UNIDADE 5

Introdução aos testes de hipóteses

Um **teste de hipóteses** é um procedimento que usa estatística amostral para testar uma alegação sobre o valor de um parâmetro populacional. Pesquisadores das mais diversas áreas contam com os testes de hipóteses para a tomada de decisões sobre novos medicamentos ou resultados de eleições, por exemplo. Observe a seguir exemplos de afirmações que poderiam ser feitas acerca da população:

- ✓ O tempo médio para a realização de um teste é de 80 minutos.
- ✓ Três por cento de toda a produção é defeituosa.
- ✓ A percentagem de desempregados em duas cidades vizinhas é igual.
- ✓ A quilometragem por litro de combustível é a mesma para três marcas de gasolina.

A finalidade de testes de hipóteses é avaliar afirmações sobre os valores de parâmetros populacionais.

Suponha, por exemplo, que um fabricante alegue que a vida média das pilhas AA é de 300 minutos. Se você suspeitasse que essa alegação não é válida, como poderia mostrar que ela é falsa? Obviamente, você não pode testar todas as pilhas. Todavia, você ainda pode tomar uma decisão razoável acerca da validade da alegação extraindo uma amostra ao acaso da população. Se a média amostral for diferente o bastante da alegação, pode-se decidir que ela é falsa.

Considere ainda uma segunda situação: Inspeciona-se uma amostra de 142 peças de uma grande remessa, encontrando-se 8% de defeituosas. O fornecedor garante que não haverá mais de 6% de peças defeituosas em cada remessa. O que devemos responder? A afirmação do fornecedor é verdadeira? Ou a diferença entre a afirmação do fornecedor e o resultado da amostra é explicada pela variação amostral?

Vamos entender os fundamentos de um teste de hipóteses para que consigamos responder perguntas como as que foram feitas acima.

5.1 Hipóteses estatísticas

Uma alegação sobre um parâmetro populacional é chamada de **hipótese estatística**. Para testar uma hipótese estatística, você deve estabelecer cuidadosamente um par de hipóteses – uma representa uma alegação e a outra, seu complemento. Quando uma dessas hipóteses é rejeitada, a outra deve ser aceita. Dessas duas hipóteses, aquela que contém uma afirmativa de igualdade é a hipótese nula. O complemento de uma hipótese nula é a hipótese alternativa.

- ✓ **Hipótese nula** (denotada por H_0) é uma hipótese estatística que contém uma afirmativa de igualdade e deve escrever-se como =, $\leq$ ou $\geq$. Para a média, temos as três formas possíveis para a hipótese nula:

$$\left.\begin{array}{l} H_0: \mu = \mu_0 \\ H_0: \mu \geq \mu_0 \\ H_0: \mu \leq \mu_0 \end{array}\right\}$$

Onde, μ_0 é algum valor que você deseja testar.

- ✓ **Hipótese alternativa** (denotada por H_a) é o complemento da hipótese nula. É uma afirmativa que deve ser verdadeira se H_0 for falsa é contém uma afirmativa de desigualdade, tal como <, $\neq$ ou >. Para a média, a hipótese alternativa comporta apenas uma das três formas:

$$H_a: \mu \neq \mu_0$$
$$H_a: \mu < \mu_0$$
$$H_a: \mu > \mu_0$$

O quadro a seguir mostra a relação entre as afirmativas verbais possíveis sobre o parâmetro μ e as hipóteses nula e alternativa correspondentes. Afirmativas semelhantes podem ser feitas para outros parâmetros populacionais, tais como *p* (proporção), σ (desvio padrão) ou σ^2 (variância).

Formulação verbal da H_0	Formulação matemática	Formulação verbal da H_a
A média é igual a algum valor (μ_0).	$H_0: \mu = \mu_0$ $H_a: \mu \neq \mu_0$	A média é diferente de algum valor (μ_0).
A média é maior ou igual a algum valor (μ_0)	$H_0: \mu \geq \mu_0$ $H_a: \mu < \mu_0$	A média é menor do que algum valor (μ_0)
A média é menor ou igual a algum valor (μ_0)	$H_0: \mu \leq \mu_0$ $H_a: \mu > \mu_0$	A média é superior a algum valor (μ_0)

Exercício – Estabelecendo as hipóteses nula e alternativa

Estabeleça as hipóteses nula e alternativa e identifique qual representa a alegação.

a) Uma universidade alega que a proporção de seus alunos formados em quatro anos é de 82%.
b) Um fabricante de torneiras alega que a taxa de fluxo médio de um determinado tipo é inferior a 2,5 galões por minuto.
c) Um fabricante de pneus alega que a variância no diâmetro de um determinado modelo é de 8,6 polegadas.
d) Uma companhia que fabrica cereais alega que o peso médio do conteúdo de suas caixas de cereais é superior a 200 gramas.
e) Uma estação de rádio alega que sua proporção de audiência local é maior do que 39%.
f) Um restaurante alega que o desvio padrão para a duração do intervalo entre um atendimento e outro é inferior a 2,9 minutos.

5.2 Tipos de erros e nível de significância

Não importando qual das hipóteses representa a alegação, você começará sempre um teste de hipóteses assumindo que a condição de igualdade na hipótese nula é verdadeira. Esse mesmo raciocínio é utilizado em processos criminais onde "Presume que o réu seja inocente até que sua culpa seja estabelecida sem sombra de dúvida", a suposição de inocência é uma hipótese nula. Assim, quando realizar um teste de hipóteses, você deve tomar uma de duas decisões:

1) Rejeitar a hipótese nula
2) Aceitar a hipótese nula

Uma vez que sua decisão baseia-se em informação incompleta (uma amostra em vez de toda a população), há sempre a possibilidade de se tomar a decisão errada. O quadro a seguir mostra os quatro resultados possíveis de um teste de hipóteses:

Decisão	A verdade real de H_0	
	H_0 é verdadeira	H_0 é falsa
Aceitar H_0	Decisão correta $(1 - \alpha)$	Erro do tipo II β
Rejeitar H_0	Erro do tipo I α	Decisão correta $(1 - \beta)$

- ✓ **Erro tipo I** (denotada por α): Ocorre se a hipótese nula for rejeitada quando ela for realmente verdadeira.
- ✓ **Erro tipo II** (denotada por β): Ocorre se a hipótese nula não for rejeitada quando ela for realmente falsa.

Uma das etapas de nosso processo de teste de hipóteses envolve a escolha do nível de significância (α), que é a probabilidade de um erro tipo I. Entretanto, não selecionamos β (a probabilidade de erro tipo II). Seria ótimo se pudéssemos ter sempre α=0 e β=0, mas isto não é possível; devemos, pois, procurar controlar as probabilidades de erro α e β. Pode-se mostrar, matematicamente, que α, β e o tamanho (n) da amostra estão todos inter-relacionados, de forma que, escolhidos quaisquer dois deles, o terceiro está automaticamente determinado. Valem as seguintes considerações de ordem prática:

1. Para α fixo, um aumento do tamanho *n* da amostra ocasiona uma redução de β; isto é, uma amostra maior reduz a chance de cometermos o erro de aceitar a hipótese nula quando ela é falsa.
2. Para um tamanho *n*, fixo, de amostra, uma diminuição de α acarreta um aumento de β; reciprocamente, um aumento de α acarreta uma diminuição de β.
3. Para reduzir α e β, devemos aumentar o tamanho da amostra.

O *nível de significância* de um teste é a probabilidade de uma hipótese nula ser rejeitada, quando verdadeira.

5.3 Estatística de teste, região crítica e valor crítico

- ✓ A **estatística de teste** é uma estatística amostral, ou um valor baseado nos dados amostrais. Utiliza-se uma estatística de teste para tomar uma decisão sobre a rejeição ou não da hipótese nula.
- ✓ A **região crítica** é o conjunto de todos os valores da estatística de teste que levam à rejeição da hipótese nula.
- ✓ O **valor crítico** é o valor, ou valores, que separa(m) a região crítica dos valores da estatística de teste que não levam à rejeição da hipótese nula. Os valores críticos dependem da natureza da hipótese nula, da distribuição amostral principal, e do nível de significância (α).

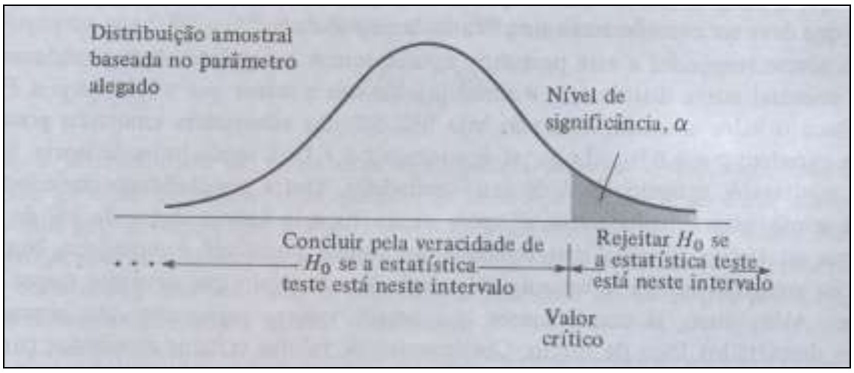


Figura 1 – A distribuição amostral é particionada em regiões de aceitação e de rejeição, com o valor crítico como ponto divisório

5.4 Teste unilaterais e testes bilaterais

Nosso interesse em detectar desvios significativos de determinado parâmetro pode envolver desvios em ambas as direções ou apenas numa direção. Observe as situações a seguir que envolvem sucessivas jogadas de uma moeda:

1. A moeda pode ser considerada não equilibrada se aparece um número muito grande, ou muito pequeno, de caras. A hipótese alternativa seria simplesmente “a moeda não é equilibrada”, ou seja, H_a: $p \neq 0{,}5$, e investigaríamos então desvios em ambas a direções.

2. Se estivéssemos apostando em caras, então nossa preocupação seria somente com um número muito pequeno de caras. A hipótese alternativa seria "aparecem muito poucas caras", isto é, H_a: $p < 0,5$, e só estaríamos interessados então nesse tipo de desvio do número esperado de caras.

Essencialmente, a hipótese alternativa é usada para indicar qual aspecto nos interessa. Há três casos possíveis: concentrar am ambas as direções, concentrar nos desvios abaixo do valor esperado ou concentrar nos desvios acima do valor esperado. Simbolicamente, no caso da jogada de uma moeda, esses três casos podem ser representados como a seguir:

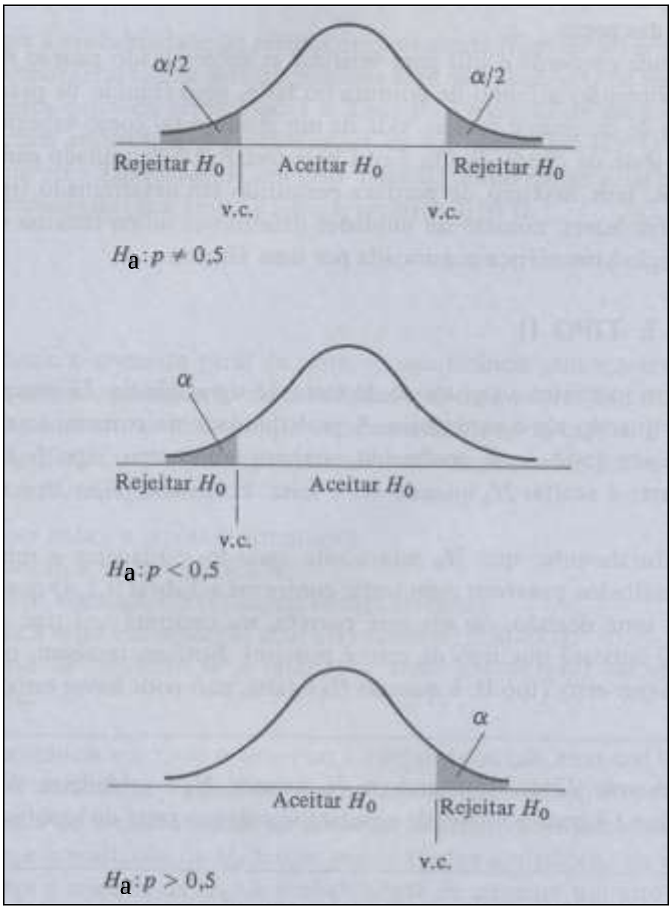
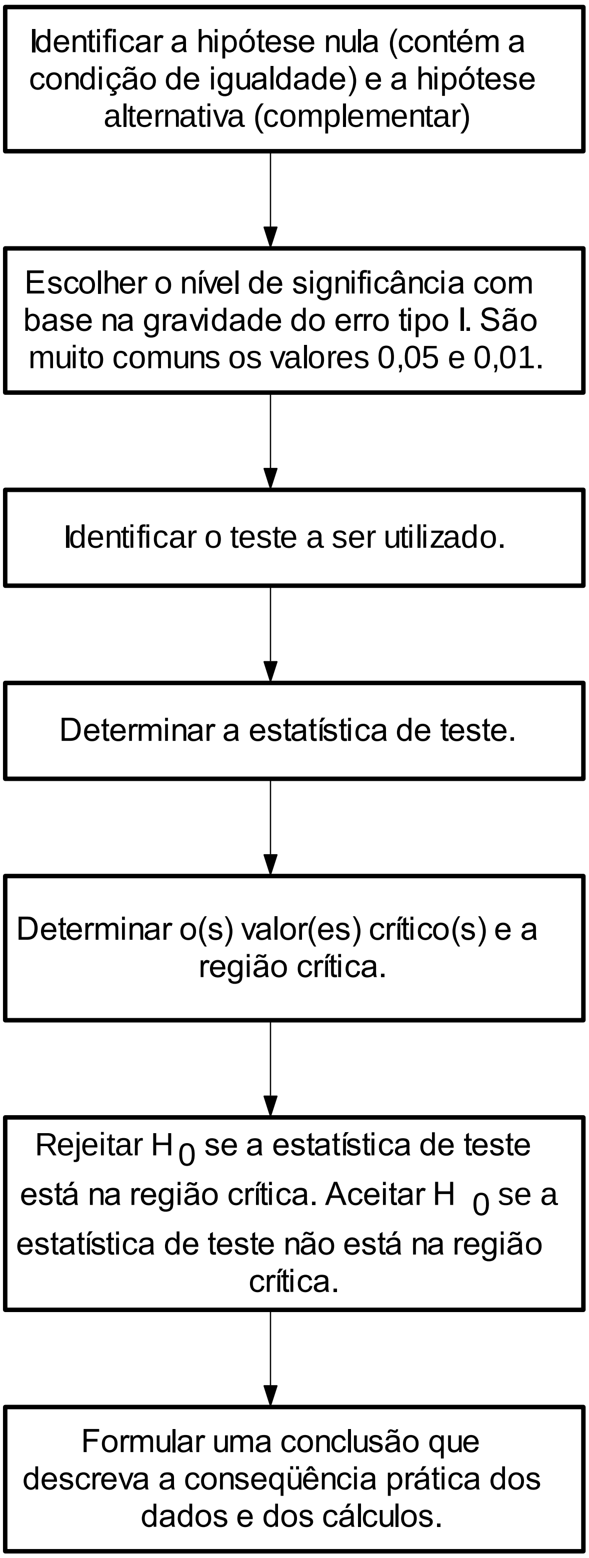


Figura 2 – Comparação da partição de uma distribuição amostral para testes unilaterais e bilaterais. Note-se, nos testes unilaterais, que o sinal > ou o sinal < aponta para a cauda utilizada.

5.5 O método clássico de testes de hipóteses



Exercícios de fixação

1. Qual é o objetivo de se realizar testes de hipóteses?

2. Explique comparativamente:
    a. Erro tipo I e erro tipo II.
    b. Hipótese nula e hipótese alternativa.
    c. Nível de significância de 0,05 e nível de significância de 0,01.
    d. Teste unilateral e teste bilateral.

3. Suponha que o leitor dispõe da seguinte informação:

    $$H_0: p = 35\% \qquad H_a: p \neq 35\%$$

    a. Explique por que a probabilidade de erro tipo II é zero se a proporção populacional é 35%.
    b. Explique por que a probabilidade de erro tipo I é zero se a proporção populacional é diferente de 35%.

4. Para cada sentença a seguir:
    ✓ Escreva as hipóteses nula e alternativa;
    ✓ Identifique se o teste é bilateral, unilateral direito ou unilateral esquerdo;
    ✓ Escreva sentenças descrevendo o erro tipo I e erro tipo II.

    a. Um vendedor de carros alega que pelo menos 24% de seus novos compradores voltarão para comprar seu próximo carro.
    b. Um clube local de xadrez alega que o período de tempo de uma partida tem desvio padrão superior a 23 minutos.
    c. Um instituto de pesquisa alega que a duração média da maioria dos sonhos é superior a 10 minutos.

5.6 Teste de hipóteses para uma média quando σ é conhecido (Teste z de uma amostra)

O objetivo dos testes de hipóteses para uma média é avaliar afirmações feitas a respeito da média populacional. Os testes de hipóteses para média exigem dados quantitativos, isto é, dados contínuos ou discretos.

O teste de hipóteses descrito nesta seção será utilizado para uma média quando o desvio padrão populacional (σ) for conhecido.

Identificadas as hipóteses estatísticas (H_0 e H_a) e definido o nível de significância (α) podemos proceder ao cálculo da estatística de teste utilizando a fórmula a seguir:

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

onde μ_0 representa o valor da média que está sendo testado, σ representa o desvio padrão populacional, n é o tamanho da amostra e $\bar{x}$ é a média da amostra.

A região crítica para o teste z é construída levando-se em consideração a distribuição z, o nível de significância adotado no teste e a hipótese alternativa. Observe as figuras a seguir:

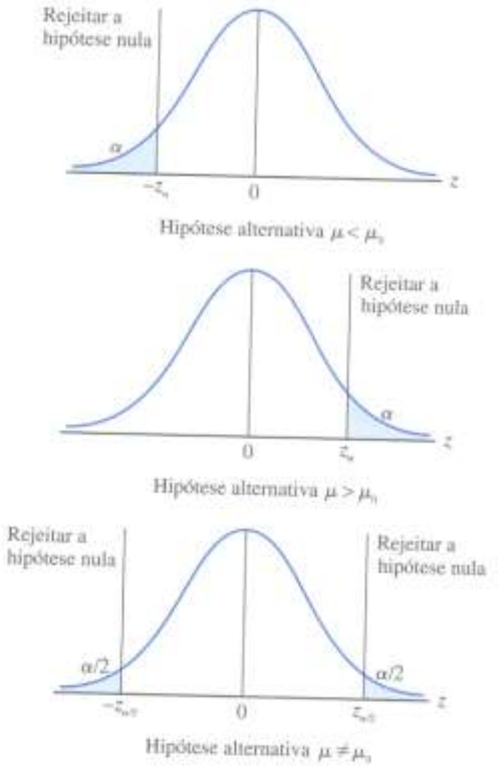


Figura 3 – Região crítica para o teste z para a média populacional.

Dessa forma, para o teste z para uma média, podemos estabelecer o seguinte critério de decisão:

Hipótese alternativa (H_a)	Rejeitar a hipótese nula se:	Não rejeitar a hipótese nula se:
H_a: $\mu < \mu_0$	$z_{teste} \leq -z_\alpha$	$z_{teste} > -z_\alpha$
H_a: $\mu > \mu_0$	$z_{teste} \geq z_\alpha$	$z_{teste} < z_\alpha$
H_a: $\mu \neq \mu_0$	$z_{teste} \leq -z_{\alpha/2}$ ou $z_{teste} \geq z_{\alpha/2}$	$-z_{\alpha/2} < z_{teste} < z_{\alpha/2}$

Vale lembrar que os valores z_α e $z_{\alpha/2}$ são obtidos diretamente da tabela da distribuição normal padrão.

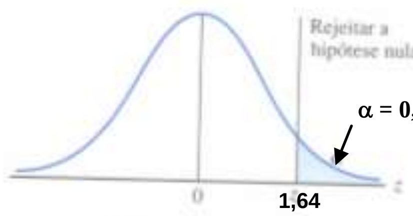
Observe o exemplo a seguir:
(Adaptado de Levine et al.) Uma das principais medidas de qualidade dos serviços oferecidos por qualquer organização corresponde à velocidade através da qual ela responde a uma reclamação feita pelos clientes. Uma grande loja de departamentos, de controle familiar passou por uma grande expansão nos últimos anos. Foi selecionada uma amostra de 50 reclamações com relação à instalação de pisos deste ano, ou seja, correspondem ao tempo de espera (em dias) entre o recebimento da

reclamação e a solução do problema relacionado com a reclamação. Nesta amostra o tempo médio de espera foi de 29 dias. Sabe-se que o desvio padrão da população é de 21 dias. O gerente afirma que o tempo médio de espera é de no máximo 20 dias. Com um nível de 5% de significância há evidências que comprovem a alegação do gerente?

Como o desvio padrão populacional é conhecido (σ=21) podemos utilizar o teste z para uma média. A partir do enunciado podemos identificar as hipóteses nula e alternativa da seguinte maneira:

H_0: $\mu \leq 20$
H_a: $\mu > 20$

Como o nível de significância (α), pré-determinado, foi de 0,05 (5%) e o teste é unilateral direito devemos localizar na tabela da distribuição normal padrão o valor de z_α, ou seja, $z_{0,05}$=1,64. Dessa forma podemos representar a região crítica do teste da seguinte maneira.



Calculando a estatística de teste, obtemos:

$$z_{teste} = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{29 - 20}{\frac{21}{\sqrt{50}}} \approx 3{,}03$$

Como a estatística de teste z_{teste}=3,03 é maior do que o valor da tabela $z_{0,05}$=1,64 deve-se rejeitar a hipótese nula, ou seja, há evidência suficiente para afirmar que o tempo médio de espera (em dias) entre o recebimento da reclamação e a solução do problema relacionado com a reclamação é superior a 20 dias, contradizendo a afirmação do gerente.

Exercício em sala

1. (Adaptado de Stevenson, 2001) Suponha que queiramos avaliar a afirmação de um fabricante, de que seus pneus radiais suportam uma quilometragem de 40000 milhas, no mínimo. Para uma amostra aleatória de 49 pneus observou-se uma média de duração de 38000 milhas. Sabe-se que o desvio padrão populacional da duração dos pneus é de 3500 milhas. Utilize um nível de significância de 0,03. O que você pode concluir?

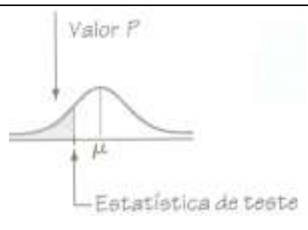
O método do valor p para o teste de hipóteses

Muitos artigos profissionais e programas de computador utilizam outra abordagem do teste de hipóteses, baseada no cálculo do valor de uma probabilidade, ou valor p. Dada uma hipótese nula e um conjunto de dados amostrais, o valor p reflete a probabilidade de se obter tais resultados no caso da hipótese nula ser, de fato verdadeira. Um valor p muito pequeno sugere que os resultados amostrais são muito improváveis sob a hipótese nula, ou seja, constitui evidência contra a hipótese nula.

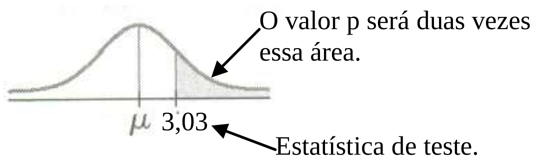
O critério de decisão baseado no valor p é feito da seguinte maneira:

- ✓ <u>Rejeitar</u> a hipótese nula (H_0) se o valor p é no máximo igual ao nível de significância (α).
- ✓ <u>Não rejeitar</u> a hipótese nula (H_0) se o valor p é maior do que o nível de significância (α).

O esquema a seguir mostra a determinação dos valores p.

Tipo de teste	Valor p	Representação gráfica
Unilateral direito	Área à direita da estatística de teste	
Bilateral	2 x a área à direita do módulo da estatística de teste.	 **Observação:** O valor da estatística de teste deverá ser utilizado em módulo para esse cálculo!
Unilateral esquerdo	Área à esquerda da estatística de teste	

Considere o exemplo citado anteriormente (Adaptado de Levine et al.). Para chegar à conclusão utilizou-se a região crítica, porém, uma outra maneira de concluir o problema seria por meio do cálculo do valor p. O teste realizado é do tipo unilateral direito e a estatística de teste encontrada foi de $z_{teste} = 3{,}03$. Dessa maneira, para calcular o valor p é necessário encontrar o valor da área à direita da estatística de teste. Observe a figura:



Utilizando a tabela da distribuição normal padrão temos que a área à direita de 3,03 é (0,5 - 0,4988) = 0,0012 e, conseqüentemente esse será o valor p. Interpretando o valor p: Se rejeitarmos H_0, ou seja, se afirmarmos que o tempo médio de espera (em dias) entre o recebimento da reclamação e a solução do problema relacionado com a reclamação é de no máximo 20 dias estamos cometendo um erro de 0,0012. Como o valor p é menor do que o nível de significância (0,0012 < 0,05) devemos rejeitar a hipótese nula, ou seja, a mesma conclusão obtida por meio da região crítica.

Exercício em sala

1. Para o exercício anterior (Adaptado de Stevenson, 2001) que mostra o caso da duração de pneus radiais, obtenha a conclusão utilizando o valor p.

2. Um representante de um grupo comunitário informa a um potencial construtor de um shopping, que a renda familiar média mensal nesta área é superior a R$6500. Suponha que, para o tipo de área envolvida, a renda familiar possa ser assumida como seguindo a distribuição normal, e que o desvio padrão pode ser aceito como sendo igual a σ = 2000, baseado em um estudo anterior. Para uma amostra aleatória de 55 residências, a renda familiar média encontrada é de R$6300. Verifique se há evidências que confirmem a hipótese do representante do grupo comunitário, considere um nível de 2% de significância. Utilize a região crítica e o valor p para concluir o teste.

Teste de hipóteses (z) com intervalos de confiança

Já vimos anteriormente o método de estimação intervalar que consiste em encontrar uma estimativa intervalar de um parâmetro populacional com contém os valores prováveis daquele parâmetro. Por conseguinte, devemos rejeitar uma afirmação de que o parâmetro populacional tenha um valor que não está compreendido no intervalo de confiança.

É importante lembrar que, como vimos o método de construção de intervalos de confiança bilaterais para um parâmetro, essa correspondência direta entre um intervalo de confiança e um teste de hipótese deve ser feita quando o teste é bilateral.

Exercício em sala

1. Um estudante de direito quer conferir a alegação de sua professora de que fraudadores condenados passam, em média 12,3 meses na cadeia. Uma amostra aleatória de 35 casos do arquivo judiciário indicou uma média de 11,5 meses. Suponha que o desvio padrão seja σ = 3,8 meses. Adote um nível de 4% de significância. O que o estudante pode concluir a partir dos dados da amostra? Utilize a região crítica, o valor p e o intervalo de confiança.

5.7 Teste de hipóteses para uma média quando σ é desconhecido (Teste t-Student de uma amostra)

O teste t-Student, assim como o teste z (visto anteriormente), é um teste de hipóteses para a média. Ele será utilizado quando σ é desconhecido e a população original tem distribuição essencialmente normal.

Quando o desvio padrão da população não é conhecido (o que é o caso no teste t-Student), usa-se o desvio padrão da amostra como estimativa, substituindo σ por *s* nas fórmulas. Isto não acarreta maiores dificuldades, pois o desvio padrão amostral dá uma aproximação bastante razoável do verdadeiro valor, na maioria dos casos. Além disso, pelo Teorema do Limite Central, sabemos que, quando o tamanho da amostra é superior ou igual a 30, a distribuição das médias é aproximadamente normal. Todavia, para amostras de menos de 30 observações, a aproximação normal não é adequada. Devemos então usar a distribuição t-Student, que é a distribuição correta quando se usa *s*.

Após a identificação do teste a ser utilizado, procedemos com a definição das hipóteses estatísticas (H_0 e H_a) e do nível de significância (α). A partir daí encontramos a estatística de teste:

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

onde μ_0 representa o valor da média que está sendo testado, s representa o desvio padrão da amostra (já que o desvio padrão da população é desconhecido), *n* é o tamanho da amostra e $\bar{x}$ é a média da amostra.

A região crítica para o teste t-Student é definida da mesma forma que no teste z, ou seja, levando-se em consideração a distribuição t-Student, o nível de significância adotado no teste e a hipótese alternativa.

O valor crítico é obtido na tabela t-Student onde:
- ✓ Graus de liberdade = n – 1.

Após determinar o número de graus de liberdade (na linha da tabela), localizar, na coluna, o valor correspondente à área da cauda que se deseja obter, o valor crítico será o valor que corresponde ao cruzamento da linha e da coluna que foram determinados. Lembrando que se um valor crítico está localizado na cauda esquerda, devemos considerá-lo negativo.

Observe o exemplo a seguir:

(Adaptado de Morettin & Bussab, 2007) A associação dos proprietários de indústrias metalúrgicas está muito preocupada com o tempo perdido com acidentes de trabalho, cuja média, nos últimos tempos tem sido da ordem de 60 horas/homem por ano. Tentou-se um programa de prevenção de acidentes, após o qual foi tomada uma amostra de nove indústrias e medido o número de horas/homens perdidas por acidente obtendo-se uma média de 50 horas e um desvio padrão de 20 horas. Você diria, considerando-se um nível de significância de 5%, que houve alguma alteração no tempo perdido com acidentes de trabalho? Suponha que a população segue aproximadamente a distribuição normal.

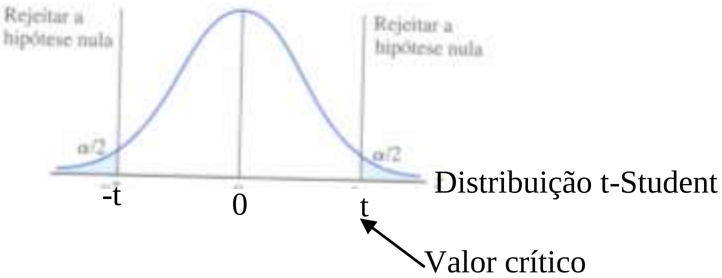
Neste exemplo não há qualquer informação sobre o desvio padrão populacional e, além disso, supõe-se que a população tem distribuição aproximadamente normal. Portanto, temos que prosseguir com o teste utilizando a distribuição *t-Student*:

Hipóteses: H_0: $\mu = 60$ versus. H_a: $\mu \neq 60$
Nível de significância: $\alpha = 0{,}05$.

Estatística de teste: $t_{teste} = \dfrac{50-60}{\dfrac{20}{\sqrt{9}}} = -1{,}5$

Graus de liberdade: n – 1 = 9 –1 = 8.

Região crítica:



Valor crítico (obtido na tabela da distribuição t-Student):

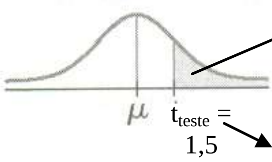
Observe que estamos na coluna que representa o valor de $\alpha/2$, já que o teste é bilateral.

Graus de liberdade	Na cauda superior área α							
	.1	.05	.025	.01	.005	.0025	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781

Valor crítico: t = 2,306

Podemos observar que a estatística de teste (t_{teste} = -1,5) está dentro do intervalo de -2,306 a 2,306, ou seja, concluímos por não rejeitar H_0. Dessa forma, podemos dizer que não há evidências de que tenha havido alteração no tempo médio perdido com acidentes de trabalho em indústrias metalúrgicas.

Se optássemos por resolver esse exemplo usando o valor p teríamos que utilizar as mesmas idéias do teste z (que já foram demonstradas), ou seja, como estamos solucionando um teste bilateral, o valor p é dado por:



O valor p será duas vezes o valor dessa área

Observe que o valor da estatística de teste está em módulo, ou seja, para encontrar o valor p em um teste bilateral deve-se utilizar o valor da estatística de teste sempre positivo.

Dessa forma temos que localizar na tabela t-Student o $\text{Valor p} = 2 \times P(t > \text{Estatística de teste})$, ou seja, na linha correspondente ao número de graus de liberdade (n – 1 = 8) devemos localizar o valor mais próximo da estatística de teste (t_{teste}=1,5, considerando-se o valor positivo), observe a seguir:

Graus de liberdade	Na cauda superior área α							
	.1	.05	.025	.01	.005	.0025	.001	.0005
1	3.078	6.314	12.706	31.821	63.657	127.32	318.31	636.62
2	1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.598
3	1.638	2.353	3.182	4.541	5.841	7.453	10.214	12.924
4	1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5	1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6	1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7	1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8	1.397	1.860	2.306	2.896	3.355	3.833	4.501	5.041
9	1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781

Valor mais próximo da estatística de teste (t_{teste} =1,255, considerando-se o valor positivo)

O valor 0,1 encontrado corresponde à área acima da estatística de teste positiva, ou seja, dessa forma o valor p = 2 x 0,1 = 0,2, ou seja, se rejeitarmos a hipótese nula estamos cometendo um erro de 0,2 = 20%, como o erro máximo permitido é de α = 0,05 = 5%, devemos portanto aceitar a hipótese nula. Concluímos que o tempo médio perdido com acidentes de trabalho é igual a 60 horas/homem, não havendo indícios suficientes para concluirmos que houve alteração.

Teste de hipóteses (t-Student) com intervalos de confiança

Já vimos anteriormente o método de estimação intervalar que consiste em encontrar uma estimativa intervalar de um parâmetro populacional com contém os valore prováveis daquele parâmetro. Por conseguinte, devemos rejeitar uma afirmação de que o parâmetro populacional tenha um valor que não está compreendido no intervalo de confiança. Vimos também que, no caso da distribuição t-Student, o desvio padrão populacional não é conhecido e em seu lugar utilizamos a estimativa por meio do desvio padrão da amostra. Dessa forma o intervalo de confiança para a média quando σ é desconhecido é obtido pela fórmula:

$$\bar{x} \pm (t) \cdot \frac{s}{\sqrt{n}}$$

Lembrando que, *s* representa o desvio padrão da amostra (já que o desvio padrão da população é desconhecido), *n* é o tamanho da amostra, $\bar{x}$ é a média da amostra e *t* é o valor crítico obtido na tabela da distribuição t-Student, levando em consideração o número de graus de liberdade é a área correspondente a $\frac{\alpha}{2}$ (já que o teste é bilateral).

Vale lembrar mais uma vez que, como vimos o método de construção de intervalos de confiança bilaterais para um parâmetro, essa correspondência direta entre um intervalo de confiança e um teste de hipótese deve ser feita quando o teste é bilateral.

Voltando ao exemplo do tempo perdido com acidentes de trabalho, como temos um teste bilateral, podemos optar por obter a conclusão com base apenas no intervalo de 95% de confiança (já que o nível de significância adotado é de 5%):

$$\bar{x} \pm (t) \cdot \frac{s}{\sqrt{n}} = 50 \pm (2{,}306) \cdot \frac{20}{\sqrt{9}} = 50 \pm 15{,}373 = [34{,}627 \quad ; \quad 65{,}373]$$

Com base no resultado do intervalo de confiança podemos observar que com 95% de confiança o tempo médio perdido com acidentes de trabalho em indústrias metalúrgicas pode variar de 34,627 horas/homem a 65,373 horas/homem. Dessa forma, como a média testada (μ_0 = 60) encontra-se dentro do intervalo de confiança, devemos aceitar a hipótese nula, concluindo que não há indícios de que tenha havido alteração.

Propriedades importantes da distribuição t-Student

1. A distribuição t-Student é diferente para cada tamanho de amostra:

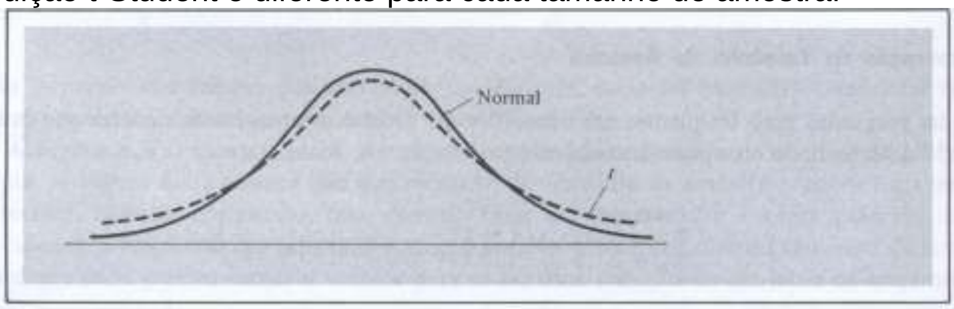


Figura 8.6 Comparação das distribuições normal e *t*. Note-se que a distribuição tem mais área nas caudas.

2. A distribuição t-Student tem a mesma forma geral de sino da distribuição normal padronizada; sua forma mais aberta reflete a maior variabilidade esperada em pequenas amostras.

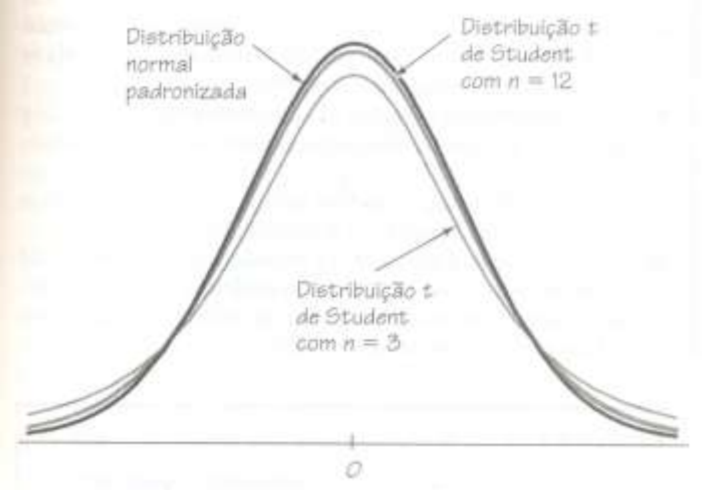


Fig. 6-5 Distribuições *t* de Student para $n = 3$ e $n = 12$. A distribuição *t* de Student tem a mesma forma geral e simetria que a distribuição normal padronizada, mas reflete a maior variabilidade esperada com pequenas amostras.

3. A distribuição t-Student tem média $t = 0$ (tal como a distribuição normal padronizada, que tem média $z = 0$).
4. O desvio padrão da distribuição t-Student varia com o tamanho da da amostra, e é maior do que 1 (ao contrário da distribuição normal padronizada, em que $\sigma = 1$).
5. À medida que o tamanho da amostra aumenta, a distribuição t-Student se aproxima da distribuição normal. Para valores de $n \geq 30$, as diferenças são tão pequenas que podemos usar os valores críticos *z* em lugar de *t*.

Exercícios em sala

Para os exercícios abaixo utilize a região crítica, o valor p e o intervalo de confiança (quando apropriado):

1. (Adaptado de Morettin & Bussab) Um escritório de investimentos acredita que o rendimento médio das diversas ações que movimenta é de 20%. Uma nova estratégia foi definida com o objetivo de aumentar o rendimento médio das diversas ações que movimenta. Para verificar esta hipótese, tomaram-se 18 empresas ao acaso obtendo-se um rendimento médio de 24,56% e um desvio padrão de 5%. Suponha que a população tenha distribuição aproximadamente normal. Faça o teste estatístico para comprovar que a nova estratégia é realmente eficiente para aumentar o rendimento médio das diversas ações. Use um nível de significância de 1%.

2. Uma amostra aleatória de oito pedidos dos arquivos de uma companhia mostra que os pedidos de certa peça de uma máquina forma despachados em: 12 10 17 14 13 18 11 9 dias. Ao nível de 5% de significância , podemos concluir que, em média, tais pedidos são despachados me menos de 10 dias?

5.8 Teste de hipóteses para uma proporção

Os testes de hipóteses para proporções são adequados quando os dados sob análise consistem de contagens ou freqüências de itens. A finalidade de tais testes é avaliar afirmações sobre a proporção (ou percentagem) de uma população.

As suposições associadas ao teste de hipóteses para uma proporção são descritas a seguir:

- ✓ Devem ser verificadas as condições para um experimento binomial. Isto é, temos um número fixo de provas independentes com probabilidade constante, e cada prova comporta dois resultados, que são designados como "sucesso" e "fracasso".
- ✓ As condições $n \cdot p \geq 5$ e $n \cdot (1-p) \geq 5$ são ambas verificadas, de modo que a distribuição binomial das proporções amostrais pode ser aproximada por uma distribuição normal com $\mu = n \cdot p$ e $\sigma = \sqrt{n \cdot p \cdot (1-p)}$.

Se as duas suposições acima forem satisfeitas podemos realizar o teste de hipóteses para uma proporção utilizando a tabela da distribuição normal padrão.

A notação que utilizaremos é a seguinte:

n = tamanho da amostra;
$\hat{p} = \frac{x}{n}$ (proporção amostral)
p_0 = proporção populacional (usada nas hipóteses)

A estatística de teste que será utilizada é obtida por:

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

Em muitos aspectos, o teste de hipóteses para uma proporção se assemelha grandemente ao teste z, principalmente pelo fato de que utilizam a mesma distribuição de probabilidades. Dessa forma, a região crítica, o valor crítico e o valor p para o teste de proporção são obtidos exatamente da mesma maneira que o teste z para uma média.

Vamos agora considerar um exemplo que ilustra a aplicação prática dos conceitos dados acima:

"Em um estudo da eficácia do air-bag em automóveis, constatou-se que, em 821 colisões de carros de tamanho médio equipados com air-bag, 46 colisões resultaram em hospitalização do motorista. Ao nível de significância de 0,01, teste a afirmação de que a taxa de hospitalização nos casos de air-bag é inferior à taxa de 7,8% para colisões de carros de tamanho médio equipados com cintos automáticos de segurança."

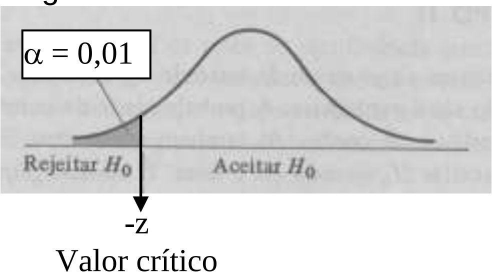
Podemos observar que a proporção amostral é de $\hat{p} = \frac{46}{821} = 0{,}056$ e a proporção populacional admitida é de 7,8% (0,078). O nível de significância a ser adotado é de 0,01 e as hipóteses podem ser definidas da seguinte maneira:

H_0: p $\geq$ 0,078
H_a: p < 0,078

Podemos verificar que as exigências do teste são ambas satisfeitas:
$n \cdot p \geq 5 \rightarrow$ 821 x 0,078 = 64,04
$n \cdot (1 - p) \geq 5 \rightarrow$ 821 x (1 – 0,078) = 756,96

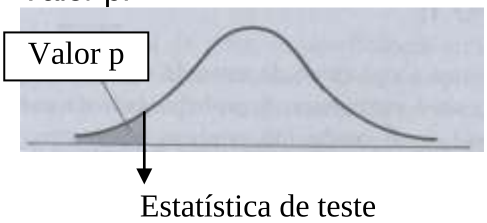
A estatística de teste é: $z_{teste} = \dfrac{\hat{p} - p_0}{\sqrt{\dfrac{p_0(1 - p_0)}{n}}} = \dfrac{0{,}056 - 0{,}078}{\sqrt{\dfrac{0{,}078 \cdot (1 - 0{,}078)}{821}}} = -2{,}35$

Região crítica:



O valor crítico deverá ser o valor de z na tabela da distribuição normal que corresponde a uma área de 0,49, ou seja, z = 2,33. Como a estatística de teste (-2,35) é menor do que o valor crítico (-2,33) devemos rejeitar a hipótese nula.

Valor p:



O valor p será a área da distribuição normal padrão abaixo da estatística de teste (-2,35), ou seja, 0,5 – 0,4906 = 0,0094. Interpretando o valor p temos que: Se rejeitarmos H_0, ou seja, se afirmarmos que a taxa de hospitalização nos casos com air-bag é inferior a 7,8% estamos cometendo um erro de 0,0094. Como o valor p (0,0094) é menor do que o nível de significância (0,01) vamos rejeitar H_0.

Concluindo: Há indícios suficientes para garantir que a taxa de hospitalização nos casos de automóveis equipados com air-bag é inferior à taxa de 7,8%.

Vale a pena lembrar que, para testes bilaterais, existe a opção de conclusão por meio do intervalo de confiança já conhecido:

$$\left[\hat{p} \pm z \times \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \right]$$

Devemos rejeitar uma afirmação de que o parâmetro populacional tenha um valor que não está compreendido no intervalo de confiança.

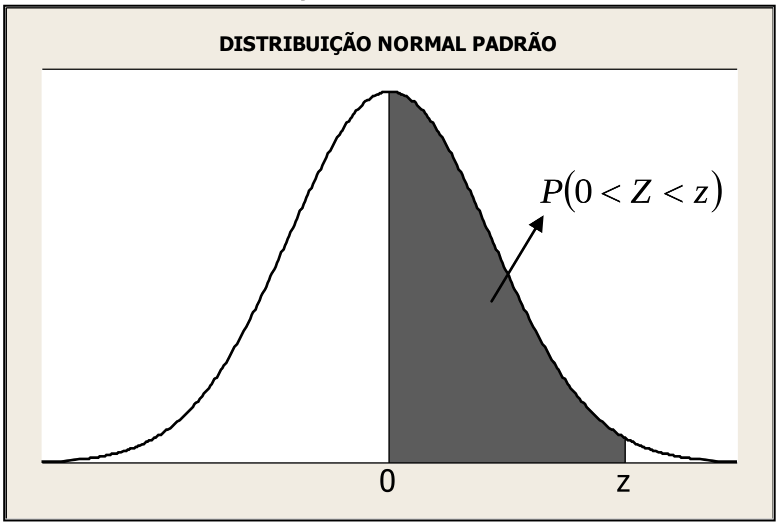
Exercícios em sala

Para os exercícios abaixo utilize a região crítica, o valor p e o intervalo de confiança (quando apropriado):

1. Um sistema de reservas da companhia Y acusa uma taxa de 7% de não-comparecimento. Adotou-se então um novo processo, pelo qual as reservas são confirmadas no dia anterior ao do vôo, fazendo-se um estudo de 5218 reservas pelo novo sistema, selecionadas aleatoriamente. Se se registraram 333 não-comparecimentos, teste a afirmação de que a taxa de não-comparecimento é menor no novo sistema. O novo sistema se afigura eficiente na redução do não-comparecimento? Utilize um nível de 2% de significância.

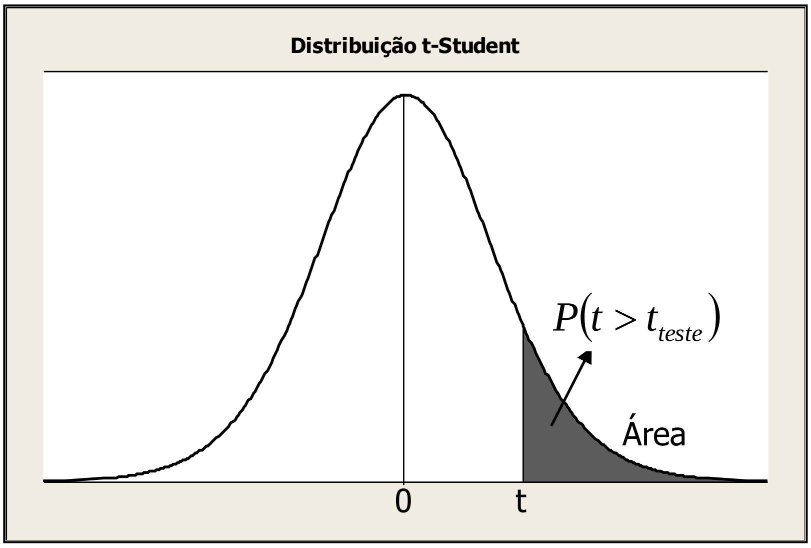
2. Em um estudo de 71 fumantes que estavam procurando deixar de fumar utilizando uma terapia especial, 32 não estavam fumando um ano após o tratamento. Ao nível de 0,10 de significância, teste a afirmação de que, dos fumantes que procuram deixar de fumar com aquela terapia, no máximo 25% voltam a fumar um ano após o tratamento. Esses resultados sugerem que a terapia não é eficaz?

DISTRIBUIÇÃO NORMAL PADRÃO



z	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998

DISTRIBUIÇÃO t-STUDENT



g.l	0,25	0,125	0,1	0,05	0,025	0,0125	0,01	0,005	0,0025	0,001	0,0005
1	1	2,414	3,078	6,314	12,71	25,45	31,82	63,66	127,3	318,3	636,6
2	0,817	1,604	1,8856	2,92	4,303	6,205	6,965	9,925	14,09	22,33	31,6
3	0,765	1,423	1,6377	2,3534	3,182	4,177	4,541	5,841	7,453	10,21	12,92
4	0,741	1,344	1,5332	2,1319	2,776	3,495	3,747	4,604	5,598	7,173	8,61
5	0,727	1,301	1,4759	2,0151	2,571	3,163	3,365	4,032	4,773	5,893	6,869
6	0,718	1,273	1,4398	1,9432	2,447	2,969	3,143	3,707	4,317	5,208	5,959
7	0,711	1,254	1,4149	1,8946	2,365	2,841	2,998	3,499	4,029	4,785	5,408
8	0,706	1,24	1,3968	1,8596	2,306	2,752	2,896	3,355	3,833	4,501	5,041
9	0,703	1,23	1,383	1,8331	2,262	2,685	2,821	3,25	3,69	4,297	4,781
10	0,7	1,221	1,3722	1,8125	2,228	2,634	2,764	3,169	3,581	4,144	4,587
11	0,697	1,214	1,3634	1,7959	2,201	2,593	2,718	3,106	3,497	4,025	4,437
12	0,695	1,209	1,3562	1,7823	2,179	2,56	2,681	3,055	3,428	3,93	4,318
13	0,694	1,204	1,3502	1,7709	2,16	2,533	2,65	3,012	3,372	3,852	4,221
14	0,692	1,2	1,345	1,7613	2,145	2,51	2,625	2,977	3,326	3,787	4,14
15	0,691	1,197	1,3406	1,7531	2,131	2,49	2,602	2,947	3,286	3,733	4,073
16	0,69	1,194	1,3368	1,7459	2,12	2,473	2,583	2,921	3,252	3,686	4,015
17	0,689	1,191	1,3334	1,7396	2,11	2,458	2,567	2,898	3,222	3,646	3,965
18	0,688	1,189	1,3304	1,7341	2,101	2,445	2,552	2,878	3,197	3,611	3,922
19	0,688	1,187	1,3277	1,7291	2,093	2,433	2,539	2,861	3,174	3,579	3,883
20	0,687	1,185	1,3253	1,7247	2,086	2,423	2,528	2,845	3,153	3,552	3,85
21	0,686	1,183	1,3232	1,7208	2,08	2,414	2,518	2,831	3,135	3,527	3,819
22	0,686	1,182	1,3212	1,7172	2,074	2,405	2,508	2,819	3,119	3,505	3,792
23	0,685	1,18	1,3195	1,7139	2,069	2,398	2,5	2,807	3,104	3,485	3,768
24	0,685	1,179	1,3178	1,7109	2,064	2,391	2,492	2,797	3,091	3,467	3,745
25	0,684	1,178	1,3164	1,7081	2,06	2,385	2,485	2,787	3,078	3,45	3,725
26	0,684	1,177	1,315	1,7056	2,056	2,379	2,479	2,779	3,067	3,435	3,707
27	0,684	1,176	1,3137	1,7033	2,052	2,373	2,473	2,771	3,057	3,421	3,69
28	0,683	1,175	1,3125	1,7011	2,048	2,368	2,467	2,763	3,047	3,408	3,674
29	0,683	1,174	1,3114	1,6991	2,045	2,364	2,462	2,756	3,038	3,396	3,659
30	0,683	1,173	1,3104	1,6973	2,042	2,36	2,457	2,75	3,03	3,385	3,646
31	0,682	1,172	1,3095	1,6955	2,04	2,356	2,453	2,744	3,022	3,375	3,633
32	0,682	1,172	1,3086	1,6939	2,037	2,352	2,449	2,738	3,015	3,365	3,622
33	0,682	1,171	1,3077	1,6924	2,035	2,348	2,445	2,733	3,008	3,356	3,611
34	0,682	1,17	1,307	1,6909	2,032	2,345	2,441	2,728	3,002	3,348	3,601
35	0,682	1,17	1,3062	1,6896	2,03	2,342	2,438	2,724	2,996	3,34	3,591