TESTES DE HIPÓTESES
T E Z

Teste de Hipóteses

Teste de hipóteses é o segundo tópico da inferência estatística.

Dado o valor de um parâmetro da população (hipótese), o teste de hipóteses usa uma amostra para avaliar se a afirmação é plausível.

Teste de Hipóteses

- Teste t para uma amostra (média)
- Teste t para duas amostras (médias)
- Teste z para uma amostra (proporção)
- Teste z para duas amostras (proporções)

COMPARANDO 1 OU 2 MÉDIAS OU PROPORÇÕES

Comparando médias

Comparamos uma média com um valor específico:

O tempo de entrega da pizza é menos de 30 minutos em média?
Teste t para uma média

Comparando médias

Comparamos duas médias entre si:

Há diferença de preços entre a loja no bairro A e a loja no bairro B?
Teste t para duas amostras independentes

Teste cego: coca-cola vs pepsi
Teste t para duas amostras pareadas (dependentes)

Comparando proporções

Comparamos uma ou duas proporções:

A taxa de devolução de produtos de uma loja online é maior do que 5%?

Teste z para uma proporção

Existe diferença entre duas campanhas de marketing para aumentar a taxa de conversão de visitantes em compradores?

Teste z para duas proporções

Teste de Hipóteses

Passo 1: Defina as hipóteses

Passo 2: Escolha o nível de significância α

Passo 3: Calcule o p-valor

Passo 4: Conclua e interprete o resultado

TESTE T
PARA 1 AMOSTRA

Teste t

Problema: O tempo de entrega da pizza é menos de 30 minutos em média?
Hipótese: Acho que leva mais do que 30 minutos, em média.

Teste: Teste t para uma média
Amostra: 30 entregas, média 32 minutos, desvio padrão 8 minutos.

Passo 1: Definir as hipóteses
Ho: $\mu \leq 30$
Ha: $\mu > 30$

Passo 2: Escolho $\alpha = 5\%$.

Teste t

Passo 3: Cálculo de t e p-valor

p-valor = $P(T > t)$

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

$\bar{x} = 32$
$\mu_o = 30$
$s = 8$
$n = 30$

$$\Rightarrow \quad t = \frac{32 - 30}{8/\sqrt{30}} = 1{,}37 \quad \Rightarrow \quad P(T > 1{,}37) = 0{,}09$$

Teste t

Ho: $\mu \leq 30$
Ha: $\mu > 30$

Passo 4: Decisão

p-valor $\geq \alpha$
Ho não é rejeitado.

Não temos evidências suficientes para processar a pizzaria, e podemos concluir que ela está correta em sua afirmação de que média de delivery é de 30 minutos ou menos.

Exemplo: Central de Atendimento

Uma empresa quer implementar um novo canal de atendimento (WhatsApp) para diminuir o volume de ligações na Central de Atendimento. Como a empresa pode avaliar se a implementação deste novo canal está sendo positiva? A empresa recebe em média 20 ligações diariamente na Central de Atendimento.

<u>Passo 1</u>: Difinir as hipóteses

Ho: $\mu \geq 20$

Ha: $\mu < 20$

<u>Passo 2</u>: Escolho $\alpha = 5\%$.

Exemplo: Central de Atendimento

Após a implementação, a empresa coleta dados da quantidade de ligações recebidas na Central de Atendimento no próximo mês. Em 30 dias, a média foi 17.8 com desvio padrão 3.

Passo 3: Cálculo de t e p-valor

p-valor = P(T < t)

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

$\bar{x} = 17.8$
$\mu_o = 20$
$s = 3$
$n = 30$

$$\Rightarrow \quad t = \frac{17.8 - 20}{3/\sqrt{30}} = -4 \quad \Rightarrow \quad P(T < -4) = 0{,}0002$$

Exemplo: Central de Atendimento

Passo 4: Decisão

p-valor = 0,0002

$\alpha = 0{,}05$

p-valor < α

Rejeitamos Ho.

Ho: $\mu \geq 20$
Ha: $\mu < 20$

Concluímos que houve uma diminuição significativa no volume de ligações após a implementação do WhatsApp.

Exemplo: Central de Atendimento

Considerações:

Período de Coleta: Certifique-se de que o período de coleta após a implementação é representativo e suficiente para capturar a variabilidade.

Controle de Variáveis: Considere outras variáveis que possam afetar o volume de ligações, como sazonalidade, outras campanhas de marketing, etc.

Validação Contínua: Continue monitorando o volume de ligações e o uso do WhatsApp para validar os resultados ao longo do tempo.

Suposições

Normalidade: A distribuição dos dados deve ser aproximadamente normal, especialmente para amostras pequenas.

Independência: As observações devem ser independentes.

Escala de Medida: Os dados devem ser contínuos e medidos em uma escala intervalar ou racional.

Ausência de Outliers: Outliers extremos devem ser identificados e tratados adequadamente.

TESTE T
PARA 2 AMOSTRAS

Comparando duas amostras

O teste de hipóteses também é muito usado para comparar se as médias de duas amostras são iguais.

- Amostras independentes: Calculamos a diferença entre duas médias
- Amostras dependentes: Calculamos a média das diferenças

AMOSTRAS INDEPENDENTES

Amostras independentes

Usamos quando as duas amostras são independentes.

$$H_0: \mu_1 \le \mu_2 \quad\quad H_0: \mu_1 - \mu_2 \le 0$$
$$H_A: \mu_1 > \mu_2 \quad \Rightarrow \quad H_A: \mu_1 - \mu_2 > 0$$

- Desvios padrão não são iguais
- Desvios padrão são iguais

Amostras independentes

Exemplo: Estes são os retornos anuais de 10 empresas na indústria do ouro e 10 na indústria do petróleo.

Ao nível de significância de 5%, as médias dos retornos são diferentes?

Assuma que as populações sejam normalmente distribuídas e que as variâncias populacionais não são iguais.

Gold	Oil
6	−3
15	15
19	28
26	18
2	32
16	31
31	15
14	12
15	10
16	15

Amostras independentes

Passo 1:

Ho: μ1 = μ2

Ha: μ1 ≠ μ2

Passo 2: α = 5%.

Passos 3 & 4: O p-valor do teste bicaudal é 0,7661. Portanto, não rejeitamos H0.

A média dos retornos na indústria do petróleo e da indústria do ouro não são estatisticamente diferentes.

	Gold	Oil
Mean	16	17.3
Variance	70.6667	114.2333
Observations	10	10
Hypothesized Mean Difference	0	
Df	17	
t Stat	**−0.3023**	
P(T ≤ t) one-tail	0.3830	
t Critical one-tail	1.7396	
P(T ≤ t) two-tail	**0.7661**	
t Critical two-tail	**2.1098**	

Suposições

Normalidade: A distribuição dos dados deve ser aproximadamente normal.

Independência: As observações devem ser independentes.

Homoscedasticidade (homogeneidade das variâncias): As variâncias das duas populações devem ser iguais.

AMOSTRAS DEPENDENTES (PAREADAS)

Amostras dependentes

Estudos "antes" e "depois" caracterizados por alguma intervenção, sobre o mesmo indivíduo.

Ho: μ antes ≤ μ depois
Ha: μ antes > μ depois

⇨

Ho: μ antes - μ depois ≤ 0
Ha: μ antes - μ depois > 0

⇨

Ho: μd ≤ 0
Ha: μd > 0

As amostras não são independentes.

O parâmetro de interesse é a diferença média μd.

Amostras dependentes

Após uma lei entrar em vigor nos EUA onde os restaurantes devem informar a quantidade de calorias de seus produtos, uma nutricionista resolveu analisar se quando as pessoas veem as calorias de um produto, elas passam a consumir menos calorias.

Amostras dependentes

A nutricionista coletou dados de 40 clientes da Starbucks, com a média de calorias consumidas antes e depois da lei ter entrado em vigor.

Saber a informação calórica dos produtos faz com que as pessoas consumam menos calorias (ao nível de significância de 5%)?

Passo 1:

Ho: μ antes ≤ μ depois ⇨ Ho: μd ≤ 0

Ha: μ antes > μ depois ⇨ Ha: μd > 0

Passo 2: $\alpha = 5\%$

Amostras dependentes

Passo 3:
O p-valor do teste unicaudal é 2,15 x 10^{-8}.

Passo 4: Rejeitamos H0.

A nutricionista pode concluir que a ingestão calórica média dos alimentos diminuiu após a nova lei.

	Before	After
Mean	400.275	391.475
Variance	49.94808	42.3583
Observations	40	40
Pearson Correlation	0.27080	
Hypothesized Mean Difference	0	
Df	39	
t Stat	**6.7795**	
P(T ≤ t) one-tail	**2.15E-08**	
t Critical one-tail	**1.6849**	
P(T ≤ t) two-tail	4.31E-08	
t Critical two-tail	2.0227	

Suposições

Normalidade: A distribuição das diferenças deve ser aproximadamente normal.

Independência: As observações devem ser independentes.

TESTANDO PROPORÇÕES

TESTE Z PARA
1 PROPORÇÃO

Teste z para uma Proporção

Exemplo: Uma empresa de e-commerce quer avaliar se a taxa de devolução de produtos é maior do que 5%. Eles coletam uma amostra aleatória de 500 pedidos e encontram que 30 deles foram devolvidos.

Ho: A proporção de devoluções é menor ou igual a 5%

Ha: A proporção de devoluções é maior do que 5%

Ho: $\rho \leq 0{,}05$

Ha: $\rho > 0{,}05$

Definimos $\alpha = 0{,}05$

Teste z para uma Proporção

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

x = 30 produtos devolvidos

n = 500 produtos analisados

A proporção da amostra é $\bar{p} = \frac{30}{500}$ = 0,06

p_0= 0,05

Teste z para uma Proporção

p-valor = $P(Z > z)$

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

$$z = \frac{0{,}06 - 0{,}05}{\sqrt{\frac{0{,}05\,(1-0{,}05)}{500}}} = 1{,}03 \quad \Rightarrow \quad P(Z > 1{,}03) = 0{,}1515$$

Teste z para uma Proporção

O p-valor é 0,1515.

$\alpha = 0,05$

Não rejeitamos Ho.

Com um nível de significância α=0,05, não há evidência suficiente para afirmar que a proporção de devoluções é maior do que 5%.

Suposições

Binomial: As observações devem seguir uma distribuição binomial (sucesso/falha).

Normalidade: A distribuição dos dados deve ser aproximadamente normal.

Independência: As observações devem ser independentes.

TESTE Z PARA 2 PROPORÇÕES

Teste z para duas Proporções

Uma empresa deseja comparar a eficácia de duas campanhas de marketing para aumentar a taxa de conversão de visitantes em compradores. Eles coletam os seguintes dados:

Campanha A: 2000 visitantes, 180 conversões

Campanha B: 2200 visitantes, 250 conversões

Ho: As proporções de conversão das duas campanhas são iguais

Ha: As proporções de conversão das duas campanhas são diferentes

Ho: $pA = pB$

Ha: $pA \neq pB$

Teste z para duas Proporções

Campanha A: 2000 visitantes, 180 conversões: $\bar{p}A = \frac{180}{2000}$ = 0,09

Campanha B: 2200 visitantes, 250 conversões: $\bar{p}B = \frac{250}{2200}$ = 0,1136

Proporção combinada: $\bar{p} = \frac{180 + 250}{2000 + 2200} = \frac{430}{4200} = 0{,}1024$

nA = 2000

nB = 2200

$$\boxed{z = \frac{\bar{p}A - \bar{p}B}{\sqrt{\bar{p}(1 - \bar{p})\left(\frac{1}{n\mathrm{A}} + \frac{1}{n\mathrm{B}}\right)}}}$$

Teste z para duas Proporções

$z = -2{,}52$

p-valor $= 2 \cdot P(Z < -2{,}52) = 0{,}0117$

Rejeitamos Ho.

Com um nível de significância α=0,05, podemos afirmar que as proporções de conversão das duas campanhas são diferentes.

Teste A/B

Email Marketing: Duas versões de um email são enviadas para diferentes grupos de usuários.

Design de Produto: Duas versões de um produto são testadas com usuários.

Preço: Duas versões de um produto são oferecidas com preços diferentes para grupos diferentes de consumidores.

Suposições

Binomial: As observações devem seguir uma distribuição binomial (sucesso/falha).

Normalidade: A distribuição dos dados deve ser aproximadamente normal.

Independência: As observações devem ser independentes.

SUPOSIÇÕES

Normalidade

Podemos aplicar o Teorema do Limite Central quando a amostra é grande.

Porém, é importante verificar com um teste de normalidade:

- Confirmar que os dados seguem uma distribuição normal, especialmente se a amostra não for extremamente grande.
- Certificar-se de que outros aspectos do modelo estatístico (como os resíduos) são normais.

Testes de Normalidade

Shapiro-Wilk: Adequado para amostras menores. É conhecido por ter alto poder em comparação com outros testes de normalidade.

Anderson-Darling: Adequado para amostras menores. É mais sensível aos dados extremos, sendo mais potente na detecção de desvios da normalidade.

Kolmogorov-Smirnov: Adequado para amostras grandes. É mais geral e pode testar contra distribuições diferentes da normal.

Ho: Os dados seguem uma distribuição normal
Ha: Os dados não seguem uma distribuição normal

Homoscedasticidade

É importante que as variâncias sejam estatisticamente iguais para termos certeza de que apenas a média é o parâmetro que está variando.

Teste de Levene

Ho: As variâncias são iguais
Ha: As variâncias não são iguais

Falhas nas suposições

<u>Dados não-normais</u>: Se os dados não seguirem uma distribuição normal, podem ser usados testes não paramétricos, que não fazem a suposição de normalidade.

<u>Heterocedasticidade</u>: Se a variância não for constante, a transformação dos dados pode ser usada.