Universidade Federal de Juiz de Fora
Departamento de Estatística

Elementos de Estatística

Prof. Marcel de Toledo Vieira, PhD

Prof. Lupércio França Bessegato, DSc

Testes de Hipóteses

Marcel de Toledo Vieira

Introdução

- **Motivação** (exemplo):
  - Suponha que tenhamos sido informados que o rendimento médio mensal dos alunos de graduação bolsistas da UFJF seja de R$ 325,00.
  - Se selecionarmos uma amostra aleatória de 25 alunos de graduação do curso de Geografia e bolsistas da UFJF e se o rendimento médio desta amostra for R$ 330,00, você acha que este valor seria compatível com a média populacional (geral) informada acima?
  - E se a média amostral calculada fosse R$ 350,00 ou R$ 400,00?
- O quanto distante a média amostral precisa estar de R$ 325,00 para podermos concluir que a média populacional do rendimento médio mensal dos alunos de graduação bolsistas da UFJF é igual a um outro valor?

Introdução

- Motivação (**continuação**):
  - Nosso objetivo é então tirar alguma conclusão sobre um parâmetro da população (μ) **a partir de** uma amostra!
  - Uma alternativa é a construção de um **intervalo de confiança** para μ (assunto para outro curso!).
  - Discutiremos hoje uma outra opção – **Testes de Hipóteses**!
  - Na condução destes testes define-se inicialmente uma **hipótese** de que a média da população é igual a algum valor estabelecido (μ_0).
  - Esta hipótese recebe o nome de: **Hipótese Nula** (H_0).

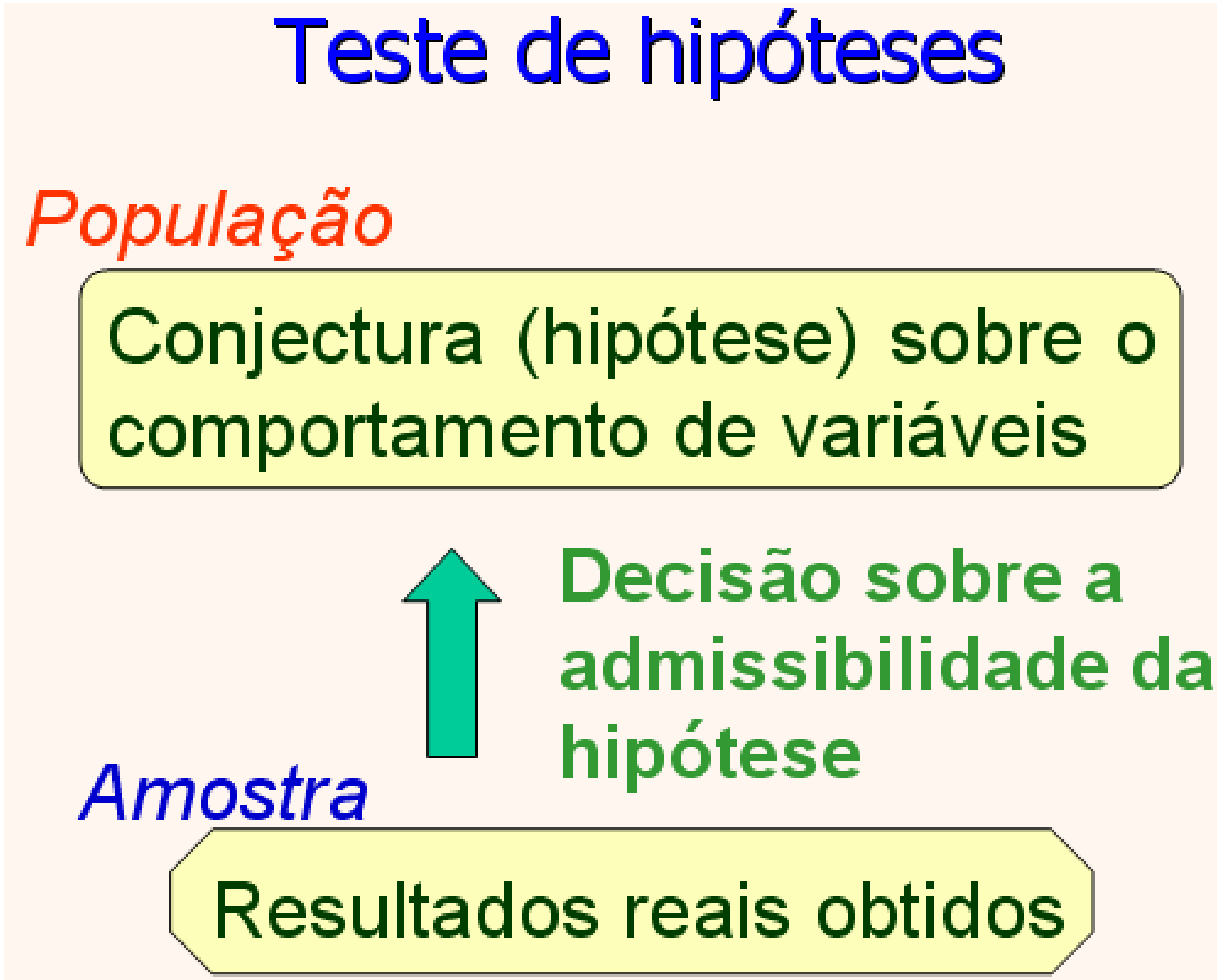
Introdução

- Voltando ao **exemplo**:
  - Podemos testar se o rendimento médio mensal dos alunos de graduação do curso de Geografia e bolsistas da UFJF é igual à média da população geral de alunos bolsistas da UFJF:

$$H_0 : \mu = \mu_0 = 325$$

$$H_1 : \mu \neq 325$$

  - Onde H_1 é a hipótese alternativa – uma segunda afirmação que contradiz H_0.
- Depois de definirmos as hipóteses, extraímos uma amostra aleatória da população de interesse, e comparamos a média dessa amostra ($\bar{x}$) com μ_0.


Teste de hipóteses
População
Conjectura (hipótese) sobre o comportamento de variáveis
Decisão sobre a admissibilidade da hipótese
Amostra
Resultados reais obtidos

Testes Significativos

- Se a diferença entre a média da amostra e a média suposta for muito grande temos evidência de que a amostra **não** pode vir de uma população com média μ_0, e assim rejeitamos H_0.
- Quando H_0 é verdadeira, isto ocorre quando a probabilidade de termos uma média da amostra tão distante quanto o $\bar{x}$ observado seja suficientemente **pequena**.
- Nestas situações pode-se dizer que o teste é estatísticamente **significativo**!

Testes Não-significativos

- Quando não há evidências suficientes para duvidarmos de H_0, a mesma não é rejeitada.
- Nestes casos admitimos que a média da população **pode** ser igual a μ_0 .
- **Atenção!** Nesta situação não podemos afirmar que aceitamos H_0. É possível que a média da população seja um valor diferente de μ_0 ...

Nível de Significância (α)

- Afirmamos acima que estaremos rejeitando H_0 quando a **probabilidade** de observarmos uma média tão ou mais extrema do que o observado for 'suficientemente pequena'.
- Mas o quanto **pequena** esta probabilidade deve ser...?
- Na **maioria** das aplicações este valor é de 0,05.
  - Nestes casos, rejeitamos H_0 quando a probabilidade de que a amostra poderia ter vindo de uma de uma população com média μ_0 for **menor ou igual a 5%**.
  - Isto significa que rejeitamos H_0 **erroneamente** 5% das vezes!

Nível de Significância (α)

- Podemos ser mais conservadores adotando uma probabilidade 0,01 **ou** menos conservadores ao adotarmos 0,10.
- Esta probabilidade é conhecida como **nível de significância** e é normalmente representada por α .

Valor *p* do Teste

- A probabilidade de obtermos uma média tão ou mais extrema do que a média da amostra observada $\bar{x}$, dado que H_0 é verdadeira recebe o nome de **valor *p*** do teste (p-*value*).
  - quando $p \leq \alpha$ rejeitamos H_0;
  - quando $p > \alpha$ não rejeitamos H_0.
- Na **literatura** da área de ciências sociais aplicadas, em geral, muito frequentemente encontramos o valor p dos testes realizados!

Teste-z

- Suponhamos que a variável aleatória contínua X tem média μ_0 e desvio-padrão σ **conhecido**. Assim,

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

tem distribuição normal padrão se n for suficientemente grande, e recebe o nome de **estatística do teste**.

  - Pode-se utilizar uma tabela da normal **ou** um pacote estatístico computacional para determinarmos a probabilidade de obtermos um valor de Z tão extremo ou mais do que o observado.

Passo a Passo...

1. Defina as hipóteses nula (H_0) e alternativa (H_1). O teste deve ser realizado com o objetivo de avaliar o quão forte é a evidência contra H_0.

2. Defina o valor de α .

3. Calcule o valor da estatística do teste.

4. Encontre o valor de p.

Teste-z – Exemplo

- Alguns estudos indicam que a média (μ_0) da pressão arterial sistólica para a população em geral de meia idade é de 128 mmHg e que o desvio padrão (σ) desta medida é de 15 mmHg.
- Será que homens de meia idade, que trabalham como executivos, em grandes empresas, possuem pressão arterial média diferente da média da população em geral na mesma faixa etária?
- Uma pesquisa foi realizada com o objetivo de dar uma resposta a esta pergunta.
- Foi selecionada uma amostra aleatória simples de 72 executivos de meia idade e encontrou que a média ($\bar{x}$) da pressão arterial sistólica era de 126,07. Será que temos evidências suficientes de que esta média é diferente o suficiente da média populacional?

Teste-z – Exemplo (Solução)

- Podemos definir inicialmente as hipóteses nula e alternativa:

$$H_0 : \mu_0 = 128$$

$$H_1 : \mu \neq 128$$

- Se trabalharmos com o pressuposto de que o desvio padrão da pressão arterial sistólica na amostra de executivos possui o mesmo valor do desvio padrão populacional, podemos calcular a estatística de teste:

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{126.07 - 128}{15 / \sqrt{72}} = -1{,}09$$

- Podemos agora calcular o valor de p:

$$[P(Z \leq -1{,}09) + P(Z \geq 1{,}09)] = 2 \cdot P(Z \geq 1{,}09) = 0{,}2758$$ **(Tabela da Normal)**

Teste-z – Exemplo (Solução)

GRÁFICO NO QUADRO

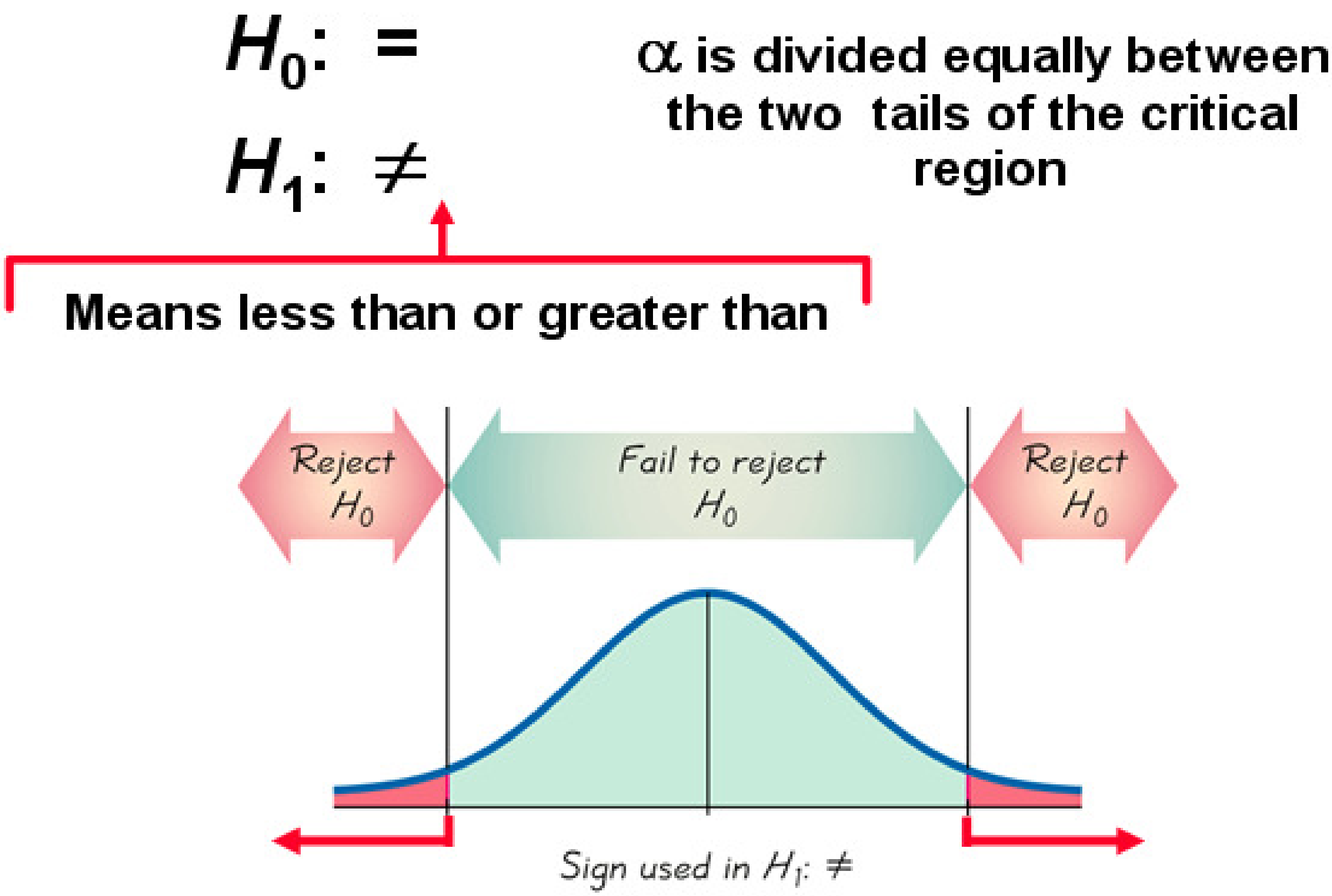
Teste-z – Exemplo (Solução)

- O valor de p calculado indica que não temos evidências suficientes de que a média da pressão arterial sistólica dos executivos seja diferente da média da população geral!

- Esta afirmativa é válida para níveis de significância de 1%, 5% ou 10%.

- Sendo assim, não temos evidências suficientes para rejeitar H_0.

Testes de Hipóteses Bilaterais



(Triola, 2004, *Elementary Statistics*)

Testes de Hipóteses Bilaterais

- Observem que **qualquer** valor da estatística de z que estiver entre -1,96 e 1,96 resultaria em um valor de p maior que 0,05.
- Alternativamente H_0 seria rejeitada para $z < -1{,}96$ ou $z > 1{,}96$. é **desconhecido** podemos substituí-lo por s, o desvio padrão amostral.
- Os valores -1,96 e 1,96 recebem o nome de **valores críticos** da estatística do teste.

Testes de Hipóteses Unilaterais

- É importante decidirmos se estamos interessados em investigar desvios de μ_0 que poderiam ocorrer em ambas as direções (teste bilateral) **ou** somente em uma direção (teste unilateral).
- Esta decisão precisa ser tomada **antes** do início do estudo!
- Se conhecimento prévio indicar que μ não pode ser menor que μ_0:

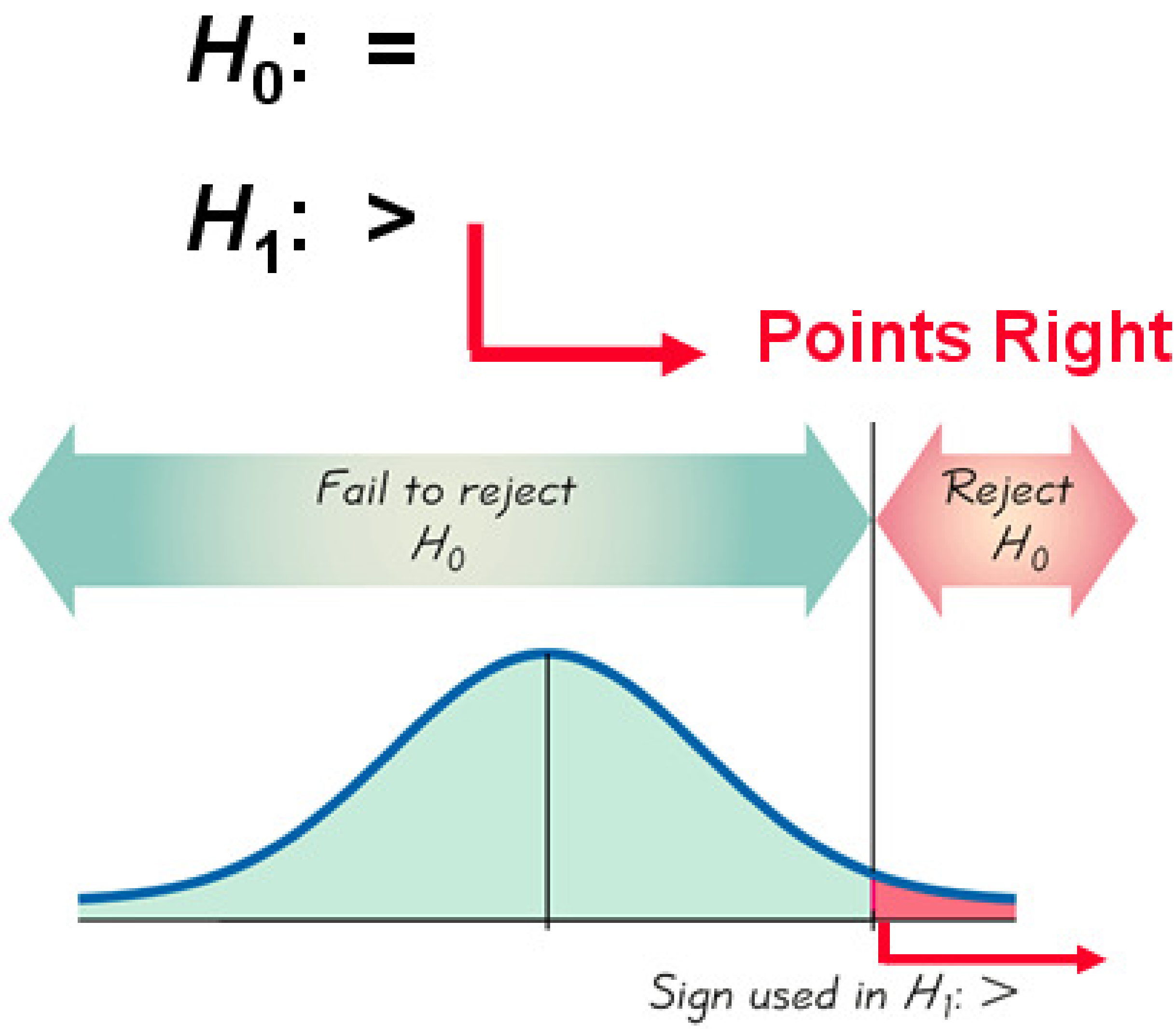
$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \geq \mu_0$$

- Ou alternativamente:

$$H_0 : \mu = \mu_0$$

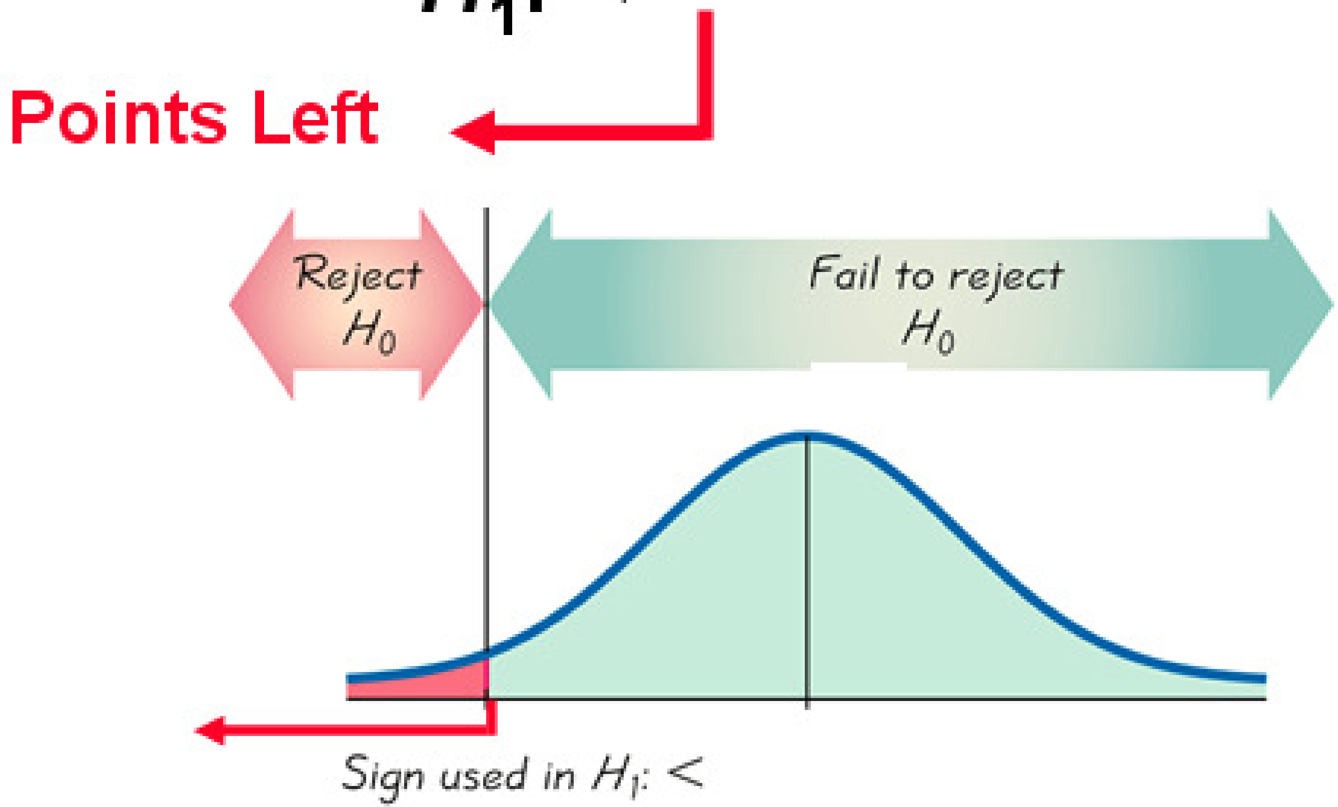
$$H_1 : \mu \leq \mu_0$$



(Triola, 2004, *Elementary Statistics*)

$$H_0: =$$

$$H_1: <$$



(Triola, 2004, *Elementary Statistics*)

Teste-z – Outro Exemplo

- Alguns estudos indicam que a média populacional (μ_0) das notas em língua portuguesa em uma prova padronizada de alunos do terceiro ano do ensino médio é 450 com desvio padrão (σ) 100.
- Será que a nota média em língua portuguesa de alunos do terceiro ano do ensino médio do Colégio João XXIII e do IFET-JF é maior do que a da população em geral?

Com o objetivo de responder a esta pergunta com realizado um estudo com uma amostra aleatória de 500 alunos e a média amostral ($\bar{x}$) encontrada foi de 461.

Teste-z – Outro Exemplo (Solução)

- Podemos definir inicialmente as hipóteses nula e alternativa:

$$H_0 : \mu_0 = 450$$

$$H_1 : \mu > 450$$

- Podemos calcular a estatística de teste:

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = \frac{461 - 450}{100 / \sqrt{500}} = 2{,}46$$

- Podemos agora calcular o valor de p:

$$P(Z \geq 2{,}46)] = 0{,}0069$$ **(Tabela da Normal)**

Teste-z – Outro Exemplo (Solução)

GRÁFICO NO QUADRO

Teste-z – Outro Exemplo (Solução)

- O valor de p calculado indica que temos evidências suficientes de que a média das notas dos alunos do Colégio João XXIII e do IFET-JF é maior do que a da população em geral!

- Esta afirmativa é válida para níveis de significância de 1%, 5% ou 10%.

- Sendo assim, temos evidências suficientes para rejeitar H_0.

Testes de Hipóteses Unilaterais *x* Bilaterais

- **Exemplo**: todos nós concordaríamos que a exposição a uma substância tóxica como monóxido de carbono ou dióxido de enxofre, poderia ser benéfica ao ser humano.
- Nesta situação ao considerarmos apenas os efeitos nocivos consideraríamos um teste **unilateral**!
- Na prática há muita discussão sobre a escolha entre testes unilaterais e bilaterais.
- Com freqüência em uma mesma situação o teste unilateral é significativo **enquanto** o bilateral não é...

EXERCÍCIOS...