

Estatística I

Prof. Fernando de Souza Bastos
fernando.bastos@ufv.br

Departamento de Estatística
Universidade Federal de Viçosa
Campus UFV - Viçosa

Testes de Hipóteses

Tipos de Erros

Procedimento geral para Testes de Hipóteses

Testes de Hipóteses

Testes de hipóteses

Muitos problemas em engenharia requerem que decidamos qual das duas afirmações competitivas acerca do valor de algum parâmetro é verdadeira. As afirmações são chamadas de **hipóteses**, e o procedimento de tomada de decisão sobre a hipótese é chamado de **teste de hipóteses**. Esse é um dos mais úteis aspectos da inferência estatística, uma vez que muitos tipos de problemas de tomada de decisão, teste, ou experimentos no mundo da engenharia podem ser formulados como problemas de teste de hipóteses.

Testes de hipóteses

Exemplo prático (Montgomery e Runger, 2016):

Suponha que um engenheiro esteja projetando um sistema de escape da tripulação de uma aeronave, que consiste em um assento de ejeção e um motor de foguete que energiza o assento. O motor de foguete contém um propelente. Para o assento de ejeção funcionar apropriadamente, o propelente deve ter uma taxa mínima de queima de 50 cm/s. Se a taxa de queima for muito baixa, o assento de ejeção poderá não funcionar apropriadamente, levando a uma ejeção não segura. Taxas maiores de queima podem implicar instabilidade no propelente ou um assento de ejeção muito potente, levando outra vez a insegurança da injeção.

Testes de hipóteses

Dessa maneira, a questão prática de engenharia que tem de ser respondida é: a taxa média de queima do propelente é igual a 50 cm/s ou é igual a algum outro valor (maior ou menor)?

Testes de hipóteses

A taxa de queima é uma variável aleatória que pode ser descrita por uma distribuição de probabilidades. Suponha que nosso interesse esteja focado na taxa média de queima (um parâmetro dessa distribuição). Especificamente, estamos interessados em decidir se a taxa média de queima é ou não 50 centímetros por segundo. Podemos expressar isso formalmente como:

$$H_0 : \mu = 50cm/s \tag{1}$$

$$H_1 : \mu \neq 50cm/s$$

Testes de hipóteses

Hipótese Estatística:

Uma hipótese estatística é uma afirmação sobre os parâmetros de uma ou mais populações.

Testes de hipóteses

A afirmação $H_0 : \mu = 50$ centímetros por segundo na Equação 1 é chamada de hipótese nula, e a afirmação $H_1 : \mu \neq 50$ centímetros por segundo é chamada de hipótese alternativa. Uma vez que a hipótese alternativa especifica valores de μ que poderiam ser maiores ou menores do que 50 centímetros por segundo, ela é chamada de hipótese alternativa bilateral.

Testes de hipóteses

A afirmação $H_0 : \mu = 50$ centímetros por segundo na Equação 1 é chamada de hipótese nula, e a afirmação $H_1 : \mu \neq 50$ centímetros por segundo é chamada de hipótese alternativa. Uma vez que a hipótese alternativa especifica valores de μ que poderiam ser maiores ou menores do que 50 centímetros por segundo, ela é chamada de hipótese alternativa bilateral.

Em algumas situações, podemos ter uma hipótese alternativa unilateral, como em:

$$\begin{cases} H_0 : \mu = 50cm/s \\ H_1 : \mu > 50cm/s \end{cases} \quad \text{ou} \quad \begin{cases} H_0 : \mu = 50cm/s \\ H_1 : \mu < 50cm/s \end{cases} \quad (2)$$

Testes de hipóteses

Sempre estabeleceremos a hipótese nula como uma reivindicação de igualdade. Entretanto, quando a hipótese alternativa for estabelecida com o sinal $<$, a reivindicação implícita na hipótese nula será \geq e quando a hipótese alternativa for estabelecida com o sinal $>$, a reivindicação implícita na hipótese nula será \leq .

Testes de hipóteses

Teste de hipóteses se apoiam no uso de informações de uma amostra aleatória proveniente da população de interesse. É importante ressaltar que a verdade ou falsidade de uma hipótese particular pode nunca ser conhecida com certeza, a menos que possamos examinar a população inteira. Testar uma hipótese envolve:

- considerar uma amostra aleatória;

Testes de hipóteses

Teste de hipóteses se apoiam no uso de informações de uma amostra aleatória proveniente da população de interesse. É importante ressaltar que a verdade ou falsidade de uma hipótese particular pode nunca ser conhecida com certeza, a menos que possamos examinar a população inteira. Testar uma hipótese envolve:

- considerar uma amostra aleatória;
- computar uma estatística de teste a partir de dados amostrais

Testes de hipóteses

Teste de hipóteses se apoiam no uso de informações de uma amostra aleatória proveniente da população de interesse. É importante ressaltar que a verdade ou falsidade de uma hipótese particular pode nunca ser conhecida com certeza, a menos que possamos examinar a população inteira. Testar uma hipótese envolve:

- considerar uma amostra aleatória;
- computar uma estatística de teste a partir de dados amostrais
- e então usar a estatística de teste para tomar uma decisão a respeito da hipótese nula.

Testes de Hipóteses Estatísticas

A hipótese nula corresponde à taxa média de queima ser igual a 50 centímetros por segundo e a alternativa corresponde a essa taxa não ser igual a 50 centímetros por segundo. Ou seja, desejamos testar

$$H_0 : \mu = 50 \text{ cm/s} \quad \text{contra} \quad H_1 : \mu \neq 50 \text{ cm/s}$$

Suponha que uma amostra de $n = 10$ espécimes seja testada e que a taxa média \bar{x} seja observada. A média amostral é uma estimativa de μ . Um valor de \bar{x} que caia próximo a $\mu = 50 \text{ cm/s}$ é uma evidência de que μ é realmente 50 cm/s. Por outro lado, uma média amostral que seja consideravelmente diferente de 50 cm/s evidencia a validade da hipótese alternativa H_1 . Assim, a média amostral é a estatística de teste nesse caso.

Testes de hipóteses

A média amostral pode assumir muitos valores diferentes. Suponha que se $48,5 \leq \bar{x} \leq 51,5$, não rejeitaremos a hipótese nula $H_0 : \mu = 50$ e se $\bar{x} < 48,5$ ou $\bar{x} > 51,5$, rejeitaremos a hipótese nula em favor da hipótese alternativa $H_1 : \mu \neq 50$. Isso é ilustrado na Figura abaixo:

Testes de hipóteses

A média amostral pode assumir muitos valores diferentes. Suponha que se $48,5 \leq \bar{x} \leq 51,5$, não rejeitaremos a hipótese nula $H_0 : \mu = 50$ e se $\bar{x} < 48,5$ ou $\bar{x} > 51,5$, rejeitaremos a hipótese nula em favor da hipótese alternativa $H_1 : \mu \neq 50$. Isso é ilustrado na Figura abaixo:

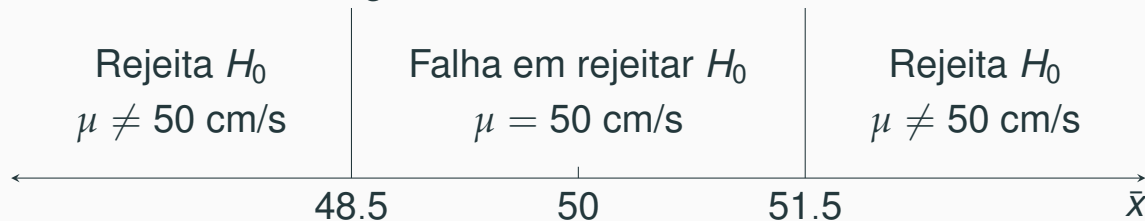


Figura 1: Critérios de decisão para testar $H_0 : \mu = 50$ cm/s versus $H_1 : \mu \neq 50$ cm/s.

Testes de hipóteses

Os valores de \bar{x} que forem menores do que 48,5 e maiores do que 51,5 constituem a **região crítica** para o teste, enquanto todos os valores que estejam no intervalo $48,5 \leq \bar{x} \leq 51,5$ formam uma região para a qual falharemos em rejeitar a hipótese nula. Por convenção, ela geralmente é chamada de **região de não rejeição**. Os limites entre as regiões críticas e a região de aceitação são chamados de valores críticos.

Testes de hipóteses

Em nosso exemplo, os valores críticos são 48,5 e 51,5. É comum estabelecer conclusões relativas à hipótese nula H_0 . Logo, rejeitaremos H_0 em favor de H_1 , se a estatística de teste cair na região crítica e falharmos em rejeitar H_0 por sua vez se a estatística de teste cair na região de aceitação.

Tipos de Erros

Testes de hipóteses

Esse procedimento pode levar a duas conclusões erradas. Por exemplo, a taxa média verdadeira de queima do propelente poderia ser igual a 50 centímetros por segundo. Entretanto, para as amostras de propelente, selecionados aleatoriamente, que são testados, poderíamos observar um valor de estatística de teste \bar{x} que caísse na região crítica. Rejeitaríamos então a hipótese nula H_0 em favor da alternativa H_1 , quando, de fato, H_0 seria realmente verdadeira. Esse tipo de conclusão errada é chamado de **erro tipo I**.

Testes de hipóteses

Esse procedimento pode levar a duas conclusões erradas. Por exemplo, a taxa média verdadeira de queima do propelente poderia ser igual a 50 centímetros por segundo. Entretanto, para as amostras de propelente, selecionados aleatoriamente, que são testados, poderíamos observar um valor de estatística de teste \bar{x} que caísse na região crítica. Rejeitaríamos então a hipótese nula H_0 em favor da alternativa H_1 , quando, de fato, H_0 seria realmente verdadeira. Esse tipo de conclusão errada é chamado de **erro tipo I**.

Erro Tipo I

A rejeição da hipótese nula H_0 quando ela for verdadeira é definida como **erro tipo I**.

Testes de hipóteses

Agora, suponha que a taxa média verdadeira de queima seja diferente de 50 centímetros por segundo, mesmo que a média amostral \bar{x} caia na região de aceitação. Nesse caso, falharíamos em rejeitar H_0 , quando ela fosse falsa. Esse tipo de conclusão errada é chamado de **erro tipo II**.

Testes de hipóteses

Agora, suponha que a taxa média verdadeira de queima seja diferente de 50 centímetros por segundo, mesmo que a média amostral \bar{x} caia na região de aceitação. Nesse caso, falharíamos em rejeitar H_0 , quando ela fosse falsa. Esse tipo de conclusão errada é chamado de **erro tipo II**.

Erro Tipo II

A falha em rejeitar a hipótese nula, quando ela é falsa, é definida como **erro tipo II**.

Testes de hipóteses

Assim, testando qualquer hipótese estatística, quatro situações diferentes determinam se a decisão final está correta ou errada. Pelo fato de a nossa decisão estar baseada em variáveis aleatórias, probabilidades podem ser associadas aos erros tipo I e tipo II. A probabilidade de cometer o erro tipo I é denotada pela letra grega α .

Decisão	H_0 é verdadeira	H_0 é falsa
Não rejeita H_0	Correta Probabilidade = $(1 - \alpha)$	Erro Tipo II Probabilidade = β
Rejeita H_0	Erro Tipo I Nível de significância α	Correta Poder = $(1 - \beta)$

Testes de hipóteses

A probabilidade do erro tipo I é chamada de **nível de significância**, ou **erro** α , ou **tamanho do teste**. No exemplo da taxa de queima de propelente, um **erro tipo I** ocorrerá quando $\bar{x} > 51,5$ ou $\bar{x} < 48,5$, quando a taxa média verdadeira de queima do propelente for $\mu = 50$ cm/s.

Testes de hipóteses

Suponha que o desvio-padrão da taxa de queima seja $\sigma = 2,5$ centímetros por segundo e que a taxa de queima tenha uma distribuição para a qual as condições do **teorema central do limite** se aplicam; logo, a distribuição da média amostral é aproximadamente normal, com média $\mu = 50$ e desvio-padrão

$$\frac{\sigma}{\sqrt{n}} = \frac{2.5}{\sqrt{10}} = 0.79.$$

Testes de hipóteses

A probabilidade de cometer o **erro tipo I** (ou o nível de significância de nosso teste) é igual à soma das áreas sombreadas nas extremidades da distribuição normal na Figura abaixo:

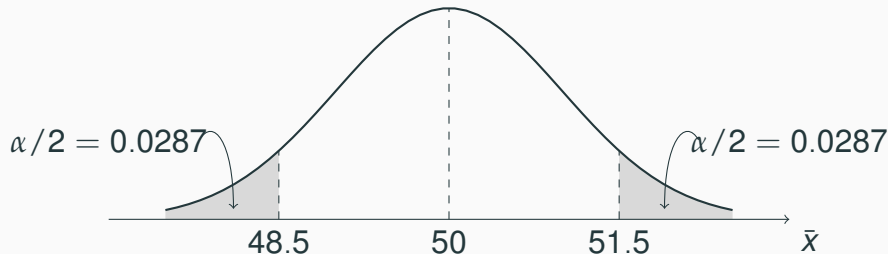


Figura 2: Região crítica para $H_0 : \mu = 50$ versus $H_1 : \mu \neq 50$ e $n = 10$

Testes de hipóteses

A probabilidade de cometer o **erro tipo I** (ou o nível de significância de nosso teste) é igual à soma das áreas sombreadas nas extremidades da distribuição normal na Figura abaixo:

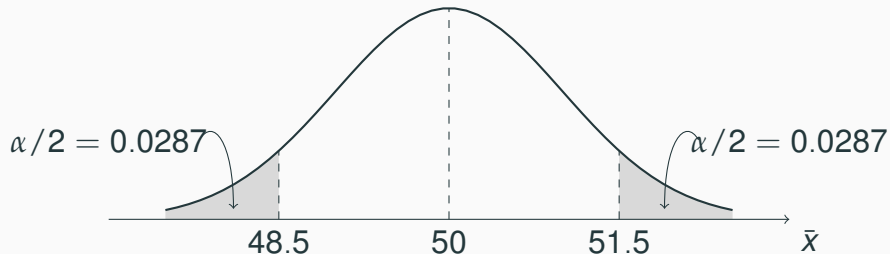


Figura 2: Região crítica para $H_0 : \mu = 50$ versus $H_1 : \mu \neq 50$ e $n = 10$

$$\alpha = P(\bar{X} < 48.5 \text{ quando } \mu = 50) + P(\bar{X} > 51.5 \text{ quando } \mu = 50).$$

Os valores de z que correspondem aos valores críticos 48,5 e 51,5 são

$$z_1 = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{48.5 - 50}{0.79} = -1.9 \quad \text{e} \quad z_2 = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{51.5 - 50}{0.79} = 1.9$$

Logo,

$$\alpha = P(z < -1.90) + P(z > 1.90) = 0.0287 + 0.0287 = 0.0574$$

Essa é a probabilidade do erro tipo I. Isso implica que 5,74% de todas as amostras aleatórias conduziram à rejeição da hipótese $H_0 : \mu = 50$ cm/s, quando a taxa média verdadeira de queima fosse realmente 50 centímetros por segundo. Da inspeção da Figura anterior, notamos que podemos reduzir α alargando a região de aceitação.

Por exemplo, se considerarmos os valores críticos 48 e 52, o valor de α será:

$$\begin{aligned}\alpha &= P\left(z < -\frac{48 - 50}{0.79}\right) + P\left(z > \frac{52 - 50}{0.79}\right) \\ &= P(z < -2.53) + P(z > 2.53) \\ &= 0.0057 + 0.0057 = 0.0114\end{aligned}$$

Por exemplo, se considerarmos os valores críticos 48 e 52, o valor de α será:

$$\begin{aligned}\alpha &= P\left(z < -\frac{48 - 50}{0.79}\right) + P\left(z > \frac{52 - 50}{0.79}\right) \\ &= P(z < -2.53) + P(z > 2.53) \\ &= 0.0057 + 0.0057 = 0.0114\end{aligned}$$

Poderíamos também reduzir α , **aumentando o tamanho da amostra**. Se $n = 16$, então $\frac{\sigma}{\sqrt{n}} = \frac{2.5}{\sqrt{16}} = 0.625$. Logo,

$$z_1 = \frac{48.5 - 50}{0.625} = -2.40 \quad \text{e} \quad z_2 = \frac{51.5 - 50}{0.625} = 2.40$$

e, $\alpha = P(z < -2.40) + P(z > 2.40) = 0.0082 + 0.0082 = 0.0164$.

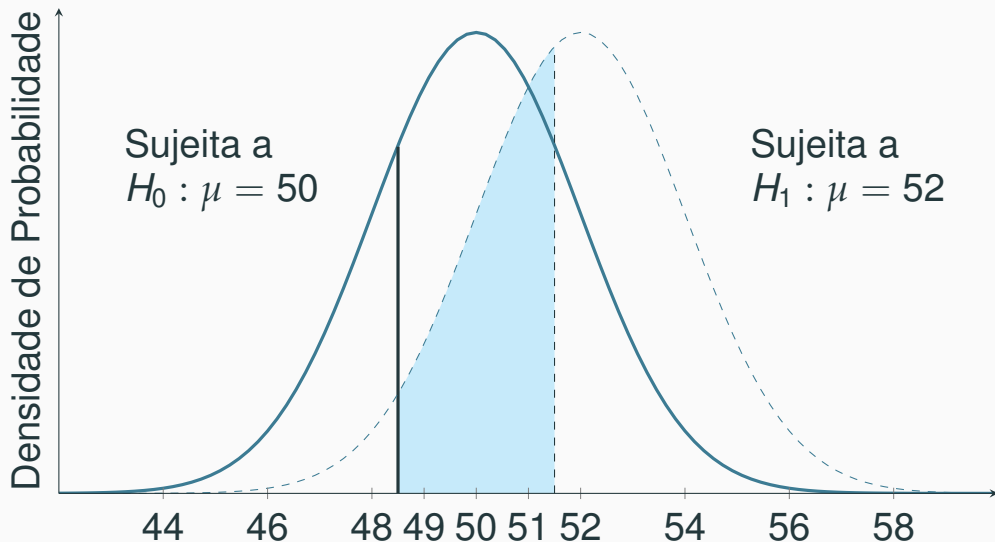
No entanto, na avaliação de um procedimento de teste de hipóteses, também é importante examinar a probabilidade do **erro tipo II**, que é denotado por β . Lembremos que,

$$\beta = P(\text{Erro tipo II}) = P(\text{Não rejeitar } H_0 \text{ dado que } H_0 \text{ é falsa})$$

Para calcular β (algumas vezes chamado de erro β), temos de ter uma hipótese alternativa específica fixada; ou seja, temos de ter um valor particular de μ . Por exemplo, suponha que seja importante rejeitar a hipótese nula $H_0 : \mu = 50$ toda vez que a taxa média de queima μ seja maior do que 52 cm/s ou menor do que 48 cm/s.

Poderíamos calcular a probabilidade de um erro tipo II, β , para os valores $\mu = 52$ e $\mu = 48$ e usar esse resultado para nos dizer alguma coisa acerca de como seria o desempenho do procedimento de teste. Por causa da simetria, só é necessário avaliar um dos dois casos. Isto é, encontrar a probabilidade de não rejeitar a hipótese nula $H_0 : \mu = 50$ cm/s, quando a média verdadeira, por exemplo, for $\mu = 52$ cm/s.

A próxima Figura nos ajudará a calcular a probabilidade do erro tipo II, β .



Um erro tipo II será cometido, se a média amostral \bar{X} cair entre 48,5 e 51,5, quando $\mu = 52$. Como visto na Figura anterior, essa é apenas a probabilidade de $48,5 \leq \bar{X} \leq 51,5$, quando a média verdadeira for $\mu = 52$, ou a área sombreada sob a distribuição normal centralizada em $\mu = 52$. Consequentemente, referindo-se à anterior, encontramos que

$$\beta = P(48.5 \leq \bar{X} \leq 51.5, \text{ quando } \mu = 52)$$

Os valores z , correspondentes a 48,5 e 51,5, quando $\mu = 52$, são

$$z_1 = \frac{48.5 - 52}{0.79} = -4.43 \quad \text{e} \quad z_2 = \frac{51.5 - 52}{0.79} = -0.63$$

logo,

$$\beta = P(-4.43 \leq z \leq -0.63) = 0.2643.$$

Geralmente, o(a) analista controla a probabilidade α do erro tipo I quando ele ou ela seleciona os valores críticos. Assim, geralmente é fácil para o analista estabelecer a probabilidade de erro tipo I em (ou perto de) qualquer valor desejado.

Geralmente, o(a) analista controla a probabilidade α do erro tipo I quando ele ou ela seleciona os valores críticos. Assim, geralmente é fácil para o analista estabelecer a probabilidade de erro tipo I em (ou perto de) qualquer valor desejado.

Uma vez que o analista pode controlar diretamente a probabilidade de rejeitar erroneamente H_0 , sempre pensamos na rejeição da hipótese nula H_0 como uma **conclusão forte**.

Uma vez que podemos controlar a probabilidade de cometer um erro tipo I (ou nível de significância), uma questão lógica é que valor deve ser usado?

Uma vez que podemos controlar a probabilidade de cometer um erro tipo I (ou nível de significância), uma questão lógica é que valor deve ser usado?

A probabilidade do **erro tipo I** é uma medida de risco, especificamente o risco de concluir que a hipótese nula é falsa quando ela realmente não é. Assim, o valor de α deve ser escolhido para refletir as consequências (econômicas, sociais etc.) de rejeitar incorretamente a hipótese nula. Frequentemente, isso é difícil de fazer, e o que tem evoluído muito na prática científica e de engenharia é usar o valor $\alpha = 0,05$ na maioria das situações, a menos que haja alguma informação disponível que indique que essa é uma escolha não apropriada.

Por outro lado, a probabilidade β do erro tipo II não é constante, mas depende do valor verdadeiro do parâmetro. Ela depende também do tamanho da amostra que tenhamos selecionado. Pelo fato de a probabilidade β do erro tipo II ser uma função do tamanho da amostra e da extensão com que a hipótese nula H_0 seja falsa, costuma-se pensar na aceitação de H_0 como uma conclusão fraca, a menos que saibamos que β seja aceitavelmente pequena.

Consequentemente, em vez de dizer “aceitar H_0 ”, preferimos a terminologia “**Não rejeitar H_0** ”. Falhar em rejeitar H_0 implica que não encontramos evidência suficiente para rejeitar H_0 , ou seja, para fazer uma afirmação forte. Falhar em rejeitar H_0 não significa necessariamente que haja uma alta probabilidade de que H_0 seja verdadeira. Isso pode significar simplesmente que mais dados são requeridos para atingir uma conclusão forte, o que pode ter implicações importantes para a formulação das hipóteses.

Procedimento geral para Testes de Hipóteses

O uso da seguinte sequência de etapas na metodologia de aplicação de testes de hipóteses é recomendado.

1. Parâmetro de interesse: A partir do contexto do problema, identifique o parâmetro de interesse.

O uso da seguinte sequência de etapas na metodologia de aplicação de testes de hipóteses é recomendado.

1. Parâmetro de interesse: A partir do contexto do problema, identifique o parâmetro de interesse.
2. Hipótese nula, H_0 : Estabeleça a hipótese nula H_0 .

O uso da seguinte sequência de etapas na metodologia de aplicação de testes de hipóteses é recomendado.

1. Parâmetro de interesse: A partir do contexto do problema, identifique o parâmetro de interesse.
2. Hipótese nula, H_0 : Estabeleça a hipótese nula H_0 .
3. Hipótese alternativa, H_1 : Especifique uma hipótese alternativa apropriada, H_1 .

O uso da seguinte sequência de etapas na metodologia de aplicação de testes de hipóteses é recomendado.

1. Parâmetro de interesse: A partir do contexto do problema, identifique o parâmetro de interesse.
2. Hipótese nula, H_0 : Estabeleça a hipótese nula H_0 .
3. Hipótese alternativa, H_1 : Especifique uma hipótese alternativa apropriada, H_1 .
4. Estatística de teste: Determine uma estatística apropriada de teste.

O uso da seguinte sequência de etapas na metodologia de aplicação de testes de hipóteses é recomendado.

1. Parâmetro de interesse: A partir do contexto do problema, identifique o parâmetro de interesse.
2. Hipótese nula, H_0 : Estabeleça a hipótese nula H_0 .
3. Hipótese alternativa, H_1 : Especifique uma hipótese alternativa apropriada, H_1 .
4. Estatística de teste: Determine uma estatística apropriada de teste.
5. Valor Tabelado ou Valor-P: Determine o valor tabelado a partir do nível de significância ou o valor-p do teste.




O uso da seguinte sequência de etapas na metodologia de aplicação de testes de hipóteses é recomendado.

1. Parâmetro de interesse: A partir do contexto do problema, identifique o parâmetro de interesse.
2. Hipótese nula, H_0 : Estabeleça a hipótese nula H_0 .
3. Hipótese alternativa, H_1 : Especifique uma hipótese alternativa apropriada, H_1 .
4. Estatística de teste: Determine uma estatística apropriada de teste.
5. Valor Tabelado ou Valor-P: Determine o valor tabelado a partir do nível de significância ou o valor-p do teste.
6. Rejeita se H_0 : Estabeleça os critérios de rejeição para a hipótese nula.




O uso da seguinte sequência de etapas na metodologia de aplicação de testes de hipóteses é recomendado.

1. Parâmetro de interesse: A partir do contexto do problema, identifique o parâmetro de interesse.
2. Hipótese nula, H_0 : Estabeleça a hipótese nula H_0 .
3. Hipótese alternativa, H_1 : Especifique uma hipótese alternativa apropriada, H_1 .
4. Estatística de teste: Determine uma estatística apropriada de teste.
5. Valor Tabelado ou Valor-P: Determine o valor tabelado a partir do nível de significância ou o valor-p do teste.
6. Rejeita se H_0 : Estabeleça os critérios de rejeição para a hipótese nula.
7. Conclusões: Decida se H_0 deve ou não ser rejeitada e reporte isso no contexto do problema.

Referências

-  Bastos, Fernando de Souza (2025). ***Apostila Interativa***. Disponível online: <https://ufvest.shinyapps.io/ApostilaInterativa/>.
-  Ferreira, Eric Batista e Marcelo Silva de Oliveira (2020). ***Introdução à Estatística com R***. Editora Universidade Federal de Alfenas. URL: https://www.unifal-mg.edu.br/bibliotecas/wp-content/uploads/sites/125/2021/12/32-EBR_Unifal.pdf.
-  Meyer, Paul L (1982). ***Probabilidade: aplicações à estatística***. Livros Técnicos e Científicos.

Referências ii

-  Montgomery, D. C. e G. C Runger (2016). ***Estatística Aplicada E Probabilidade Para Engenheiros***. 6ª ed. São Paulo: Grupo Gen-LTC.
-  Morettin, P.A. e W.O Bussab (2023). ***Estatística básica***. 10ª ed. São Paulo: Editora Saraiva.
-  Peternelli, Luiz Alexandre (s.d.). ***Apostila (EST 106)***. Formato slide Disponível no PVANet - Moodle.