

INFERÊNCIA ESTATÍSTICA: TESTE DE HIPÓTESES

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

- Objetivo do teste de hipóteses
- Tipos de testes
- Definindo as hipóteses nula e alternativa
- Nível de significância α
- Erros tipo I e tipo II
- p-valor
- Tomada de decisão

TESTE DE HIPÓTESES

O que é?

Uma técnica estatística que permite verificar se uma afirmação (hipótese) sobre uma população é plausível com base em uma amostra dos dados observados.

Usados nas mais diversas áreas, como saúde, economia, psicologia e marketing para verificar a validade de afirmações e tomar decisões com base em dados coletados.

Tipos de testes

Podemos testar:

- Comparação entre valores
- Relação entre valores

Podemos fazer testes para:

- Dados quantitativos (média)
- Dados qualitativos (proporção)

Comparando médias

Por que preciso fazer um teste estatístico para comparar médias?

“Há diferença entre a média de preços da loja no bairro A e da loja no bairro B?”

Não posso simplesmente comparar as médias da loja A com a loja B?

Comparando médias

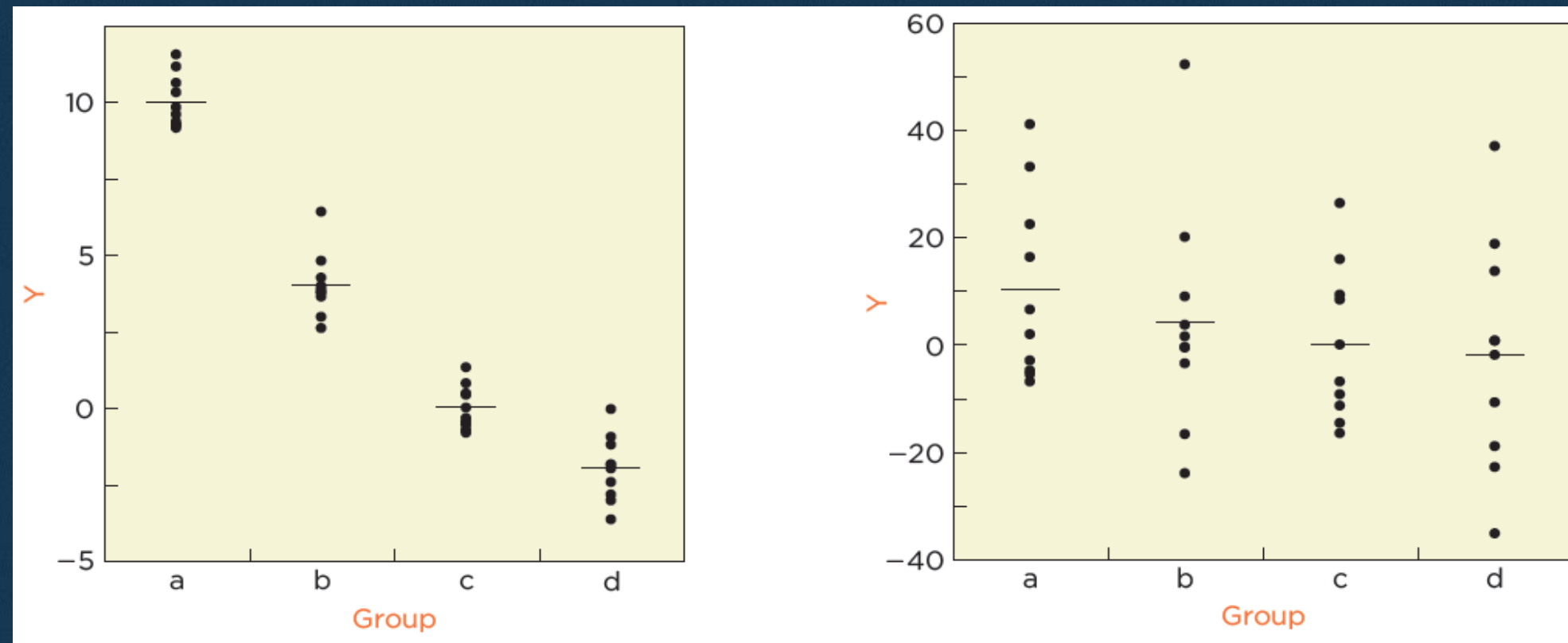
Considere as seguintes médias:

Category	Mean
a	10
b	4
c	0
d	-2

Essas médias são estatisticamente diferentes?

Para responder a esta pergunta, precisamos saber a variação dentro de cada categoria.

Comparando médias



TESTES DE HIPÓTESES E SUAS APLICAÇÕES

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 médias

Comparamos três ou mais médias entre 1 ou 2 grupos:

Há diferença de salários entre os cargos de analista de dados, cientista de dados e engenheiro de dados em uma empresa?

Teste ANOVA de 1 via

Há diferença de salários entre os cargos de analista de dados, cientista de dados e engenheiro de dados em uma empresa por nível de escolaridade?

Teste ANOVA de 2 vias

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

Relação

Relação entre duas variáveis qualitativas (categóricas):

Existe relação entre tipo de filme e a compra de pipoca?

Existe relação entre músculo e flexibilidade?

Teste qui-quadrado de independência (Pearson)

Relação

Relação entre duas variáveis quantitativas:

O número de horas de treinamento está associado à quantidade de vendas?

Qual o impacto da temperatura nas vendas?

Correlação

Análise de Regressão

Obs. A regressão linear pode ser usada para testar a relação entre uma variável dependente e várias variáveis independentes.

Exercício

Separei minha base de clientes em duas partes e, para cada uma, mandei um email com assuntos diferentes. Quero saber qual assunto gera mais cliques, em média. Qual teste devo realizar?

Estou testando diariamente qual foto tem mais cliques no meu site, em média, entre 5 fotos. Qual teste devo realizar?

TESTE DE HIPÓTESES EM 4 PASSOS

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

**Passo 1: Defina as
hipóteses**

Parâmetros e estatísticas

	Parâmetros (população)	Estatísticas (amostra)
Média	μ	\bar{x}
Desvio Padrão	σ	s
Variância	σ^2	s^2
Proporção	ρ	\bar{p}
	Letras gregas	Letras do latim

Passo 1: Defina as hipóteses

Hipótese: afirmação sobre um parâmetro da população.

Hipótese nula (H_0): especifica uma ação padrão, preserva o status quo, “não há mudanças”.

Hipótese alternativa (H_a ou H_1): contradiz a hipótese nula.
Mostra o que queremos testar.

Hipóteses nula e alternativa

As hipóteses são duas afirmações exaustivas e mutuamente exclusivas.

Exemplo: Quero testar se a média é menor que 3.

H_0 : a média é maior ou igual a 3

H_a : a média é menor que 3

$H_0: \mu \geq 3$ O sinal de igual sempre está na H_0

$H_a: \mu < 3$

Hipóteses

O teste pode ser:

- a. unicaudal à esquerda
- b. unicaudal à direita
- c. bicaudal

$$H_0: \mu \geq \mu_0$$

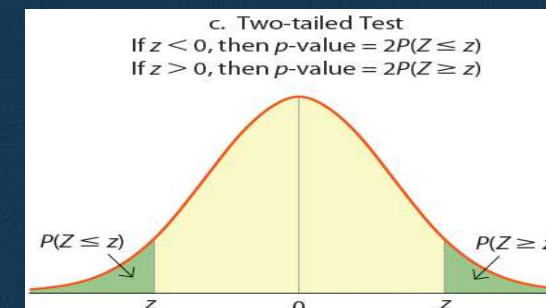
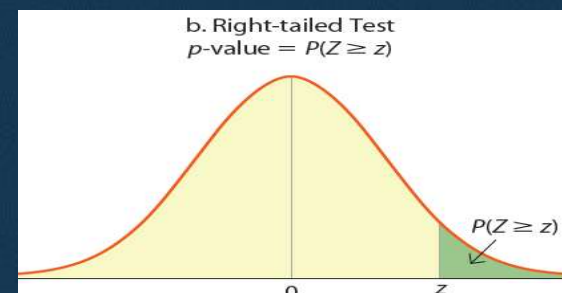
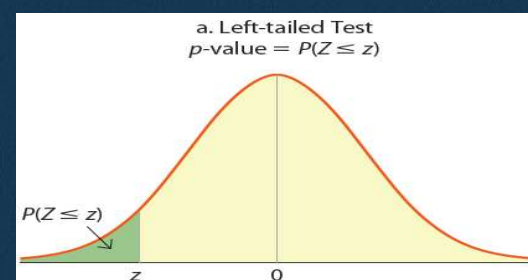
$$H_a: \mu < \mu_0$$

$$H_0: \mu \leq \mu_0$$

$$H_a: \mu > \mu_0$$

$$H_0: \mu = \mu_0$$

$$H_a: \mu \neq \mu_0$$



Tomada de Decisão

Ao conduzir um teste de hipóteses, existem duas conclusões possíveis (que serão decididas ao final da análise):

- Rejeitar H_0
- Não rejeitar H_0

Aplicação: teste t



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

Problema



Uma pizzaria afirma que o tempo de entrega delivery é de 30 minutos ou menos em média, mas você acha que é mais que isso.

Você coleta uma amostra, na qual anotou os minutos de várias entregas, para realizar um teste de hipóteses.

Analisando 30 entregas, a média é de 32 minutos, com desvio padrão de 8 minutos.

Passo 1 no Problema



Defina suas hipóteses nula e alternativa.

Comece pelo que você quer testar.

$H_0: \mu \leq 30$

$H_a: \mu > 30$

O teste é unicaudal à direita.

Tomada de Decisão

$$H_0: \mu \leq 30$$

$$H_a: \mu > 30$$



Ao final dos 4 passos, eu vou concluir o problema com uma destas afirmações:

- Rejeitar H_0

O tempo médio de entrega é maior que 30 minutos, vou processar essa pizzaria!

- Não rejeitar H_0

Não tenho evidências para concluir que o tempo médio de entrega é maior que 30 minutos.

**Passo 2: Escolha o nível de
significância α**

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

Tipos de Erros

Erro tipo I e Erro tipo II

Erro tipo I: Rejeitar H_0 incorretamente.

Processar a pizzeria injustamente.

Erro tipo II: Manter H_0 incorretamente.

Concluir que a pizzeria está correta quando não está.

Tipos de Erros

Erro tipo I: Rejeitar H_0 quando é verdadeira.

Probabilidade = α

Erro tipo II: Não rejeitar H_0 quando é falsa.

Probabilidade = β

Decisão com base na amostra	H_0 é verdadeira	H_0 é falsa
Não rejeitar H_0	Decisão Correta ($1 - \alpha$)	Erro tipo II (β)
Rejeitar H_0	Erro tipo I (α)	Decisão Correta ($1 - \beta$)

Erro Tipo I
(falso positivo)



Erro Tipo II
(falso negativo)



Erro Tipo I

Ocorre quando incorretamente rejeitamos H_0 quando é verdadeira.

O erro tipo I é o nível de significância α .

É definido antes de realizar o teste de hipóteses.

Valores usados para α são 0,01 (1%), 0,05 (5%) ou 0,10 (10%).

Erro Tipo I

Selecionar o erro tipo I de 0,05 indica que você assume uma chance de 5% de estar errado ao rejeitar a hipótese nula.

Você pode reduzir esse erro, escolhendo um outro valor para α , como 1%.

Porém, quanto menor é α , mais difícil é afirmar que um resultado é estatisticamente significativo, pois mais criterioso fica nosso parâmetro (mais específico e menos sensível).

Poder do Teste

Poder do teste: a probabilidade de rejeitar corretamente a hipótese nula quando ela é falsa (e H_a é verdadeira).

É a capacidade do teste de detectar um efeito real quando ele realmente existe.

O poder do teste é $1-\beta$.

Erro Tipo II

O erro tipo II ocorre quando não rejeitamos H_0 quando é falsa.

É o erro de não detectar um efeito ou diferença que realmente existe.

A probabilidade de cometer um erro de tipo II é β , que está relacionada ao poder do teste.

Erro Tipo II

Podemos diminuir o risco de cometer um erro do tipo II:

- Aumentando o tamanho da amostra;
- Aumentando o nível de significância;
- Reduzindo a variabilidade nos dados (por exemplo, melhorando a precisão das medições ou controlando variáveis de confusão);
- Usando testes estatísticos mais adequados.

Passo 2 no problema



O α é um valor escolhido pelo pesquisador.

Vamos definir $\alpha = 5\%$.

Erro do tipo I: Rejeitar H_0 incorretamente (processar a pizzaria injustamente).

Existe uma probabilidade de até 5% que eu vou processar a pizzaria injustamente.

Exemplo: Phishing

Exemplo: Phishing



Phishing (trocadilho com “fishing”, que significa pesca em inglês).

É a tentativa de obter dados pessoais como senhas e números de cartão (e, indiretamente, dinheiro) disfarçado como uma empresa confiável.

As empresas estão trabalhando em otimizar um limite que identifique um email como falso.

A fatura falhou - conta bloqueada

NETFLIX

Oi [REDACTED]

Estamos tendo problemas com suas informações de faturamento atuais. Tentaremos novamente, mas por enquanto você pode atualizar seu MASTERCARD em seus detalhes de pagamento.

ATUALIZAR CONTA AGORA

Estamos aqui para ajudar quando você precisar. Visite a Central de [Ajuda](#) para mais informações ou [entre em contato conosco](#).

Seus amigos no Netflix

Programa de reembolso de impostos: seu valor de desconto

Oi [REDACTED]

Devido à escalada recente, o governo tomou medidas urgentes para listar a COVID-19 como uma doença de notificação obrigatória na lei.

Como medida de precaução, o governo do Reino Unido cooperou com o National Insurance e com o National Health Service (NHS) para estabelecer um novo programa de reembolso de impostos para lidar com o COVID-19 em seu plano de ação.

Você agora tem direito a um reembolso de imposto de
213,64 [REDACTED]

Obtenha seu reembolso

Esses fundos podem ser usados para ajudar a se proteger do COVID-19, de acordo com as recomendações do NHS para ficar seguro.



Verifique sua conta

Olá, querido assinante!

Precisamos do seu Mithlife para fazer sua conta do PayPal funcionar novamente. Até esta derrubada a conexão com sua conta do PayPal é limitada.

Onde você pode encontrar este problema?

Recentemente, descobrimos que houve uso potencialmente ilegal do cartão de crédito dessa conta do PayPal. Para sua proteção, limitamos o acesso à sua conta do PayPal.

O que você pode fazer agora?

Como resultado dessa atividade suspeita em sua conta, pedimos que você aja como um titular legítimo do cartão

confirmar. Então você pode usar sua conta do PayPal como antes.

O PayPal deseja a você muitas economias ao comprar com custos seguros

Como verificar sua conta do PayPal em apenas algumas etapas:

1

2

3

[TRI] [*****SPAM*****] Atendimento BradescoDiaNoite - Thunderbird

File Edit View Go Message OpenPGP Tools Help

Get Mail Write Address Book Decrypt Reply Reply All Forward Delete Junk Print Stop


Subject: [TRI] [*****SPAM*****] Atendimento BradescoDiaNoite

From: BradescoDiaNoite <Atendimento@bradesco.com.br>

Sender: tri-bounces@ufrgs.br

Date: 06/22/2010 01:52 AM

To: security@ufrgs.br

 **Bradesco** | Rede de Atendimento

Autenticação: E98817

Prezado(a) Cliente,

O Bradesco trabalha continuamente para manter o mais alto nível de segurança em suas transações contábeis. Por precauções de otimização da segurança online pedimos a você Cliente para refazer algumas confirmações de dados.

Verificamos que seu Cartão de Segurança encontra-se prestes a ser desativado em nosso sistema, para sua segurança recomendamos seu recadastramento.

O recadastramento é simples, rápido e seguro, sendo realizado a partir de nosso site:

<http://www.netbankingbradesco.com/>

Atenção: Caso sua atualização não seja realizada, o acesso via Caixas-Eletrônicas e Internet ficará suspenso e seu Cartão junto com Chaves de Segurança serão cancelados, impossibilitando acessos e movimentações. O desbloqueio pode ser feito em qualquer agência (necessário CPF, RG e comprovante de residência).

Dica de Segurança: O Bradesco nunca envia emails com arquivos para download.

Em caso de dúvida, contatar a central Bradesco, pelo e-mail ou pelo [Fone Fácil Bradesco](#), de segunda a sexta-feira das 07:00 às 20:00 horas

© 2010 Banco Bradesco S.A. Todos os direitos reservados.

Part 1.2

<http://gpsvo.net/cache/Cadastro>

Endereço de e-mail falso. Não confie nos dados do cabeçalho. Eles são facilmente forjados.

Logos e dados de entidades conhecidas são usados para dar maior credibilidade.

A ISCA ! É o que motiva o usuário a efetuar alguma ação.

Link falso ! Olhe na parte inferior da imagem o endereço que aparece ao passar o cursor por cima.

Ameaças ou tom de urgência. É o que induz o usuário a agir primeiro e pensar depois!

Esse site hospeda uma página maliciosa, semelhante à do site original. É ali que os dados do usuário serão capturados. (OBS: Esse é um site que fica na Holanda)

Exemplo: Phishing



Quais são as hipóteses?

- H_0 : o email não é phishing (é verdadeiro)
- H_a : o email é phishing

Quais são os erros tipo I e tipo II?

- Tipo I (rejeito H_0 incorretamente): o email é verdadeiro, mas vai para o spam.
- Tipo II (mantenho H_0 incorretamente): o email é phishing, mas vai para a caixa de entrada.

Passo 3: Calcule o p-valor

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

p-valor

O p-valor é a probabilidade de encontrarmos um resultado (ou um mais extremo), dado que a H_0 é verdadeira.

Como o que você está testando é a H_a , um p-valor pequeno é uma evidência contra H_0 .

Lançamento de uma moeda

Você joga uma moeda 100 vezes e obtem cara 30 vezes, então você suspeita que pode ser uma moeda “viciada”.

H_0 : a moeda não é viciada (é normal)

H_a : a moeda é viciada

Lançamento de uma moeda

O p-valor é a probabilidade de encontrarmos um resultado (ou um mais extremo), dado que a H_0 é verdadeira.

O p-valor é a probabilidade de obtermos 30 caras (ou menos), dado que a moeda é normal.

- Um p-valor alto significaria que o resultado obtido (30 caras) é comum em uma moeda normal.
- Um p-valor baixo sugeriria que esse resultado é incomum e que a moeda pode ser “viciada”.

Teste Estatístico

Os testes estatísticos se baseiam na distribuição amostral da média que estima o parâmetro especificado nas hipóteses nula e alternativa.

Teste estatístico para a média:

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

Problema



Uma pizzaria afirma que o tempo de entrega delivery é de 30 minutos ou menos em média, mas você acha que é mais que isso.

Você tira uma amostra, anotando minutos de entrega, para realizar um teste de hipóteses.

Analizando 30 entregas, a média é de 32 minutos, com desvio padrão de 8 minutos.

Problema



O p-valor é a probabilidade de encontrarmos um resultado (ou um mais extremo), dado que a H_0 é verdadeira.

Pizzaria: é a probabilidade da entrega levar uma média de 32 minutos ou mais, dado que a pizzaria está correta (média é menor que 30).

Passo 3 no problema



$$\text{p-valor} = P(T > t)$$

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

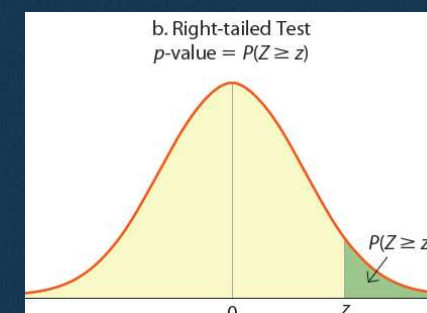
$$\begin{aligned} \mu_0 &= 30 \\ \bar{x} &= 32 \\ s &= 8 \\ n &= 30 \end{aligned}$$



$$t = \frac{32 - 30}{8/\sqrt{30}} = 1,37$$



$$P(T > 1,37) = 0,09$$



**Passo 4: Conclua e
interprete o resultado**

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

Tomada de Decisão

Regras de decisão:

- Rejeite H_0 se $p\text{-valor} < \alpha$
- Não rejeite H_0 se $p\text{-valor} \geq \alpha$

Conclua e interprete os resultados no contexto do problema.

Passo 4 no Problema

$$H_0: \mu \leq 30$$
$$H_a: \mu > 30$$



p-valor = 0,09

p-valor $\geq \alpha = 0,05$

H_0 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.

Suposições

Teste t para uma amostra e duas amostras:

- Normalidade dos dados (em amostras pequenas; o TLC pode ser usado em amostras grandes).
- Homogeneidade das variâncias (para o teste t de duas amostras com variâncias iguais).
- Independência das observações.

Teste de hipóteses x Intervalo de Confiança

O intervalo de confiança cria um intervalo e te dá uma margem de erro, o que o teste de hipóteses não faz.

Mas ele te dá uma “decisão”: baseado no seu α e comparando com o p-valor encontrado, você vai rejeitar a hipótese nula ou não.

Os dois fazem coisas diferentes, mas no fundo o resultado é o mesmo, então fica a critério do pesquisador.



**“NÃO REJEITA HO” x
“ACEITA HO”**

Exemplo: Julgamento

Em um julgamento, o réu é tido como inocente e a acusação tenta mostrar que o réu é culpado.



Na linguagem de teste de hipóteses, quais são as duas hipóteses possíveis?

Lembrando que as hipóteses devem ser exaustivas e mutuamente exclusivas.

Exemplo: Julgamento



Quais são as hipóteses nula e alternativa?

H_0 : O réu não é culpado (é inocente)

H_a : O réu é culpado

O que significa rejeitar a hipótese nula?

Rejeito que o réu é inocente, logo ele é declarado culpado.

O que significa se nós não rejeitamos a hipótese nula?

Não há provas para julgar o réu culpado
(o pedido é improcedente).

Exemplo: Julgamento



Não declaramos que o réu é inocente.

As provas são coletadas para tentar provar a hipótese alternativa (rejeitando a nula).

Na falta de provas, dizemos que não temos evidências para rejeitar a hipótese nula.

Da mesma forma fazemos com o teste de hipóteses.

Não podemos “aceitar” a nula. Concluimos que não temos provas para rejeitá-la.

O PIOR TIPO DE ERRO

Tipos de Erros

Erro tipo I: Rejeitar H_0 quando é verdadeira.

Probabilidade = α

Erro tipo II: Não rejeitar H_0 quando é falsa.

Probabilidade = β

Decisão com base na amostra	H_0 é verdadeira	H_0 é falsa
Não rejeitar H_0	Decisão Correta ($1 - \alpha$)	Erro tipo II (β)
Rejeitar H_0	Erro tipo I (α)	Decisão Correta ($1 - \beta$)

Exemplo: Medicamento

Um pesquisador médico deseja comparar a eficácia de dois medicamentos.

As hipóteses nula e alternativa são:

H_0 : A eficácia dos dois medicamentos são iguais

H_a : A eficácia dos dois medicamentos são diferentes

Exemplo: Medicamento

Erro tipo I: Rejeitar H_0 quando verdadeira

Concluir que os dois medicamentos são diferentes quando, de fato, eles têm a mesma eficácia.

Erro do tipo II: Não rejeitar H_0 quando falsa

Concluir que os medicamentos têm a mesma eficácia quando, de fato, eles são diferentes.

Exemplo: Julgamento

Em um julgamento, o réu é tido como inocente e a acusação tenta mostrar que o réu é culpado.

As hipóteses nula e alternativa são:

H_0 : O réu não é culpado (é inocente)

H_A : O réu é culpado



Exemplo: Julgamento

Erro Tipo I: Rejeito H_0 quando verdadeira
Concluo que o réu é culpado mas ele não é (é condenado).

Erro Tipo II: Não rejeito H_0 quando falsa
Concluo que o réu não é culpado quando ele é (fica em liberdade).



Conclusão

Conforme você conduz seu teste de hipóteses, considere os riscos de cometer os erros do tipo I e do tipo II.

Se as consequências de um tipo de erro forem mais graves do que o outro tipo de erro, escolha um nível de significância e uma poder do teste que refletirá a gravidade relativa dessas consequências.

P-HACKING

p-hacking

p-hacking é um termo relativamente novo, criado em 2014, e é usado para descrever a manipulação dos dados para obter um p-valor desejado.

Quando queremos provar uma hipótese (alternativa), queremos rejeitar a hipótese nula.

Ou seja, o “sonho” dos pesquisadores é encontrar um p-valor $< \alpha$.

p-hacking

Essas manipulações acontecem ao tomar certas decisões somente quando é conveniente, como:

- Trocar o valor do α no final da análise
- Parar de coletar dados
- Transformar os dados

p-hacking

O p-hacking pode ser feito de forma inconsciente.

Pesquisas mostram que pequenas manipulações, sendo conscientes ou não, podem produzir um p-valor $< 0,05$ em 61% das vezes.

Exemplo: Ao encontrar o p-valor=0,06, pode existir a tentação de alterar o teste estatístico, testar algumas transformações, ou excluir uma observação que não segue o padrão, a fim de diminuir esse p-valor.

Porém isso é errado.

p-hacking

Como prevenir o p-hacking?

- Decida seus parâmetros antes de analisar os dados e reporte qualquer mudança.
- Decida quando irá parar de coletar dados e o que será considerado um outlier.
- Replique seus próprios resultados.