 \\
V(\hat{p}_1 - \hat{p}_2) &= V(\hat{p}_1) + V(\hat{p}_1) \\
&= \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2},
\end{aligned}
$$

ou seja,

$$
(\hat{p}_1 - \hat{p}_2) \overset{\text{aprox}}{\sim} N\left(p_1 - p_2, \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}\right).
$$

Teste de hipótese para a proporção de duas populações

Se a hipótese nula for verdadeira, as proporções populacionais são iguais. Denotando seu **valor comum** por $\overline{p}$, temos

$$p_1 = p_2 = \overline{p}.$$

Podemos obter um **estimador** para $\overline{p}$, através da ponderação dos estimadores $\hat{p}_1$ e $\hat{p}_2$, obtendo a proporção **combinada**

$$\overline{p} = \frac{y_1 + y_2}{n_1 + n_2},$$

em que y_1 e y_2 são os números de sucessos em cada amostra.

Substituindo os valores de p_1 e p_2 na expressão da $\mathrm{V}(\hat{p}_1 - \hat{p}_2)$, temos que a **estatística de teste** para a diferença de duas proporções é

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\overline{p}(1-\overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \overset{\text{aprox}}{\sim} \mathrm{N}(0, 1).$$

Etapas do teste

Procedimentos gerais

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste

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

   onde $\overline{p}$ é calculado pela equação apresentada anteriormente.
5. Determinar a região crítica (região de rejeição), com base no nível de significância α.
6. Conclusão do teste.

Exercício: celular no trânsito

- Em um estudo com 2870 motoristas, 1210 afirmaram ter o hábito de **mexer no celular com o carro em movimento**. Depois de sancionada uma multa, foi realizado outro estudo com 2200 motoristas, dos quais 725 afirmaram ter ainda o hábito.
- Usando um nível de significância de 10%, é possível verificar a alegação de que a proporção de motoristas com hábito de mexer no celular no trânsito **diminuiu significativamente após a criação da multa**?

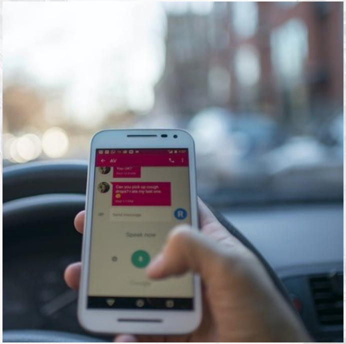

Figura 36. Foto de Roman Pohorecki no Pexels.

Solução

1. Hipóteses: $H_0 : p_1 = p_2 \quad \text{vs} \quad H_a : p_1 > p_2$ (unilateral à direita).
2. Proporções

$$\hat{p}_1 = \frac{y_1}{n_1} = \frac{1210}{2870} = 0.422 \quad \text{e} \quad \hat{p}_2 = \frac{y_2}{n_2} = \frac{725}{2200} = 0.33$$

$$\overline{p} = \frac{y_1 + y_2}{n_1 + n_2} = \frac{1210 + 725}{2870 + 2200} = 0.382.$$

Estatística de teste

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\overline{p}(1 - \overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{0.422 - 0.33}{\sqrt{0.382(1 - 0.382)\left(\frac{1}{2870} + \frac{1}{2200}\right)}} = 6.682.$$

Solução

3. Nível de significância
   $\alpha = 0.1 \rightarrow \text{RC} = \{z > 1.282\}$.
4. Conclusão do teste:
   - $z \in \text{RC}$, portanto **rejeita** H_0.
   - **p-valor** $= \text{P}(Z > 6.682) \approx 0$: existem evidências de que a criação da multa teve efeito.

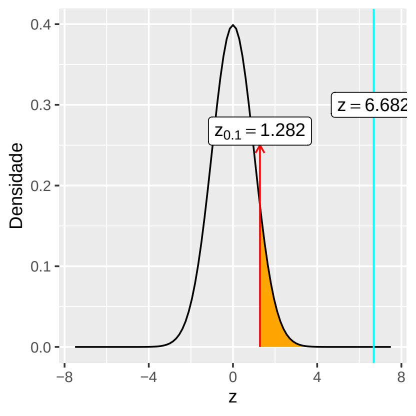

Figura 37. Resultado do teste de hipótese.

Teste de hipótese para a razão de variâncias de duas populações

Ideia geral

Considerando duas populações Y_1 e Y_2 com médias μ_1 e μ_2, e variâncias σ_1^2 e σ_2^2, ou seja

$$Y_1 \sim \mathrm{N}(\mu_1, \sigma_1^2) \qquad \text{e} \qquad Y_2 \sim \mathrm{N}(\mu_2, \sigma_2^2).$$

Já vimos que a **distribuição amostral** da razão de duas variâncias amostrais (s_1^2 e s_2^2) possui distribuição F com $n_1 - 1$ graus de liberdade no numerador e $n_2 - 1$ graus de liberdade no denominador.

Intuitivamente:

- Se a **razão das duas variâncias** for próxima de 1, então elas são aproximadamente iguais.
- Em um teste de hipótese para a **igualdade de variâncias** entre duas populações, verifica-se então se a razão das variâncias está ou não próxima de 1.

Condições para o teste

Quando temos os seguintes requisitos:

- Temos uma AAS.
- As duas populações são independentes.
- As duas populações tem, **cada uma**, distribuição Normal (essa é uma exigência estrita).

Sendo assim, usamos a **estatística de teste**

$$F = \frac{s_1^2}{s_2^2} \sim F_{\nu_1, \nu_2},$$

em que $\nu_1 = n_1 - 1$ graus de liberdade no numerador e $\nu_2 = n_2 - 1$ graus de liberdade no denominador.

Importante: s_1^2 deve ser sempre a **maior** das duas variâncias amostrais.

Etapas do teste

Procedimentos gerais para um teste de hipótese para a diferença de duas variâncias

1. Definir a hipótese nula (H_0) e a alternativa (H_a).
2. Definir o nível de significância α.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Calcular a estatística de teste

$$F = \frac{s_1^2}{s_2^2}.$$

5. Determinar a região crítica (região de rejeição), com base no nível de significância α.
6. Conclusão do teste.

Exemplo: variação em moedas de quarto de dólar

- Nos EUA, as moedas de quarto de dólar sofreram alterações no peso depois de 1964.
- Uma amostra de 40 moedas fabricadas *antes* de 1964 resultou em um desvio-padrão de 0.087 g.
- Uma amostra de 40 moedas fabricadas *depois* de 1964 resultou em um desvio-padrão de 0.06194 g.
- Ao se projetar uma máquina de vendas com moedas, deve-se considerar os desvios-padrão antes e depois de 1964.
- Use o nível de significância de 5% para testar a afirmativa de que os pesos de quarto de dólar antes e depois de 1964 **são provenientes de populações com o mesmo desvio-padrão**.

Solução

1. Hipóteses: $H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_a : \sigma_1^2 \neq \sigma_2^2$ (bilateral).
2. Estatística de teste

$$F = \frac{s_1^2}{s_2^2} = \frac{0.087^2}{0.06194^2} = \frac{0.0076}{0.0038} = 1.973.$$

3. Nível de significância
   $\alpha = 0.05 \rightarrow \text{RC} = \{F < 0.529 \text{ ou } F > 1.891\}$.
4. Conclusão do teste:
   - $F \in \text{RC}$, portanto **rejeita** H_0.
   - **p-valor** $= 2 \times \text{P}(F > 1.973) = 2 \times 0.018 = 0.036$: existem evidências de que a variação dos pesos de quarto de dólar feitos depois de 1964 é significativamente diferente da variação entre os quartos de dólar feitos antes de 1964.

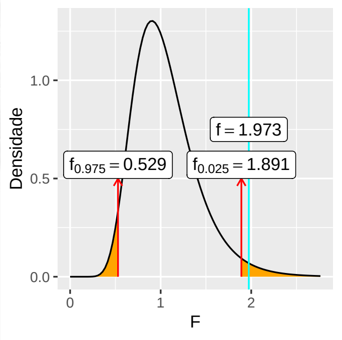

Figura 38. Resultado do teste de hipótese.

Resumo

- Fundamentos de testes de hipóteses.
- Testes para uma população.
- **Testes para duas populações**.
  - Testes de hipótese para a diferença de média de duas populações: σ^2 conhecido.
  - Testes de hipótese para a diferença de média de duas populações: σ^2 desconhecido.
    - Variâncias iguais $\sigma_1^2 = \sigma_2^2$.
    - Variâncias diferentes $\sigma_1^2 \neq \sigma_2^2$.
  - **Testes de hipótese para a diferença de média de duas populações: σ^2 desconhecido e amostras emparelhadas**.
  - Teste de hipótese para a diferença de proporção de duas populações.
  - Teste de hipótese para a razão de variâncias de duas populações.