Engenharias

M109

Estatística

Aula 5 – Testes de Hipóteses

Créditos:
Profª. Rosimara Beatriz Arci Salgado
Profª. Karina Perez Mokarzel Carneiro
Prof. Rausley Adriano Amaral de Souza

Inatel | CAMINHOS QUE CONECTAM COM O FUTURO

Bibliografia básica:

Arango HG. Bioestatística: teórica e computacional. 3ªed. Rio de Janeiro: Guanabara Koogan; 2011.

SPIEGEL, Murray Ralph; FARIA, Alfredo Alves De Probabilidade e estatística. São Paulo, SP: McGraw-Hill, 1978.

RBAS1

TESTE DE HIPÓTESES

Inferência estatística é um ramo da Estatística cujo objetivo é fazer afirmações a partir de um conjunto de valores representativo (amostra) sobre um universo (população), assume-se que a população é muito maior do que o conjunto de dados observados, a amostra. Tal tipo de afirmação deve sempre vir acompanhada de uma medida de precisão sobre sua veracidade. Para realizar este trabalho, o estatístico coleta informações de dois tipos, experimentais (as amostras) e aquelas que obtém na literatura.

RBAS2

Teste de Hipóteses trata-se de uma técnica para se fazer a inferência estatística sobre uma população a partir de uma amostra

Source:https://pt.wikipedia.org/wiki/Infer%C3%AAncia_estat%C3%ADstica#:~:text=Infer%C3%AAncia%20estat%C3%ADstica%20%C3%A9%20um%20ramo,de%20dados%20observados%2C%20a%20amostra.

Slide 3

RBAS1 Veremos a definição do que é um teste de hipótestes, como ele deve ser feito e, obviamente, a importância de se fazer um teste de hipóteses.

Rosimara Beatriz Arci Salgado; 31/03/2022

RBAS2 Essa amostra é representativa da população. A partir das observações realizadas da amostra, vamos tirar conclusão para toda a população. Isso é inferência estatística, ou seja, concluir para uma população a partir de uma amostra. Em uma eleição, você deseja inferir qual vai ser o resultado final. Essa pesquisa fazemos apenas com uma amostra e não com toda a população. Para fazer a inferência estatística nós fazemos um teste de hipóteses. Construímos a estrutura de um teste de hipóteses, testamos uma hipótese e chegamos à conclusão de um determinado problema.

Rosimara Beatriz Arci Salgado; 31/03/2022

CONCEITOS BÁSICOS

RBAS3

Hipótese: é uma pressuposição a respeito de um determinado problema.

O mecanismo de comprovação (verificação) é denominado **teste de hipóteses**. Assim, testar uma hipótese quer dizer verificar se um pressuposto é verdadeiro ou não.

A veracidade ou não do pressuposto é chamada de **conclusão**.

RBAS3 Suposição antecipada, ou seja, você faz uma afirmação (pressuposição) a respeito de um problema que você quer estudar e testa se aquela pressuposição é verdadeira ou ela é falsa. Você tem algum problema e você quer responder alguma pergunta ou observação que foi feita, você determina/monta uma hipótese e verifica se essa hipotese é verdadeira ou ela é falsa. Você comprova que aquela hipótese é verdadeira ou você comprova que aquela hipótese é falsa a partir do seu problema, então, você faz o que nós chamamos de teste de hipóteses com base na hipótese montada. Após verificar a veracidade ou não da sua hipótese, você tira a sua conclusão a respeito de um determinado problema/observação que você fez. Então, basicamente, hipótese é esta pressuposição que você monta.

Rosimara Beatriz Arci Salgado; 31/03/2022

POR EXEMPLO: RBAS4

Níveis elevados de bilirrubina em recém-nascidos afetam a capacidade auditiva deles?

Para se chegar a uma conclusão sobre essa questão é necessário **formular uma hipótese e testá-la**.

A formulação da hipótese está relacionada com a forma de conduzir a experiência (**desenho experimental**).

Inatel CAMINHOS QUE CONECTAM COM O FUTURO

RBAS4 Alguém observou que alguns recém-nascidos estavam com problema de capacidade auditiva e, também, observou que, a princípio, parecia que existia uma co-relação entre aqueles recém-nascidos com problemas de capacidade auditiva e o nível elevado dessa substância, que é a bilirrubina, uma substância produzida na bílis, que causa icterícia em crianças, a criança fica "amarelinha". Essa substância produzida pelo corpo, que se chama bilirrubina e o médico (alguém) observou que parece que existe uma relação entre produção elevada de bilirrubina e a capacidade auditiva, agora eu preciso responder esta pergunta, que na verdade é provar se o nível elevado de bilirrubina afeta ou não a capacidade auditiva dos recém-nascidos que, obviamente, foi observado em recém-nascidos, ou seja, para responder a esta pergunta, para se ter uma conclusão a respeito dessa pergunta, nós vamos formular uma hipótese e testar essa hipótese. E a maneira até como eu faço a pergunta e formulo a hipótese vai determinar a forma como eu vou conduzir a minha experiência, o meu teste de fato. Tudo isso é o desenho experimental. Então, eu tenho um problema e agora eu vou formular uma hipótese para testar essa questão (nesse caso, foi uma pergunta) e avaliar esse problema, que é um problema recorrente.

Rosimara Beatriz Arci Salgado; 02/04/2022

TEORIA DOS TESTES DE HIPÓTESES

Neste exemplo, poderia ser selecionada uma amostra de *n* recém-nascidos e:

- Efetuar uma dosagem dos níveis de bilirrubina. RBAS5
- Aferição da capacidade auditiva (quantitativa).

A partir dos dados, seriam constituídos dois grupos:

- A = taxa de bilirrubina normal
- B = taxa de bilirrubina elevada

Seriam então comparadas as capacidades auditivas médias dos grupos A e B.

Inatel | CAMINHOS QUE CONECTAM COM O FUTURO

Slide 6

RBAS5 Então, como seria a questão do desenho experimental, como eu conduziria essa experiência de teste de hipótese? Eu vou selecionar alguns recém-nascidos com base em alguns critérios, que foram mencionados na 1ª aula: proporcionalidade, aleatoriedade... Então, faço duas ações:

1) Efetuo a dosagem de bilirrubina de todos os "n" recém-nascidos;

2) Faço aferição da capacidade auditiva de todos os "n" recém-nascidos.

A partir dessas duas ações, que vão direcionar à montagem da minha hipótese, eu vou pegar os "n" recém-nascidos e dividi-los em dois grupos: Grupo A (c/ taxa de bilirrubina normal) e Grupo B (c/ taxa de bilirrubina elevada), separação dos grupos tendo como critério a taxa de bilirrubina normal (padrão) informada na literatura. Baseado na aferição da capacidade auditiva, eu vou comparar esses dois grupos (com taxa normal de bilirrubina e com taxa elevada de bilirrubina) no que diz respeito à capacidade auditiva. Aí, então, eu estou formulando a(s) minha(s) hipóteses. Como seria isso? (ver próximo slide).

Rosimara Beatriz Arci Salgado; 02/04/2022

TEORIA DOS TESTES DE HIPÓTESES

Construção das Hipóteses

- A = taxa de bilirrubina normal
- B = taxa de bilirrubina elevada

Hipótese de Nulidade → H_o RBAS6 $H_o \rightarrow \mu_A = \mu_B$	**Igualdade entre os grupos**
Hipótese Alternativa → H_1 RBAS7 $H_1 \rightarrow \mu_A \neq \mu_B$	Diferença (não igualdade) entre os grupos

RBAS8

Média da capacidade auditiva do grupo A

Média da capacidade auditiva do grupo B

RBAS6 Por hipótese de nulidade (H0), a capacidade auditiva do grupo A (c/ taxa de bilirrubina normal) é igual à capacidade auditiva do grupo B (c/ taxa de bilirrubina elevada).
Rosimara Beatriz Arci Salgado; 02/04/2022

RBAS7 Por hipótese alternativa (H1), a capacidade auditiva do grupo com taxa de bilirrubina normal (Grupo A) é diferente da capacidade auditiva do grupo com taxa de bilirrubina elevada (Grupo B). Pode ser maior ou menor, independente disso, mas é diferente.
Rosimara Beatriz Arci Salgado; 02/04/2022

RBAS8 Então, aqui eu tenho uma situação binária. Quando eu aceito uma hipótese, consequentemente, a outra é falsa. Então, eu faço teste em uma das hipóteses, eu aceito essa hipótese como verdadeira e, consequentemente, a outra hipótese, eu rejeito. Por padrão, nós sempre testamos H0. Se eu aceitar essa hipótese via um teste estatístico, eu consequentemente, rejeito H1. Se eu rejeitar H0, consequentemente, eu aceito H1. O que significa aceitar H0 neste exemplo via teste estatístico? Quando eu aceito H0, eu estou aceitando/provando a igualdade entre os grupos A e B, ou seja, em termos de capacidade auditiva os dois grupos (taxa normal e taxa elevada de bilirrubina) não têm diferenças. A conclusão do experimento é que os dois grupos são iguais em termo de capacidade auditiva. A variação na taxa de bilirrubina não afeta a capacidade auditiva dos recém-nascidos. Por outro lado, se eu rejeitar H0, consequentemente, eu estou aceitando H1, como hipótese verdadeira, eu estou dizendo que em termos de capacidade auditiva os dois grupos são diferentes. Portanto, a taxa de bilirrubina influencia na capacidade auditiva. Em termos de capacidade auditiva, os grupos A e B são diferentes. Sempre eu testo uma hipótese, essa hipótese é sempre H0, eu vou aceitar ou rejeitar essa hipótese. Ao aceitar, tendo como foco o parâmetro da capacidade auditiva, os grupos são iguais. Se eu rejeito, eles são diferentes. Nesse caso, estou analisando apenas um aspecto: taxa elevada de bilirrubina, mas podem existir outros aspectos, que serão objetos de investigação. A observação/pergunta que foi feita foi relacionada entre a taxa de bilirrubina e a capacidade auditiva, mas nada impede de que outros aspectos sejam investigados em outros estudos, após a conclusão deste estudo.
Existe uma probabilidade de erro associada à minha conclusão, ou seja, ao se afirmar que H0 é verdadeira, existe uma probabilidade dessa afirmação estar errada. Isso temos que mensurar. É possível testar mais de 2 hipóteses, mas não é tão comum, mas é possível. O mais comum é testar 2 hipóteses e, depois de fazer o teste, avaliam-se outros aspectos e testam-se mais duas hipóteses sobre esses outros aspectos. A questão é avaliar um problema de cada vez. H0 é antagônico à H1, eu posso fazer igualdade em H0 e "midi" A > "midi" B ou o contrário, "midi" B > "midi A" em H1. Neste caso aqui, muito provavelmente, como estou estudando que a taxa de bilirrubina elevada afeta a capacidade auditiva, então quer dizer que, normalmente, a capacidade auditiva com taxa normal de bilirrubina vai ser maior que a capacidade auditiva do grupo com taxa elevada de bilirrubina. Porém, nesse caso do slide, eu estou testando a diferença.
Rosimara Beatriz Arci Salgado; 02/04/2022

RBAS10

TEORIA DOS TESTES DE HIPÓTESES

Assim:

Aceitar H_0 ou dizer que ela é verdadeira por meio de um teste significa afirmar que níveis de bilirrubina **não** estão relacionados com a perda da capacidade auditiva.

Rejeitar H_0 ou dizer que ela é falsa, implica comprovar que os níveis de bilirrubina afetam a capacidade auditiva.

Naturalmente, aceitar H_1 implica rejeitar H_0 e vice-versa.

Hipótese de Nulidade → H_o
$H_o \rightarrow \mu_A = \mu_B$
Hipótese Alternativa → H_1
$H_1 \rightarrow \mu_A \neq \mu_B$

Slide 8

RBAS10 Logicamente, existe uma probabilidade da minha conclusão estar errada e a gente precisa mensurar isso. Sempre a minha conclusão virá acompanhada de um erro. Não é possível concluir com 100% de probabilidade de acerto, existe um erro sempre.

Rosimara Beatriz Arci Salgado; 02/04/2022

RBAS9

TEORIA DOS TESTES DE HIPÓTESES

As hipóteses devem ser **binárias** e **excludentes**.

Permitem que apenas uma das hipóteses seja testada para se obter uma conclusão consistente.

Por convenção, testa-se sempre H_0. Dessa forma, aceitar H_0 implica comprovar a igualdade (nulidade das diferenças); rejeitar H_0 significa comprovar a diferença entre os grupos testados.

Slide 9

RBAS9 Binárias: hipóteses verdadeiras ou falsas.
Excludentes: se uma hipótese é verdadeira, consequentemente, a outra é falsa e vice-versa.
Rosimara Beatriz Arci Salgado; 02/04/2022

RBAS11

TEORIA DOS TESTES DE HIPÓTESES

Regra de decisão

É um procedimento estatístico para se decidir entre **aceitar ou rejeitar** a hipótese de nulidade (H_0).

Normalmente, formula-se a regra de decisão para testar uma hipótese a partir do resultado de um **teste estatístico**.

Ao se tomar uma decisão sobre se a hipótese de nulidade deve ser aceita ou não, existirá a possibilidade de se estar cometendo um **erro**. Por esse motivo, as regras de decisão são construídas seguindo **critérios** que permitam reduzir os erros a elas associados.

Inatel CAMINHOS QUE CONECTAM COM O FUTURO

Slide 10

RBAS11 Eu preciso implementar um teste estatístico (fazer cálculos) para testar as hipóteses, para verificar se H0 é verdadeiro ou falso, que é a regra de decisão.

Rosimara Beatriz Arci Salgado; 02/04/2022

TEORIA DOS TESTES DE HIPÓTESES

Erros de decisão

RBAS12

Erro tipo I → EI	**Rejeitar H_0 quando ela é verdadeira**
Erro tipo II → EII	**Aceitar H_0 quando ela é falsa** RBAS13

Hipótese de Nulidade → H_0
$H_o \rightarrow \mu_A = \mu_B$
Hipótese Alternativa → H_1
$H_1 \rightarrow \mu_A \neq \mu_B$

Nível de significância (α) → probabilidade de cometer EI RBAS19

Portanto EI: Erro ao se afirmar que existem diferenças entre os grupos que estão sendo comparados

Slide 11

RBAS12 Eu implemento um teste estatístico e chego à conclusão de que H0 não é verdadeira (rejeito H0), a igualdade não é verdadeira ou, em nosso exemplo, a capacidade auditiva dos grupos A e B não é igual, eu rejeitei H0, aceitei H1. Eu afirmo isso, mas de fato, se eu fosse consultar toda a população de recém-nascidos, eu iria verificar que essa hipótese que eu rejeitei (H0), ela é, de fato, verdadeira, ou seja, eu cometi um erro: eu cheguei à conclusão de que H0 é falsa, quando de fato, ela é verdadeira. Este erro é chamado Erro do tipo 1 (E1).
Rosimara Beatriz Arci Salgado; 02/04/2022

RBAS13 Eu chego à conclusão de que os grupos são iguais (H0), mas, de fato, eles não são iguais. Também cometi um erro. Este erro é chamado Erro tipo 2 (EII).
Rosimara Beatriz Arci Salgado; 02/04/2022

RBAS19 Rejeito H0 e aceito H1, quando, na verdade, H0 é verdadeira.
Rosimara Beatriz Arci Salgado; 03/04/2022

TEORIA DOS TESTES DE HIPÓTESES

Erros de decisão

Erro tipo I → EI	Rejeitar H_0 quando ela é verdadeira
Erro tipo II → EII	Aceitar H_0 quando ela é falsa

		Fato	
		Verdadeiro(+)	Falso(-)
Decisão	Aceitação (+)	Decisão correta(+,+)	**Erro II**
	Rejeição (-)	**Erro I**	Decisão correta(-,-)

RBAS15

Inatel CAMINHOS QUE CONECTAM COM O FUTURO

Slide 12

RBAS15 Eu aceitei H0 e ela é, de fato, verdadeira: decisão correta.
Eu rejeitei H0 e ela é, de fato, falsa: decisão correta.
Eu aceitei H0 e ela é, de fato, falsa: Erro II.
Eu rejeitei H0 e ela é, de fato, verdadeira: Erro I.
Rosimara Beatriz Arci Salgado; 02/04/2022

Exemplo:

Suponha que a proporção da população com mais de 60 anos afetada pela doença de Alzheimer seja, segundo a literatura, de 20%.

RBAS16

Se um pesquisador efetuasse um levantamento de dados na região onde ele trabalha e verificasse que, dentre 50 pessoas na faixa etária considerada, os portadores da síndrome seria de 16.

Slide 13

RBAS16 Um observador diagnosticou que na região dele a incidência de pessoas acima de 60 anos com Alzheimer é maior (de 32% e não de 20%, tal como prevê a literatura) do que a proporção nacional. Pela literatura, nesse grupo de 50 pessoas, deveriam haver 10 pessoas afetadas com doença de Alzheimer e foram observadas, de fato, 16. Então, parece que nessa região a incidência de pessoas afetadas pela doença de Alzheimer é maior do que na literatura. Será que existe um motivo dessa incidência ser maior nessa região? Matematicamente, 10 é diferente de 16, os nºs são diferentes, óbvio, mas será que essa diferença é realmente significativa estaticamente? É isso que precisa provar. Então o pesquisador formula uma hipótese e realiza o seu teste, para verificar se a incidência da doença de Alzheimer nessa região é diferente (nesse caso, é maior) da incidência nacional. Precisa ser investigado se de repente o sexo dessas pessoas da região em estudo é só de mulheres ou só de homens, enfim, precisa ser investigado.

Rosimara Beatriz Arci Salgado; 02/04/2022

Como a proporção por ele encontrada na experiência (16/50 = 32%) difere da proporção citada na literatura, o pesquisador pode imaginar que, por alguma razão, a proporção de indivíduos afetados pela síndrome na sua região de atuação é diferente/maior que a da população em geral.

Contudo, essa diferença pode ser casual, decorrente do fato de se estar trabalhando com uma amostra, isto é, tratar-se de um erro de amostragem, o que levaria a tirar uma conclusão errada, fazer uma afirmação errada.

RBAS17

RBAS17 Se eu pegar grupos diferentes de 50 pessoas, podemos encontrar 10 pessoas em um grupo, 11 pessoas no outro, 12 pessoas no outro, 15 pessoas em outro... com Alzheimeir devido ao erro de amostragem, pode ser uma diferença casual. É preciso montar uma hipótese e fazer o teste.

De repente, por algum motivo, essas 50 pessoas não representam a população como um todo. De repente essas 50 pessoas são do sexo feminino, será que o grupo feminino é igual ao masculino em termos de Mal de Alzheimer? Não sabemos, precisa ser investigado.

Rosimara Beatriz Arci Salgado; 02/04/2022

Exemplo:

RBAS18

O problema é decidir:

A partir de que valor é razoável começar a pensar que as taxas sejam diferentes.

Inatel CAMINHOS QUE CONECTAM COM O FUTURO

Slide 15

RBAS18 A partir de que valor é razoável começar a pensar que as taxas sejam diferentes? No caso específico, eu tenho 16 pessoas (valor observado) e eu espero 10 (20% de 50). 16 eu posso considerar que é diferente de 10? 15 é diferente de 10? 9? 11? Qual é a faixa de valores que eu posso afirmar que esses números são diferentes? Nós estamos tentando encontrar essa faixa de valores, para aceitar H0 ou rejeitar H0.

Rosimara Beatriz Arci Salgado; 02/04/2022

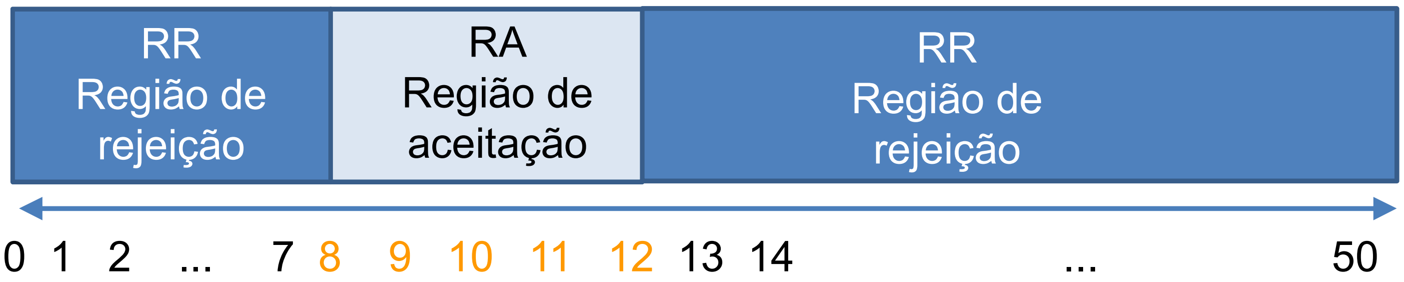
Exemplo:

Na amostra com 50 indivíduos, o valor esperado de doentes de Alzheimer é de 10 indivíduos. RBAS21

RBAS20

$$\mu = E = n.p = 50 \times 0,2 = 10$$

Podemos criar uma regra que aceite H_0 quando o número de indivíduos com a síndrome esteja entre 8 e 12 inclusive, porém **não** é um critério de referência (conhecido).



Slide 16

RBAS20 "midi" é a média da população, que é igual ao valor esperado (Esperança), que é igual ao nº de indivíduos (tamanho da amostra: n) * a taxa de proporcionalidade (p).

Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS21 Eu poderia estabelecer uma regra empírica, não muito objetiva, mas como eu espero 10 pessoas, eu posso criar/estabelecer uma regra, quando o nº de indivíduos com a síndrome esteja, por exemplo, entre 8 e 12. Peguei 20% para mais e 20% para menos, isso vai dar entre 8 e 12. A minha regra empírica, sem muito critério seria o seguinte: eu considero que se num grupo de 50 pessoas eu encontrar 8, 9, 10, 11 ou 12 pessoas com Alzheimer está igual à média nacional, ou seja, na minha região a média é igual à média nacional. Foi uma casualidade essa diferença. Agora, se tiver 13 pessoas, 14 pessoas, 50 pessoas, 1 pessoa, 2 pessoas, 7 pessoas, aí, eu vou concluir que aquela minha região tem uma incidência diferente da incidência nacional. Então é a faixa de valores, nós estamos tentando descobrir isso. Neste caso, esta regra criada não teve nenhum critério objetivo, uma referência, de fato, para a gente construir a faixa, o intervalo de valor (RA - Região de Aceitação), para que ao apresentar os meus resultados todos saibam qual é a referência para determinar a minha faixa de valores, onde eu vou aceitar a igualdade entre os dois grupos comparados ou rejeitar a igualdade entre os dois grupos comparados. Neste nosso exemplo, estou comparando o meu grupo de 50 pessoas com o resto da população, que eu tenho a referência na literatura.

Rosimara Beatriz Arci Salgado; 03/04/2022

Exemplo:

RBAS22

Na prática, define-se o erro máximo tolerado.

Por exemplo: se for definido um nível de significância (probabilidade de cometer EI) de 5%, a decisão tomada aceitará um erro tipo I de no máximo 5% e a confiança (grau de confiança) desta decisão será de 95%.

No exemplo dos portadores de Alzheimer, quais seriam os limites de aceitação da hipótese de nulidade ao nível de significância de 5%?

RBAS23

$$H_o \rightarrow \bar{x} = \mu$$

$$H_1 \rightarrow \bar{x} \neq \mu$$

RBAS22 Como nós precisamos de uma referência, nós vamos determinar o erro máximo tolerado. É a partir da determinação do erro máximo tolerado, que a gente vai determinar aquela faixa de valores e esse erro máximo tolerado é o nosso nível de significância, ou seja, todos sabem o que é o nível de significância (probabilidade de se cometer o erro tipo I - rejeitar H0, quando ela é verdadeira).

Eu vou fazer um teste estatístico e o nível de significância aceitável é de 5% ou, a probabilidade de se cometer um erro tipo I é de no máximo 5%. A partir desse valor, eu determino a faixa de valores do nº de pessoas acometidas do Mal de Alzheimer em torno da média 10, em que eu vou aceitar a hipótese de nulidade. Agora eu tenho uma referência, estou tentando estabelecer uma referência, o erro máximo tolerável. Quando eu digo que o nível de significância é de 5%, probabilidade de se cometer um erro tipo I é de 5%, o complemento disso é o grau de confiança que, nesse caso, é de 95%.

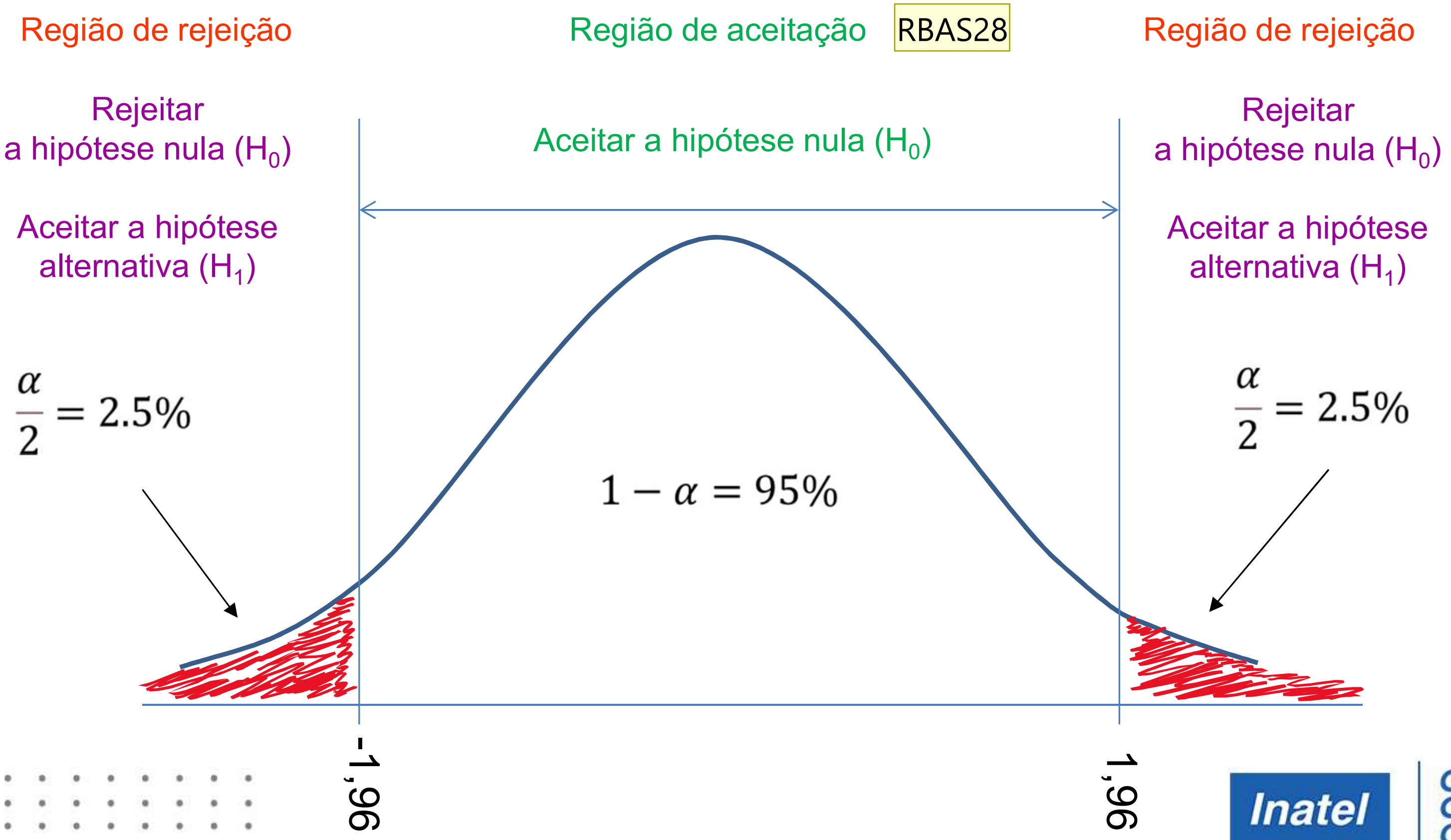
Fiz o meu experimento com 50 pessoas, encontrei 16 pessoas com a síndrome de Alzheimer, eu aceito ou rejeito a iguadade: 16=10? Se eu faço o teste em 50 idosos, na média, no máximo 5% (5 em cada 100) das minhas afirmações vão estar erradas. Se achar isso muito, precisa diminuir esse nível de significância para 3%, 2%, 1%... obviamente que dminuir o nível de significância, você precisa alterar o tamanho da sua amostra. Ao invés de trabalhar com 50 pessoas, você precisa de 100, 200 pessoas, para que se tenha o nível de significância menor.
Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS23 Qual é a faixa de valores que eu aceito ou rejeito H0? É isso que nós vamos descobrir.
Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS25

Exemplo: Nível de significância α = 5%

Uma distribuição estatística é uma função que define uma curva, e a área sob essa curva determina a probabilidade de ocorrer o evento por ela correlacionado. Admitir que a distribuição amostral das médias é normal, Gaussiana.



Inatel | CAMINHOS QUE CONECTAM COM O FUTURO

RBAS25 Breve revisão: área da função densidade de probabilidade (fdp).
Imaginando que se tenham vários grupos de 50 pessoas e a média de idosos com Alzheimer de cada grupo seja ora, 16, ora 10, ora 12, ora 11, ora 10.... Para esse conjunto muito grande de médias, eu faço o histograma. A distribuição dessas médias (valores) é Gaussiana (lá pelo Teorema do Limite Central, de Probabilidade). Quando você tem um número de amostras muito grande a soma dessa variável vai dar Gaussiana e, aí, eu tenho, tradicionalmente, a nossa distribuição Gaussiana que é no formato de sino.
Imaginemos que a nossa fdp tenha média, por exemplo, 10 e um determinado desvio padrão, que determina a dispersão dessa Gaussiana.
Nós vamos admitir que a distribuição que estamos trabalhando, ela é Gaussiana e é bastante razoável isso.
A área abaixo da curva (fdp) é 1. A soma das probabilidades é 100%, 1. Como a distribuição é simétrica, a área à esquerda de 10 (média, "ponto central do pico do sino") é 0,5 (metade). Se eu pegasse o ponto 12, eu posso dizer que a área à esquerda de 12, com certeza, é maior que 0,5. Não sei exatamente, quanto é, mas sei que com certeza é maior do que 0,5. Se eu pegar outra referência, por exemplo, 7. A área à esquerda de 7, com certeza, é menor do que 0,5.
Não existe fórmula para Gaussiana, para calcular a área exata e, também, mesmo sabendo o valor da área, querendo saber o valor da ordenada, não tem uma fórmula para isso, mas eu preciso saber. Para isso, eu tenho uma tabela: dada a área, encontre Z e dado Z, encontre a área. Obviamente, se Z for igual à média e a média for 0, o desvio padrão é 1. Normalmente, essas tabelas trabalham com o valor normalizado: Z = (x-"midi")/sigma, onde Z é normalizado, ou seja é igual a 0, a média é 0. Então, à esquerda de 0 a área é 0,5 e à direita de 0 a área é 0,5, mas para qualquer valor diferente de 0, eu preciso da tabela.
Pela tabela, quanto maior Z (ordenada), maior é a área. Para esses valores positivos de Z da tabela, a área é sempre maior que 0,5, visto que Z>0, a curva normal é padronizada.
Rosimara Beatriz Arci Salgado; 03/04/2022

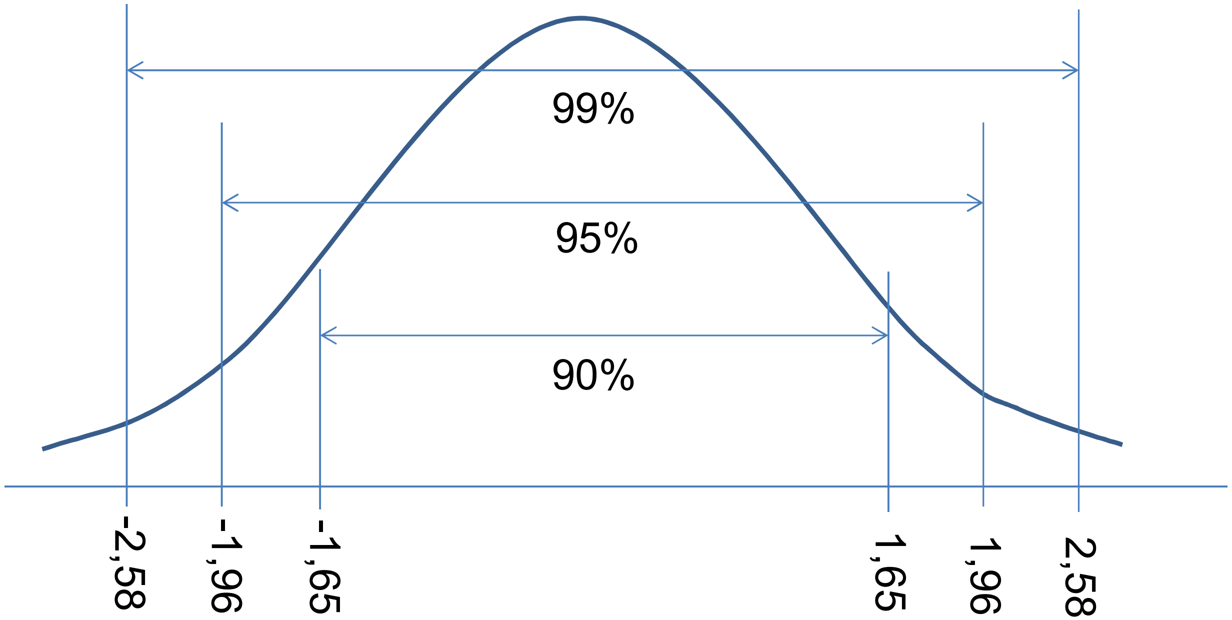
RBAS28 Eu vou calcular a faixa de valores em torno de 10, que é o valor esperado, que estarei aceitando (região de aceitação), H0 (10=16)? Os valores fora dessa faixa eu estarei rejeitando.
Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS26

Exemplo:

5% = 0,05

0,05/2 = 0,025



RBAS26 Os níveis de significância de 1%, 5% e 10% são os mais usados.
Para esses níveis de significância, os valores de Z serão SEMPRE estes tabelados, fixos.
Nível de significância de 5%, Z=1,96; para Nível de significância de 10%, Z=1,65 e para Nível de significância de 1%, Z=2,58.
Dentre esses níveis de significância, o de 5% é o mais utilizado.
Rosimara Beatriz Arci Salgado; 03/04/2022

Exemplo:

No exemplo dos portadores de Alzheimer, quais seriam os limites de aceitação da hipótese de nulidade ao nível de significância de 5%?

5% = 0,05

0,05/2 = 0,025

Na tabela padronizada, para A = 0,025, tem-se z = -1,96

RBAS30

RBAS29

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = n.p = 50.0,2 = 10$$

RBAS31

$$\sigma = \sqrt{n.p.q} = 2,8284$$

$$\boxed{q = 1\text{-}p = 1\text{-}0,2 = 0,8}$$

Inatel | CAMINHOS QUE CONECTAM COM O FUTURO

Slide 20

RBAS29 Z é a variável padronizada (média 0 e desvio padrão 1). A média (midi) esperada do problema é 50 (tamanho da amostra) *0,2 (incidência, probabilidade de ocorrência do problema) =10. Se eu alterar o tamanho da amostra ou a incidência do problema (p), a média será outra.
Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS30 midi = média populacional
Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS31 Desvio padrão é a raiz quadrada de n * p * q, onde q = 1-p (distribuição binomial). Portanto, q=80%, ou seja, 0,8.
Desvio padrão = raiz quadrada de 8, valor calculado assim: n.p.q = 50*0,2*0,8
Rosimara Beatriz Arci Salgado; 03/04/2022

Exemplo:

5% = 0,05

0,05/2 = 0,025

Na tabela, para A = 0,025, tem-se z = -1,96

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = n.p = 50.0{,}2 = 10$$

$$\sigma = \sqrt{n.p.q} = 2{,}8284$$

$$\pm 1{,}96 = \frac{x - 10}{2{,}8284}$$

RBAS32

$$4{,}46 \leq x \leq 15{,}54$$

Slide 21

RBAS32 Acabei de descobrir aqui a faixa de valores em que eu vou aceitar a hipótese de nulidade (H0) dentro do meu problema de portadores de Alzheimer. O meu X tem que ser maior que 4,46 e menor que 15,54. O menor nº inteiro maior que 4,46 é 5 e o menor nº inteiro menor que 15,54 é 15.

As fórmulas usadas nesse exemplo se adequam a este cenário, mas eu posso ter cenários diferentes com fórmulas diferentes.

Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS34

Exemplo:

5% = 0,05
0,05/2 = 0,025

Na tabela, para A = 0,025, tem-se z = -1,96

$$z = \frac{x - \mu}{\sigma}$$

$$\mu = n.p = 50.0{,}2 = 10$$

$$\sigma = \sqrt{n.p.q} = 2{,}8284$$

$$\pm 1{,}96 = \frac{x - 10}{2{,}8284}$$

$$4{,}46 \le x \le 15{,}54$$

$16 \neq 10$, rejeito H_0

$$H_o \rightarrow \bar{x} = \mu$$
$$H_1 \rightarrow \bar{x} \neq \mu$$

RBAS33

Conclusão:

Para nível de significância de 5%, H_o deve ser aceita para $5 \le x \le 15$

Inatel CAMINHOS QUE CONECTAM COM O FUTURO

RBAS33 Conclusão: H0 deve ser aceita quando você encontrar na sua experiência entre 5 e 15 pessoas com Alzheimer.
O pesquisador pegou 50 pessoas. Se dessas 50 pessoas, 5 pessoas tem Alzheimer ou 10 ou 14 ou 15 pessoas com Alzheimer isso significa que este grupo de 50 pessoas reflete os 20% lá da população como um todo, não existe diferença. Eu posso considerar que está tudo normal. Agora, no exemplo, 16 pessoas tinham pessoas Alzheimer e 16 não está dentro da faixa de aceitação. Portanto, este grupo de 50 pessoas tem alguma coisa, que não sabemos exatamente o quê, que está fazendo com que, de fato, a incidência de Alzheimer para esse grupo seja diferente da população como um todo. No nosso caso, 16 realmente é diferente de 10. Se eu tivesse encontrado, 14 idosos com problemas de Alzheimer no estudo do grupo de 50 pessoas, 14, estatisticamente, não é diferente de 10. 14, 15, 8, 10, 9... estatisticamente falando não é diferente de 10. Encontrei ali a faixa de valores aceitável.
Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS34 O meu critério foi baseado no nível de significância, um erro tipo I, que nesse caso, é de 5%.
No caso em estudo, teve uma incidência de 16 pessoas. A minha conclusão rejeita H0, porque 16 não está na faixa de aceitação. Se eu pegar um outro grupo de 50 pessoas e imagine que eu encontre 10 pessoas, nesse caso, eu aceito H0. Outro grupo de 50, imagina que eu tenha encontrado 3 pessoas com Alzheimer, nesse caso, eu rejeito H0 (3 não está na faixa) e assim por diante. De vez em quando, eu rejeito H0 e, às vezes, eu posso aceitar H0. Na média, no máximo 5% das vezes que eu rejeitar H0, H0 é verdadeira. Isso é um erro, rejeitar quando não se deveria. Se eu nunca quero errar, é preciso avaliar todos os idosos do Brasil e do Mundo para não errar, não trabalhar com amostras e, sim, com toda a população. Veja que estamos trabalhando com hipóteses e, portanto, existe uma probabilidade de erro associada a sua afirmação.
A conclusão revela que existe uma diferença significativa entre os grupos testados.
Eu posso estar errada na minha conclusão, mas estabeleci um critério (nível de significância). Seguindo os resultados encontrados, precisaria ser investigado o porquê desses 16 idosos com síndrome de Alzheimer no grupo de 50 pessoas, estudar outros aspectos relacionados a esse grupo para se chegar a novas conclusões dessa diferença.
Rosimara Beatriz Arci Salgado; 03/04/2022

RBAS1

NÍVEL DE SIGNIFICÂNCIA DE UM TESTE, *P*

Podemos determinar também a probabilidade de erro tipo um (nível de significância) do resultado.

Por exemplo: no caso da experiência com portadores de Alzheimer, o pesquisador tinha encontrado 16 casos da síndrome em 50 indivíduos com mais de 60 anos. Rejeitar H_0 com este resultado implica uma probabilidade de erro de que ordem?

Slide 23

RBAS1 Na aula passada, vimos o nível de significância "alfa", ou seja, a probabilidade de erro máxima que eu aceitava. Então, eu aceitava 5% de probabilidade de errar, ou seja, a cada 100 testes com 50 indivídulos diferentes cada e verificasse quantas pessoas têm incidência de Alzheimer, a minha afirmação iria estar errada 5 vezes em um conjunto de 100 medidas na média. O nível de significância "alfa" era a probabilidade máxima de aceitar H0. Agora, o nível o nível de significância do teste (resultado), determinado pelo letra "p" é relacionado ao teste, ao valor calculado.

Rosimara Beatriz Arci Salgado; 15/04/2022

RBAS17

NÍVEL DE SIGNIFICÂNCIA DE UM TESTE, *P*

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

Na tabela, para $z_{calculado}$ = 2,12, tem-se: A = 0,98299 (área da esquerda)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Área acumulada sob a curva normal padronizada (valores positivos de z)													
2	z	A	z	A	z	A	z	A	z	A	z	A	z	A
3	0	0,50000	0,5	0,69146	1	0,84134	1,5	0,93319	2	0,97725	2,5	0,99379	3	0,99865
4	0,01	0,50399	0,51	0,69497	1,01	0,84375	1,51	0,93448	2,01	0,97778	2,51	0,99396	3,01	0,99869
5	0,02	0,50798	0,52	0,69847	1,02	0,84614	1,52	0,93574	2,02	0,97831	2,52	0,99413	3,02	0,99874
6	0,03	0,51197	0,53	0,70194	1,03	0,84849	1,53	0,93699	2,03	0,97882	2,53	0,99430	3,03	0,99878
7	0,04	0,51595	0,54	0,70540	1,04	0,85083	1,54	0,93822	2,04	0,97932	2,54	0,99446	3,04	0,99882
8	0,05	0,51994	0,55	0,70884	1,05	0,85314	1,55	0,93943	2,05	0,97982	2,55	0,99461	3,05	0,99886
9	0,06	0,52392	0,56	0,71226	1,06	0,85543	1,56	0,94062	2,06	0,98030	2,56	0,99477	3,06	0,99889
10	0,07	0,52790	0,57	0,71566	1,07	0,85769	1,57	0,94179	2,07	0,98077	2,57	0,99492	3,07	0,99893
11	0,08	0,53188	0,58	0,71904	1,08	0,85993	1,58	0,94295	2,08	0,98124	2,58	0,99506	3,08	0,99896
12	0,09	0,53586	0,59	0,72240	1,09	0,86214	1,59	0,94408	2,09	0,98169	2,59	0,99520	3,09	0,99900
13	0,1	0,53983	0,6	0,72575	1,1	0,86433	1,6	0,94520	2,1	0,98214	2,6	0,99534	3,1	0,99903
14	0,11	0,54380	0,61	0,72907	1,11	0,86650	1,61	0,94630	2,11	0,98257	2,61	0,99547	3,11	0,99906
15	0,12	0,54776	0,62	0,73237	1,12	0,86864	1,62	0,94738	2,12	0,98300	2,62	0,99560	3,12	0,99910
16	0,13	0,55172	0,63	0,73565	1,13	0,87076	1,63	0,94845	2,13	0,98341	2,63	0,99573	3,13	0,99913
17	0,14	0,55567	0,64	0,73891	1,14	0,87286	1,64	0,94950	2,14	0,98382	2,64	0,99585	3,14	0,99916
18	0,15	0,55962	0,65	0,74215	1,15	0,87493	1,65	0,95053	2,15	0,98422	2,65	0,99598	3,15	0,99918
19	0,16	0,56356	0,66	0,74537	1,16	0,87698	1,66	0,95154	2,16	0,98461	2,66	0,99609	3,16	0,99921
20	0,17	0,56749	0,67	0,74857	1,17	0,87900	1,67	0,95254	2,17	0,98500	2,67	0,99621	3,17	0,99924
21	0,18	0,57142	0,68	0,75175	1,18	0,88100	1,68	0,95352	2,18	0,98537	2,68	0,99632	3,18	0,99926
22	0,19	0,57535	0,69	0,75490	1,19	0,88298	1,69	0,95449	2,19	0,98574	2,69	0,99643	3,19	0,99929
23	0,2	0,57926	0,7	0,75804	1,2	0,88493	1,7	0,95543	2,2	0,98610	2,7	0,99653	3,2	0,99931

Slide 24

RBAS17 Afirmamos que a média padrão amostral é Gaussiana, por isso, estamos usando essa fórmula de "z", que é a normalização, ou seja, "z" é a variável aleatória normalizada. Portanto "z" tem média 0 e desvio padrão ou variância unitário (igual a 1).
Primeiro eu vou calcular o "z" dado um valor de "x", dada a média e dado o desvio padrão, para depois calcular o nível de significância do teste.
O valor de "x" no exemplo é 16. Eu tenho 16 pessoas num grupo de 50 pessoas com Alzheimer. A média esperada é 10 (50 * 0,2). O desvio padrão calculado (raiz quadrada de n*p*q, onde "q" é "1-p".
Para o valor fornecido de 16 pessoas, eu tenho um valor "z", que é o nível crítico, o valor da variável aleatória gaussiana normalizada, que é de 2,12. Agora eu tenho que calcular a probabilidade associada a 2,12.
Rosimara Beatriz Arci Salgado; 18/09/2022

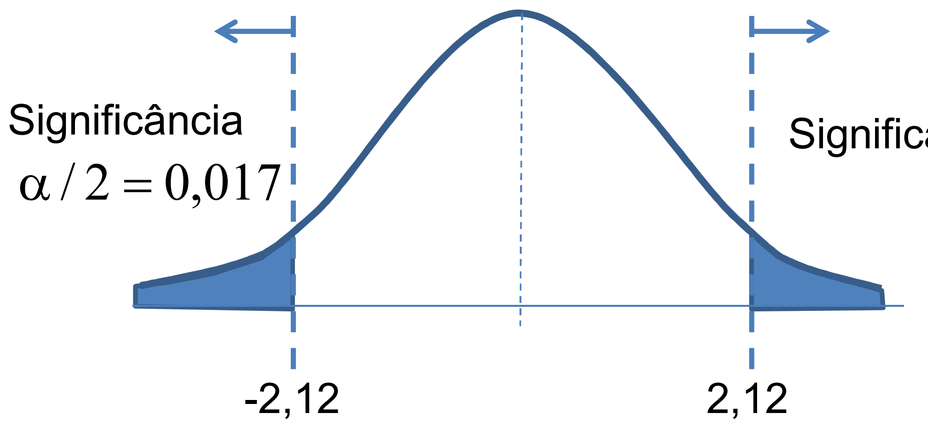
NÍVEL DE SIGNIFICÂNCIA DE UM TESTE, *P*

Cálculo de Z $\rightarrow z = \dfrac{x-\mu}{\sigma} = \dfrac{16-10}{2{,}8284} \cong 2{,}12$

RBAS3

Na tabela, para $z_{calculado}$ = 2,12, tem-se A = 0,98299 (área da esquerda). A área da direita é $1-0{,}98299 = 0{,}017$

Para $z_{calculado}$ = -2,12, tem-se A = 0,017 (área à esquerda de -2,12)

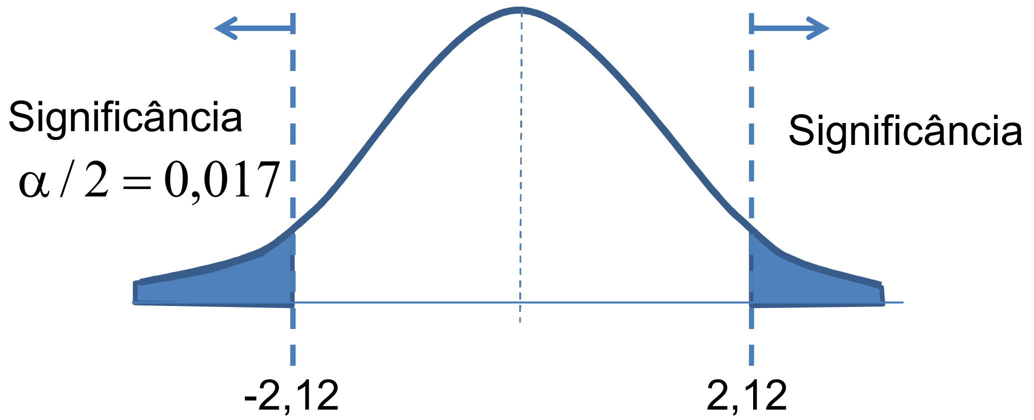


RBAS3 Pelo fato da média amostral ser Gaussiana e, portanto, simétrica, o nível de significância é igual para os dois lados, a área é igual para os dois lados em relação à sua média. O nível de significância do teste será a soma das duas áreas (à direita de 2,12 e à esquerda de 2,12)

Rosimara Beatriz Arci Salgado; 15/04/2022

RBAS4

NÍVEL DE SIGNIFICÂNCIA DE UM TESTE, *P*



Como é bilateral, $p = 2 \times 0{,}017 = 0{,}034 \quad \text{ou} \quad 3{,}4\%$

Conclusão: Como *p* está entre 1 e 5%, pode-se afirmar que a diferença encontrada pelo pesquisador é significante. Rejeita-se H_0.

$$H_o \rightarrow \bar{x} = \mu$$

$$H_1 \rightarrow \bar{x} \neq \mu$$

A probabilidade de rejeição de H_0 associada ao resultado da experiência é denominada **nível de significância do teste**, ou simplesmente ***p***

RBAS4

O nível de significância do teste, que é o "p", é 2 vezes a área à direita de 2,12 ou à esquerda de -2,12.
O "p" é diferente do "alfa", pois o "alfa" foi calculado para estabelecer uma faixa (valor mínimo e máximo) em que eu aceito H0 (se o resultado do teste estiver dentro dessa faixa) ou rejeito H0 (se o resultado do teste estiver fora dessa faixa).
Agora eu peguei o valor encontrado no teste, que foi de 16 e calculei o nível de significância do teste, que foi de 3,4%.
O valor de "p", nesse caso, é menor que 5% (1%<p<5%). Portanto, como "p" é menor que 5%, eu rejeito H0, porque 3,4% é menor que "alfa" - ambos se referem ao Erro tipo I.

Rosimara Beatriz Arci Salgado; 15/04/2022

NÍVEL CLÁSSICOS DE SIGNIFICÂNCIA

De um modo geral as hipóteses são testadas em três níveis de significância (α): 1, 5 e 10%

Nível de Significância *p*	**Conclusão**
Menor que 1%	Diferença **altamente significante** (certeza máxima para se rejeitar H_0)
Entre 1 e 5%	Diferença **significante** (maior certeza para se rejeitar H_0)
Entre 5 e 10%	Diferença **provavelmente significante** (provavelmente, maior certeza para aceitar H_0)
Maior que 10%	Diferença **não significante** (certeza máxima para aceitar H_0)

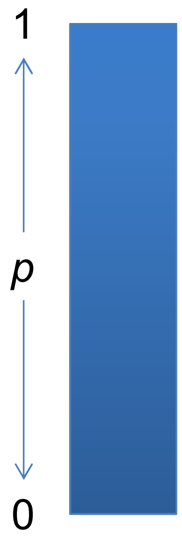
RBAS5 Eu tenho uma certeza maior de rejeitar H0, pois existe uma diferença altamente significante entre os grupos testados (rejeito H0, aceito H1 - hipótese alternativa, das diferenças).
Rosimara Beatriz Arci Salgado; 15/04/2022

RBAS6 Que foi o caso do exemplo, 3,4%.
Rosimara Beatriz Arci Salgado; 15/04/2022

RBAS7 Normalmente, a referência é de 5%, ou seja, quando eu calculo um valor de nível de significância menor do que 5%, eu rejeito H0; quando o valor de p é maior do que 5%, eu aceito H0 como verdadeiro (não rejeito H0).
O parâmetro fundamental (principal) para o nosso teste estatístico é o nível de significância p. A partir dele eu faço uma comparação com o meu "alfa" (referência). Se o "p" for menor do que 'alfa", eu rejeito H0, caso contrário, eu aceito H0 como sendo verdadeira. Existe, obviamente, uma probabilidade de erro na minha afirmação. Nesse caso, a probabilidade é de 3,4%, eu tenho que aceitar esse erro. Só não vai existir erro quando tivermos acesso a toda a população, porque sua afirmação será 100% verdadeira.
Rosimara Beatriz Arci Salgado; 15/04/2022

RBAS8

NÍVEL CLÁSSICOS DE SIGNIFICÂNCIA



Certeza máxima para <u>não rejeitar</u> H_0 ou certeza máxima para aceitar H_0

Menor certeza de rejeitar H_0 ou maior certeza para aceitar H_0

Maior certeza de rejeitar H_0 ou maior certeza para aceitar H_1

Certeza máxima para rejeitar H_0 ou certeza máxima para aceitar H_1

Quanto menor *p*, maior a evidência de que **existem** diferenças, então rejeita-se H_0 com maior certeza

Quanto maior *p*, maior a evidência de que <u>**não existem**</u> diferenças, então diminui a certeza da rejeição de H_0

RBAS8 A probabilidade de erro do teste, que é o nível de significância do teste varia entre 0 e 1. Se o "p" for muito próximo de 0 ou igual a 0, eu tenho certeza máxima para rejeitar H0. Você tem maior certeza para rejeitar H0. Se o "p" for próximo de 1, você tem uma certeza máxima para não rejeitar H0, portanto, você estará aceitando H0 (igualdade entre os grupos). No meio da escala ao lado, quanto mais o "p" for próximo de 0, maior certeza para se rejeitar H0. À medida que o "p" cresce, é mais próximo de 1, maior certeza de aceitar H0 ou menor certeza de rejeitar H0.

CONCLUSÃO: Se o cálculo de "p" for menor que 5%, eu rejeito H0 (os grupos testados não são iguais, eles são diferentes). Se o "p" for maior que 5%, os grupos testados, provavelmente, são iguais, eu não posso rejeitar H0, eu não tenho certeza de rejeitar H0.

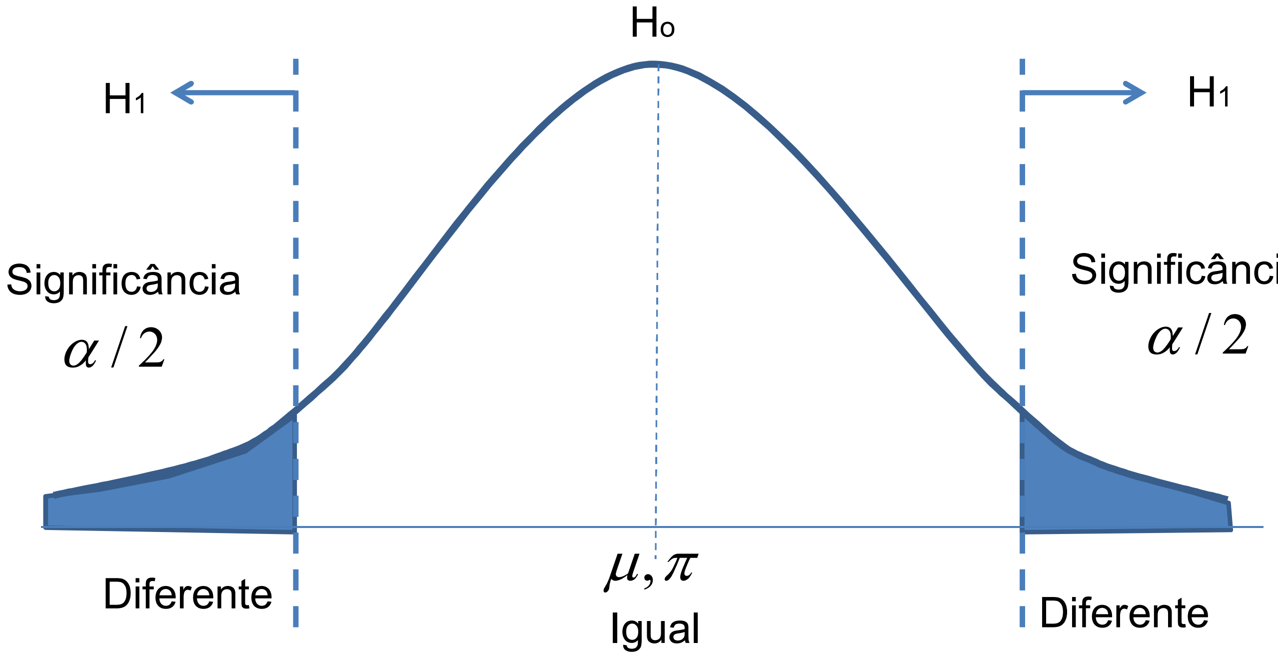
Rosimara Beatriz Arci Salgado; 15/04/2022

RBAS9

TESTES UNILATERAIS E BILATERAIS

A) Teste Bilateral (Testes Para Diferenças)

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

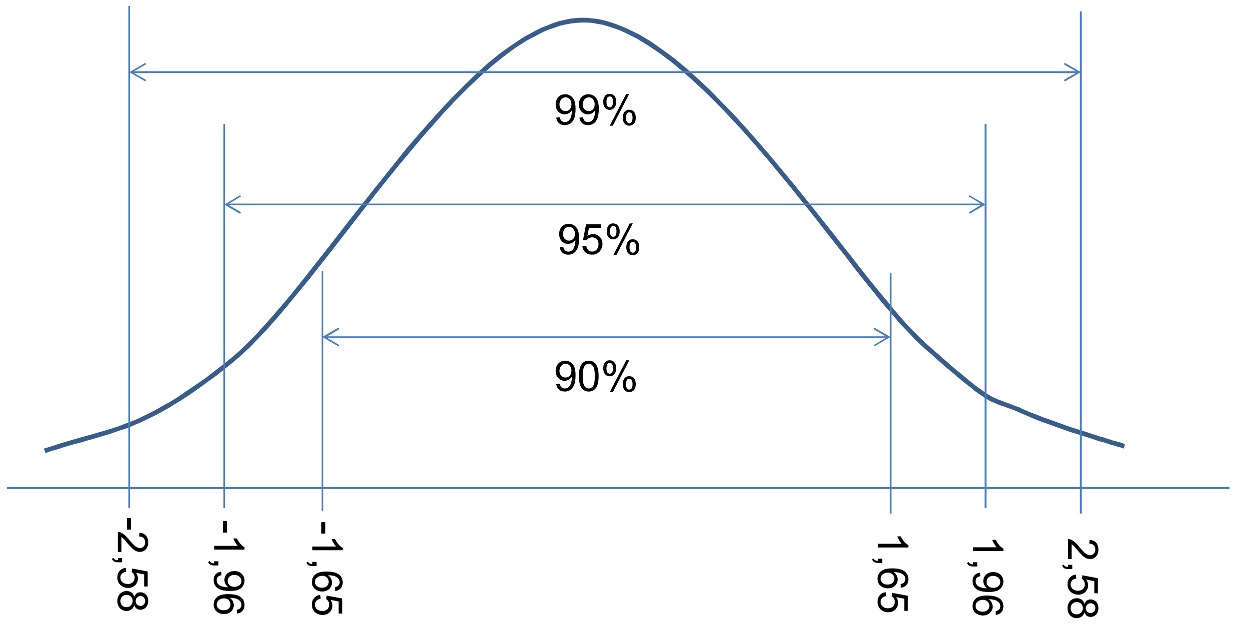


RBAS9 "x barrado" é a média amostral e "midi" é a média populacional.
Nível de significância de 5% (alfa)
Rosimara Beatriz Arci Salgado; 15/04/2022

RBAS10

TESTES UNILATERAIS E BILATERAIS

A) Teste Bilateral (Testes Para Diferenças)



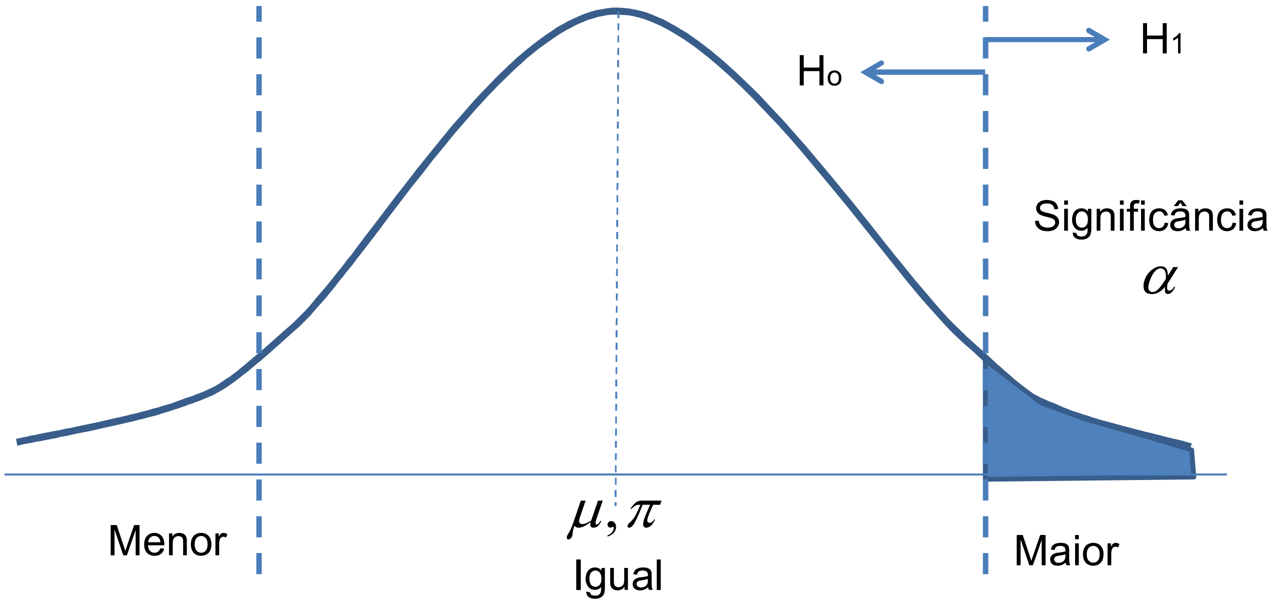
Slide 30

RBAS10 Grau de confiança de 99%, "alfa"=1%, Z=2,58;
Grau de confiança de 95%, "alfa" = 5%, Z=1,96, que é o mais comum;
Grau de confiança de 90%, "alfa" = 10%, Z=1,65.
Rosimara Beatriz Arci Salgado; 15/04/2022

TESTES UNILATERAIS E BILATERAIS

B) Teste Unilateral (Verificar se uma estatística é maior que...)

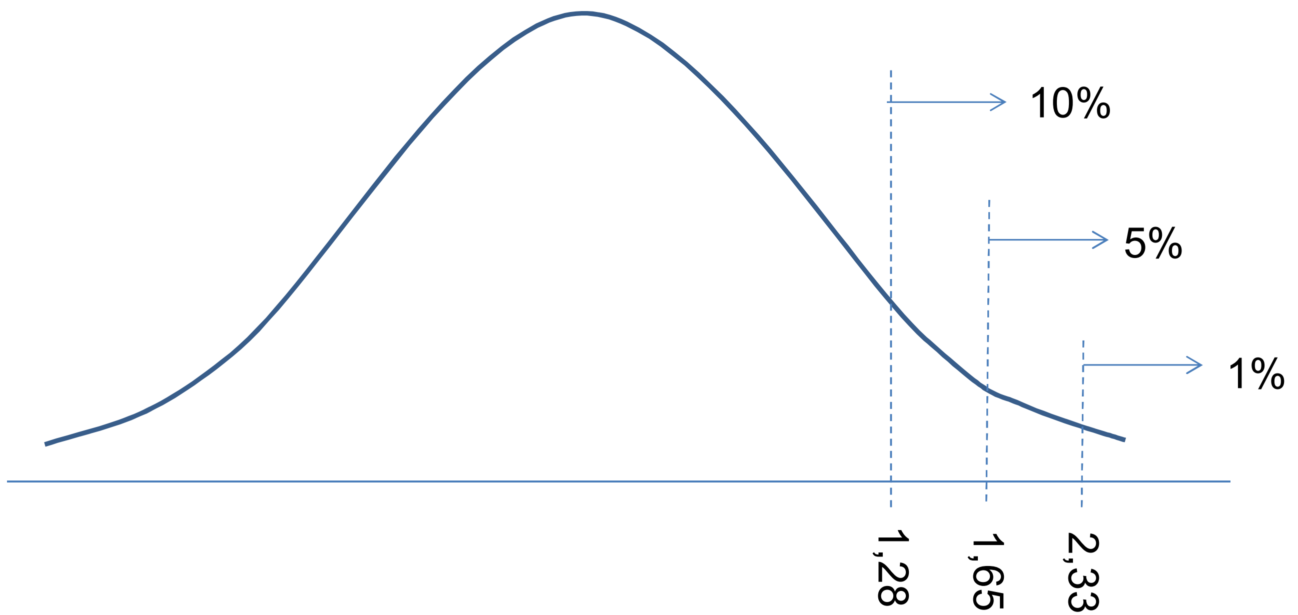
$$H_o \rightarrow \bar{x} \leq \mu \quad ou \quad p \leq \pi \quad \rightarrow \quad \text{sempre tem a igualdade}$$
$$H_1 \rightarrow \bar{x} > \mu \quad ou \quad p > \pi$$



RBAS12 Em testes unilaterais, normalmente, a igualdade sempre está em H0. No caso, o nível de significância "alfa" está todo na cauda da lateral direita, não tem a área nas duas caudas, é apenas à direita. Isso deve ao fato de se ter um teste unilateral. Não apenas igualdade, temos "maior do que ou igual".
Rosimara Beatriz Arci Salgado; 15/04/2022

TESTES UNILATERAIS E BILATERAIS

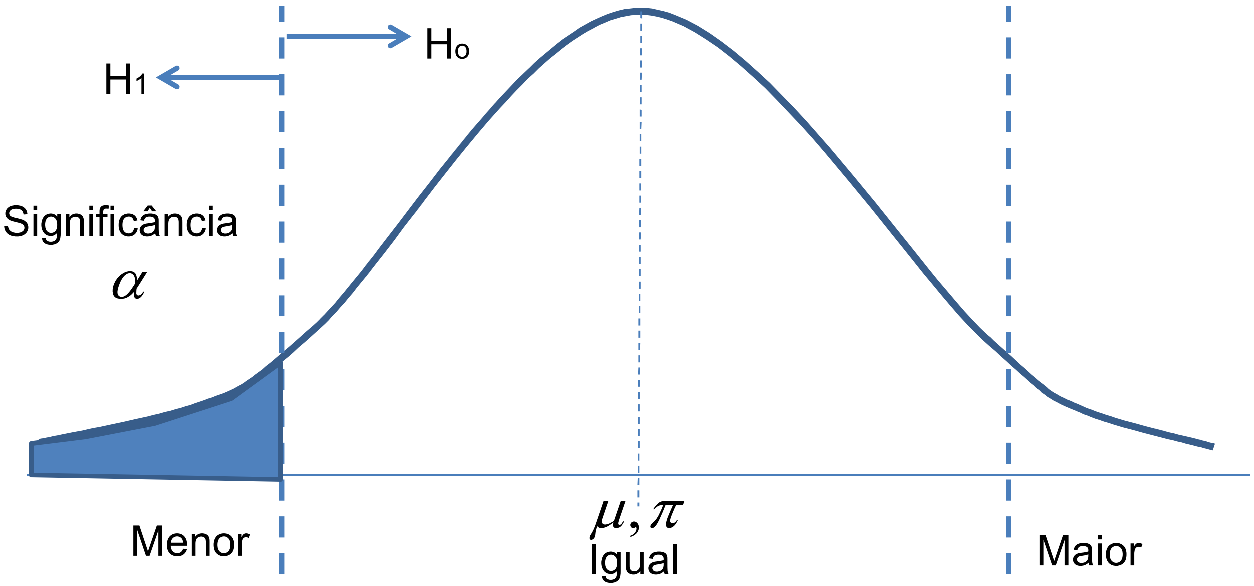
B) Teste Unilateral (Verificar se uma estatística é <u>maior</u> que...)



TESTES UNILATERAIS E BILATERAIS

C) Teste Unilateral (Verificar se uma estatística é <u>menor</u> que...)

$$H_o \rightarrow \bar{x} \geq \mu \quad ou \quad p \geq \pi \quad \rightarrow \quad \text{sempre tem a igualdade}$$
$$H_1 \rightarrow \bar{x} < \mu \quad ou \quad p < \pi$$



RBAS13

Exemplo 2:

Suponha que um estudo em determinada região mostra que a ingestão diária média de calorias em adultos é de 2.400kcal. Considere que um grupo de 25 adultos desta população apresentou um consumo médio de 3.000kcal, com um desvio padrão populacional de 1.250kcal. Para testar se o consumo calórico deste grupo é diferente do padrão de consumo da população, pode ser efetivado o teste da Normal para médias, como mostrado a seguir:

RBAS13 Vamos calcular o nível de significância do teste, para saber se essa diferença é significante ou não.
Rosimara Beatriz Arci Salgado; 15/04/2022

Exemplo 2:

1º Passo: Hipótese:

$$H_o \rightarrow \bar{x} = \mu$$
$$H_1 \rightarrow \bar{x} \neq \mu$$

RBAS18

2º Passo: Cálculo de Z, onde EP é o Erro Padrão da Média (medida de variação da média amostral ($\bar{x}$) em relação à média da população (μ). Sendo assim o EP ajuda a avaliar a confiabilidade da média amostral calculada):

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

RBAS15

$$EP = \frac{\sigma}{\sqrt{n}} = \frac{1250}{\sqrt{25}} = 250 \quad \rightarrow \quad Z = \frac{3000 - 2400}{250} = 2{,}4$$

Inatel | CAMINHOS QUE CONECTAM COM O FUTURO

RBAS15 Sempre vou adotar essa fórmula. O desvio padrão populacional (sigma) é dado no enunciado. Quando eu tenho uma amostra maior que 30, eu posso utilizar o desvio padrão amostral (s) dividido pela raiz quadradad de n. Para amostras, menor que 30, eu devo utilizar o desvio padrão populacional, pois é mais confiável. Em caso dele não ser fornecido, usa-se o teste t (de student), pois a variância é desconhecida, ao invés do teste da normalidade. Na prática, o desvio padrão populacional, normalmente, é desconhecido.

"Erro Padrão da Média": é uma medida de variação de uma média amostral em relação à média da população. É uma medida que ajuda a verificar a confiabilidade da média amostral calculada. Para obter uma estimativa de Erro Padrão basta dividir o Desvio Padrão pela raiz quadrada do tamanho amostral.

https://youtu.be/DTF0jt3ROVU

Em relação ao Desvio Padrão, os valores coletados distam em média 1250 kcal da média amostral.

Assim como a média amostral é uma estimativa da média populacional, o desvio padrão também é uma estimativa do desvio padrão real.

Em nosso exemplo, EP, as médias amostrais calculadas a partir de amostras de tamanho igual a 25, distam em média 250 kcal do parâmetro populacional real.

Rosimara Beatriz Arci Salgado; 15/04/2022

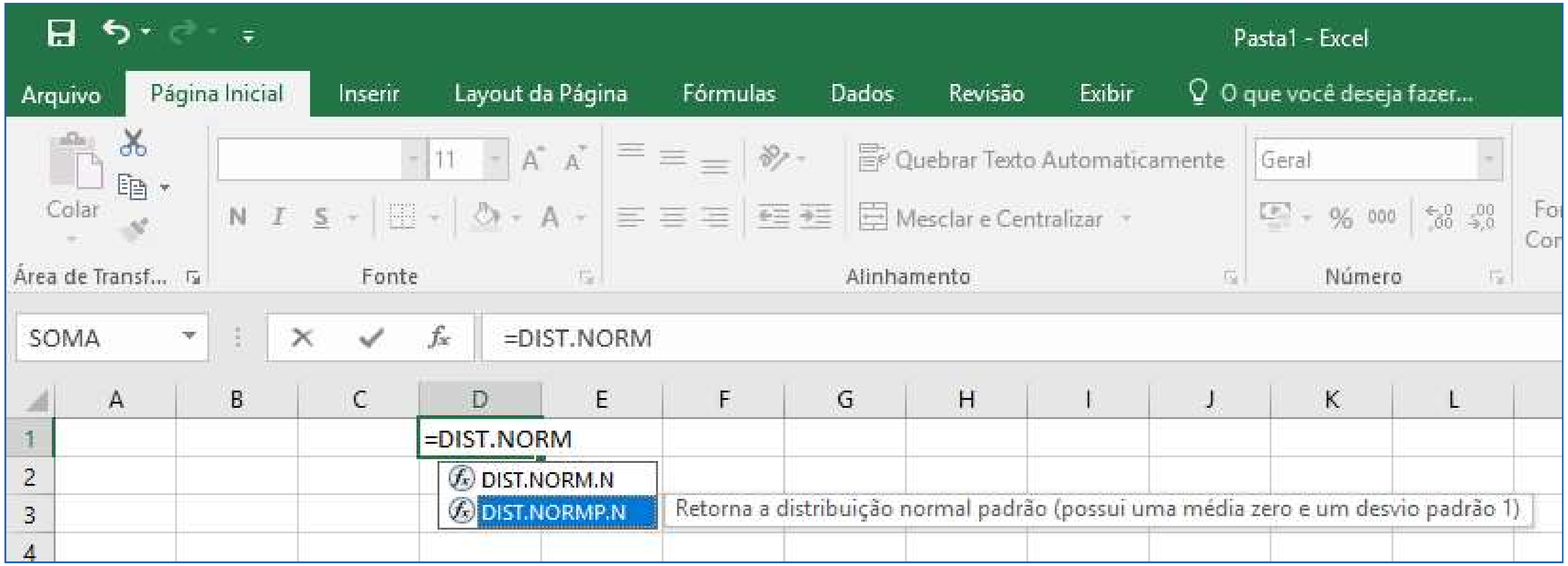
RBAS18 Trata-se de um teste bilateral, pois quero verificar se há diferença ou não entre a média amostral e a média populacional.

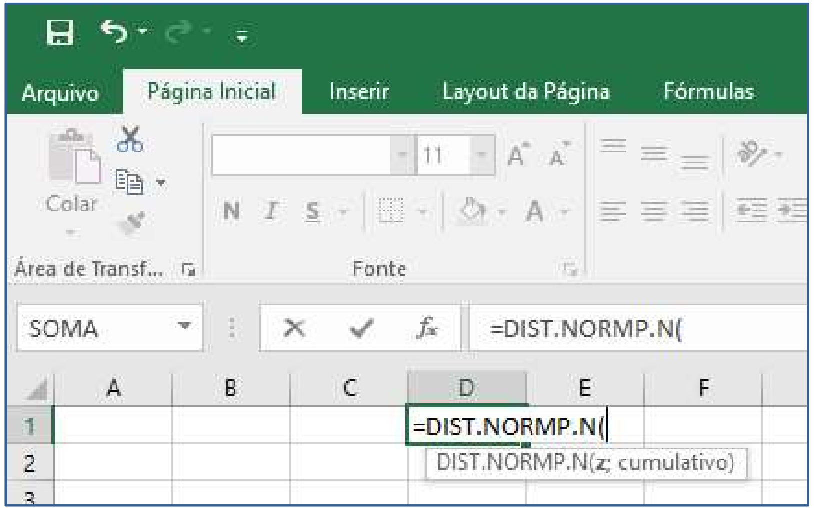
Rosimara Beatriz Arci Salgado; 18/09/2022

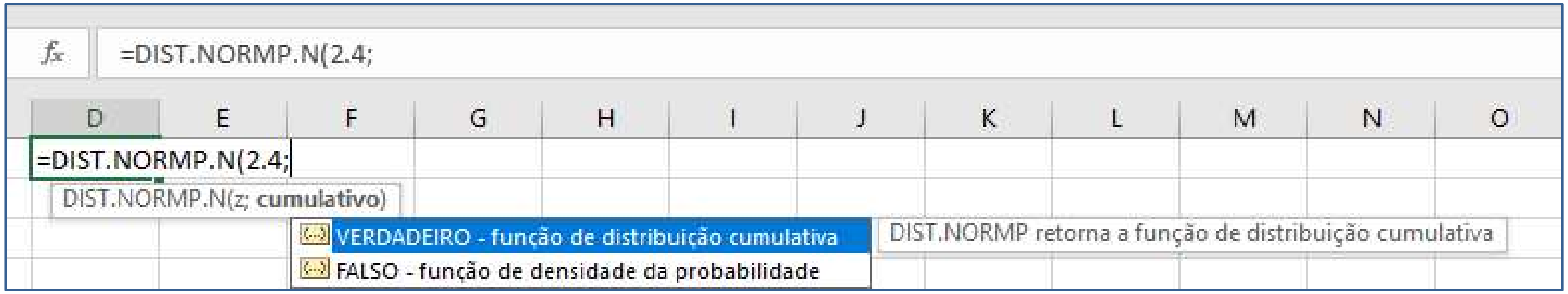
3º Passo: Determinação de *p*:

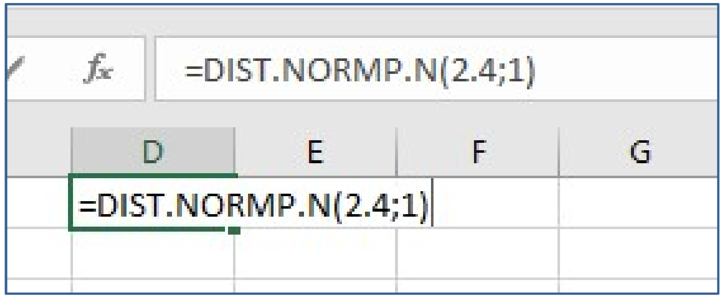
Na tabela, para Z = 2,4, tem-se A = 0,9918

$$p = 2 \times (1 - 0{,}9918) = 0{,}0164 \quad \text{ou} \quad 1{,}64\%$$




Arquivo
Página Inicial
Inserir
Layout da Página
Fórmulas
Colar
Área de Transf...
Fonte
SOMA
=DIST.NORMP.N(
=DIST.NORMP.N(
DIST.NORMP.N(z; cumulativo)


=DIST.NORMP.N(2.4;
=DIST.NORMP.N(2.4;
DIST.NORMP.N(z; cumulativo)
VERDADEIRO - função de distribuição cumulativa
FALSO - função de densidade da probabilidade
DIST.NORMP retorna a função de distribuição cumulativa


=DIST.NORMP.N(2.4;1)
=DIST.NORMP.N(2.4;1)

A = 0,991802

RBAS16

4º Passo: Decisão:

Rejeita-se H_0 com um nível de significância de 1,64%. Ou seja, $p < 5\%$.

5º Passo: Conclusão:

Conclui-se que o consumo calórico da amostra pode ser considerado diferente do da população, a um nível de significância de 1,64%.

Slide 38

RBAS16 Para casa: calcular a faixa de valores em que H0 é aceitável ao nível de significância de 5% ("alfa" = 5%).

Para teste bilateral, $Z = 1{,}96$

$-1{,}96 = (x-2400)/250$
$x = 1910$

$1{,}96=(x-2400)/250$
$x=2890$

$1910 \le x \le 2890$

Portanto, 3000 está fora da faixa de aceitação de H0. Rejeito H0, aceito H1, há diferença significante entre os valores.

Rosimara Beatriz Arci Salgado; 15/04/2022

F I M

Inatel

CAMINHOS
QUE CONECTAM
COM O FUTURO