PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TRANSPORTES E GESTÃO TERRITORIAL – PPGTG
DEPARTAMENTO DE ENGENHARIA CIVIL – ECV

DISCIPLINA: TGT410026 – FUNDAMENTOS DE ESTATÍSTICA

9ª AULA: TESTES DE HIPÓTESES

INTRODUÇÃO

Os Testes de Hipóteses em conjunto com a estimação por intervalo de parâmetros populacionais o objetivo central da Inferência estatística.
Veremos nesta aula como o Teste de Hipóteses pode ser usado para determinar se uma afirmação sobre o valor de um parâmetro populacional deve ou não ser rejeitada.
Ao testar hipóteses, inicia-se pela criação de uma hipótese experimental a respeito de um parâmetro da população. A essa hipótese experimental dá-se o nome de **HIPÓTESE NULA – H_0**. A partir daí, define-se uma outra hipótese, denominada **HIPÓTESE ALTERNATIVA – H_a**, a qual é o OPOSTO daquilo que foi formulado na hipótese nula.
O procedimento de teste de hipóteses usa dados de uma amostra aleatória para testar as duas afirmações antagônicas, H_0 e H_a.
Revisaremos nessa aula os testes de hipóteses envolvendo a média e a proporção populacional.

COMO DESENVOLVER AS HIPÓTESES ESTATÍSTICAS

Em geral, pode não ser claro à primeira vista como as hipóteses nula e alternativa devem ser formuladas. Na sequência, são apresentadas diretrizes para o estabelecimento das hipóteses nula e alternativa para três tipos de situação nas quais comumente se empregam os procedimentos de testes de hipótese.

a) Como testar hipóteses de pesquisa

Em estudos de pesquisa as hipóteses nula e alternativa devem ser formuladas de tal maneira que a rejeição da hipótese nula corrobore a conclusão da pesquisa, isto é, a hipótese alternativa deve ser validada. As hipóteses de pesquisa, portanto, devem ser expressas como a hipótese alternativa.

Vejamos um exemplo: Considere um modelo de automóvel que atinge uma eficiência média de 10,21 km/l em termos de consumo de combustível. Uma equipe de pesquisa de produto desenvolveu um novo sistema de injeção de combustível, projetado para aumentar a taxa de quilômetros por litro de combustível.
O objetivo da equipe de pesquisa é comprovar que o novo sistema ***AUMENTA*** a taxa média de quilômetros por litro de combustível. Nesse caso, a hipótese de pesquisa é que o novo sistema de injeção de combustível produzirá uma taxa média de quilômetros por litro superior a 10,21; ou seja, $\mu > 10{,}21$. A formulação das hipóteses básica e alternativa relativas ao estudo são:

$$H_0 : \mu \leq 10{,}21$$
$$H_a : \mu > 10{,}21$$

Note que as duas hipóteses são antagônicas; isso é proposital para que, sendo rejeitada a hipótese básica, automaticamente fica validada a hipótese alternativa, que é a hipótese de pesquisa da equipe de produto.

b) Como testar a validade de uma afirmação

Em situações que envolvem testar a validade de uma afirmação, a hipótese nula geralmente se baseia no pressuposto que a afirmação é VERDADEIRA.

Vejamos um exemplo: Considere o caso de um fabricante de refrigerantes que declara que os frascos de dois litros dos seus produtos contêm, no mínimo, uma média de 1,99 litros.
Nesse tipo de situação, é normal presumir que a informação do fabricante é VERDADEIRA, a menos que a evidência da amostra seja contraditória.
A formulação das hipóteses básica e alternativa relativas à informação do fabricante será:

$$H_0 : \mu \geq 1{,}99$$
$$H_a : \mu < 1{,}99$$

Se os resultados da amostra indicarem que $\mathbf{H_0}$ não pode ser rejeitada, a afirmação do fabricante NÃO SERÁ CONTESTADA. Entretanto, se os resultados da amostra indicarem que $\mathbf{H_0}$ pode ser rejeitada, a inferência é que $\mathbf{H_a}$ seja verdadeira. Com essa conclusão, a evidência estatística indica que a informação do fabricante é incorreta e que os

frascos de refrigerantes são preenchidos com uma média MENOR que a anunciada de 1,99 litros. Medidas cabíveis contra o fabricante devem ser consideradas.
Em situações que envolvem a validação de uma afirmação, a hipótese alternativa deve ser formulada de tal maneira que, em sendo rejeitada a hipótese nula, haja evidência estatística de que a hipótese declarada é incorreta. Nesse caso, iniciativas para corrigir a afirmação (declaração) devem ser consideradas sempre que $\mathbf{H_0}$ for rejeitada.

c) Como testar hipóteses em situação de Tomada de Decisão

Em muitas situações, há a necessidade de se tomar providências tanto quando $\mathbf{H_0}$ não pode ser rejeitada como quando $\mathbf{H_0}$ pode ser rejeitada. Esse é o caso típico de um tomador de decisão que precisa escolher entre dois cursos de ação: um associado à hipótese nula ($\mathbf{H_0}$) e outro, à hipótese alternativa ($\mathbf{H_a}$).

Vejamos um exemplo: Considere uma amostra de peças de uma remessa recém-recebida, um inspetor de qualidade precisa decidir se aceitará a remessa ou se devolverá ao fornecedor porque ela (a remessa) não cumpre as especificações de contrato.
Suponha que as especificações de uma peça em particular exijam um tamanho médio de duas polegadas por peça. Se o tamanho médio for maior ou menor que o padrão de duas polegadas, as peças causarão problemas de qualidade na operação de montagem.
Nesse tipo de situação, as hipóteses básica e alternativa são formuladas da seguinte maneira:

$$H_0 : \mu = 2"$$
$$H_a : \mu \neq 2"$$

Se os resultados da amostra indicarem que $\mathbf{H_0}$ não pode ser rejeitada, o inspetor de controle de qualidade não terá nenhuma razão para duvidar que a remessa esteja de acordo com a especificação, e a remessa será aceita.
Entretanto, se os resultados da amostra indicarem que $\mathbf{H_0}$ deve ser rejeitada, a conclusão será de que as peças não cumprem as especificações. Nesse caso, o inspetor de qualidade terá suficientes evidências para devolver a remessa ao fornecedor.

Resumo das formas das Hipóteses Nula e Alternativa

A forma geral do teste de hipóteses envolvendo a média populacional é dada por:

$$H_0 : \mu \leq \mu_0 \qquad H_0 : \mu \geq \mu_0 \qquad H_0 : \mu = \mu_0$$
$$H_a : \mu > \mu_0 \qquad H_a : \mu < \mu_0 \qquad H_a : \mu \neq \mu_0$$

onde μ_0 é um valor hipotético a ser testado na inferência.

As duas primeiras formas de teste de hipóteses mostradas acima denominam-se testes unicaudais; a terceira forma é denominada teste bicaudal.

NOTA:
EM MUITAS SITUAÇÕES, A ESCOLHA DE $\mathbf{H_0}$ E $\mathbf{H_a}$ NÃO É CLARA, E É NECESSÁRIO DISCERNIMENTO PARA SELECIONAR A FORMA APROPRIADA. ENTRETANTO, COMO MOSTRAM AS FORMAS APRESENTADAS, O TERMO DE IGUALDADE (≤, ≥ ou =) ***SEMPRE*** APARECE NA HIPÓTESE NULA.

Exemplos:

1) O gerente do Praia Mole Resort Hotel afirmou que o valor médio da conta dos hóspedes em um final de semana é igual a US$ 600. Um membro da equipe de contabilidade do hotel observou que o total nas contas dos hóspedes se elevou nos últimos meses. O contador usará uma amostra de contas de hóspedes em fins de semana para testar a afirmação do gerente.
a) qual forma de hipótese deverá ser utilizada para testar a afirmação do gerente? Explique.

$$H_0 : \mu \geq 600 \qquad H_0 : \mu \leq 600 \qquad H_0 : \mu = 600$$
$$H_a : \mu < 600 \qquad H_a : \mu > 600 \qquad H_a : \mu \neq 600$$

b) qual conclusão é apropriada quando $\mathbf{H_0}$ não pode ser rejeitada?
c) qual conclusão é apropriada quando $\mathbf{H_0}$ pode ser rejeitada?

2) O gerente de uma concessionária de automóveis está pensando em um novo plano de bonificações visando aumentar o volume de vendas. Atualmente, o volume médio de vendas é de 14 automóveis por mês. O gerente quer desenvolver um estudo para verificar se o novo plano de bonificações aumenta o volume de vendas. Para coletar dados sobre o plano, uma amostra da equipe de vendas será

autorizada a vender sob o novo plano de bonificações durante um mês.
a) Desenvolva as hipóteses nula e alternativa mais apropriadas a essa situação de pesquisa;
b) Comente a conclusão da situação: H_0 (hipótese nula) não pode ser rejeitada;
c) Comente a conclusão da situação: H_0 (hipótese nula) pode ser rejeitada.

3) Em virtude do tempo e dos custos elevados de produção e transformação, um diretor de manufatura precisa convencer a administração de que um novo método de manufatura proposto reduz os custos, antes de o novo método ser implementado. O método de produção atual opera com um custo médio de R$ 220 por hora. Um estudo e pesquisa medirão o custo do novo método ao longo de um período de produção amostral.
a) Desenvolva as hipóteses nula e alternativa mais apropriadas para o estudo em questão;
b) Comente a conclusão da situação: H_0 (hipótese nula) não pode ser rejeitada;
c) Comente a conclusão da situação: H_0 (hipótese nula) pode ser rejeitada.

ERROS DO TIPO I E DO TIPO II

Ao realizar testes de hipótese nos estudos de inferência, estamos sujeitos à dois tipos de erros: ERRO TIPO I e ERRO TIPO II. Isso ocorre porque as hipóteses nula (H_0) e alternativa (H_a) são afirmações excludentes a respeito da população: ou H_0 é verdadeira ou H_a é verdadeira, mas não ambas.
Do ponto de vista ideal, o procedimento de teste de hipóteses deve levar à ACEITAÇÃO DE H_0 quando H_0 É VERDADEIRA, e a REJEIÇÃO DE H_0 quando H_a É VERDADEIRA.

O resumo dessa situação é mostrada a seguir:

Erros e conclusões no Teste de Hipóteses

		Situação da População	
		H_0 verdadeira	**H_a verdadeira**
Conclusão	Aceitar H_0	Conclusão correta	Erro do Tipo II
do Teste	Rejeitar H_0	Erro Tipo I	Conclusão correta

É fácil ver que nem sempre é possível tomar as decisões corretas. Isso ocorre porque, como os testes de hipóteses baseiam-se em dados

de amostras, devemos admitir a possibilidade de erros. A primeira linha da tabela acima revela o que pode acontecer se a DECISÃO for aceitar H_0. Se H_0 for verdadeira, a decisão do teste está correta. Entretanto, se H_a for verdadeira, cometemos um erro do TIPO II, ou seja, aceitamos H_0 quando ela é falsa.
A segunda linha mostra o que pode acontecer se a DECISÃO for rejeitar H_0. Se H_0 for verdadeira, nesse caso cometemos um erro do TIPO I, ou seja, rejeitamos H_0 quando ela é verdadeira. Finalmente, se H_a for verdadeira, rejeitar H_0 será a decisão correta.

Nível de significância

Podemos apresentar agora uma importante definição: NÍVEL DE SIGNIFICÂNCIA. Num teste de hipóteses, recebe essa denominação, a probabilidade de se cometer um erro do TIPO I quando a HIPÓTESE NULA é verdadeira.

O símbolo grego α (alfa) é utilizado para denotar o nível de significância do teste. Tem tudo a ver com o nível de confiança de uma estimação, onde este é denotado por $(1 - \alpha)$. Como ambos os termos são probabilidades, a soma de α com $(1 - \alpha)$ deve ser igual a 1.
A seguir apresentamos algumas relações entre o nível de confiança e o nível de significância em estudos estatísticos:

Nível de confiança	Nível de significância
$1 - \alpha$	α
90% = 0,90	10% = 0,10
95% = 0,95	5% = 0,05
99% = 0,99	1% = 0,01

NOTA IMPORTANTE:

Conforme vimos, nos testes de hipóteses os erros do TIPO I são controlados ao se adotar um nível de significância α. Entretanto, a probabilidade de se cometer um erro do TIPO II nem sempre é controlada diretamente.
Portanto, se decidirmos aceitar H_0, não poderemos determinar quão confiantes podemos estar a respeito dessa decisão. Em razão da incerteza associada à probabilidade de se cometer o erro TIPO II quando se realizam testes de significância, os estatísticos frequentemente recomendam usar a afirmação “não rejeitar H_0” em vez de “aceitar H_0”.
Isso parece sensato, pois, ao não aceitar H_0 diretamente, o decisor evita o risco de cometer um erro do TIPO II.

Aplicações:

1) O rótulo de um frasco de 2,83 litros de suco de laranja afirma que o suco de laranja contém em média 1 grama ou menos de gordura. Responda as questões a seguir considerando um teste de hipóteses que possa ser usado para testar a afirmação constante no rótulo.

a) desenvolva as hipóteses nula e alternativa apropriadas;

$$H_0 : \mu \leq 1$$

$$H_a : \mu > 1$$

b) qual é o erro do TIPO I nessa situação? Quais são as consequências de cometer esse erro?

Resp: Afirmar que $\mu > 1$ quando isso não é verdade

c) qual é o erro do TIPO II nessa situação? Quais são as consequências de cometer esse erro?

Resp: Afirmar que $\mu \leq 1$ quando isso não é verdade

2) Suponha que um novo método de produção seja implementado se um teste de hipóteses sustentar a conclusão de que o novo método reduz a média de custo operacional por hora.

a) Estabeleça as hipóteses nula e alternativa apropriadas considerando que o custo médio de produção atual seja igual a R$ 220 por hora.

$$H_0 : \mu \geq 220$$

$$H_a : \mu < 220$$

b) qual é o erro do TIPO I nessa situação? Quais são as consequências de cometer esse erro?

Resp: Afirmar que $\mu < 220$ quando isso não é verdade

c) qual é o erro do TIPO II nessa situação? Quais são as consequências de cometer esse erro?

Resp: Afirmar que $\mu \geq 220$ quando isso não é verdade

3) Foi divulgado recentemente que os jovens dos EUA assistem a 56,2 minutos de TV diariamente no horário nobre. Um pesquisador alemão acredita que os jovens alemães do sexo masculino passam mais tempo assistindo à TV no horário nobre. Uma amostra de jovens da Alemanha será selecionada pelo pesquisador e o tempo que eles passam assistindo à TV em um dia será registrado. Os resultados da amostra serão usados para testar as hipóteses nula e alternativa seguintes:

$$H_0 : \mu \leq 56,2$$

$$H_a : \mu > 56,2$$

a) qual é o erro do TIPO I nessa situação? Quais são as consequências de cometer esse erro?

b) qual é o erro do TIPO II nessa situação? Quais são as consequências de cometer esse erro?

MÉDIA DA POPULAÇÃO: σ conhecido

Até agora vimos como se elaboram as hipóteses nula e alternativa que servirão para a realização dos testes de hipóteses. Neste curso trabalharemos os testes de médias e de proporções populacionais.

Os testes de hipóteses envolvendo uma média populacional podem ser desenvolvidos de duas formas: testes unicaudais e testes bicaudais.

Teste Unicaudal

Os testes unidaudais sobre a média de uma população assumem uma das duas formas seguintes:

Teste da cauda inferior

$H_0 : \mu \geq \mu_0$

$H_a : \mu < \mu_0$

Teste da cauda superior

$H_0 : \mu \leq \mu_0$

$H_a : \mu > \mu_0$

Estatística de teste

A estatística a ser adotada para a realização do teste de hipóteses de uma média populacional com desvio padrão conhecido é:

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

A interpretação da função da estatística de teste é a seguinte: quão pequena deve ser a estatística de teste z antes de optarmos por rejeitar a hipótese nula?

Dois critérios podem ser utilizados para resolver esta questão: critério do ***valor p*** e critério do ***valor crítico***.

Valor p

É uma probabilidade calculada, usando-se a estatística de teste, que mede o apoio (ou falta de apoio) proporcionado pela amostra à hipótese nula.

A regra de rejeição da hipótese nula quando se usa o critério do ***valor***

p é: REJEITAR H_0 se o valor p $\leq \alpha$, onde α é o nível de significância do teste.

Valor crítico

É outro critério para decidir a respeito de um teste de hipóteses. O valor crítico é o valor da estatística de teste que corresponde a uma área α (o nível de significância) localizada na cauda inferior da distribuição amostral da estatística de teste. Simplificando o discurso, o valor crítico é o maior valor da estatística de teste que resultará na REJEIÇÃO de H_0.

Operacionalmente determinamos o valor crítico em função do nível de significância α adotado. Em seguida vamos comparar o valor crítico com o valor da estatística de teste (calculada em função dos dados da amostra).

A regra geral de rejeição para um teste unicaudal (inferior) pelo critério do valor crítico é:

Rejeitar H_0 *se $z \leq -z_\alpha$ em que $-z_\alpha$ é o valor crítico, ou seja,*
o valor que produz uma área $\boldsymbol{\alpha}$ *na cauda inferior*
da distribuição normal padrão.

Vamos apresentar um exemplo onde o teste é efetuado e a conclusão é tomada usando os dois critérios.

Exemplo da *Hilltop Coffee*:

A Federação Americana dos Direitos do Consumidor (FADC) realiza periodicamente, estudos estatísticos concebidos para testar as afirmações feitas por fabricantes a respeito de seus produtos.

Como exemplo, vamos discutir o caso do café *Hilltop*. O rótulo de uma lata grande informa que a lata contém 3 libras (1,36 kg) de café.

A FADC sabe que o processo de produção da *Hilltop* não consegue colocar exatamente 3 libras de café em cada lata, mesmo que o peso médio de enchimento da população de todas as latas cheias seja de, no mínimo 3 libras por lata. No entanto, considerando que o peso médio populacional seja de, no mínimo, 3 libras por lata de café, os direitos dos consumidores estão garantidos.

A FADC interpreta a informação contida no rótulo de uma lata grande de café como uma afirmação da parte da empresa *Hilltop* de que o peso médio populacional de enchimento é, no mínimo, 3 libras por lata de café. As discussões a seguir mostrarão como a FADC pode checar a afirmação da *Hilltop* realizando um teste de hipóteses.

A primeira providência (etapa) consiste em desenvolver as hipóteses nula e alternativa para o teste. Se o peso médio populacional de enchimento das latas grande for de, no mínimo, 3 libras por lata, a afirmação da *Hilltop* está correta. Essa premissa estabelece a hipótese nula (H_0)para o teste.

Entretanto, se o peso médio da população for inferior a 3 libras por lata, a afirmação da *Hilltop* está incorreta. Essa premissa estabelece a hipótese alternativa (H_a) para o teste.

Considerando μ como o peso médio de enchimento da população, as hipóteses nula e alternativa são as seguintes:

$$H_0 : \mu \geq 3$$
$$H_a : \mu < 3$$

ENQUANTO preparamos uma amostra para a realização do teste, tente visualizar os desdobramentos das seguintes situações:
a) os dados amostrais indicam que H_0 não pode ser rejeitada;
b) os dados amostrais indicam que H_0 pode ser rejeitada.

Uma amostra de 36 latas de café foi selecionada e apresentou um valor para $\bar{x}$ = 2,92 libras. Esse resultado foi utilizado para estimar a média da população μ.

Em relação a empresa *Hilltop Coffee*, testes anteriores realizados pela FADC mostram que o desvio padrão da população pode ser considerado conhecido e com valor σ = 0,18 libras.

O próximo passo na preparação do teste é decidir pelo nível de significância α. No estudo da *Hilltop Coffee*, o diretor do programa de testes fez a seguinte afirmação: "*Se a empresa está cumprindo suas especificações de peso, com μ = 3, não quero mover nenhum processo contra eles. Não obstante, estou disposto a arriscar uma chance de 1% de cometer esse erro*".

Diante da afirmação do diretor, o nível de significância ficou estabelecido em α = 0,01. Desse modo, devemos projetar o teste de hipóteses de forma que a probabilidade de cometermos um erro do

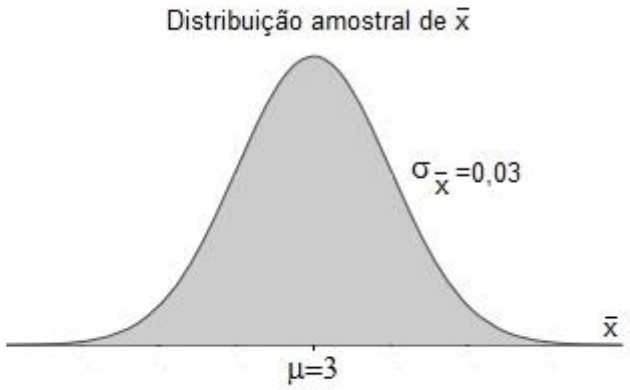
TIPO I seja 0,01.

E a distribuição amostral de $\bar{x}$?

Variáveis como pesos de enchimento, alturas, etc. são tipicamente do tipo contínuas e possuem distribuição normal. Como a amostra foi selecionada dessa população, a variável $\bar{x}$ também possui distribuição normal.

Assim, para o estudo da *Hilltop Coffee*, a distribuição amostral de $\bar{x}$ está normalmente distribuída com $\sigma = 0{,}18$ e tamanho amostral $n = 36$. A figura a seguir mostra a distribuição amostral de $\bar{x}$ quando a hipótese nula é verdadeira, ou seja, $\mu = \mu_0 = 3$ e desvio padrão amostral igual a

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0{,}18}{\sqrt{36}} = 0{,}03$$



Cálculo da estatística de teste: com *n =36*, $\sigma = 0{,}18$, $\bar{x} = 2{,}92$ e $\mu=3$, a estatística de teste *z* é:

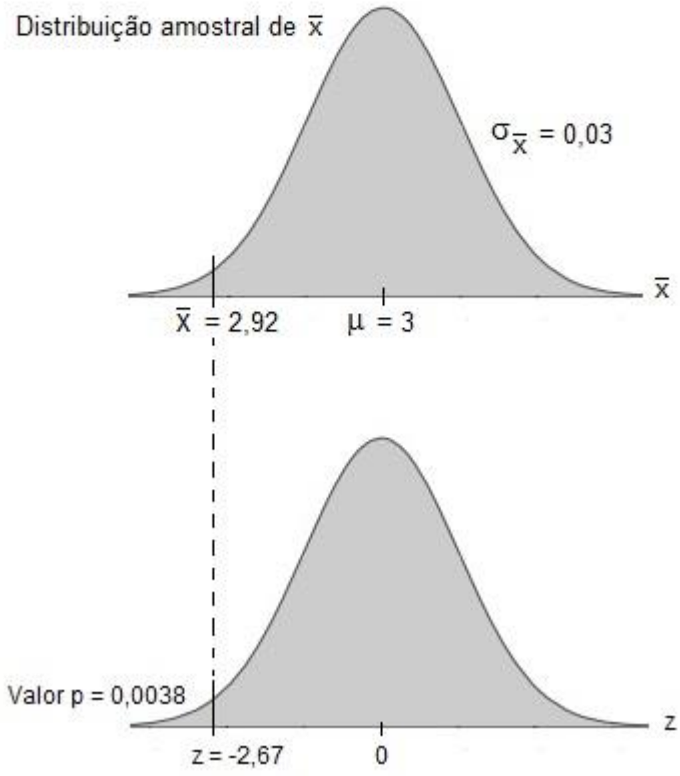
$$z = \frac{\bar{x} - \mu_0}{\sigma\sqrt{n}} = \frac{2{,}92 - 3}{0{,}18\sqrt{36}} = -2{,}67$$

Realizando o teste:

a) Usando o critério do ***valor p***

O ***valor p*** é a probabilidade de a estatística de teste z ser menor que -2,67 (a área sob a curva normal padrão à esquerda da estatística de teste).

Usando a tabela normal, descobrimos que a área entre a média e o valor -2,67 é 0,4962. Assim, o valor p é 0,500 – 0,4962 = 0,0038. A figura a seguir ilustra essa situação.

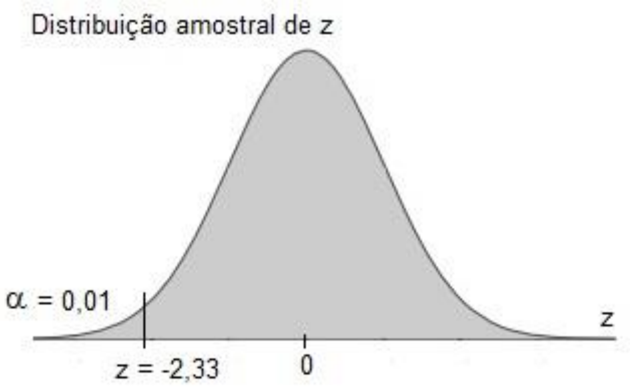


A decisão do teste usando o critério do ***valor p*** é: *"rejeitar H_0 desde que $p \leq \alpha$"*.

Com o ***valor p*** igual a 0,0038 e o valor de $\alpha = 0{,}01$, REJEITAMOS a hipótese nula; existem suficientes evidências estatísticas para rejeitar a hipótese nula dado o nível de significância de 1%.

b) Usando o critério do ***valor crítico***

No caso em que σ é conhecido, a distribuição amostral *z* da estatística de teste é uma distribuição normal padrão. Portanto, o ***valor crítico*** é o valor da estatística de teste que corresponde a uma área α na cauda inferior de uma distribuição normal padrão.
Com $\alpha = 0{,}01$, descobrimos na tabela de distribuição normal padrão que o valor $z = -2{,}33$ produz uma área igual a 0,01 na cauda inferior. A figura a seguir ilustra a situação em estudo.



Aplicando o critério do ***valor crítico***, observamos que z_c = *-2,33* é maior que z = *-2,67*, o que evidencia a rejeição da hipótese nula.

Ambos os critérios, ***valor p*** e ***valor crítico***, levarão à mesma conclusão do teste. A vantagem do critério do ***valor p*** é que este nos diz quão significativos são os resultados (pois, sabemos qual o nível de significância).

Teste Bicaudal

Nos teste de hipóteses, a forma para se desenvolver um teste bicaudal a respeito de uma média populacional é expressa da seguinte maneira:

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

A exemplo do teste unicaudal, valem também aqui os critérios de decisão dos teste de hipóteses: critério do ***valor p*** e critério do ***valor crítico***.

Estatística de teste: é calculada da mesma maneira que no teste unicaudal, ou seja:

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

Onde $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$ é o desvio padrão amostral.

Critério do valor p

Na prática, para calcular o ***valor p*** devemos encontrar a probabilidade de obtermos um valor para a estatística de teste, que seja, *no mínimo, tão improvável quanto valor de z_c calculado (fornecido pela amostra)*. Assim, valores de $z \geq z_c$ são também tão improváveis quanto o valor z_c. Como o teste é bicaudal, valores de $z \leq z_c$ são, também tão improváveis quanto o valor z_c.

Determinado o ***valor p*** usando a tabela de probabilidade normal comparamos com o nível de significância α para saber se a hipótese nula deve ser rejeitada ou não.

ORIENTAÇÃO:

O cálculo do ***valor p*** de um teste bicaudal pode parecer confuso em comparação ao teste unicaudal; entretanto ele é simplificado, seguindo-se as três etapas seguintes:
a) calcule o valor da estatística de teste *z*;
b) se o valor da estatística de teste estiver na cauda superior ($z > 0$), encontre a área sob a curva normal padrão à direita de z; se o valor da estatística de teste estiver na cauda inferior ($z < 0$), encontre a área sob a curva normal padrão à esquerda de *z*.
c) duplique o valor da área da cauda, ou probabilidade, obtida em (b) para obter o ***valor p***.

Critério do valor crítico

Como o teste é bicaudal, existirão dois valores críticos, um na cauda superior e outro na cauda inferior da distribuição normal padrão. Com um nível de significância α, a área em cada cauda além dos valores críticos é α/2.

Assim, usando a tabela de distribuição normal padrão descobrimos que os valores críticos da estatística de teste são $-z_{\alpha/2}$ e $+z_{\alpha/2}$.

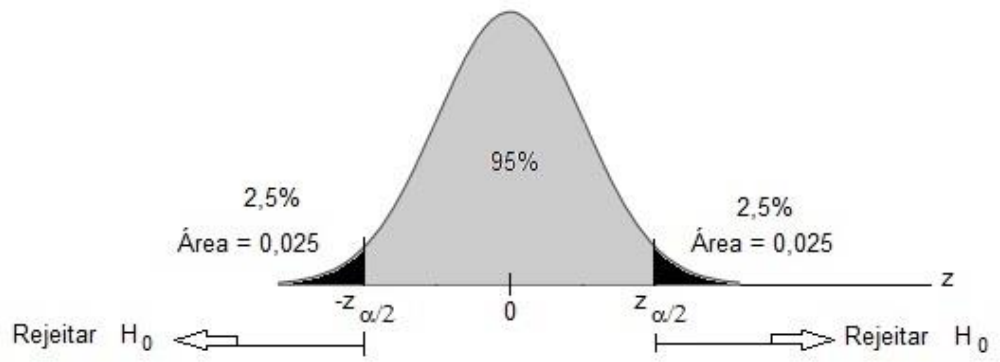
Com isso, o critério do teste pelo valor crítico é: Rejeitar H_0 se:
$z \leq -z_{\alpha/2}$ ou se $z \geq +z_{\alpha/2}$.

Relação entre a Estimação por intervalo e o Teste de Hipóteses

Quando estudamos a estimação por intervalo de uma média populacional com desvio padrão σ conhecido, correspondente a um índice de confiança (1 – α), vimos que é criado um intervalo de confiança $-z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$ a $+z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$, onde estima-se que a média da população esteja contida com aquela confiança estabelecida.

Podemos realizar agora um teste de hipóteses para saber se com uma confiança de $(1 - \alpha)$ e uma correspondente significância α, realmente a média da população está contida no intervalo especificado.

Para realizar o teste de hipóteses de uma média populacional estimada por intervalo e com desvio padrão conhecido, devemos utilizar o teste bicaudal. As áreas extremas (direita e esquerda) da cauda da distribuição normal padrão $(\pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}})$ correspondem às áreas críticas de rejeição da hipótese nula. Isso porque a área de $(1 - \alpha)$ sob a curva normal nos define quão provável esse intervalo contém a média estimada. Veja a figura a seguir e interprete essa situação.



Assim, as hipótese nula e alternativa podem ser estabelecidas para a realização do teste:

$$H_0 : \mu = \mu_0$$
$$H_a : \mu \neq \mu_0$$

CRITÉRIO DO INTERVALO DE CONFIANÇA para testar uma hipótese da forma dada acima:

a) selecione uma amostra simples da população e use o valor da média amostral $\bar{x}$ para desenvolver o intervalo de confiança da média populacional μ:

$$\mu \Rightarrow \bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

b) se o intervalo de confiança contiver o valor hipotético μ_0, não rejeite H_0. Caso contrário, rejeite H_a.

Vamos ilustrar esse processo de teste de hipóteses apresentando um caso real:

A Associação Americana de Golfe estabelece normas que os fabricantes de equipamentos devem cumprir para que seus produtos sejam aceitos e usados nos eventos da AAG. A MaxFlight utiliza um processo de manufatura de alta tecnologia para produzir bolas de golfe que atingem uma distância média de arremesso de 295 jardas (269,7 m). Às vezes, porém, o processo se desajusta e produz bolas de golfe que atingem uma distância média de arremesso diferente de 295 jardas.
Quando a distância média cai abaixo de 295 jardas, a empresa se preocupa em perder vendas pelo fato de as bolas não atingirem a distância anunciada. Quando a distância média passa de 295 jardas, as bolas de golfe da MaxFlight podem ser rejeitadas pela AAG em virtude de excederem o padrão de distância estabelecido.
O programa de controle de qualidade da MaxFlight envolve extrair amostras periódicas de 50 bolas de golfe para monitorar o processo de manufatura. Para cada amostra é realizado um teste de hipóteses com o objetivo de determinar se o processo se desajustou.

Como se realiza o teste:

Primeiramente deve-se estabelecer as hipóteses nula e alternativa. Supondo que o processo de manufatura esteja funcionando corretamente, é natural que a MaxFlight esteja confiante que as bolas produzidas atingem a distância média de 295 jardas; com isso fica estabelecida a hipótese nula: $H_0 : \mu = 295$.
Cabe à amostra apresentar evidências que a distância média alcançada pelas bolas da MaxFlight é diferente de 295 jardas; com isso fica estabelecida a hipótese alternativa: $H_a : \mu \neq 295$.
Teste:

$$H_0 : \mu = 295$$
$$H_a : \mu \neq 295$$

Testes anteriores, realizados quando se sabia que o processo estava ajustado, mostram que o desvio padrão populacional pode ser considerado conhecido e igual a σ = *12* jardas. Selecionada uma amostra de 50 bolas, a empresa determinou a distância média amostral: $\bar{x} = 297{,}6$ jardas, com um desvio padrão amostral de $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = 1{,}7$ jardas.

Decisão:

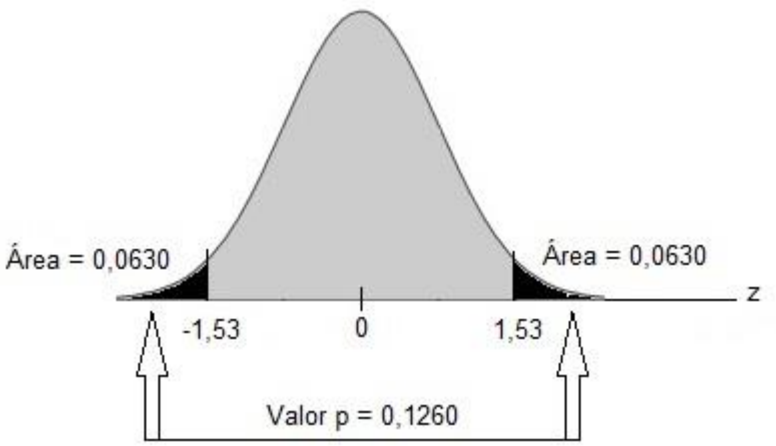
Se a distância média amostral $\bar{x} = 297{,}6$ for significativamente diferente de 295, rejeitar-se-á a hipótese nula. Nesse caso serão tomadas medidas corretivas para ajustar o processo de manufatura. No entanto, se a média amostral $\bar{x} = 297,6$ não se afastar significativamente do valor 295, H_0 não poderá ser rejeitada e nenhuma ação será encaminhada para ajustar o processo de manufatura.

Cálculo da estatística de teste:

Já sabemos como calcular a estatística de teste de uma média populacional normalmente distribuída:

$$z = \frac{\bar{x} - \mu_0}{\sigma\sqrt{n}} = \frac{297,6 - 295}{12\sqrt{50}} = 1,53$$

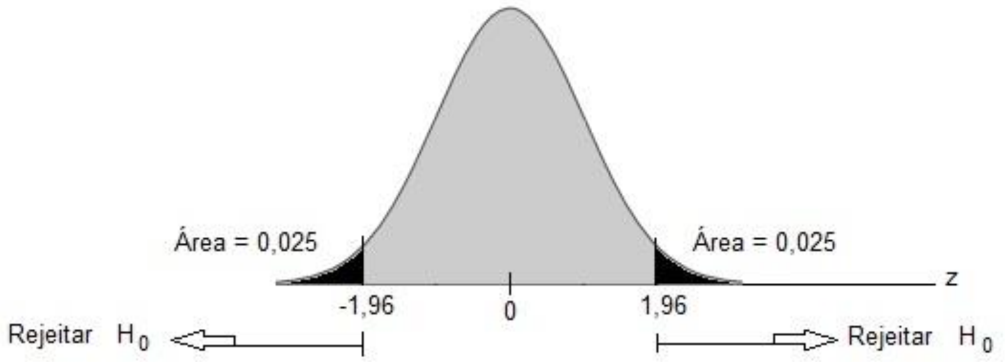
Como z > 0, vamos determinar o ***valor p*** na cauda superior da distribuição normal padrão. Entrando com o valor de z = 1,53 retiramos da tabela normal o valor da área à direita de 1,53 igual a 0,500 – 0,4370 = 0,0630. Sendo o teste bicaudal, esse valor tem que ser considerado em dobro, o que resulta uma área igual a 0,1260. Com isso, o ***valor p = 0,1260***. O nível de significância adotado foi de α = 5%. A figura a seguir ilustra essa situação.



Finalmente resulta para o teste não rejeitar a hipótese nula, pois o ***valor p*** é maior que 0,05.

Testando, agora com o critério do ***valor crítico***:

Como α = 0,05, α/2 = 0,025; procurando na tabela de distribuição normal padrão um valor de z que produza uma área igual ou maior que 0,025 é *z = 1,96*. Assim os valores z = -1,96 e z = 1,96 são os valores críticos (z_c) para o problema em questão. A figura a seguir ilustra essa situação.



O critério de teste usando o valor crítico estabelece que a hipótese nula H_0 deve ser rejeitada se: $z \leq -z_c$ ou se $z \geq z_c$. Como a estatística de teste é z = 1,53 e está contida no intervalo dado pelos valores críticos, a hipótese nula não deverá ser rejeitada.

RESUMO

Resumo dos testes de hipóteses a respeito de uma média populacional: caso em que σ é conhecido

	Teste da cauda superior	Teste da cauda inferior	Teste bicaudal
Hipótese	$H_0: \mu \geq \mu_0$ $H_a: \mu < \mu_0$	$H_0: \mu \leq \mu_0$ $H_a: \mu > \mu_0$	$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$
Estatística de teste	$z = \frac{\bar{x} - \mu_0}{\sigma \sqrt{n}}$	$z = \frac{\bar{x} - \mu_0}{\sigma \sqrt{n}}$	$z = \frac{\bar{x} - \mu_0}{\sigma \sqrt{n}}$
Regra de rejeição: ***Valor p***	Rejeitar H_0 se Valor p $\leq \alpha$	Rejeitar H_0 se Valor p $\leq \alpha$	Rejeitar H_0 se Valor p $\leq \alpha$
Critério de rejeição: ***Valor crítico***	Rejeitar H_0 se $z \leq -z_\alpha$	Rejeitar H_0 se $z \geq z_\alpha$	Rejeitar H_0 se $z \leq z_{\alpha/2}$ ou $z \geq z_{\alpha/2}$

MÉDIA DA POPULAÇÃO: σ desconhecido

O caso em que σ é desconhecido, não se pode estimar o desvio padrão antes de se fazer uma amostragem; a amostra deverá ser usada para estimar μ e σ. Utilizamos, então a média amostral $\bar{x}$ como uma estimativa de μ e usamos o desvio padrão amostral *s* da amostra como uma estimativa de σ.

A exemplo da estimação por intervalo quando não se conhece σ, usamos a distribuição de Student t, aqui nos testes de hipóteses também usaremos a distribuição t. A estatística de teste baseia-se na distribuição t com (n – 1) graus de liberdade. A expressão é:

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

Se a amostra for suficientemente grande já sabemos que podemos supor que a distribuição amostral é normal.

Teste Unicaudal

Para o teste de hipóteses quando σ é desconhecido valem as mesmas etapas já vistas anteriormente. O teste unicaudal pode ser à direita (cauda superior) ou à esquerda (cauda inferior).

Para o teste da cauda inferior: $\begin{array}{l} H_0 : \mu \leq \mu_0 \\ H_a : \mu > \mu_0 \end{array}$

Para o teste da cauda superior: $\begin{array}{l} H_0 : \mu \geq \mu_0 \\ H_a : \mu < \mu_0 \end{array}$

Teste Bicaudal: $\begin{array}{l} H_0 : \mu = \mu_0 \\ H_a : \mu \neq \mu_0 \end{array}$

Resumo

Resumo dos testes de hipóteses a respeito de uma média populacional: caso em que σ é desconhecido

	Teste da cauda superior	Teste da cauda inferior	Teste bicaudal
Hipótese	$H_0 : \mu \geq \mu_0$ $H_a : \mu < \mu_0$	$H_0 : \mu \leq \mu_0$ $H_a : \mu > \mu_0$	$H_0 : \mu = \mu_0$ $H_a : \mu \neq \mu_0$
Estatística de teste	$t = \frac{\bar{x} - \mu_0}{s\sqrt{n}}$	$t = \frac{\bar{x} - \mu_0}{s\sqrt{n}}$	$t = \frac{\bar{x} - \mu_0}{s\sqrt{n}}$
Regra de rejeição: ***Valor p***	Rejeitar H_0 se Valor p ≤ α	Rejeitar H_0 se Valor p ≤ α	Rejeitar H_0 se Valor p ≤ α
Critério de rejeição: ***Valor crítico***	Rejeitar H_0 se $t \leq -t_\alpha$	Rejeitar H_0 se $t \geq t_\alpha$	Rejeitar H_0 se $t \leq -t_{\alpha/2}$ ou $t \geq t_{\alpha/2}$

PROPORÇÃO DA POPULAÇÃO

Usando p0 para denotar o valor hipotético da proporção populacional, as três formas de testes de hipótese a respeito de uma proporção populacional são as seguintes:

$H_0: p \leq p_0$ $H_a: p > p_0$	$H_0: p \geq p_0$ $H_a: p < p_0$	$H_0: p = p_0$ $H_a: p \neq p_0$

A primeira forma é chamada teste da cauda superior, a segunda é denominada teste da cauda inferior e a terceira forma é designada teste bicaudal.

Estatística de teste

Quando a hipótese nula é verdadeira enquanto igualdade, o valor esperado de $\bar{p}$ equivale ao valor hipotético p_0; ou seja, $E(\bar{p}) = p_0$. O desvio padrão de $\bar{p}$ é dado por:

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

Sendo $np \geq 5$ e $n(1-p) \geq 5$, já sabemos que a distribuição amostral de $\bar{p}$ pode ser aproximada pela distribuição normal padrão. Nesse caso a estatística de teste a ser utilizada no teste é:

$$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}}$$

Resumo

Resumo dos testes de hipóteses a respeito de uma proporção populacional

	Teste da cauda superior	Teste da cauda inferior	Teste bicaudal
Hipótese	$H_0: p \geq p_0$ $H_a: p < p_0$	$H_0: p \leq p_0$ $H_a: p > p_0$	$H_0: p = p_0$ $H_a: p \neq p_0$
Estatística de teste	$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}}$	$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}}$	$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}}$
Regra de rejeição: ***Valor p***	Rejeitar H_0 se Valor p $\leq \alpha$	Rejeitar H_0 se Valor p $\leq \alpha$	Rejeitar H_0 se Valor p $\leq \alpha$
Critério de rejeição: ***Valor crítico***	Rejeitar H_0 se $z \leq -z_\alpha$	Rejeitar H_0 se $z \geq z_\alpha$	Rejeitar H_0 se $z \leq z_{\alpha/2}$ ou $z \geq z_{\alpha/2}$

Aplicações:

1) No ano passado a Associação Brasileira de Golfe constatou que apenas cerca de 20% dos jogadores em atividade eram mulheres. Em um esforço para aumentar a proporção de mulheres jogadoras, a Associação implementou uma promoção especial, idealizada para atrair mulheres golfistas. Um mês depois de implementada a ação, o gerente da Associação solicitou um estudo estatístico para determinar se a proporção de mulheres golfistas havia aumentado. O gerente da Associação determinou que um nível de significância de 5% deveria ser usado para a execução do teste.

Hipóteses para o teste: o objetivo do estudo é verificar se houve aumento na proporção de mulheres golfistas com a promoção efetuada. Um teste da cauda superior, com $H_a: p > p_0$ se mostra apropriado. As hipóteses são:

$$H_0 : p \leq p_0$$
$$H_a : p > p_0$$

O valor de p_0 é que será usado para saber se a promoção foi benéfica, ou não. Assim, $p_0 = 0,20$.

$$H_0 : p \leq 0,20$$
$$H_a : p > 0,20$$

Foi selecionada uma amostra aleatória de 400 jogadores de golfe e constatou-se que 100 deles eram mulheres. A proporção amostral de mulheres golfistas é: $\bar{p} = \frac{100}{400} = 0,25$.

Estatística de teste:

Primeiramente calculemos o desvio padrão amostral ou erro padrão:

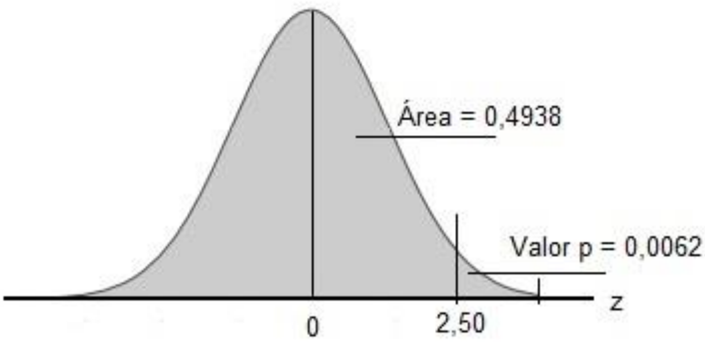
$$\sigma_{\bar{p}} = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0,20(1-0,20)}{400}} = 0,02$$

A estatística de teste será:

$$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{0,25 - 0,20}{0,02} = 2,50$$

Como o teste é da cauda superior, o ***valor p*** é a probabilidade de z ser maior ou igual a $z = 2,50$. Assim, o ***valor p*** = 0,500 – 0,4938 = 0,0062.

A figura a seguir apresenta o cálculo do ***valor p***.



Lembrando que foi especificado um nível de significância de α = 5%. Um ***valor p*** = 0,0062 < 0,05 fornece suficiente evidência estatística para rejeitarmos a hipótese nula ao nível de 5%. Assim, podemos concluir que a promoção especial aumentou o número de jogadoras de golfe.

Usando, agora, o critério do valor crítico; o valor crítico correspondente a área de 0,05 na cauda superior é $z_{0,05} = 1,645$. Desse modo a regra de rejeição usando-se o critério do valor crítico é rejeitar H_0 se $z \geq 1,645$. Como $z = 2,50$ é > 1,645, H_0 é rejeitada.

2) Considere o seguinte teste de hipóteses:

$$H_0 : p \geq 0,75$$
$$H_a : p < 0,75$$

Uma amostra de 300 itens foi selecionada. Calcule o valor p e apresente a sua conclusão a respeito de cada um dos resultados amostrais. Use α = 0,05.

a) $\bar{p}$ = 0,68   b) $\bar{p}$ = 0,72   c) $\bar{p}$ = 0,70   d) $\bar{p}$ = 0,77

Resposta:

1.a) Cálculo do desvio padrão: $\sigma_{\bar{p}} = \sqrt{\frac{p_0(1-p_0)}{n}}$

$= \sqrt{\frac{0,75(1-0,75)}{300}} = 0,025$

Estatística de teste: $z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{0,68 - 0,75}{0,025} = -2,80$

Na tabela normal padrão o valor p = 0,500 – 0,4974 = 0,0026.
Conclusão: Rejeitar H_0, pois valor p < α.

1.b) Cálculo do desvio padrão: o valor não muda; $\sigma_{\bar{p}} = 0,025$

Estatística de teste: $z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{0,72 - 0,75}{0,025} = -1,20$

Na tabela normal padrão o valor p = 0,500 – 0,3849 = 0,1151.

Conclusão: Não rejeitar H_0, pois valor p > α.

1.c) Desvio padrão: $\sigma_{\bar{p}} = 0,025$

Estatística de teste: $z = \dfrac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \dfrac{0,70 - 0,75}{0,025} = -2,0$

Na tabela normal padrão o valor p = 0,500 – 0,4772 = 0,0228.
Conclusão: Rejeitar H_0, pois valor p < α.

1.d) Desvio padrão: $\sigma_{\bar{p}} = 0,025$

Estatística de teste: $z = \dfrac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \dfrac{0,77 - 0,75}{0,025} = 0,8$

Na tabela normal padrão o valor p = 0,500 + 0,2881 = 0,7881.
Conclusão: Não rejeitar H_0, pois valor p > α.

3) Em uma matéria de capa, a *Business Week* (Janeiro/2004) publicou informações a respeito dos hábitos de dormir dos norte-americanos. O artigo afirmou que a privação do sono leva a uma série de problemas e apontou que o deixar de dormir provoca acidentes fatais nas estradas. Cinquenta e um por cento dos motoristas adultos admitem dirigir enquanto estão sonolentos. Um pesquisador aventou a hipótese de que essa questão era um problema ainda maior para as pessoas que trabalham no turno da noite.
a) Formule as hipóteses que podem ser usadas para ajudar a determinar se mais de 51% da população de trabalhadores do turno da noite admitem dirigir enquanto estão sonolentos.
b) Uma amostra de 500 trabalhadores do turno da noite revelou que 232 admitiram dirigir enquanto estavam sonolentos. Qual é a proporção amostral? Qual é o valor p?
c) Com α = 0,01, qual é a sua conclusão?

4) De acordo com o Departamento do Censo dos EUA, a razão principal que levam as pessoas que mudam de residência a escolherem determinada região é o fato de a localização ser conveniente para o trabalho. Com base nos dados do Censo de 1990, 24% das pessoas que mudaram de residência indicaram "localização conveniente para o trabalho" como a razão principal para escolherem a nova região. Suponha que uma amostra de 300 pessoas que se mudaram durante 2003 tenha revelado que 93 o fizeram com o objetivo de morar mais perto do trabalho. Os dados da amostra dão suporte à conclusão de pesquisa segundo a qual em 2003 um número maior de pessoas escolheu onde morar baseando-se em quão perto estarão do trabalho? Qual é a estimação por ponto da proporção de pessoas que se mudaram em 2003 que escolheram a nova região porque a localização é conveniente para o trabalho? Qual é a sua

conclusão de pesquisa? Use $\alpha = 0{,}05$.

5) A CNN foi durante muito tempo a líder de audiência em jornalismo da TV a cabo. Uma pesquisa realizada em março de 2003 indicou que a CNN apresentou uma média de telespectadores de 600 mil pessoas por dia em 2002. Suponha que, para uma amostra de 40 dias, durante o primeiro semestre de 2003, o público médio tenha sido 612 mil telespectadores, com um desvio padrão de 65 mil pessoas.
a) quais são as hipóteses adequadas se a gerência da CNN quisesse obter informações sobre quaisquer alterações no público telespectador da CNN.
b) qual é o ***valor p***?
c) escolha seu próprio nível de significância. Qual é a sua conclusão?
d) qual recomendação você faria à gerência da CNN nessa aplicação?

6) La Belle Garden é uma empresa especialista em paisagismo personalizado para áreas residenciais. O custo de mão-de-obra estimado associado a uma proposta de paisagismo em particular baseia-se no número de plantações de árvores, arbustos, etc. Pata fins de estimação do custo, os gerentes utilizam duas horas de mão-de-obra para o plantio de uma árvore de tamanho médio. Os tempos reais de uma amostra de dez plantações durante o mês passado são apresentados a seguir (expresso em horas).

1,7	1,5	2,6	2,2	2,4	2,3	2,6	3,0	1,4	2,3

Com um nível de significância de 0,05 (5%), teste se a média de tempo de plantio das árvores difere de duas horas.
a) estabeleça as hipóteses nula e alternativa.
b) calcule a média da amostra.
c) calcule o desvio padrão da amostra.
d) qual é o ***valor p***?
e) qual é a sua conclusão?

7) Os funcionários de escritório da Shell Oil foram solicitados a responder qual programação de trabalho seria a mais atraente: trabalhar cinco dias de oito horas por semana ou trabalhar quatro dias de dez horas por semana. Admitamos que *p = a proporção da população de funcionários de escritório que preferem trabalhar quatro dias de dez horas por semana*.
a) estabeleça as hipóteses para o caso de a gerência da Shell estar interessada em obter evidências estatísticas que mostrem que mais de 50% dos funcionários de escritório preferem trabalhar quatro dias de dez horas por semana.
b) qual é a proporção amostral de uma amostra de 105 funcionários

de escritório tiver revelado que 67 prefeririam a programação de quatro dias de dez horas?
c) qual o valor p? Use $\alpha = 0{,}01$. Qual é a sua conclusão?

8) Uma estação de rádio de Balneário Beach anunciou que pelo menos 90% dos hotéis e motéis estariam lotados no fim de semana da Páscoa. A estação aconselhou os ouvintes a fazerem reservas antecipadamente, caso planejassem passar o fim de semana no balneário. No sábado à noite, uma amostra de 58 hotéis e motéis revelou que 49 exibiam o anúncio "sem vagas" e 9 "com vagas". Qual é a sua reação à afirmação da estação de rádio depois de ver a evidência da amostra? Use $\alpha = 0{,}05$ ao realizar o teste estatístico. Qual é o ***valor p***?

9) Em 2000, o rendimento médio por ação da população de corporações de serviços financeiros foi de US$ 3. Em 2001, uma amostra de 10 corporações de serviços financeiros forneceu os seguintes dados de rendimento por ação:

1,92	2,16	3,63	3,16	4,02	3,14	2,20	2,34	3,05	2,38

a) formule as hipóteses nula e alternativa que podem ser usadas para determinar se o rendimento médio por ação da população em 2001 difere dos US$ 3 registrados em 2000.
b) calcule a média amostral.
c) calcule o desvio padrão da amostra.
d) qual é o valor p?
e) Use $\alpha = 0{,}05$. Qual a sua conclusão?

10) Uma promoção de uma empresa aérea para pessoas que fazem viagens de negócios baseia-se na suposição de que dois terços dessas pessoas usam *laptops* em viagens de negócios noturnas.
a) estabeleça as hipóteses que podem ser usadas para testar a suposição.
b) qual é a proporção amostral de uma pesquisa patrocinada pela American Express que revelou que 355 de 546 pessoas que fazem viagens de negócios usam um laptop em viagens de negócio noturnas?
c) qual é o valor p?
d) Use $\alpha = 0{,}05$. Qual é a sua conclusão?