Universidade Federal de Santa Maria
Departamento de Estatística

TESTES DE HIPÓTESES

Estatística Inferencial

Professora: Laís Helen Loose

Introdução aos testes de hipóteses

Fundamentos dos testes de hipóteses

- Uma **hipótese estatística** é uma afirmação sobre os parâmetros de uma ou mais populações.
- Normalmente são formuladas **duas hipóteses:**
  - A hipótese nula (H_0)
  - A hipótese alternativa (H_1)
- Os testes de hipóteses fornecem procedimentos **(regras de decisão)** que nos permitem rejeitar ou não rejeitar a hipótese nula **com base nas informações fornecidas pela amostra.**

Fundamentos dos testes de hipóteses

Exemplo

- Considere que desejamos estudar a proporção de peixes machos e fêmeas de uma certa espécie em uma lagoa.
- Sem nenhuma informação prévia, supõe-se que a proporção de machos é igual à proporção de fêmeas $(50\%$ ou $p = 0,5)$.
- Se em uma amostra, coletada aleatoriamente, de 100 peixes:
  - 53 forem fêmeas.
  - 64 forem fêmeas.
  - 95 forem fêmeas.
- Qual seria a evidência necessária para concluir que a proporção de fêmeas é maior que a de machos nessa população?

Fundamentos dos testes de hipóteses

- No exemplo anterior, se a hipótese inicial for verdadeira ($p = 0,5$), a probabilidade de se observar 95 fêmeas é **altamente improvável.**

- A ocorrência de resultados altamente improváveis pode ser explicada de uma das duas formas:
  - Ou um evento raro realmente ocorreu.
  - Ou a suposição inicial 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 para o teste de hipóteses

Procedimentos para o teste de hipóteses

1. Estabelecer a hipótese nula (H_0) e a hipótese alternativa (H_1) com base no problema.
2. Definir um nível de significância α do teste.
3. Definir o tipo de teste, com base na hipótese alternativa.
4. Identificar a estatística de teste e caracterização da sua distribuição amostral.
5. Determinar a região de rejeição (região crítica) com base no nível de significância α.
6. Concluir o teste com base na estatística de teste observada e valor crítico ou no nível descritivo (p-valor).

1. Definição de hipóteses: tipos de hipótese

Hipótese nula - H_0

- A hipótese nula engloba o valor do parâmetro que se assume como verdadeiro para a população.
- É a hipótese considerada verdadeira até que se prove o contrário.
- Geralmente representa o contrário do que queremos provar.

Hipótese alternativa - H_1

- É uma afirmativa de que o parâmetro tem um valor que, de alguma forma, difere da hipótese nula.

1. Definição de hipóteses: exemplo

Considere que no 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 (p) é maior do que a proporção de machos.

- Podemos definir as hipóteses de interesse:

$$\mathrm{H}_0 : p \leq 0,5 \quad \text{versus} \quad \mathrm{H}_1 : p > 0,5 \text{ unidades/ml}$$

Observações:

- A hipótese nula sempre abrange a igualdade.
- Em um teste estatístico, o teste é planejado para avaliar a força da evidência contra a hipótese nula. Logo, esta é a afirmativa testada e a decisão do teste será em relação à essa hipótese.

1. Definição de hipóteses: decisão

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

- **Não rejeitar H_0**. Ou seja, a amostra não contém evidências suficientes para a rejeitar a hipótese nula.

- **Rejeitar H_0** em favor da hipótese alternativa H_0. Ou seja, amostra contém evidências suficientes para rejeitar a hipótese nula em favor da hipótese alternativa H_1.

- **Observação:** não rejeitar H_0, **não implica** que H_0 seja realmente a hipótese verdadeira, mas que os dados não estão mostrando evidências suficientes para rejeitá-la.

2. Nível de significância: tipos de erro

Quando realizamos um teste de hipótese estamos sujeitos a cometer 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).

	Situação	
Decisão	H_0 Verdadeira	H_0 Falsa
Rejeitar H_0	Erro tipo I	Decisão correta
Não rejeitar H_0	Decisão correta	Erro tipo II

2. Nível de significância: tipos de erro

As probabilidades de cometer os erros do tipo I e II são definidas por α e β, respectivamente:

- $\alpha =$ P(erro do tipo I) $=$ P(rejeitar H_0 quando H_0 é verdadeira)
- $\beta =$ P(erro do tipo II) $=$ P(não rejeitar H_0 quando H_0 é falsa)

Observações:

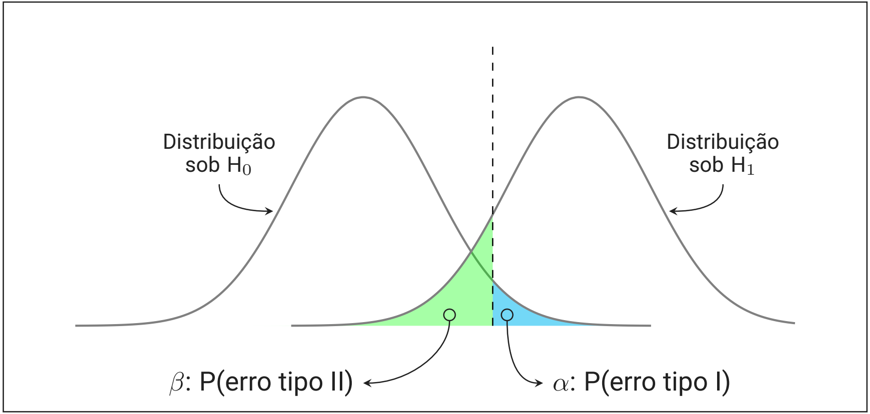
- A probabilidade do erro do tipo I (α) é chamado de nível de significância ou ainda tamanho do teste.

2. Nível de significância: tipos de erro

- Em nosso exemplo, temos que:
  - $\alpha = \mathrm{P}_{\mathrm{H}_0}$(concluir que a proporção de fêmeas é maior quando na verdade não é).
  - $\beta = \mathrm{P}_{\mathrm{H}_1}$(concluir que a proporção sexual é igual ou menor quando na verdade não é).

- **O ideal seria reduzir ao mínimo as probabilidades dos dois tipos de erros.** Uma maneira de alcançar esse objetivo é aumentando o tamanho da amostral, o que aumenta o custo da pesquisa.
- Para um tamanho de amostra fixo, **à medida que diminuímos uma dessas probabilidades, a outra tende a aumentar.**
- Geralmente o pesquisador tenta **controlar o erro do tipo I.**

2. Nível de significância: tipos de erro



3. Tipos de testes

$$\mathrm{H}_0 : \mu = \mu_0 \qquad \mathrm{H}_0 : \mu \geq \mu_0 \qquad \mathrm{H}_0 : \mu \leq \mu_0$$

$$\mathrm{H}_1 : \mu \neq \mu_0 \qquad \mathrm{H}_1 : \mu < \mu_0 \qquad \mathrm{H}_1 : \mu > \mu_0$$

4. Estatística de teste

- A **estatística de teste** mede a distância entre o que foi observado na amostra e o que seria esperado se a hipótese nula fosse verdadeira.
- Ela é **usada para tomar a decisão** sobre a hipótese nula, supondo que ela seja verdadeira.
- De acordo com o tipo de teste de hipóteses feito, **uma distribuição de probabilidade é associada à estatística de teste.**
- Quando H_0 é verdadeira, a estatística de teste segue a distribuição de referência, ou seja, se H_0 é verdadeira, então o resultado da estatística de teste deve ser um **valor não extremo** da distribuição de referência.

4. Estatística de teste

Para a média (μ)

$$Z_{\mathsf{H}_0} = \frac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim \mathsf{N}(0,1) \qquad \text{ou} \qquad \mathsf{T}_{\mathsf{H}_0} = \frac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)}$$

Para a proporção (p)

$$Z_{\mathsf{H}_0} = \frac{\widehat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim \mathsf{N}(0,1)$$

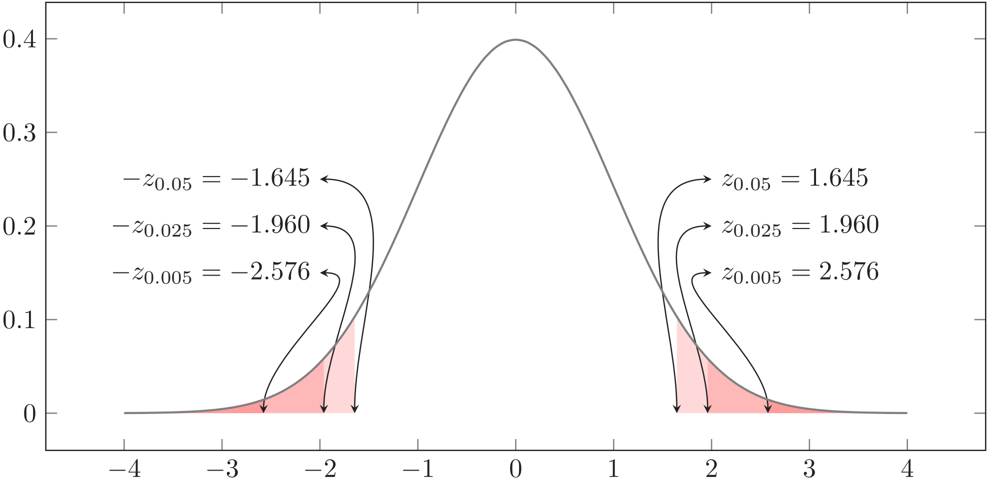
Para a variância (σ^2)

$$\chi^2_{\mathsf{H}_0} = \frac{(n-1)\cdot S^2}{\sigma_0^2} \sim \chi^2_{(n-1)}$$

5. Região critica ou região de rejeição

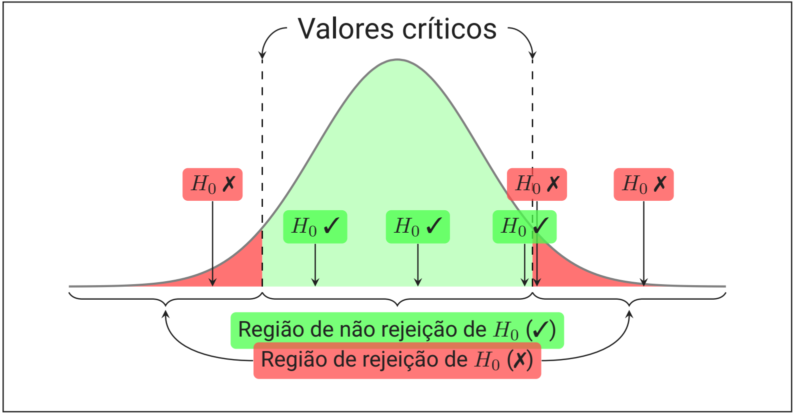
- A regra para decidir se rejeitamos ou não a hipótese nula é expressa em termos de um conjunto denominado **região critica ou região de rejeição.**
- É necessário comparar a estatística de teste com algum valor de referência, que nos informe o quão extrema é a estatística de teste observado para rejeição de H_0.
- O valor que divide a região de rejeição da região de não rejeição é chamado de **valor crítico.**
- O **valor crítico** depende:
  - Da distribuição amostral da estatística de teste sob H_0.
  - Do nível de significância σ.

5. Região critica ou região de rejeição



6. Conclusão do teste

- Comparar o valor observado da estatística de teste com a região de rejeição de H_0.
  - Se a estatística de teste estiver **dentro** da região crítica **devemos rejeitar** H_0:
  - Se a estatística de teste estiver **fora** da região crítica **não devemos rejeitar** H_0.



Nível descritivo ou valor-p

- Outra forma de chegar a uma conclusão em um teste de hipótese é utilizando o **nível de descritivo do teste**, também chamado de **valor-p.**

- A ideia é calcular, supondo que a hipótese nula é verdadeira, a probabilidade de ser **observado em outro experimento uma estatística de teste igual ou mais extrema que àquela observada no experimento atual.**

- **Quanto menor o valor-p, mais evidência temos contra H_0**, ou seja, temos mais evidência para rejeitar a hipótese nula em favor de hipótese alternativa.

- Dessa forma, vamos rejeitar a hipótese nula se o valor-p for menor que o nível de significância α do teste.

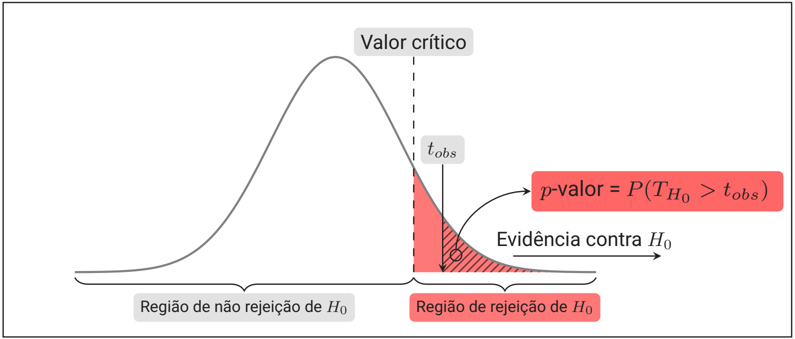
- Um **valor-p muito pequeno significa que é muito improvável que a hipótese nula H_0 seja verdadeira** e, por isso, a rejeitamos.

Nível descritivo para teste unilateral

$$\text{valor-p} = P(\mathrm{T}_{\mathrm{H}_0} < t_{\text{obs}}) \quad \text{para testar } \mathrm{H}_0 : \mu = \mu_0 \text{ vs } \mathrm{H}_1 : \mu < \mu_0$$

$$\text{valor-p} = P(\mathrm{T}_{\mathrm{H}_0} > t_{\text{obs}}) \quad \text{para testar } \mathrm{H}_0 : \mu = \mu_0 \text{ vs } \mathrm{H}_1 : \mu > \mu_0$$

em que $\mathrm{T}_{\mathrm{H}_0}$ é a estatística de teste sob hipótese nula e t_{obs} o seu valor observado.



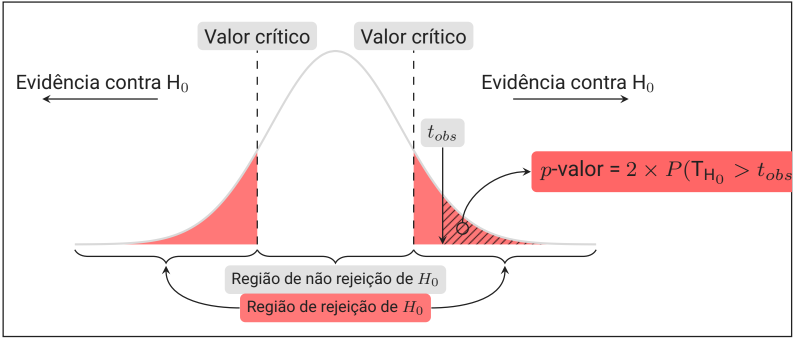
Nível descritivo para teste bilateral

A definição depende da relação entre t_{obs} e μ_0:

$$\text{valor-p} = 2 \cdot P(\text{T}_{\text{H}_0} < t_{\text{obs}}) \quad \text{se } \overline{x} < \mu_0$$

$$\text{valor-p} = 2 \cdot P(\text{T}_{\text{H}_0} > t_{\text{obs}}) \quad \text{se } \overline{x} > \mu_0$$

em que T_{H_0} é a estatística de teste e t_{obs} o seu valor observado.



Exemplos

Exercício 1: Teste de hipóteses para a média (variância conhecida)

Uma máquina enche sacos de açúcar com um desvio-padrão de 10 gramas. Inicialmente a máquina estava regulada para, em média, encher pacotes com 1000 gramas. Suspeita-se que ela se desregulou e para verificar essa hipótese, foi coletada uma amostra de 50 pacotes, que resultou em uma média de 987 gramas. Teste a hipótese de que o peso médio dos pacotes é igual 1000 gramas, com um nível de significância de 5%.

Exemplos

Resolução

- **Definindo as hipóteses:** $H_0 : \mu = 1000$ versus $H_1 : \mu \neq 1000$
- **Nível de significância:** $\alpha = 0,05$
- **Tipo de teste:** teste bilateral
- **Estatística de teste:** $Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \Rightarrow z_{obs} = \dfrac{987 - 1000}{10/\sqrt{50}} = -9,19$
- **Região crítica** para $\alpha = 0,05$ (teste bilateral): RC $= \{z_{obs} < -1,96$ ou $z_{obs} > 1,96\}$
- **Conclusão:** $z_{obs} \in$ RC, portanto temos evidências para rejeitar H_0, ou seja, temos evidências de que o peso médio dos pacotes seja diferente de 1000 gramas.

Exemplos

Exercício 2: Teste de hipóteses para a média (variância desconhecida)

Uma indústria farmacêutica especifica que em certo analgésico a quantidade média de um tipo de ácido deve ser $5,5$ mg por comprimido. A indústria suspeita que houve problemas na produção de um determinado lote e que, nesse lote, a quantidade média dessa substância está maior que a especificada. Para verificar essa suspeita, a indústria selecionou uma amostra aleatória de 50 comprimidos desse lote, observando uma quantidade média do ácido igual a $5,8$ mg e um desvio-padrão de $0,85$ mg. Teste a hipótese de que a quantidade média do ácido é maior que $5,5$ mg por comprimido ao nível de significância de 5%.

Exemplos

Resolução

- **Definindo as hipóteses:** $H_0 : \mu \leq 5,5 \quad$ versus $\quad H_1 : \mu > 5,5$
- **Nível de significância:** $\alpha = 0,05$
- **Tipo de teste:** teste bilateral
- **Estatística de teste:** $T = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1} \Rightarrow t_{obs} = \dfrac{5,8 - 5,5}{0,85/\sqrt{50}} = 2,49$
- **Região crítica** (teste bilateral): RC $= \{t_{obs} > 1,676\}$
- **Conclusão:** $t_{obs} \in$ RC, temos evidências para rejeitar H_0. Temos evidências de que a quantidade de ácido é diferente da especificada.
- **p-valor:** Temos um teste unilateral à direita, então:

$$\alpha^* = P(T_{\nu=49} > t_{obs}) = P(T_{\nu=49} > 2,49) = 0,008$$

Observe que $\alpha^* < \alpha = 0,05$, como é esperado, a conclusão é a mesma, devemos rejeitar H_0.

Exemplos

Exercício 3: Teste de hipóteses para a proporção

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%. Calcule o valor-p.

Exemplos

Resolução

- **Hipóteses:** $H_0: p \leq 0,5 \quad$ versus $\quad H_1: p > 0,5$ (teste unilateral à direita)
- $\widehat{p} = \overline{x} = \dfrac{668}{726} = 0,92$
- **Estatística de teste:** $Z = \dfrac{\widehat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1) \Rightarrow z_{obs} = \dfrac{0,92 - 0,5}{\sqrt{\frac{0,5 \cdot 0,5}{726}}} = 22,63$
- **Região crítica** (teste unilateral à direita): RC $= \{z_{obs} > 1,645\}$
- **Conclusão:** Como $z_{obs} \in$ RC, temos evidências para rejeitar H_0, ou seja, temos evidências de que a proporção de imunizados é maior do que 50%.
- **valor-p** $= P(Z > z_{obs}) = P(Z > 22,63) \approx 0$

Exemplos

Exercício 4: Teste de hipóteses para a variância

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 à concentração de princípio ativo na solução. O lote é rejeitado se claramente ultrapassar o limite de $\sigma^2 = 0,0009$. A variância amostral foi de $0,0013$ e 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 $\alpha = 5\%$.

Exemplos

Resolução

- **Hipóteses:** $H_0 : \sigma^2 \leq 0,0009$ versus $H_1 : \sigma^2 > 0,0009$ (teste unilateral à direita)
- **Estatística de teste:** $\chi^2 = \dfrac{(n-1) \cdot S^2}{\sigma_0^2} \rightarrow \chi^2_{obs} = \dfrac{(16-1) \cdot 0,0013}{0,0009} = 21,67$
- **Região crítica** (teste unilateral à direita - $\alpha = 5\%$): RC $= \{\chi^2_{obs} > 24,99\}$
- **Conclusão:** Como $\chi_{obs} \notin$ RC, não rejeitamos H_0. Ao nível de 5% de significância, não se rejeita a hipótese de que a variância seja igual a $0,0009$.