Probabilidade e Estatística

Professores:

Otton Teixeira da Silveira Filho

Regina Célia Paula Leal Toledo

Probabilidade e Estatística

Livro Texto:

[1] "Noções de Probabilidade e Estatística"
Marcos Nascimento Magalhães e Antonio Carlos Pedroso de Lima, Edusp (2005).

[2] "Probabilidade: Um Curso Introdutório"
Carlos A. B. Dantas, Edusp (2004).

Aula 12

Professores:

Otton Teixeira da Silveira Filho
Regina Célia Paula Leal Toledo

Inferência Estatística - Testes de Hipóteses

Conteúdo:

12.1 Introdução

12.2 Teste para a Média Populacional

12.3 Teste para a Média com Variância Desconhecida

12.1 Introdução

Ao trabalharmos com probabilidade estamos trabalhando com o que não conhecemos.

Para trabalharmos temos, então, que fazer hipóteses sobre

A média da população

O modelo

Etc.

12.1 Introdução

Nas aulas 10 e 11 vimos:

Estimação

Agora veremos como avaliar as hipóteses necessárias para a obtenção de informações.

Procedimento Clássico de testes de hipóteses.

12.1 Introdução

Um exemplo:

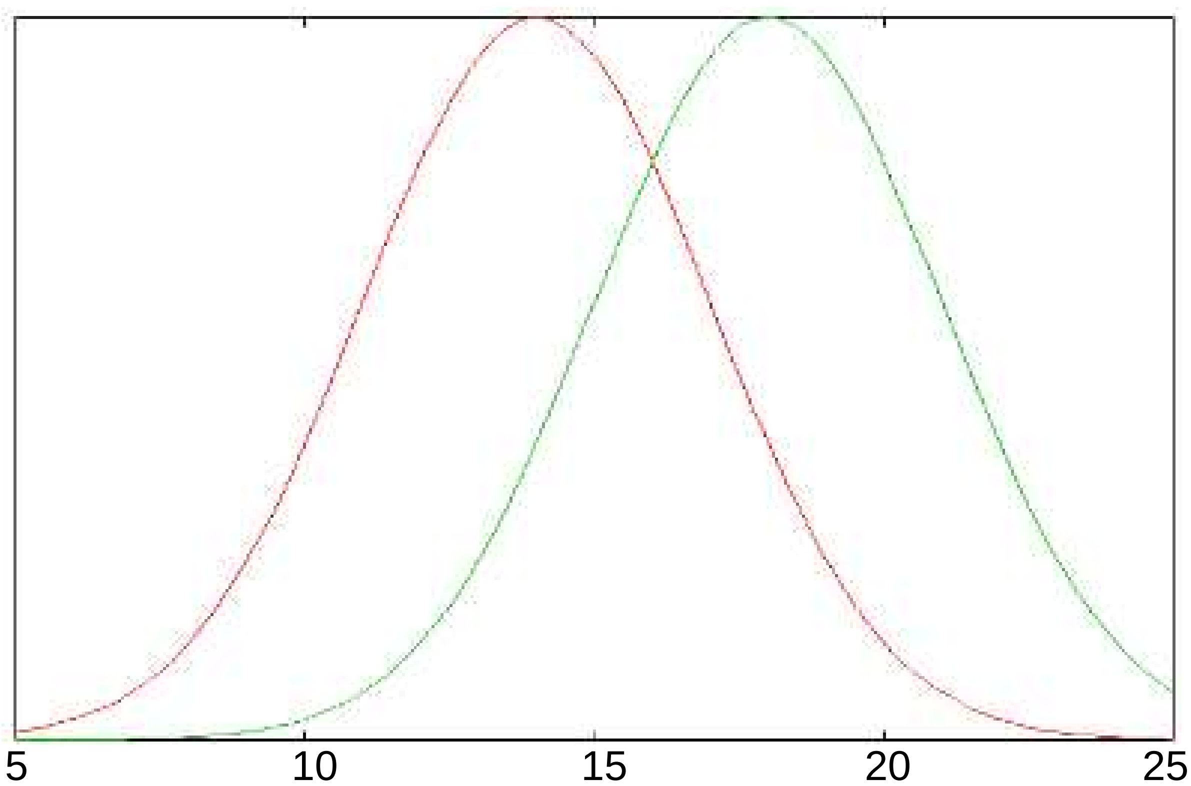
Em pessoas sadias, uma determinada substância encontrada no sangue se comporta segundo um modelo Normal com média 14 unidades/ml com desvio padrão 6 unidades/ml.
Portadores de uma determinada doença, tem a mesma substância com valores alterados para uma média de 18 unidades/ml.

Hipóteses:

- A população doente também segue uma distribuição Normal
- O desvio padrão dos doentes é também de 6 unidades/ml.

12.1 Introdução

O exemplo num gráfico:



12.1 Introdução

Como foi visto, as duas distribuições tem uma área em comum, ou seja, existem valores dentro da normalidade média que podem estar dentro dos parâmetros de anormalidade média.

Como distinguir, então, os doentes dos sãos?

12.1 Introdução

Vamos dizer que queremos testar um tratamento para combater esta doença. Tomemos uma amostra aleatória de 30 pessoas e representemos as amostras $\boldsymbol{X}_i$ como $(\boldsymbol{X}_1, ..., \boldsymbol{X}_{30})$. Temos então que

$$X_i \sim N(\mu, 36)$$

podendo termos 14 ou 18 para a média.

Se o valor obtido pela amostra tiver um valor médio "alto e próximo" de 18, diremos que o tratamento não funciona mas se a amostra tiver um valor médio "baixo e próximo" de 14 então o tratamento é efetivo.

Mas o que quer dizer "alto", "baixo" e "próximo"?

12.1 Introdução

Quando temos situações similares a esta, temos que fazer um estudo probabilístico do problema o que inclui um teste das hipóteses que estamos fazendo.

12.1 Introdução

Um exemplo:

Queremos estudar a tolerância de um equipamento eletrô-nico. Pelas características do equipamento se admite que a probabilidade de falha é constante, ou seja, após cada teste existe uma probablidade ***p*** que ele falhe.

Pretende-se verificar se o modelo geométrico com ***p***=0,4 é adequado para caracterizar a variável aleatória ***X*** que re-presenta o número de testes antes de se apresentar uma falha.

Note que:
Estamos testando o modelo e não a variável

12.1 Introdução

Um exemplo:

Testando o modelo

Supondo que temos uma amostra de 80 equipamentos para os testes e, assim, temos 80 realizações da variável X dadas pelo vetor amostral ($\boldsymbol{X}_1$, ..., $\boldsymbol{X}_{80}$). Assim $\boldsymbol{X}_i$ representa o número de impactos até i-ésimo equipamento falhar e podemos construir uma tabela de frequência agrupando, ou não, os valores da variável ***X*** em categorias.

Se a hipótese de que a distribuição ser do tipo geométrico, devemos ter uma frequência não muito diferente da obtida diretamente da distribuição.

Mas o que quer dizer não muito diferente?

12.1 Introdução

Um exemplo:

Mas o que quer dizer não muito diferente?

Podemos avaliar a "distância" entre os valores esperados (e_i) e os observados (o_i) usando, por exemplo,

$$Q^2 = \frac{\sum_{i=1}^{k} (o_i - e_i)^2}{e_i}$$

com ***k*** sendo o total das categorias que atribuimos para a variável ***X***.

12.1 Introdução

Nas duas situações que foram apresentadas estabelecemos procedimentos, ainda incompletos para dois tipos de teste:

O primeiro se refere a média populacional

O segundo é sobre a distribuição de probabilidade

Faça os exercícios da seção 8.1

12.1 Introdução

Voltando ao primeiro exemplo:

Em pessoas sadias, uma determinada substância encontrada no sangue se comporta segundo um modelo Normal com média 14 unidades/ml com desvio padrão 6 unidades/ml.
Portadores de uma determinada doença, tem a mesma substância com valores alterados para uma média de 18 unidades/ml.

Hipóteses:

- A população doente também segue uma distribuição Normal
- O desvio padrão dos doentes é também de 6 unidades/ml.

12.2 Teste para Média Populacional

Aqui é suposto que é conhecida a variância populacional.

Embora um caso particular, existem muitas aplicações em que é possível trabalhar com esta suposição:

- Se podemos assegurar que uma máquina fornece medidas com uma precisão constante;

- Usar informações de estudos estatísticos de casos similares ao que estamos tratando.

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Vamos usar a média amostral $\overline{X}$, um estimador não viciado e consistente com μ. Com o valor observado de $\overline{X}$, denominado $\boldsymbol{X}_{obs}$ avaliaremos o tratamento proposto.

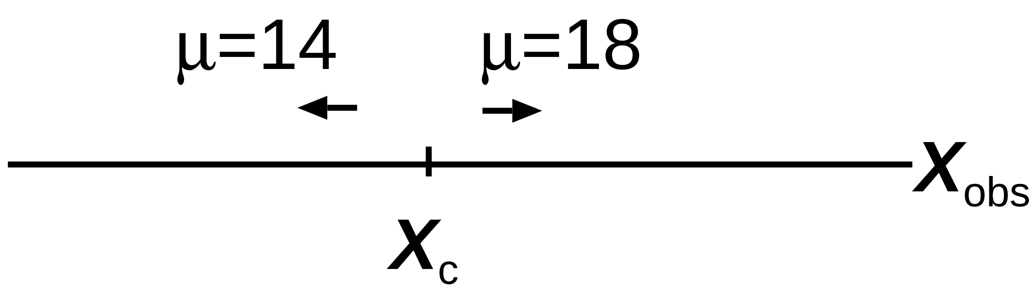
- Como $\overline{X}$ é variável aleatória, ela pode apresentar valores maiores que 14 mesmo que μ = 14.
- Um critério para decidir sobre o valor de μ é determinar um valor crítico $\boldsymbol{X}_c$, tal que se $\overline{X}$ for maior que $\boldsymbol{X}_c$, concluímos que a amostra da população com média 14, sendo o tratamento eficaz.

Obs: sendo $\overline{X}$ aleatória, podemos incluir a igualdade a $\boldsymbol{X}_c$ em qualquer decisão.

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Graficamente



A questão é que temos não só que determinar X_c como quantificar os erros associados às conclusões.

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Investigaremos se um determinado tratamento será capaz de promover a cura da população doente, ou seja, transferir esta população (de média 18 unidades/ml) para a de sãos (de média 14 unidades/ml).

Temos duas hipóteses:

H_0 : Hipótese nula (tratamento não eficaz)

H_a : Hipótese alternativa (tratamento eficaz)

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

O tratamento é eficaz se fizer os indivíduos da amostra mudarem para uma população de média menor que 18 unidades/ml. Caso contrário a média não se alteraria. Assim as hipóteses de interesse seriam escritas como

$$H_0 : \mu = 18 \text{ versus } H_a : \mu < 18$$

Aqui temos uma hipótese composta unilateral.

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Para verificarmos se o tratamento produz algum efeito, seja benéfico ($\mu < 18$) ou danoso ($\mu > 18$), devemos construir um teste de hipóteses bilateral:

$$H_0 : \mu = 18 \text{ versus } H_a : \mu \neq 18$$

Por uma conveniência técnica deixamos a igualdade na hipótese nula.

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Os dois erros que podem ser cometidos ao realizar um teste de hipóteses são:

i) Rejeitar a hipótese H_0 quando tal hipótese é verdadeira (Erro tipo I);

ii) Não rejeitar a hipótese H_a quando ela deveria ser rejeitada (Erro tipo II).

Num resumo temos

Decisão	H_0 Verdadeira	H_0 Falsa
Rejeitar H_0	Erro Tipo I	Sem Erro
Não rejeitar H_0	Sem Erro	Erro Tipo II

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Uma parte importante do teste de hipóteses é controlar a probabilidade de cometermos um erro do tipo I. Denotaremos esta probabilidade por α e denotaremos por β a probablidade de erro do tipo II.

α = P(erro tipo I) = P(rejeitar H_0 e H_0 verdadeira)

β = P(erro tipo II) = P(não rejeitar H_0 e H_0 falsa)

No caso de nosso exemplo temos

α = P(concluir que o tratamento é eficaz quando na verdade ele não é);

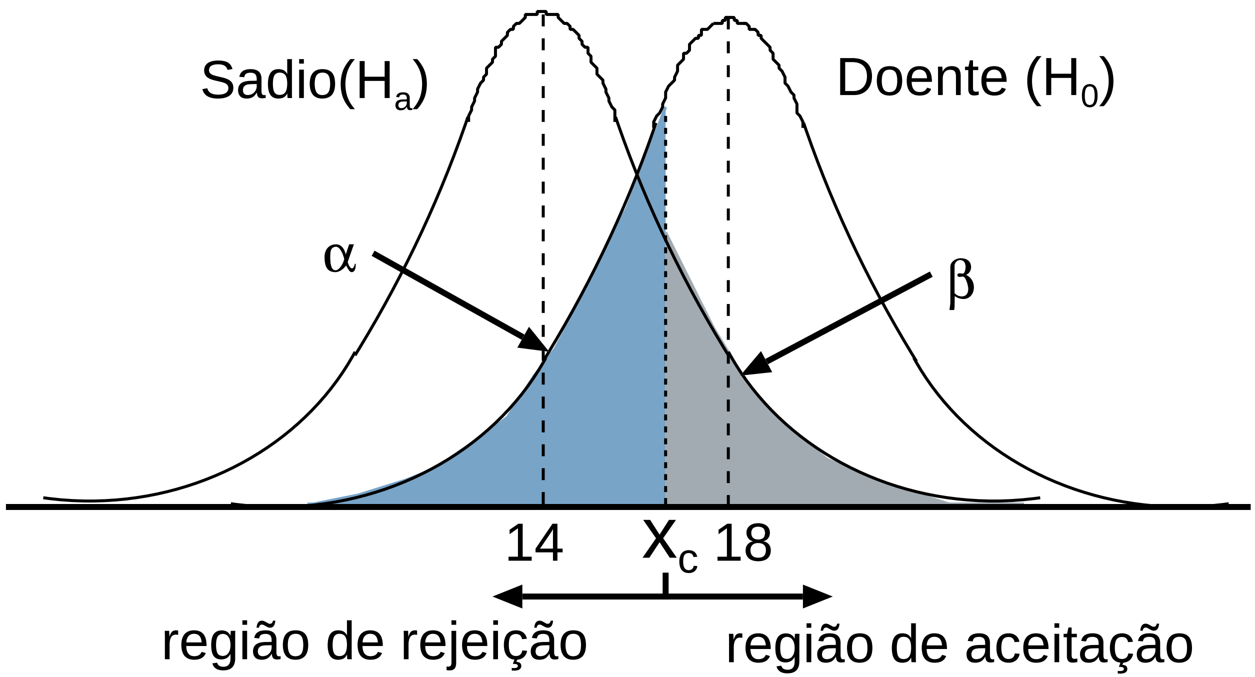
β = P(concluir que o tratamento não é eficaz quando na verdade ele é);

A situação ideal é quando estas probabilidades são próximas de zero, mas....

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Observe que ao deslocarmos X_c , variamos as probabilidades α e β de forma que ao diminuirmos uma, aumentamos a outra, como pode ser visto na figura abaixo:



Portanto, devemos evitar o erro do tipo I, ou seja, α deve ser o menor possível.

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

Supondo α conhecido vamos descrever como determinar o valor crítico $\boldsymbol{X}_c$.

Notando que

α = P(erro tipo I) = P(rejeitar H_0 e H_0 verdadeira) ou

$$\alpha = P(\overline{X} < x_c \,|\mu = 18) = P\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} < \frac{x_c - 18}{6/\sqrt{30}}\right) = P(Z < z_c)$$

com Z ~ N(0,1). Se dado α obtemos z_c na tabela Normal.

Nos resta calcular x_c ....

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

calculando x_c ...

$$z_c = \frac{x_c - 18}{6/\sqrt{30}} \Rightarrow x_c = 18 + z_c \frac{6}{\sqrt{30}}$$

Se fizermos α = 0,05, 5% de probabilidade para dar erro do tipo I, temos

$$0,05 = P(Z < z_c) \Rightarrow z + c = -1,64$$

logo

$$x_c = 18 - 1,64 \frac{6}{\sqrt{30}} = 16,20$$

assim o método é eficaz se a média observada for menor que 16,20 .

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:

A região dada pelo conjunto de números reais menores que 16,20 é denominada Região de Rejeição ou Região Crítica (RC). No caso no qual estamos trabalhando ela será

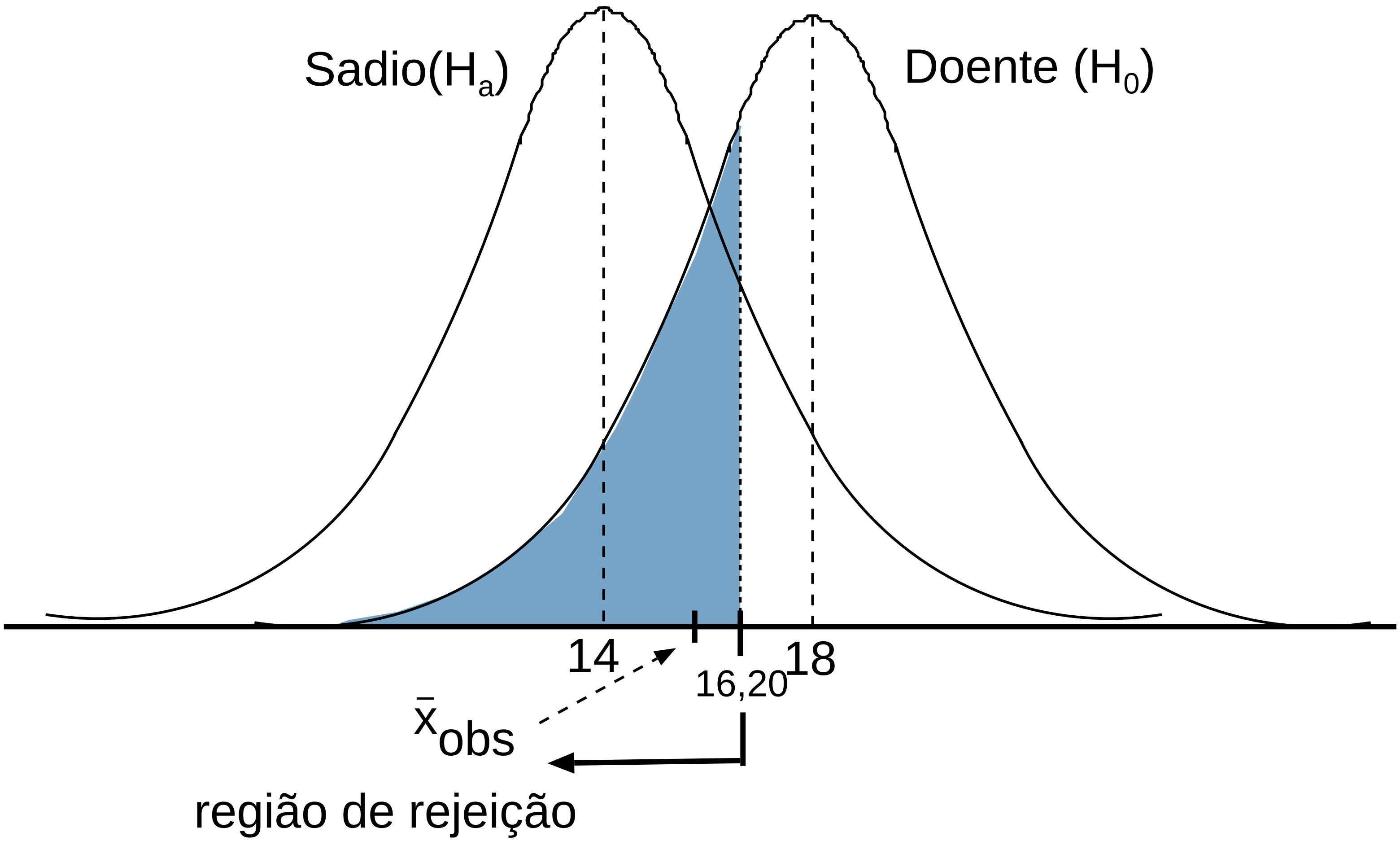
$$RC = \{x \in \mathbb{R} : x < 16,20\}$$

Denominamos Região de Aceitação (RA) ao complementar de RC. Se a amostra obtida nos deu como estimativa $\overline{X}_{obs} = 16,04$ pertencente à RC, então rejeitamos H_0 no nível de significância $\alpha = 0,05$.

Graficamente...

12.2 Teste para Média Populacional

Voltando ao primeiro exemplo:



12.2 Teste para Média Populacional

Hipóteses Bilaterais

De forma similar do que no caso unilateral, podemos construir testes de hipóteses bilaterais.

A diferença está em termos que considerar uma Região de Rejeição composta de duas partes disjuntas.

12.2 Teste para Média Populacional

Hipóteses Bilaterais

Exemplificando, suponha que μ é conhecida e que as hipóteses numa e alternativa são expressas como

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

A Região de Crítica (RC) será

$$RC = \{x \in \mathbb{R} : x < x_{c1} \; ou \; x > x_{c2}\},$$

e, para um valor α fixo, determinamos x_{c1} e x_{c2} de modo que

$$RC = \{x \in \mathbb{R} : \overline{X} < x_{c1} \; ou \; \overline{X} > x_{c2}\} = \alpha$$

12.2 Teste para Média Populacional

Hipóteses Bilaterais

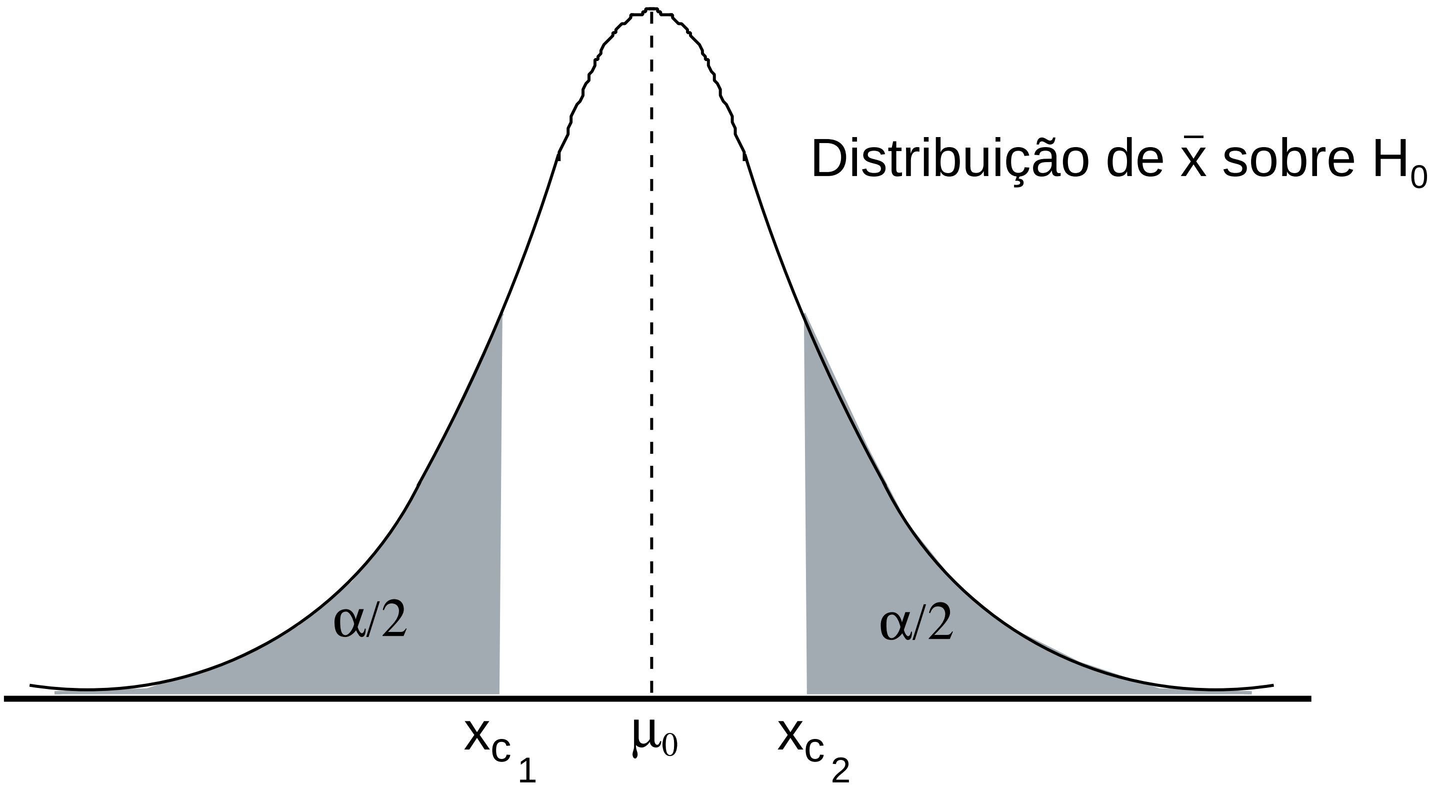
Como a densidade Normal é simétrica, distribuimos a massa α igualmente entre as duas partes da Região de Rejeição, ou seja,

$$P(\overline{X} < x_{c1}) = \frac{\alpha}{2} \; e \; P(\overline{X} > x_{c2}) = \frac{\alpha}{2}$$

Graficamente fica mais fácil de perceber...

12.2 Teste para Média Populacional

Hipóteses Bilaterais



Escolha dos valores críticos

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

Um pesquisador deseja estudar o efeito de uma substância no tempo de reação de certas cobaias submetidas a um estímulo elétrico. Os valores de tempo de reação foram 9,1; 9,3; 7,2; 7,5; 13,3; 10,9; 7,2; 9,9; 8,0; 8,6. Admite-se que o tempo de reação segue, em geral, o modelo Normal com média 8 e desvio padrão 2 segundos. O pesquisador desconfia que o tempo médio sobre alteração pela substância.

Neste caso as hipóteses de interesse são

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

Neste caso as hipóteses de interesse são

H_0 : as cobaias apresentam tempo de reação padrão;

H_a : as cobaias apresentam tempo de reação alterado.

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

Em termos estatísticos, as hipóteses envolvem μ e podem ser escritas como

$H_0 : \mu = 8{,}0;$ $H_a : \mu \neq 8{,}0;$

É natural usarmos a média amostral $\overline{X}$ para construir a estatística de teste e usamos que $\overline{X} \sim N(\mu, \frac{4}{10})$ já que a amostra é de tamanho 10 e o desvio padrão vale 2. Pela forma que H_a foi especificado, teremos a região crítica como

$$RC = \{x \in \mathbb{R} : x < x_{c1} \ ou \ x > x_{c2}\},$$

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

Fixando $\alpha = 0{,}06$ termos:

ou 0,06 = P(erro tipo I) = P(rejeitar H_0 e H_0 verdadeira)

$$0,06 = P(\overline{X} \in RC | \mu = 8,0) = P(\overline{X} < x_{c1} \ ou \ \overline{X} > x_{c2} \ \mu = 8,0)$$

ou ainda

$$0,06 = P\left(\frac{\overline{X} - 8,0}{\sqrt{4/10}} < \frac{x_{c1} - 8,0}{\sqrt{4/10}} \text{ ou } \frac{\overline{X} - 8,0}{\sqrt{4/10}} > \frac{x_{c2} - 8,0}{\sqrt{4/10}}\right)$$

$$= P(Z < z_{c1} \text{ ou } Z > z_{c2})$$

onde

$$z_{c_j} = \frac{(x_{c_j} - 8,0)}{\sqrt{4/10}}, j = 1,2 \text{ e } Z \sim N(0,1).$$

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

onde

$$z_{c_j} = \frac{(x_{c_j} - 8,0)}{\sqrt{4/10}}, j = 1,2 \text{ e } Z \sim N(0,1).$$

Da tabela da distribuição Normal obtemos z_{c1} = -1,88 e z_{c2} = 1,88. Logo,

$$x_{c1} = 8 - 1,88\sqrt{4/10} = 6,8 \qquad x_{c2} = 8 + 1,88\sqrt{4/10} = 9,2$$

donde podemos expressar a Região Crítica como

$$RC = \{x \in \mathbb{R} : x < 6,8 \ ou \ x > 9,2\}$$

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

Região Crítica

$$RC = \{x \in \mathbb{R} : x < 6,8 \; ou \; x > 9,2\}$$

Ao calcular a média amostral achamos 9,1 que está fora de RC, aceitamos a hipótese H_0 com significância 6%, valor de α .

Este resultado indica que o tempo de reação das cobaias não fica alterado pela substância pesquisada.

Podemos fazer um pouco mais....

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

Qual é a probabilidade do erro tipo II?

$$RC = \{x \in \mathbb{R} : x < 6,8 \text{ ou } x > 9,2\}$$

O erro tipo II é a probabilidade de aceitar incorretamente H_0. Mas isto não é tão simples pois existem vários valores possíveis para μ já que a hipótese alternativa H_a é composta e temos que a probabilidade do erro tipo II é função de μ, ou seja, $\beta(\mu)$

Calculemos o caso

$$\mu = 9,0.$$

12.2 Teste para Média Populacional

Hipóteses Bilaterais

...e mais um exemplo

Calculemos o caso μ = 9,0.

$$\beta(9,0) = \text{ P(erro tipo II)} = \text{P(não rejeitar } H_0 \text{ e } H_0 \text{ falsa)}$$

$$\beta(9,0) = P(\overline{X} \notin RC|\mu = 9,0) = P(6,8 \leq \overline{X} < 9,2|\mu = 9,0)$$

$$0,06 = P\left(\frac{6,8-9,0}{\sqrt{4/10}} < \frac{\overline{X}-9,0}{\sqrt{4/10}}\right) =$$

$$= P(-3,48 \leq Z \leq 0,32) = 0,4497 + 0,1255 = 0,6252$$

Para o caso μ = 9 e com probabilidade de 62,52% estaríamos concluindo, de forma equivocada, que H_0 é verdadeira.

12.2 Teste para Média Populacional

Hipóteses Bilaterais

Função poder $\pi(\mu)$

Definimos a função poder como

$$\pi(\mu) = P(\text{rejeitar } H_0 \mid \mu)$$

Se a média for a de H_0, a função poder é igual ao nível de significância α;

Se μ é um dos valores de H_a, a função poder será dada por $\pi(\mu) = 1 - \beta(\mu)$;

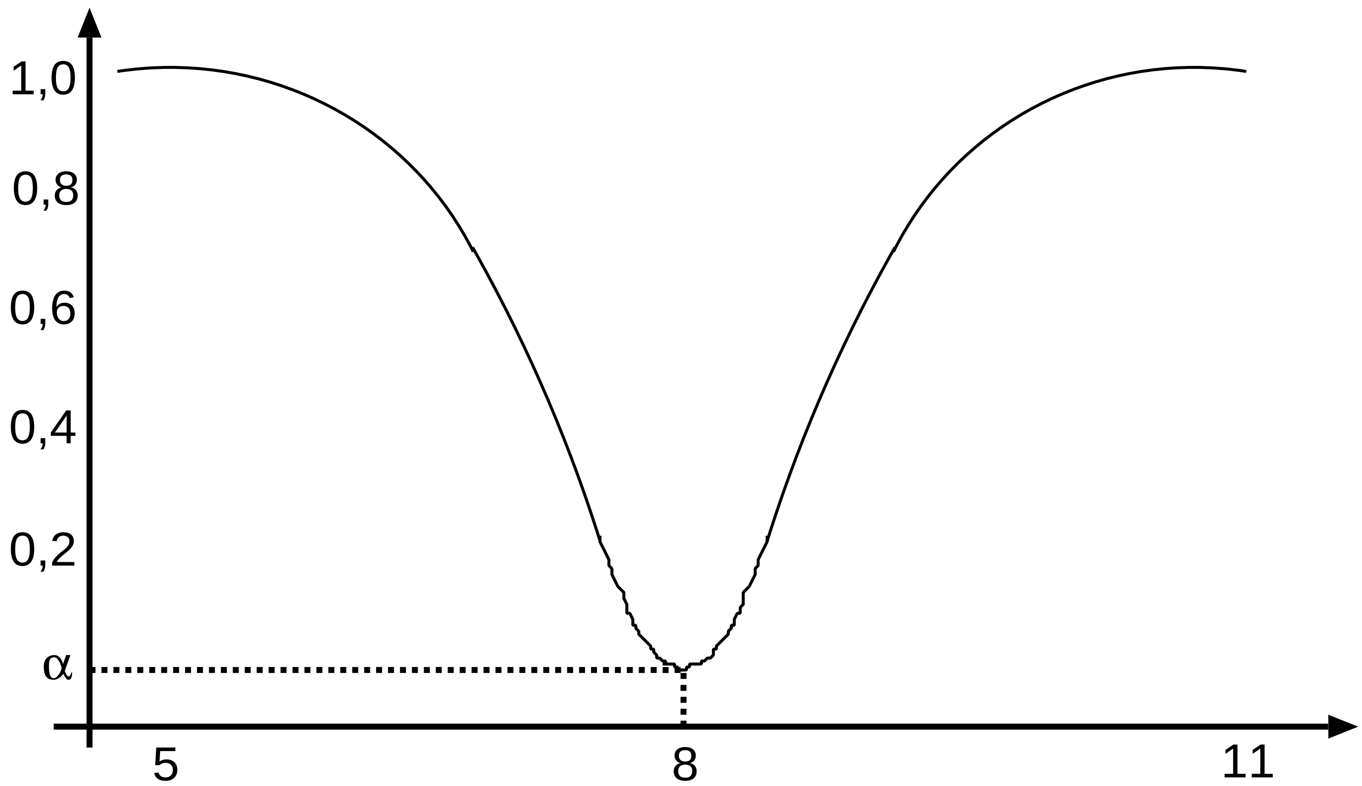
Para um mesmo nível de significância, quanto maior o poder melhor o teste.

No caso do exercício tratado anteriormente temos o gráfico da função poder dado por...

12.2 Teste para Média Populacional

Hipóteses Bilaterais

Função poder $\pi(\mu)$



A medida que nos afastamos da região de hipótese nula, $\pi(\mu)$ aumenta.

12.2 Teste para Média Populacional

Resumo

Não é possível diminuir simultaneamente os dois erros num mesmo teste;

Podemos diminuir simultaneamente os dois erros aumentando o tamanho da amostra;

Mesmo que não se conheça a distribuição da população, podemos contornar este problema usando o Teorema Central do Limite.

12.2 Teste para Média Populacional

Sumário para um teste de hipóteses

1. Estabelecer as hipóteses nula e alternativa;

2. Definir a forma da região crítica, com base na hipótese alternativa;

3. Identificar a distribuição do estimador e obter sua estimativa;

4. Fixar α e obter a região crítica;

5. Concluir o teste com base na estimativa e na região crítica.

12.2 Teste para Média Populacional

Mais um exemplo

Um relatório afirma que 40% da água obtida em poços artesianos no nordeste é salobra. Tal relatório foi colocado em discussão e, para redimir dúvidas foram examinados 400 poços e 120 deles tinham água salobra. Qual seria a conclusão que se chega para um nível de 3%?

Vamos construir o teste...

12.2 Teste para Média Populacional

Mais um exemplo

Definindo as hipóteses

• O parâmetro de interesse é a proporção ***p*** de poços com água salobra. Pela informação fornecida, temos um teste bilateral com

$$H_0 : \boldsymbol{p} = 0{,}40; \qquad H_a : \boldsymbol{p} \neq 0{,}40.$$

12.2 Teste para Média Populacional

Mais um exemplo

- Sabemos que o melhor estimador para $\boldsymbol{p}$ é a proporção amostral $\hat{\boldsymbol{p}}$ cuja distribuição pode ser aproximada por um modelo Normal, ou seja,

$$\hat{\boldsymbol{p}} \sim N(\boldsymbol{p}, \boldsymbol{p}(1-\boldsymbol{p})/n)$$

- Dado que o testes é bilateral, a Região Crítica é da forma

$$RC = \{x \in \mathbb{R} : x < p_{c1} \text{ ou } x > p_{c2}\},$$

12.2 Teste para Média Populacional

Mais um exemplo

Para α = 0,03, os valores $\boldsymbol{p}_{c1}$ e $\boldsymbol{p}_{c2}$ são calculados através de

$$P(\widehat{p} < p_{c1} | H_0) = \frac{0,03}{2} \qquad P(\widehat{p} > p_{c2} | H_0) = \frac{0,03}{2}$$

- Sob a hipótese H_0 , $\boldsymbol{p}$ vale 0,40 e, portanto,

$$\widehat{p} \sim N\left(0,4; \frac{0,24}{400}\right)$$

Calculando a probabilidade ...

12.2 Teste para Média Populacional

Mais um exemplo

Calculando a probabilidade

$$P(\widehat{p} < p_{c1}|H_0) = P\left(\frac{\widehat{p} - 0,40}{\sqrt{0,24/400}} < \frac{p_{c1} - 0,4}{\sqrt{0,24/400}}\right) = 0,015$$

Da tabela N(0,1) segue

$$\frac{p_{c1} - 0,40}{\sqrt{0,24/400}} = -2,17$$

e isto resulta em....

12.2 Teste para Média Populacional

Mais um exemplo

E isto resulta em $\boldsymbol{p}_{c1}$ = 0,347. De maneira análoga obtemos $\boldsymbol{p}_{c2}$ = 0,453.

$$P(\widehat{p} < p_{c1}|H_0) = P\left(\frac{\widehat{p} - 0,40}{\sqrt{0,24/400}} < \frac{p_{c1} - 0,4}{\sqrt{0,24/400}}\right) = 0,015$$

Finalmente obtemos a Região Crítica

$$RC = \{x \in \mathbb{R} : x < 0,347 \ ou \ x > 0,453\}$$

A amostra forneceu $\hat{\boldsymbol{p}}_{obs}$ = 120/400 = 0,300 que pertence à região crítica. Com isto a hipótese nula deve ser rejeitada ao nível de 3%, ou seja,

O relatório não está correto.

12.3 Teste para Média com Variância Desconhecida

Neste caso mais geral vamos começar trabalhando com a suposição de que a distribuição é Normal.

Vamos supor que nossa amostra seja representada pelo vetor amostral $(X_1, ..., X_n)$, todas com densidade Normal de média μ e variância σ^2. Usaremos o estimador S_2 que já vimos na Aula 10:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i^2 - n\overline{X}^2)$$

12.3 Teste para Média com Variância Desconhecida

E definimos a variável padronizada T

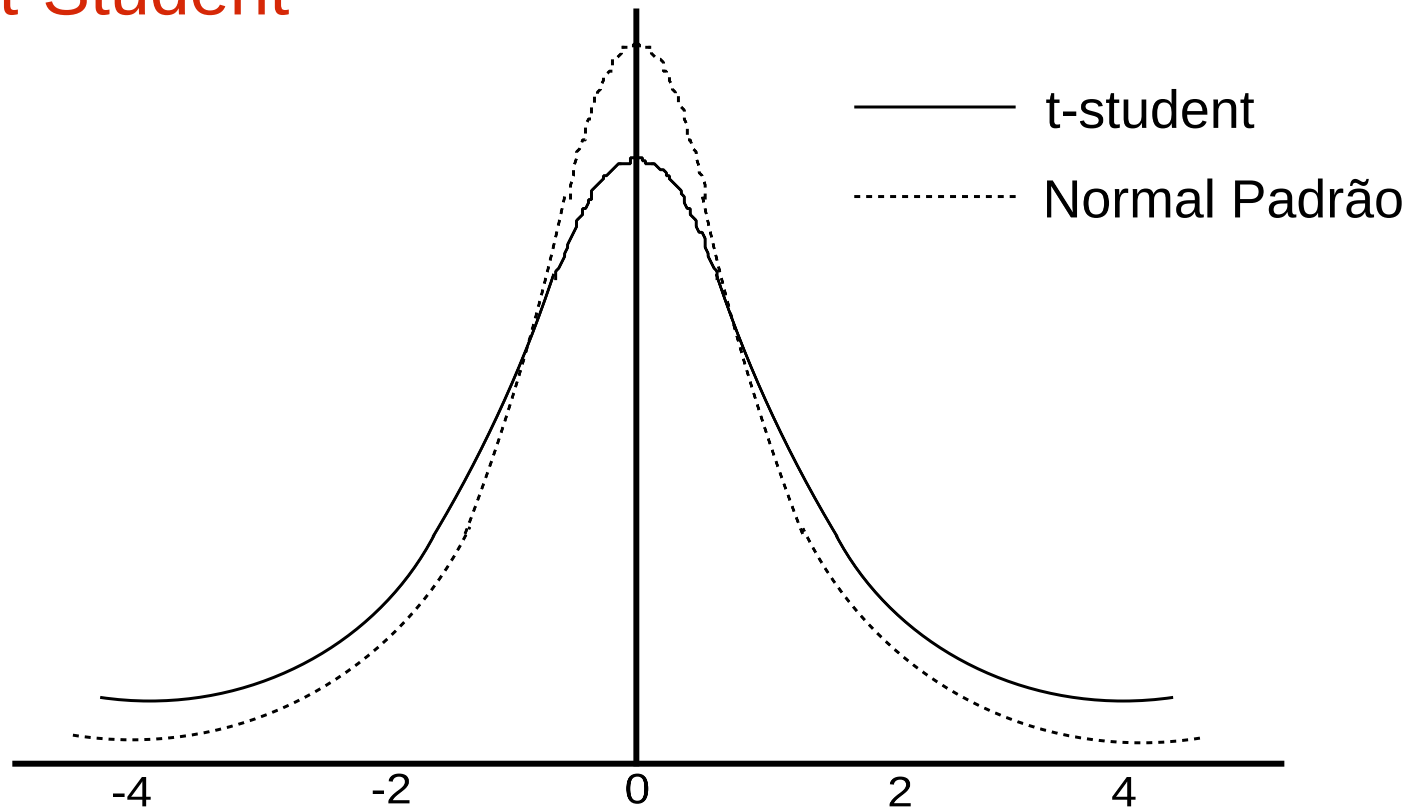
$$T = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} = \frac{\overline{X} - \mu}{S/\sqrt{n}}$$

que é aleatória. Mas, devido ao denominador acima envolver S, a distribuição de T não é Normal como $\overline{X}$. A distribuição de T é denominada densidade t de Student e seu parâmetro é chamado de graus de liberdade, que neste caso corresponde ao número de dados menos 1.

- A notação desta distribuição é $t_{(n-1)}$.
- Para os cálculos manuais são utilizadas tabelas.

12.3 Teste para Média com Variância Desconhecida

Modelo t-Student



O modelo t-Student tem caudas de maior massa do que N(0, 1).

12.3 Teste para Média com Variância Desconhecida

Modelo t-Student

- A medida que o tamanho da amostra aumenta, t-Student tende para a Normal padrão;

- As tabelas de t-Student se limitam a 120 graus de liberdade;

- Para graus superiores a 120, as probablidades são obtidas da tabela da distribuição Normal.

12.3 Teste para Média com Variância Desconhecida

Modelo t-Student

- A Região Crítica neste caso envolve S^2 que sendo aleatória faz com que amostras diferentes forneçam, em geral, regiões críticas diferentes;

- Optamos, então, por usar na região crítica valores da quantidade padronizada T.

12.3 Teste para Média com Variância Desconhecida

Um exemplo

Se investiga se uma doença dos rins altera o consumo de oxigênio destes orgãos. No caso de indivíduos sadios admite-se que o consumo tem distribuição Normal com média $12cm^3$ /minuto. Uma amostra de cinco pacientes com a doença obteve os valores (14,4; 12,9; 15,0; 13,7; 13,5). Qual seria a conclusão ao nível de significância de 1%?

12.3 Teste para Média com Variância Desconhecida

Um exemplo

H_0 : a doença não altera a média de consumo renal de oxigênio;

H_a : indivíduos portadores da doença tem a média alterada.

Em termos da média populacional, estamos testando as hipóteses

$H_0 : \mu = 12;$ versus $H_a : \mu \neq 12;$

e a região crítica será $RC = \{t \in \mathbb{R} : t < t_1 \text{ ou } t > t_2\}$

12.3 Teste para Média com Variância Desconhecida

Um exemplo

Sendo o desvio padrão desconhecido, usaremos o estimador

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i^2 - n\overline{X}^2)$$

e a quantidade t.

Sendo H_0 verdadeira

$$T = \frac{\overline{X} - 12}{S/\sqrt{5}} \sim t_{(4)}$$

logo

12.3 Teste para Média com Variância Desconhecida

Um exemplo

Logo,

$$P(T < t_1) = 0,01/2 \Rightarrow t_1 = -4,604;$$

$$P(T > t_2) = 0,005 \Rightarrow t_2 = 4,604;$$

valores obtidos tabela da distribuição t-Student com 4 graus de liberdade.

$$RC = \{t \in \mathbb{R} : t < -4,604 \ ou \ t > 4,604\}.$$

12.3 Teste para Média com Variância Desconhecida

Um exemplo

Sendo, $\overline{X}_{obs} = 13,90$ $\overline{S}^2_{obs} = 0,67$ calculamos o valor padronizado

$$t_{obs} = \frac{\overline{X}_{obs} - 12}{S_{obs}/\sqrt{5}} = \frac{13,90 - 12}{0,82\sqrt{5}} = 5,18$$

Como $t_{obs} \in RC$, decidimos pela rejeição da hipótese nula, ou seja,

A doença tem influência no consumo renal médio de oxigênio ao nível de 1%.

12.3 Teste para Média com Variância Desconhecida

Intervalo de confiança para μ com variância desconhecida

Supondo que uma amostra aleatória $X_1, ..., X_n$ obtida de uma população com distribuição Normal e média e variância desconhecidas, temos que

$$\frac{\overline{X} - \mu}{S_{obs}/\sqrt{n}} \sim t_{(n-1)}$$

Fixando um coeficiente de confiança $\gamma (0 < \gamma < 1)$ e utilizando a tabela de distribuição t-Student com n - 1 graus de liberdade, podemos obter $t_{\gamma/2}$ tal que

$$P\left(-t_{\gamma/2} < \frac{\overline{X} - \mu}{S_{obs}/\sqrt{n}} < t_{\gamma/2}\right) = \gamma$$

que nos dá o intervalo de confiança....

12.3 Teste para Média com Variância Desconhecida

Intervalo de confiança para μ com variância desconhecida

Intervalo de confiança

$$IC(\mu, \gamma) = \left[\overline{X} - t_{\gamma/2}\frac{S}{\sqrt{n}}; \overline{X} + t_{\gamma/2}\frac{S}{\sqrt{n}}\right]$$

12.3 Teste para Média com Variância Desconhecida

Intervalo de confiança para μ com variância desconhecida

continuando um exemplo

No último exemplo, rejeitamos a hipótese nula. Vamos estabelecer um intervalo de confiança para a média populacional. No exemplo foram obtidos $\overline{X}_{obs} = 13,90 \quad \overline{S}^2_{obs} = 0,67$
Com γ = 0,90 obtemos da tabela da distribuição t-Student com 4 graus de liberdade o valor $t_{\gamma/2}$ = 2,132. Assim,

$$\overline{X}_{obs} = 13,90 \quad \overline{S}^2_{obs} = 0,67$$

$$IC(\mu, 90) = [13,90 - 2,132\sqrt{0,67/5}; 13,90 + 2,132\sqrt{0,67/5}] =$$
$$= [13,90; 14,71]$$

12.3 Teste para Média com Variância Desconhecida

Caso não Normal

Neste caso é necessário utilizar técnicas não-paramétricas para realizar o teste da média;

Uma abordagem é considerar uma amostra suficientemente grande, usando o Teorema Central do Limite.

Aula 12

Professores:

Otton Teixeira da Silveira Filho
Regina Célia Paula Leal Toledo

Inferência Estatística - Testes de Hipóteses

Conteúdo:

- 12.1 Introdução
- 12.2 Teste para a Média Populacional
- 12.3 Teste para a Média com Variância Desconhecida

$$\sum_{i=1}^{n}(X_i - \bar{X})^2$$

$$\sigma_x^2 \qquad \bar{Y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$A \cap B = \varnothing$$