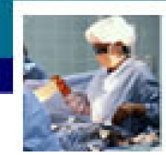
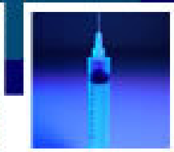
Bioestatística
Aula 7
Teoria dos Teste de Hitóteses

Prof. Tiago A. E. Ferreira

Hipóteses





- Hipótese
  - É uma **pressuposição** de um determinado problema.
    - Uma vez formulada, a hipótese estará sujeita a uma **comprovação**
    - O mecanismo de comprovação de uma hipótese é denominado **teste de hipótese**.
      - Este teste irá verificar se o pressuposto é verdadeiro ou não.
    - A verificação ou não do pressuposto é chamada de **conclusão**

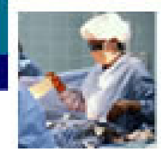
Testando uma Hipótese

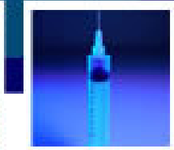
- Imagine que estamos interessados em responder a questão:
  - Níveis elevados de bilirrubina em recém-nascidos afetam a capacidade auditiva?
    - Dado uma amostra de n recém-nascidos, construímos dois grupos:
      - A = taxa de bilirrubina normal
      - B = taxa de bilirrubina elevada
    - Gera-se a **hipótese de nulidade (ou nula) H_0**:

$$H_0 \rightarrow \mu_A = \mu_B$$

      - μ_A = média da capacidade auditiva do grupo A
      - μ_B = média da capacidade auditiva do grupo B

Testando Hipótese



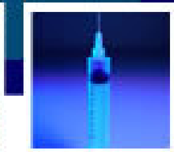


- Desta forma, a hipótese nula está testando:
  - Se as médias auditivas dos dois grupos são iguais!
    - **Aceitar** H_0 significa afirmar que os níveis de bilirrubinas não afetam a audição!
    - **Rejeitar** H_0 significa comprovar que os níveis de bilirrubina afetam a capacidade auditiva!
- A hipótese oposta à H0 é chamada de **Hipóteses Alternativa (H_1),** que para o exemplo:

$$H_1 \rightarrow \mu_A \neq \mu_B$$

Testando Hipótese

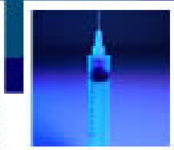




- Desta forma,
  - Aceitando-se H_0, automaticamente rejeita-se H_1
  - Rejeitando-se H_0, automaticamente aceita-se H_1
- Observe ainda que,
  - As hipóteses são **binárias e excludentes**!

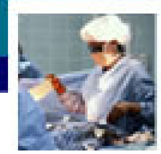
Regra de Decisão

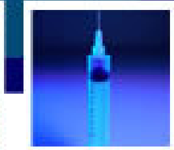




- Para se decidir se aceitamos ou se rejeitamos uma hipótese, algum critério deve ser utilizado.
  - Este critério chama-se de **regra de decisão**
    - Tal regra de decisão deve ser montada a partir de resultados ou conhecimentos prévios
    - Deve-se também permitir a redução de erros.

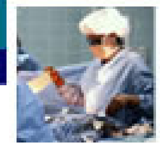
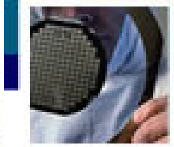
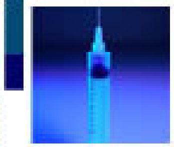
Erros de Decisão





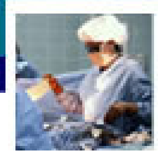
- Toda regra de decisão está sujeita a erro!
  - Dado que uma certa característica de uma população tem uma média de 20%.
    - Quando realizamos uma experiência com uma amostra, esperamos encontrar um valor para esta característica em torno dos 20%.
      - Se encontramos 22%, os 2% de desvio deve ser de erro caudado pela amostragem?
      - E se encontramos 32%?
  - A questão é: quando podemos afirma que o desvio encontra trata-se de um efeito amostral ou um real modificação do valor da média devido algum efeito local?

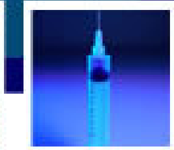
Erros de Decisão

- Desta forma, pode-se ter dois erros:
  - Aceitar que a média medida é igual a média populacional, quando não é!
  - Rejeitar que a média medida é igual a média populacional, quando é!
- Ou seja, chamamos estes erros de:
  - **Erro Tipo I**: rejeitar H_0 quando é verdadeira
  - **Erro Tipo II:** aceitar H_0 quando é falsa
- Assim, o Erro Global (EG) é a composição dos erros I e II

Níveis de Significância (α)

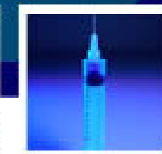




- Dado um certo um teste de hipótese, qual a probabilidade de termos um Erro Tipo I?
  - A esta probabilidade chamamos **Nível de Significância.**
    - Assim, se for definido um nível de significância de 5%, implica em afirmar que a decisão tomada aceitará um erro do tipo I de no máximo de 5%

Níveis de Significância (α)





- Dado que a proporção média da população com mais de 60 anos afetada pela doença de Alzheimer seja de 20% com $\sigma = 2.8284$.
  - Em um grupo de 50 pessoas, quais os limites de aceitação da hipótese nula ao nível de significância de 5%?
    - Supondo normalidade e distribuindo o erro a ambos os lados da normal (2,5% para cada lado)

$$z = \frac{x-\mu}{\sigma} = \mp 1{,}96 \therefore \frac{x-10}{2{,}8284} = \mp 1{,}96 \therefore \begin{cases} x_1 \cong 4{,}46 \\ x_2 \cong 15{,}54 \end{cases}$$

Níveis de Significância (α)

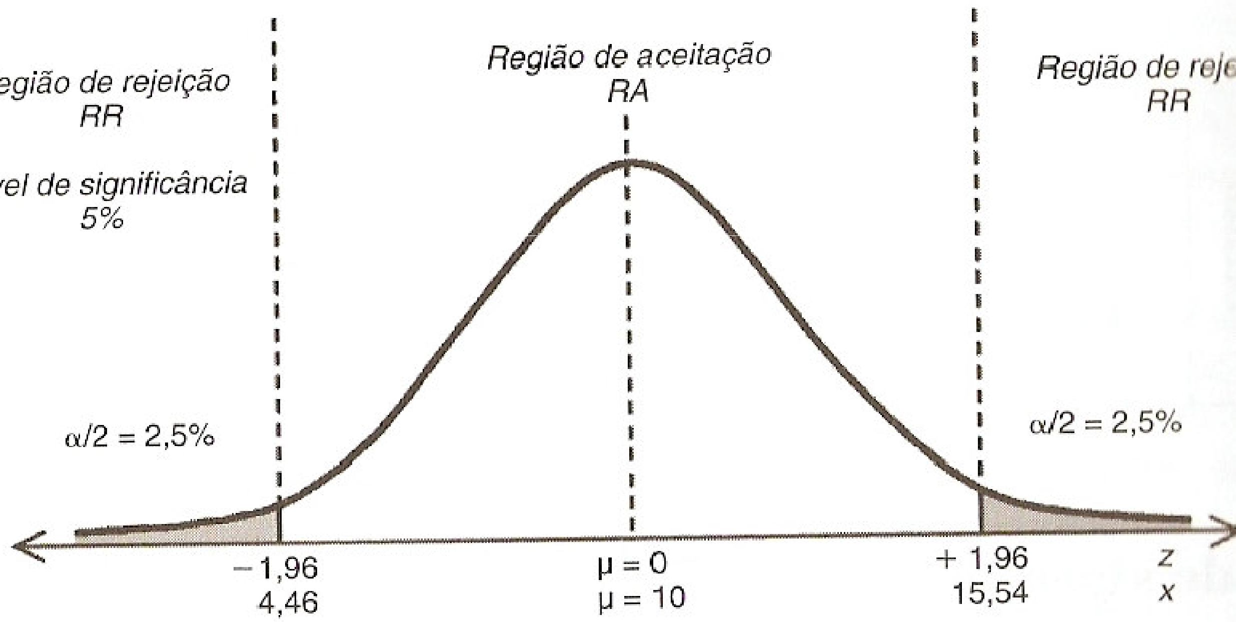
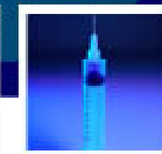


Fig. 8.3 Limites de decisão para diferenças (teste bilateral), a um nível de significância de 5%.

Níveis de Significância de um Teste (p)



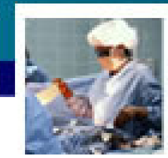


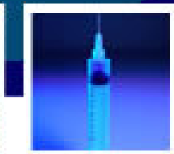
- A probabilidade de se rejeitar H_0 é chamada de **nível de significância o teste**, ou simplesmente ***p*.**
  - Dado o exemplo anterior, em 50 indivíduos foram encontrados 16 portadores de Alzheimer, assim:

$$z = \frac{x-\mu}{\sigma} = \frac{16-10}{2{,}8284} \cong 2{,}12$$

    - O valor da área da normal de $-\infty$ até z é de 0,98299, deixando uma área a direita de 0,017.
    - Como o erro está distribuído em ambos os lados, então p=0,034 ou p = 3,4%

Níveis de Significância de um Teste (p)





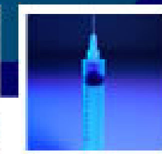
- Seguindo o quadro:

Nível de Significância	Conclusão
Menor que 1%	Diferença altamente significante
Entre 1 e 5%	Diferença significante
Entre 5 e 10%	Diferença provavelmente significante
Maior que 10%	Diferença não significante

– O valor de p está entre 1% e 5%, o que é possível a afirma que a diferença é significante!

Níveis de Significância de um Teste (p)



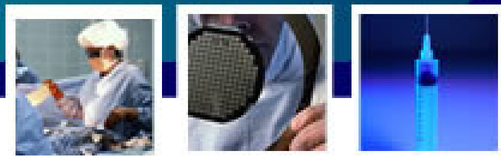


- Assim, é possível construir uma regra simples:
  - Quanto menor *p*, maior a evidência de existam diferenças, então rejeitas-se H_0 com maior certeza;
  - Quanto maior *p*, maior a evidência de que não existam diferenças, então diminui a certeza de rejeição de H_0.

Testes Unilaterais e Bilaterais

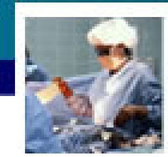
- No exemplo anterior (Alzheimer) aplicou-se um teste bilateral (ou bicaudal)
  - Visto que estávamos verificando se duas grandezas diferiam uma da outra
    - Esta diferença tanto pode ser positiva (ou para mais) como pode ser negativa (ou para menos)!
- Entretanto o problema poderia ter sido formulado de maneira diferente:
  - Dado que o valor indicado é 10 e o valor medido foi 16, o pesquisador poderia ter formulado:
    - A taxa medida não é maior que a taxa indicada?

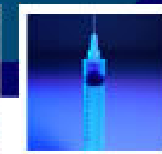
Testes Unilaterais e Bilaterais



- Para esta segundo formulação, aplica-se um teste **unilateral** (ou **unicaudal**)
  - Visto que para este caso o erro para menos não tem sentido!
- Assim, teste bilaterais
  - O erro é dividido nas extremidades da distribuição.
- Testes unilaterais
  - O erro é concentrado em um dos erros da distribuição.

Teste Bilateral (Teste para Diferenças)



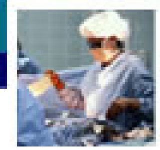
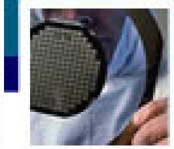
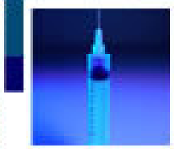


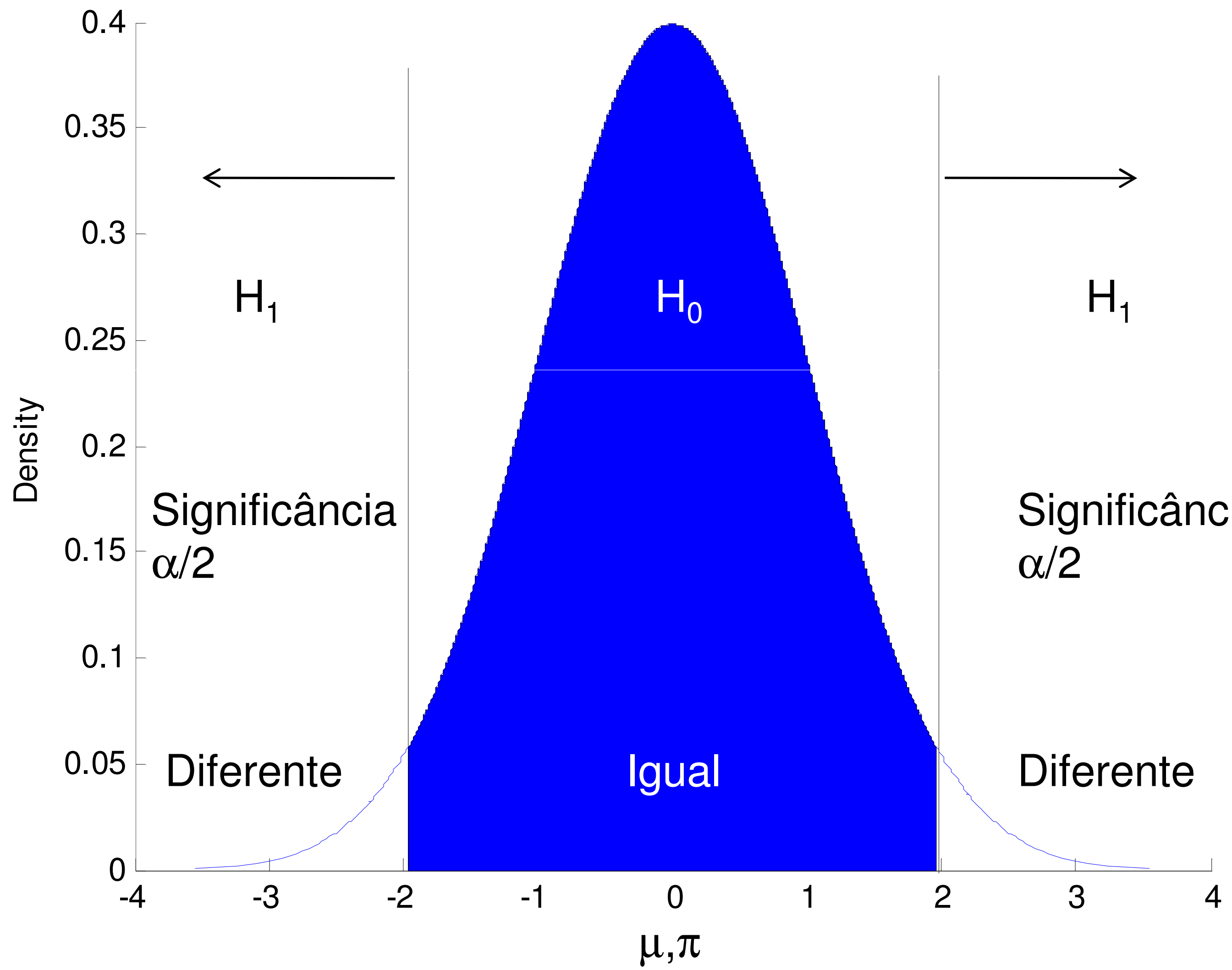
- Este tipo de teste verifica se uma estatística experimental (dados amostrais) dada por ($\bar{x}$,p) é diferente de uma estatística populacional (μ,π)
- Hipóteses:

$$H_0 \rightarrow \bar{x} = \mu \quad ou \quad p = \pi$$

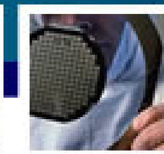
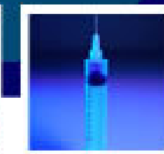
$$H_1 \rightarrow \bar{x} \neq \mu \quad ou \quad p \neq \pi$$

Graficamente



Teste Unilateral

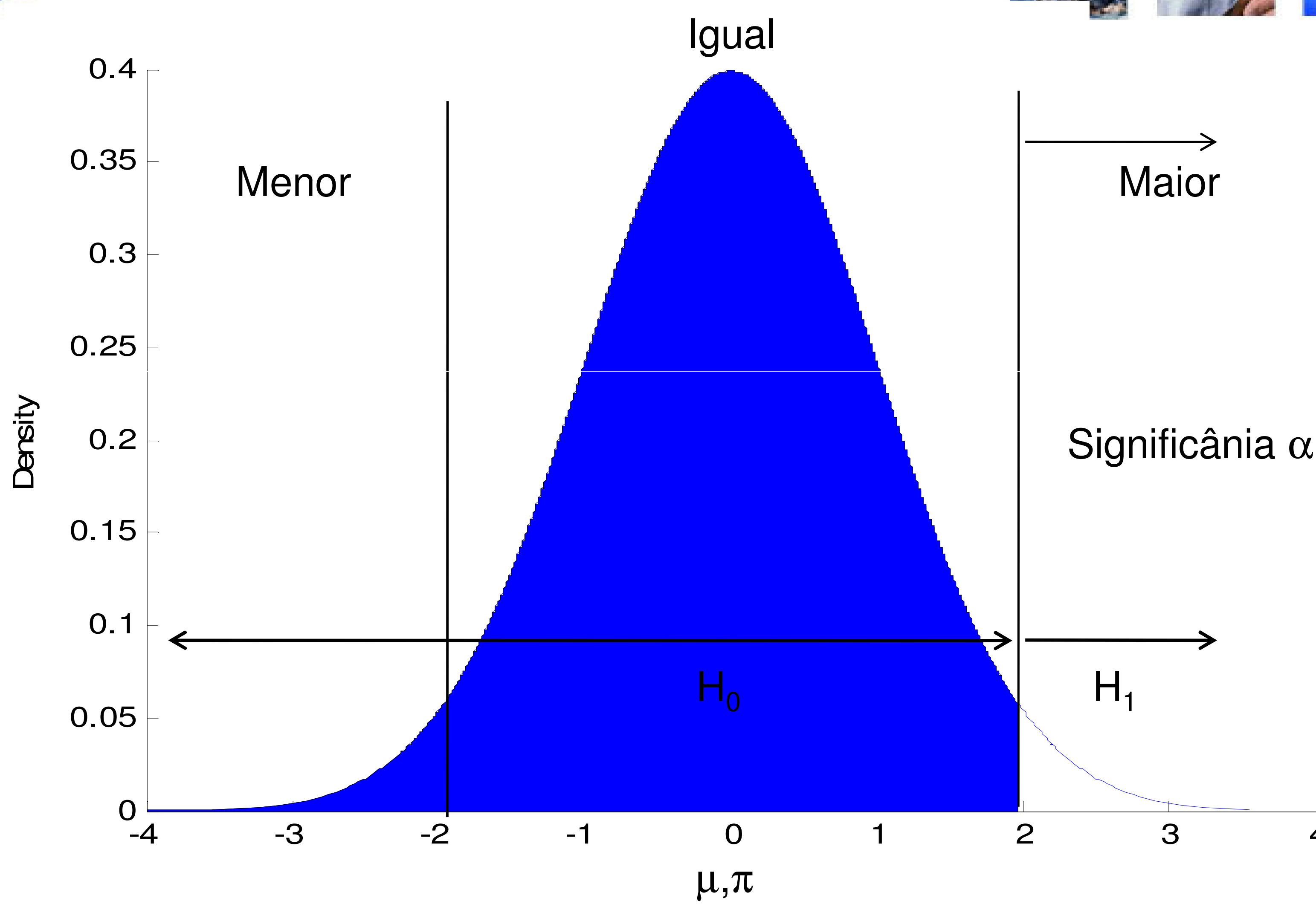




- Ao verificar se uma dada estatística é maior que outra!
- Hipóteses:

$$H_0 \rightarrow \bar{x} = \mu \; ou \;\; p = \pi \; \left(ou \; \bar{x} \leq \mu \, , \, p \leq \pi \right)$$

$$H_1 \rightarrow \bar{x} > \mu \; ou \;\; p > \pi$$

Graficamente



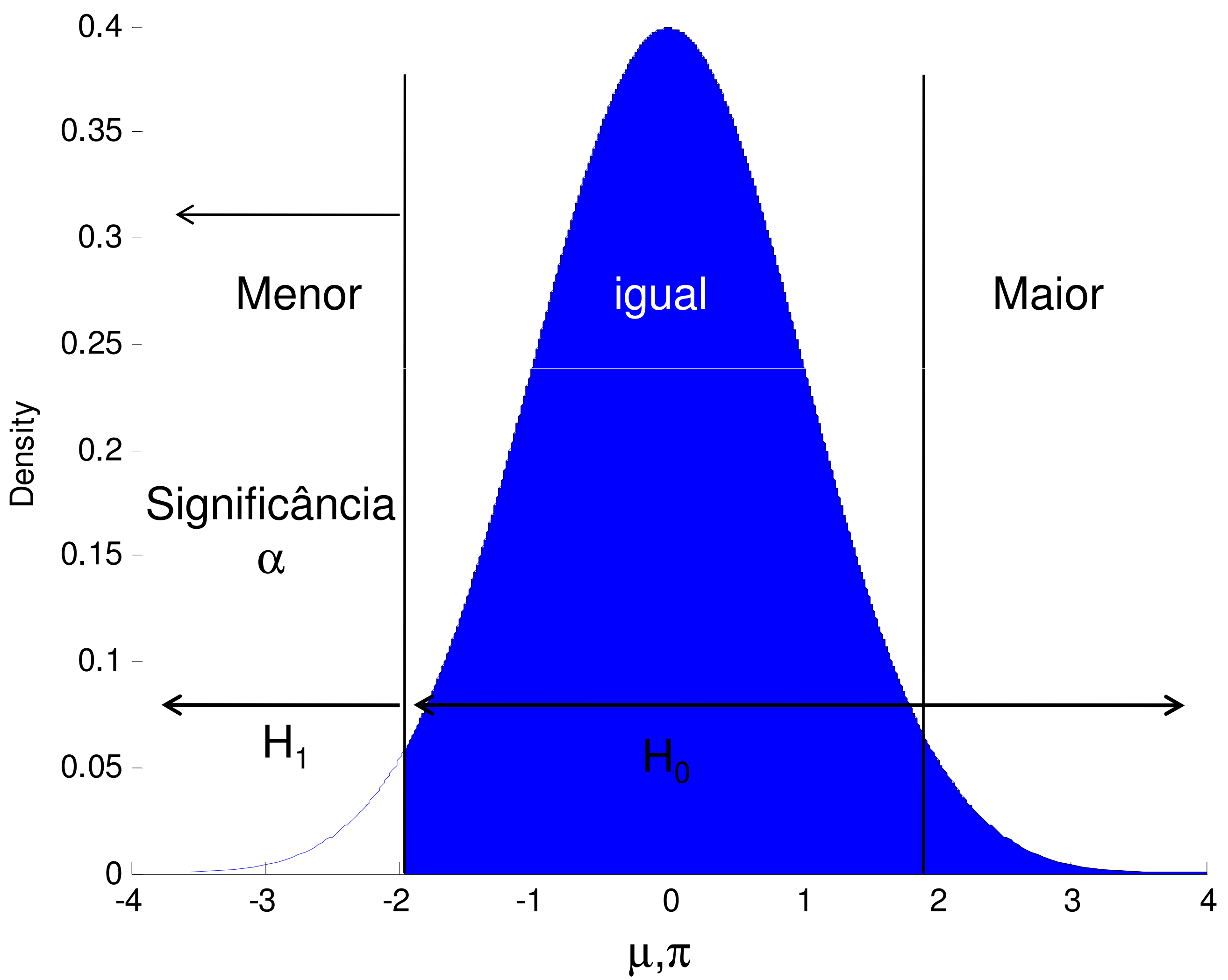
Teste Unilateral

- Ao verificar se uma dada estatística é menor que outra!
- Hipóteses:

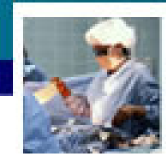
$$H_0 \rightarrow \bar{x} = \mu \; ou \;\; p = \pi \; \left(ou \; \bar{x} \geq \mu \, , \, p \geq \pi \right)$$

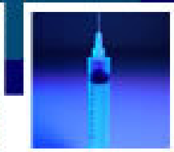
$$H_1 \rightarrow \bar{x} < \mu \; ou \;\; p < \pi$$

Graficamente



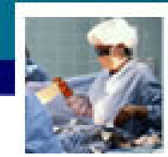
Exemplo

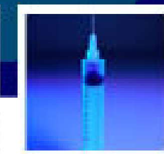




- Suponha que um estudo mostra que a ingestão média diária de calorias em adultos é de 2400Kcal. Considere um que grupo de 25 adultos desta população apresentou um consumo de 3000Kcal, com um desvio padrão de 1250Kcal.
  – Para testar se o consumo calórico deste grupo é diferente do consumo populacional, pode ser efetivado um teste normal para as médias!

Exemplo



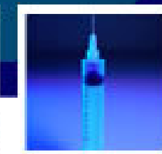


- Inicialmente,defini-se as hipótese:
  - $H_0 \rightarrow \bar{x} = \mu$
  - $H_1 \rightarrow \bar{x} \neq \mu$
    - Este teste é bilateral! Assim, devemos procurar na tabela o valor $\alpha/2$
    - Observando a tabela da normal:

α(%)	Z_t
10	1,65
5	1,96
1	2,58

Exemplo





- A decisão para aceitação ou rejeição de H_0 é:
  - $Z_c < Z_t$, então **aceita-se H_0**
  - $Z_c \geq Z_t$, então **rejeita-se H_0**

  - Assim,

$$\mu_{DAM} = \mu = 2400Kcal$$

$$\sigma_{DAM} = \frac{\sigma}{\sqrt{n}} = \frac{1250}{\sqrt{5}} = 250$$

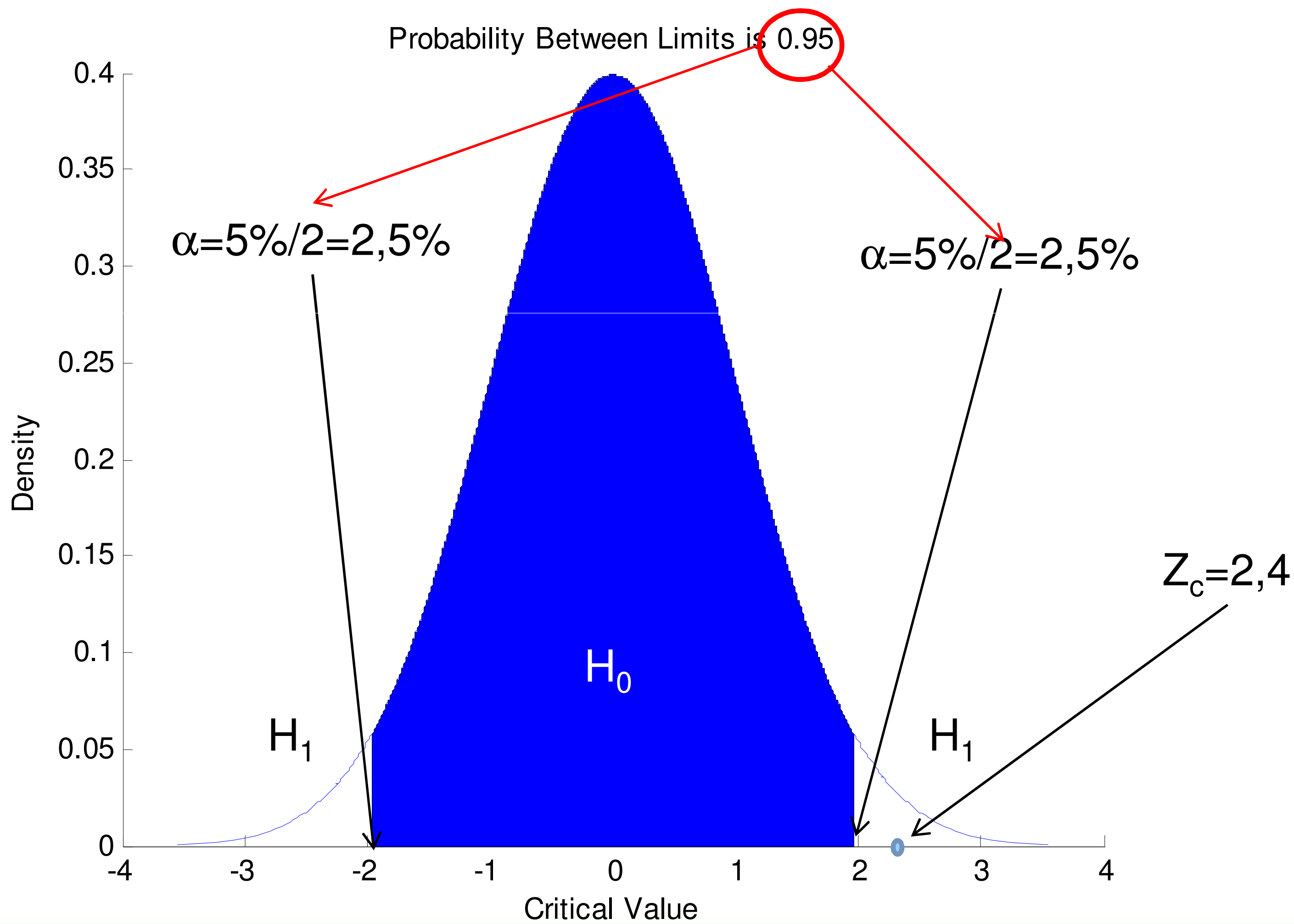
Exemplo

- Então, como x = 3000Kcal

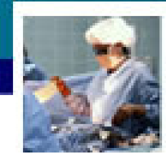
$$z_c = \frac{3000 - 2400}{250} = 2{,}4$$

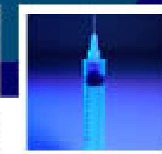
  – Assim, com um nível de significância em 5% (α=0,05), a decisão estatística seria rejeitar H_0, pois z_c=2,4 > z_t(5%) = 1,96.
    - Conclui-se que o padrão de consumo de calorias do grupo (amostra) é diferente do normal (população).

Exemplo - graficamente



Exemplo





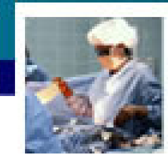
- Outra forma de abordar o problema é calcular o valor máximo da v.a. padronizada.
  - Assim, considerando um nível de significância de 5% para o teste bilateral, z=+1,96, então:

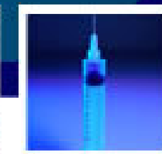
$$z = \frac{x-\mu}{\sigma} \therefore +1{,}96 = \frac{x-2400}{250}$$

$$x = 1{,}96 \cdot 250 + 2400 = 2890$$

  - O valor 2890 é o limite superior para a aceitação de H_0

Exemplo

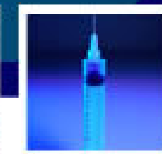




- Uma terceira forma de decidir sobre H_0 é determinar o nível de significância do teste (p)
  - Como já visto 3000Kcal equivale a z=2,4
  - z=2,4 implica em uma área de 0,99180
  - Assim, a significância para o teste bilateral será:
    - p=(1-0,99180).2=0,0164 ou 1,64%
    - Este resultado pode ser interpretado como:
      - A probabilidade de rejeitar H_0 estando ela correta (erro tipo I) é de 1,64%. Desta forma, a este nível H_0 pode ser rejeitada!

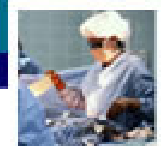
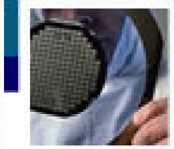
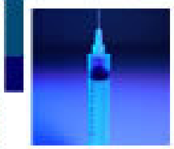
Exemplo





- O registro de vacinação de uma localidade informou que na última campanha 10% da população deixou de ser imunizada. Entretanto, em uma amostra de 130 pessoas foram detectados 18 casos de não vacinação.
  - Para se testar, com um nível de significância de 5%, se a proporção de indivíduos não imunizados na amostra é maior que a proporção verificada na população deve ser realizado um teste unilateral normal para proporções

Exemplo

- Hipóteses:

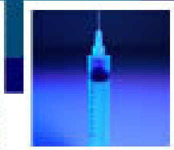
$$H_0 \rightarrow p = \pi \left(ou \;\; p \leq \pi \right)$$
$$H_1 \rightarrow p > \pi$$

- Onde:
  - p = proporção amostral = 18/130 = 0,1385 ou 13,85%
  - π = proporção populacional = 10%
- Para o caso seria esperado uma média verdadeira para a amostra de
  - μ = 130 . 0,10 = 13 pessoas
  - E o desvio padrão:

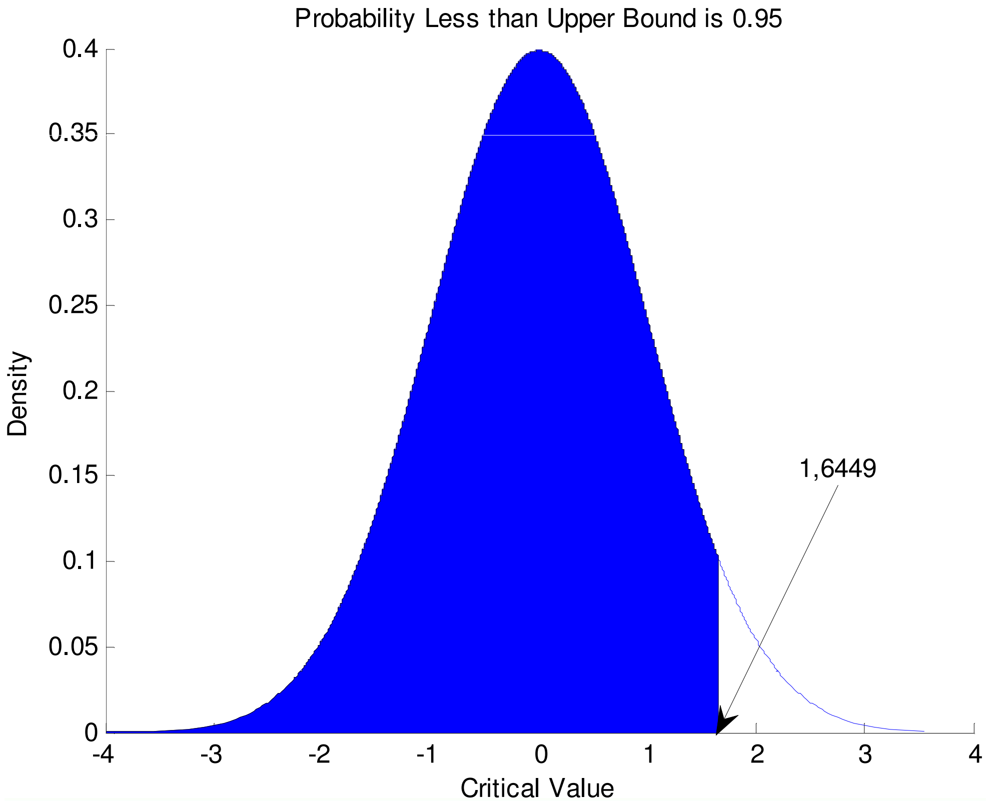
$$\sigma = \sqrt{npq} = \sqrt{130 \cdot 0{,}10 \cdot 0{,}90} = 3{,}42$$

Exemplo





- Então: $$z = \frac{x - \mu}{\sigma} = \frac{18 - 13}{3{,}42} = +1{,}46$$



Aceita-se H_0

Poder de um Teste

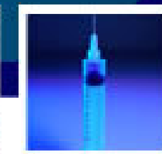
- Imagine agora que sabemos que a média abdominal de adultos é de 80cm (literatura) e que em um grupo de 36 indivíduos encontramos uma média de 83,2cm com um desvio padrão de 8cm.
  - Pergunta-se, a medida experimental é igual a medida da literatura?
    - Hipóteses:

$$H_0 \rightarrow \mu_{\exp} = \mu_{lit} = 80cm$$

$$H_1 \rightarrow \mu_{\exp} \neq 80cm$$

Poder de um Teste





- O nível de significância do teste seria:

$$z_c = \frac{83{,}2 - 80}{8 / \sqrt{36}} = +2{,}4 \rightarrow \acute{A}rea = 0{,}99180$$

  – E como o teste é bilateral, p=0,0164 ou 1,64%
    - Esta é a probabilidade de se cometer o **erro tipo I**, o que permite decidir se H_0 deve ser aceita ou rejeitada.
      – **Observe que para α=5% H_0 seria rejeitada!**

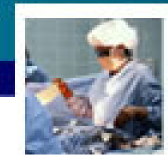
Poder de um Teste

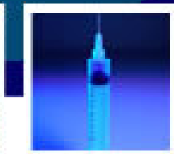
- Observe que o que foi testado foram as hipóteses: H_0 ser igual a 80cm contra H_1 ser diferente de 80cm!
  - Um vez que H1 defende que o diâmetro não é de 80cm, caberia a pergunta:
    - Qual é o diâmetro?
    - Se estabelecermos esta valor par 85 cm, por exemplo, poderíamos ter as hipóteses:

$$H_0 \to \mu_{\exp} = 80cm \left(\mu_{lit}\right)$$

$$H_1 \to \mu_{\exp} = 85cm$$

Poder de um Teste

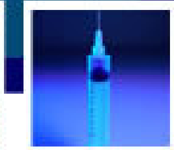




- Observe que aceitar o valor 80cm significa negar o valor 85cm, e vice-versa!
  - Neste caso, a probabilidade de rejeitar H_0 quando ela é verdadeira (erro tipo I) é o nível de significância (α).
  - Mas, a probabilidade de rejeitar H1quando ela é verdadeira é a probabilidade do erro tipo II (β), e o sei complemento ($1- \beta$) é o que se conhece como **poder (ou potência) de um teste.**

Poder de um Teste





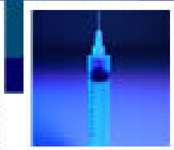
- Assim, o **poder de um teste** é a probabilidade de rejeitar H_0 quando ela é falsa!
  - Assim, no exemplo dos diâmetros abdominais
    - O valor de para um teste unilateral é p=0.0082 ou 0,82%
    - A probabilidade do erro tipo II

$$z_c \frac{83,2-85}{8/\sqrt{36}} = +1,35 \; ou \; Área = 0,9115$$

  - O que dá β=0,0885 ou 8,85%

Poder de um Teste

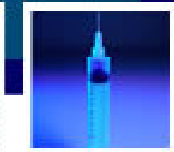




- Assim, o poder do teste é (1-β) = 0,9115 ou 91,15%
- Conclusão:
  - Pode ser rejeitada a hipótese de que o diâmetro abdominal da amostra é igual a 80cm (H_0) contra a hipótese de que ele é 85cm (H_1), para um nível de significância do teste $p < 0,01$ e um poder do teste (1-β) > 90%

Poder de um Teste

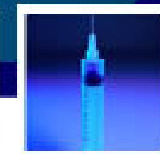




- Conclusão alternativa:
  - Uma probabilidade de rejeitar H_0 equivocadamente de menos de 1% e de aceitar H corretamente de mais de 90%.
    - Ou seja, uma probabilidade de errar na decisão de rejeitar μ=80cm de menos de 1% e de acertar na decisão de aceitar μ=85cm de mais de 90%.

Função de Poder

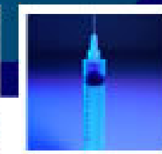




- É possível definir um conjunto de possíveis valores para a hipótese alternativa!
  - Calculando-se o poder do teste para cada um destes valores, gera-se uma função matemática conhecida como **função poder** ou **função de poder.**
    - Esta função irá mostrar os resultados da decisão estatística decorrentes de uma faixa de valores alternativos à hipótese nula!.

Referências





- Livro Texto: Bioestatística. Teoriaia e Computacional (Héctor G. Arango). Guanaba
  - Capítulo 8
- Leituras Complementares:
  - Livro: Bioestatística. Princípios e Aplicações (Sidia M. Callegari-Jaqcques)
    - Capítulo 6.
  - Livro: Estatística Básica (Wilton de O. Bussab e Pedro A. Morettin). Saraiva
    - Capítulo: 12
  - Livro: Probabilidade. Aplicações à Estatística (Paul L. Meyer). LTC
    - Capítulo 15