Aula 04. Introdução aos Testes de Hipóteses

Estatística Inferencial

MBA CDIA

ENAP - Escola Nacional de Administração Pública

2025

Teste de hipóteses

- Na prática, frequentemente estamos interessados em testar uma afirmação ou declaração sobre um parâmetro (isto é, uma característica da população).
- Em estatística, este procedimento é chamado de "teste de hipóteses".
- No teste de hipóteses, existem duas declarações concorrentes sobre o parâmetro populacional: H_0 (hipótese nula) vs H_1 (hipótese alternativa).
- Uma vez que as hipóteses são formuladas, usamos dados coletados de uma amostra para testar se rejeitamos H_0, apoiando assim H_1.

Hipóteses nula e alternativa

- Uma **hipótese estatística** é frequentemente uma declaração sobre parâmetro(s) populacional(is).
- A **hipótese nula**, denotada H_0, afirma que o parâmetro é igual a um valor específico.
  ex. $H_0 : \mu = 35$
- A **hipótese alternativa**, denotada H_1, afirma que o valor do parâmetro difere do valor especificado pela hipótese nula.
  ex. $H_1 : \mu < 35$
  $H_1 : \mu > 35$
  $H_1 : \mu \neq 35$

Tipos de hipótese alternativa

Existem três tipos de hipótese alternativa. Por exemplo, considere a hipótese nula $H_0 : \mu = 35$.

1. $H_1 : \mu < 35 \Rightarrow$ chamada hipótese alternativa **unilateral à esquerda**.
2. $H_1 : \mu > 35 \Rightarrow$ chamada hipótese alternativa **unilateral à direita**.
3. $H_1 : \mu \neq 35 \Rightarrow$ chamada hipótese alternativa **bilateral**.

As hipóteses unilaterais à esquerda e à direita são chamadas de **hipóteses unilaterais**.

Exemplo 1

No ano passado, o aluguel médio mensal de um apartamento em uma determinada cidade era de R$ 800,00. Um corretor de imóveis acredita que o aluguel está mais alto este ano. Enuncie as hipóteses nula e alternativa apropriadas para testar a crença do corretor.

Solução:

- $H_0 : \mu = 800$ (o aluguel médio não mudou)
- $H_1 : \mu > 800$ (o aluguel médio aumentou)

Exemplo 2

Um grupo de comércio prevê que os gastos com volta às aulas serão em média R$ 606,40 por família este ano. Um modelo econômico diferente será necessário se a previsão estiver errada. Especifique as hipóteses nula e alternativa para determinar se um modelo econômico diferente é necessário.

Solução:

- $H_0 : \mu = 606,40$ (a previsão está correta)
- $H_1 : \mu \neq 606,40$ (a previsão está errada)

Exemplo 3

Um determinado modelo de carro pode ser encomendado com motor grande ou pequeno. A média de quilômetros por litro para carros com motor pequeno é 25,5. Um engenheiro automotivo acha que a média para carros com motor maior será menor que isso. Enuncie as hipóteses nula e alternativa apropriadas.

Solução:

- $H_0 : \mu = 25,5$ (mesmo consumo)
- $H_1 : \mu < 25,5$ (consumo menor com motor maior)

Observação sobre H_0 e H_1

- Uma hipótese nula é geralmente pensada como um estado padrão da natureza (ex. conhecimento existente).
- Uma hipótese alternativa, por outro lado, contradiz o estado padrão (ex. novo conhecimento).
- Na maioria dos casos, o que desejamos estabelecer é colocado na hipótese alternativa.

Após construir H_0 e H_1

- Após desenvolver H_0 e H_1, coletamos um conjunto de dados.
- Com base nos dados, construímos uma **estatística de teste**[1] para chegar a uma das seguintes decisões:
  1. Rejeitar H_0
  2. Falhar em rejeitar H_0.
- Se rejeitarmos H_0, concluímos que H_1 é verdadeira.
- Se falharmos em rejeitar H_0, concluímos que os dados não fornecem evidência suficiente para rejeitar H_0.

[1] A ser discutido a seguir.

Exemplo 4

Caixas de um determinado tipo de cereal são rotuladas como contendo 500 gramas. Um inspetor acha que o peso médio pode ser menor que isso.

a. Enuncie as hipóteses nula e alternativa apropriadas.
b. Se o inspetor realizar um teste e rejeitar a hipótese nula, qual é uma conclusão apropriada?
c. Se o inspetor realizar um teste e não rejeitar a hipótese nula, qual é uma conclusão apropriada?

Erros Tipo I e Tipo II

- Como nossa tomada de decisão no teste de hipóteses é baseada em informações amostrais, temos chance de cometer erros.
- Existem dois tipos de erro no teste de hipóteses:
  1. Um **erro Tipo I** ocorre quando H_0 é verdadeira na realidade, mas rejeitamos H_0.
  2. Um **erro Tipo II** ocorre quando H_1 é verdadeira na realidade, mas não rejeitamos H_0.

	Realidade	
Decisão	H_0 Verdadeira	H_0 Falsa
Rejeitar H_0	Erro Tipo I	Decisão correta
Não rejeitar H_0	Decisão correta	Erro Tipo II

Exemplo 5

O reitor de uma escola de negócios quer determinar se o salário inicial médio dos graduados de sua escola é maior que R$ 50.000,00. Ele realizará um teste de hipóteses com as seguintes hipóteses nula e alternativa:

$H_0 : \mu = R\$50.000 \qquad H_1 : \mu > R\$50.000$

a. Suponha que a média verdadeira seja $\mu = R\$50.000$, e o reitor rejeita H_0. Isso é um erro Tipo I, um erro Tipo II ou uma decisão correta?
b. Suponha que a média verdadeira seja $\mu = R\$55.000$, e o reitor rejeita H_0. Isso é um erro Tipo I, um erro Tipo II ou uma decisão correta?
c. Suponha que a média verdadeira seja $\mu = R\$55.000$, e o reitor não rejeita H_0. Isso é um erro Tipo I, um erro Tipo II ou uma decisão correta?

Observações sobre Erros Tipo I e II

- A probabilidade de ter erro Tipo I é denotada por α.
- A probabilidade de ter erro Tipo II é denotada por β.
- Minimizar tanto α quanto β é impossível devido a um trade-off entre α e β *ex.* Se diminuímos α, então β tende a aumentar.
- Em geral, controlar o nível α (isto é, chance de cometer erro Tipo I) é mais importante, pois o erro Tipo I leva à aceitação de conhecimento novo incorreto.

Teste de Hipóteses para Média Populacional

- Agora estudaremos um teste de hipóteses para média populacional μ quando o desvio padrão populacional σ é conhecido.
- **Premissa requerida:** O tamanho da amostra é grande $(n > 30)$, ou a população é normal (pelo menos aproximadamente).
- Existem duas maneiras de realizar testes de hipóteses que produzem os mesmos resultados:
  - Método do valor crítico
  - Método do p-valor

Procedimento: Método do p-valor

Passo 1. Enunciar as hipóteses nula e alternativa. A hipótese nula tem a forma $H_0 : \mu = \mu_0$. A hipótese alternativa está em uma das três formas:

- Unilateral à esquerda: $H_1 : \mu < \mu_0$
- Unilateral à direita: $H_1 : \mu > \mu_0$
- Bilateral: $H_1 : \mu \neq \mu_0$

Passo 2. Escolher um nível de significância α (níveis comuns são $\alpha = 0,05$, $\alpha = 0,01$).

Passo 3. Calcular a estatística de teste:

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

(Isso mede o quão distante a média amostral $\bar{x}$ está do valor hipotético μ_0 em H_0.)
Passo 4. Calcular o p-valor da estatística de teste z.

- **Unilateral à esquerda:** p-valor = área sob a distribuição normal padrão à esquerda de z, isto é, $P(Z < z)$.
- **Unilateral à direita:** p-valor = área sob a distribuição normal padrão à direita de z, isto é, $P(Z > z)$.
- **Bilateral:** p-valor = soma das áreas sob a distribuição normal padrão à esquerda de $-|z|$ e à direita de $|z|$, isto é, $2 \cdot P(Z \leq -|z|)$.

Procedimento: Método do p-valor (cont.)

Passo 5. Determinar se deve rejeitar H_0: - Rejeitar H_0 se p-valor $\leq \alpha$.
Passo 6. Enunciar uma conclusão.

Exemplo ilustrativo

Suponha:

- $H_0 : \mu = 530,\ H_1 : \mu > 530$
- $\alpha = 0,05$
- $z = 2,76$

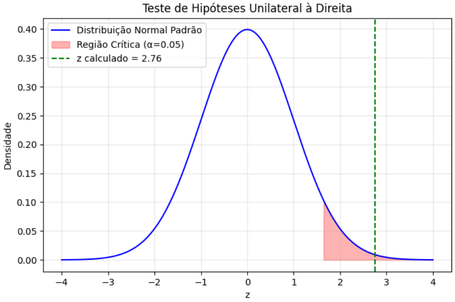
Passo 4. Determinar o p-valor.
Como o teste é unilateral à direita: p-valor $= P(Z > 2,76) = 0,0029$
Passo 5. Como p-valor $= 0,0029 \leq \alpha = 0,05$, rejeitamos H_0 ao nível $\alpha = 0,05$.
Passo 6. Concluímos que a nota média para alunos que completam o programa de preparação online é maior que 530.

Implementacao em Python

```
from scipy import stats
z_stat = 2.76
alpha = 0.05
# Função automática para p-valor
p_valor = stats.norm.sf(z_stat)  # sf = 1 - cdf (survival function)
#p_valor = 1 - stats.norm.cdf(z_stat)
print(f"p-valor = {p_valor:.4f}")
print(f"Decisão: {'Rejeitar H0' if p_valor < alpha else 'Não rejeitar H0'}")

import matplotlib.pyplot as plt
import numpy as np
# Gráfico da distribuição e região crítica
x = np.linspace(-4, 4, 1000)
y = stats.norm.pdf(x)
z_critico = stats.norm.ppf(1 - alpha)

plt.figure(figsize=(8, 5))
plt.plot(x, y, 'b-', label='Distribuição Normal Padrão')
plt.fill_between(x[x > z_critico], y[x > z_critico], alpha=0.3,
                 color='red', label=f'Região Crítica (={alpha})')
plt.axvline(z_stat, color='green', linestyle='--',
            label=f'z calculado = {z_stat:.2f}')
plt.xlabel('z')
plt.ylabel('Densidade')
plt.title('Teste de Hipóteses Unilateral à Direita')
plt.legend()
plt.grid(True, alpha=0.3)
plt.show()
```



Implementacao em Python (cont.)

Solução 2 - Quando temos os dados da amostra:

```
# Supondo dados que geraram z = 2.76
# (média amostral que resulta nesse z)
import numpy as np

n = 100
sigma = 50  # desvio padrão populacional conhecido
x_barra = 530 + 2.76 * (sigma / np.sqrt(n))  # média que gera z=2.76

# Criar amostra artificial com essas características
amostra = np.random.normal(x_barra, sigma, n)

# Teste automático
from statsmodels.stats.weightstats import ztest
estatistica, p_valor = ztest(amostra, value=530, alternative='larger')

print(f"Estatística z: {estatistica:.2f}")
print(f"p-valor: {p_valor:.4f}")
print(f"Conclusão: A nota média é significativamente maior que 530")
```

Resultado: p-valor 0.05 $\rightarrow$ Rejeitar H_0

Observações sobre p-valor

- Um p-valor pequeno ($\leq \alpha$) implica que a estatística de teste observada é muito incomum quando H_0 é verdadeira.
- Então, ou:
  1. Tivemos muita sorte e observamos algo raro.
  2. H_0 é falsa.

  Tendemos a acreditar que não somos tão sortudos, então afirmamos que H_0 é falsa.
- Dizemos que H_0 é rejeitada ao nível α, ou que o resultado é estatisticamente significativo ao nível α.

Teste de Hipóteses quando σ é Desconhecido

- Na prática, frequentemente não conhecemos o desvio padrão populacional σ.
- Quando σ é desconhecido, usamos o desvio padrão amostral s e a distribuição t de Student.
- **Premissa requerida:** O tamanho da amostra é grande ($n > 30$) ou a população é aproximadamente normal.
- Usaremos o método do p-valor.

Procedimento quando σ é desconhecido

Passos 1-2: Iguais ao caso quando σ é conhecido.
Passo 3. Calcular a estatística de teste:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

- Como σ é desconhecido, substituímos por s. - Usamos a estatística t, que vem de uma distribuição t de Student com $gl = n - 1$.
Passo 4. Calcular o p-valor usando a distribuição t com $gl = n - 1$.
Passos 5-6: Iguais ao caso quando σ é conhecido.

Exemplo: Teste com σ desconhecido

Em um estudo médico recente, 76 indivíduos foram colocados em uma dieta com baixo teor de gordura. Após 12 meses, sua perda média de peso amostral foi $\bar{x} = 2,2$ kg, com desvio padrão amostral de $s = 6,1$ kg. Use o nível de significância $\alpha = 0,05$ para testar a afirmação de que a perda média de peso é maior que 0.

Solução:

- $H_0 : \mu = 0$, $H_1 : \mu > 0$ (unilateral à direita)
- $\alpha = 0,05$
- $t = \frac{2,2-0}{6,1/\sqrt{76}} \approx 3,144$
- $gl = 75$

Exemplo (cont.)

Passo 4. Usar a tabela t para calcular o p-valor.

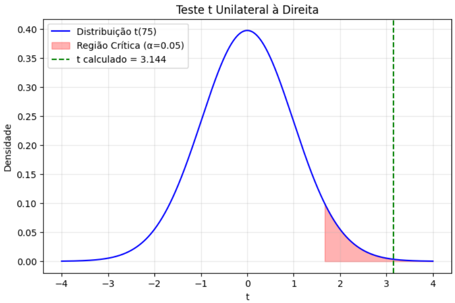
- Como $H_1 : \mu > 0$ é unilateral à direita, p-valor $= P(T > 3,144)$.
- $gl = 75$ não aparece na tabela t, então arredondamos para baixo para $gl = 60$.
- Na tabela t com $gl = 60$, o p-valor está entre 0,001 e 0,0025.

Passo 5. Como p-valor $< \alpha = 0,05$, rejeitamos H_0.

Passo 6. Concluímos que a perda média de peso de pessoas que foram colocadas em uma dieta com baixo teor de gordura por 12 meses é maior que 0.

Implementação em Python

```
from scipy import stats
t_stat = 3.144
gl = 75
alpha = 0.05
# Função automática para p-valor
p_valor = stats.t.sf(t_stat, df=gl)  # sf = 1 - cdf
#p_valor = 1 - stats.t.cdf(t_stat, df=gl)
print(f"p-valor = {p_valor:.4f}")
print(f"Decisão: {'Rejeitar H0' if p_valor < alpha else 'Não rejeitar H0'}")
import matplotlib.pyplot as plt
import numpy as np
# Gráfico da distribuição t e região crítica
x = np.linspace(-4, 4, 1000)
y = stats.t.pdf(x, df=gl)
t_critico = stats.t.ppf(1 - alpha, df=gl)
plt.figure(figsize=(8, 5))
plt.plot(x, y, 'b-', label=f'Distribuição t({gl})')
plt.fill_between(x[x > t_critico], y[x > t_critico], alpha=0.3,
                 color='red', label=f'Região Crítica (={alpha})')
plt.axvline(t_stat, color='green', linestyle='--',
            label=f't calculado = {t_stat:.3f}')
plt.xlabel('t')
plt.ylabel('Densidade')
plt.title('Teste t Unilateral à Direita')
plt.legend()
plt.grid(True, alpha=0.3)
plt.show()
```



Solução 2 - Quando temos os dados da amostra:

```
# Dados do problema
n = 76
x_barra = 2.2  # média amostral
s = 6.1        # desvio padrão amostral
mu_0 = 0       # valor sob H0
alpha = 0.05
# Criar amostra artificial com essas características
import numpy as np
np.random.seed(42)
# Gerar amostra com média e desvio próximos aos desejados
amostra = np.random.normal(x_barra, s, n)
# Ajustar para ter exatamente a média e desvio desejados
amostra = (amostra - amostra.mean()) * (s / amostra.std()) + x_barra

# Teste t automático
from scipy import stats #(scipy >= 1.6.0):
resultado = stats.ttest_1samp(amostra, mu_0, alternative='greater')
p_valor = resultado.pvalue
estatistica = resultado.statistic
print(f"Estatística t: {estatistica:.3f}")
print(f"p-valor: {p_valor:.4f}")
print(f"Conclusão: A perda média de peso é significativamente maior que 0 kg")
```

Resultado: p-valor $= 0{,}0012 < 0{,}05 \rightarrow$ Rejeitar H_0

Teste de Hipóteses para Proporção Populacional

- Agora estudaremos como testar uma hipótese para proporção populacional p.
- Para nosso teste de hipóteses ser válido, uma premissa requerida é:

$$np_0 \geq 10 \text{ e } n(1 - p_0) \geq 10$$

  onde $p_0 =$ proporção populacional especificada por H_0.
- Usaremos o método do p-valor para realizar o teste de hipóteses.

Procedimento para teste de proporção

Passos 1-2: Similares aos testes anteriores, mas agora $H_0 : p = p_0$.
Passo 3. Calcular a estatística de teste:

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

Passo 4. Calcular o p-valor usando a distribuição normal padrão.
Passos 5-6: Determinar se deve rejeitar H_0 e enunciar conclusão.

Exemplo: Teste de proporção

Suponha que 67% de todos os pedidos de seguro de automóvel no Brasil são feitos por solteiros com menos de 25 anos. Em uma amostra aleatória de 53 pedidos em uma cidade brasileira, 42 foram feitos por solteiros com menos de 25 anos.
Teste ao nível de 5% de significância se a proporção de pedidos feitos por solteiros com menos de 25 na cidade brasileira é diferente da proporção nacional dos EUA.
Solução:

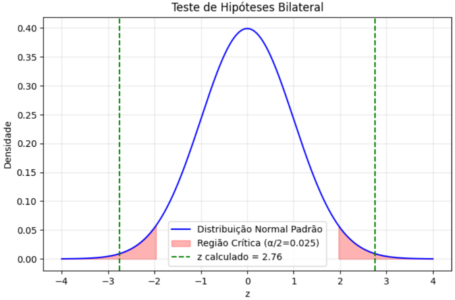
- $H_0 : p = 0,67,\ H_1 : p \neq 0,67$ (bilateral)
- $\hat{p} = \frac{42}{53} \approx 0,7925$
- $z = \frac{0,7925 - 0,67}{\sqrt{\frac{0,67(1-0,67)}{53}}} \approx 1,90$

Exemplo (cont.)

Passo 4. Determinar o p-valor.
p-valor (bilateral) $= 2 \cdot P(Z < -1,90) = 2(0,0287) = 0,0574$
Passo 5. Como p-valor $> \alpha$ $(= 0{,}05)$, falhamos em rejeitar H_0.
Passo 6. Não há evidência suficiente para concluir que a proporção de pedidos feitos na cidade brasileira por solteiros com menos de 25 anos é diferente da proporção nacional dos EUA.

Implementação em Python

```
n = 53          # tamanho da amostra
x = 42          # sucessos observados
p0 = 0.67       # proporção sob H0
alpha = 0.05
from statsmodels.stats.proportion import proportions_ztest
estatistica, p_valor = proportions_ztest(count=x, nobs=n, value=p0,
                                        alternative='two-sided')
print(f"Estatística z: {estatistica:.2f}")
print(f"p-valor: {p_valor:.4f}")
import matplotlib.pyplot as plt
import numpy as np
x = np.linspace(-4, 4, 1000)
y = stats.norm.pdf(x)
z_critico = stats.norm.ppf(1 - alpha/2)
plt.figure(figsize=(8, 5))
plt.plot(x, y, 'b-', label='Distribuição Normal Padrão')
plt.fill_between(x[x > z_critico], y[x > z_critico], alpha=0.3,
                 color='red', label=f'Região Crítica (/2={alpha/2})')
plt.fill_between(x[x < -z_critico], y[x < -z_critico], alpha=0.3,
color='red')
plt.axvline(z_stat, color='green', linestyle='--',
            label=f'z calculado = {z_stat:.2f}')
plt.axvline(-z_stat, color='green', linestyle='--')
plt.xlabel('z')
plt.ylabel('Densidade')
plt.title('Teste de Hipóteses Bilateral')
plt.legend()
plt.grid(True, alpha=0.3)
plt.show()
```



Exercício Extra: Avaliação de Programa Social

Contexto: O governo implementou um programa de capacitação profissional para aumentar a empregabilidade de jovens adultos. Antes do programa, a taxa de emprego dos participantes era 45%. Após 6 meses, você tem dados de 200 participantes.

`programa_social_dataset.csv`

Tarefas:

- Teste se o programa aumentou significativamente a taxa de emprego
- Calcule o intervalo de confiança para a nova taxa

Relação entre Testes de Hipóteses e Intervalos de Confiança

Quando são equivalentes:

- Para testes **bilaterais** sobre média ou proporção:
  - Rejeitar $H_0 : \mu = \mu_0$ ao nível $\alpha \Leftrightarrow \mu_0$ não está no IC de $(1 - \alpha)100\%$
  - Exemplo: Se IC 95% $= [48, 52]$ e $H_0 : \mu = 45$, rejeitamos H_0 a $\alpha = 0,05$

Quando NÃO são equivalentes:

- Para testes **unilaterais**:
  - IC bilateral não corresponde diretamente ao teste unilateral
  - Exemplo: $H_0 : \mu = 50$ vs $H_1 : \mu > 50$ com IC 95% $= [48, 56]$
  - IC contém 50, mas ainda podemos rejeitar H_0 no teste unilateral!
- Solução: usar IC unilateral (menos comum na prática)

Vantagem do IC: Fornece toda a faixa de valores plausíveis, não apenas sim/não.

Exemplos Práticos: Testes vs Intervalos

Exemplo 1 - Teste Bilateral (Equivalentes):

- Tempo médio de atendimento no SUS: $\bar{x} = 45$ min, IC 95% $= [42, 48]$
- $H_0 : \mu = 40$ min (meta governamental)
- Como $40 \notin [42, 48]$, rejeitamos H_0 a $\alpha = 0,05$
- p-valor do teste bilateral também seria $< 0,05$

Exemplo 2 - Teste Unilateral (Não equivalentes):

- Taxa de aprovação em concursos: $\hat{p} = 0,35$, IC 95% $=$ [0,30, 0,40]
- $H_0 : p = 0,32$ vs $H_1 : p > 0,32$
- IC contém 0,32, mas teste unilateral pode dar p-valor $< 0,05$!
- IC bilateral é conservador para testes unilaterais

Recomendação: Use IC para estimar magnitude do efeito e teste de hipóteses para decisões específicas.

Conclusões e Próximos Passos

O que aprendemos hoje:

- Conceitos fundamentais de teste de hipóteses
- Erros Tipo I e Tipo II
- Método do valor crítico e método do p-valor
- Testes para média (com σ conhecido e desconhecido)
- Teste para proporção
- Relação entre testes de hipóteses e intervalos de confiança
- Aplicação prática em avaliação de políticas públicas

Próxima aula:

- Testes para duas amostras
- ANOVA - Análise de Variância
- Testes não-paramétricos

Dúvidas?