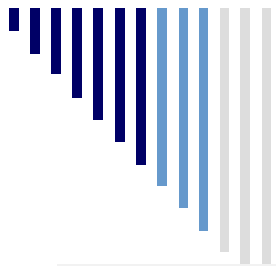


Probabilidade e Estatística

Aula 9

Fundamentos de Testes de Hipóteses

Leitura: Devore, Capítulo 8



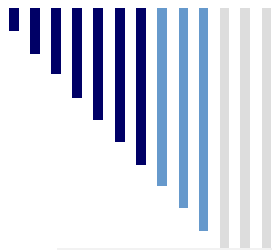
Objetivos

Neste capítulo, vamos aprender:

- Os princípios básicos de testes de hipóteses
 - Estabelecer as hipóteses
 - Os erros possíveis de um teste de hipóteses
 - Avaliar qual hipótese é mais condizente com os dados

- 3 tipos de testes de hipóteses:

- 2 abordagens para testes de hipóteses



Objetivos

Neste capítulo, vamos aprender:

- Os princípios básicos de testes de hipóteses
- 3 tipos de testes de hipóteses (bicaudal, de cauda inferior e de cauda superior) para os parâmetros:
 - a média populacional, μ
 - a proporção populacional, p
- Duas abordagens para realizar um teste de hipóteses:
 - Valor crítico
 - P-valor



As Hipóteses

"uma hipótese designa **qualquer suposição** de algo verosímil, **possível de ser verificado**, a partir da qual se extrai uma conclusão."



As Hipóteses

- Uma hipótese é **uma afirmação** sobre **o verdadeiro valor de um parâmetro populacional**:

Definição!

- Média populacional: μ

Exemplo: O valor médio de uma conta de celular nesta cidade é $\mu = \text{R\$}52$

- Proporção populacional: p

Exemplo: A proporção de adultos desta cidade que possuem telefones celulares é $p = 0.68$



As Hipóteses

- Um teste possui **duas hipóteses**:
 - H_0 : a hipótese nula
 - H_1 : a hipótese alternativa
- As hipóteses representam afirmações complementares, isto é, uma é o contrário da outra.
- O teste vai decidir, baseado em informações de uma amostra, qual das hipóteses é mais condizente com os dados.



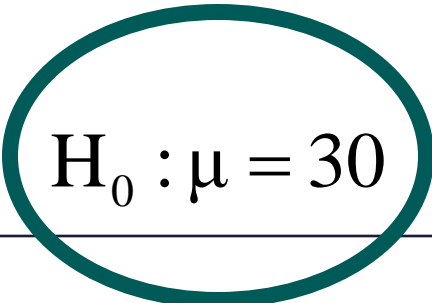
A Hipótese Nula, H_0

- Apresenta uma afirmação (numérica) a ser testada.

Exemplo: O diâmetro médio de parafusos produzidos em uma fábrica é igual a 30 mm.

$$H_0 : \mu = 30$$

- Sempre é uma afirmação a respeito do parâmetro populacional e não sobre a estatística amostral.


$$H_0 : \mu = 30$$


$$H_0 : \bar{X} = 30$$

A Hipótese Nula, H_0



- O teste começa assumindo que a hipótese nula é verdadeira.
 - Similar a noção de inocente até que seja provado o contrário
 - Se refere ao valor aceito historicamente ou socialmente (o que a maioria das pessoas acha que é verdadeiro)
- Pode ou não ser rejeitada
 - Apenas é rejeitada se houver evidência muito forte de que não é verdadeira.
- Sempre contém sinais com igualdade:
“=”, ou “ \leq ”, ou “ \geq ”

A Hipótese Alternativa, H_1



- É o oposto da hipótese nula
 - Ex: O diâmetro médio de parafusos produzidos na fábrica não é igual a 30mm ($H_1: \mu \neq 30$)
- Desafia o senso comum, ou seja, aquilo que as pessoas acreditam.
 - Rejeitamos a hipótese nula a favor da hipótese alternativa apenas quanto existe uma evidência muito forte.
- Em geral é a hipótese que o pesquisador está tentando provar!
- Apenas contém sinais de desigualdade: “ \neq ”, “ $>$ ” e “ $<$ ”

Procedimento de um Teste

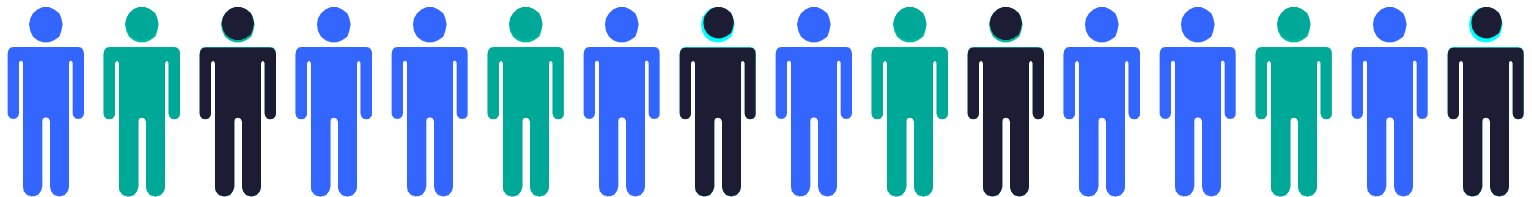
- Defina a afirmação a ser testada:

Exemplo: a idade média da população é 50 anos.

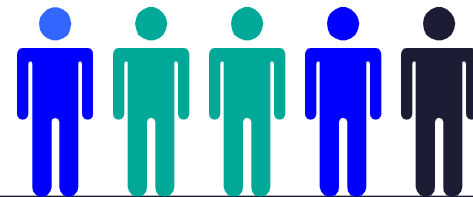
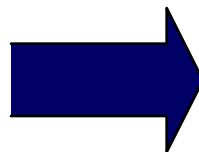
$$H_0: \mu = 50, \quad H_1: \mu \neq 50$$

- Pegue uma amostra da população e encontre a média amostral.

População:



Amostra:

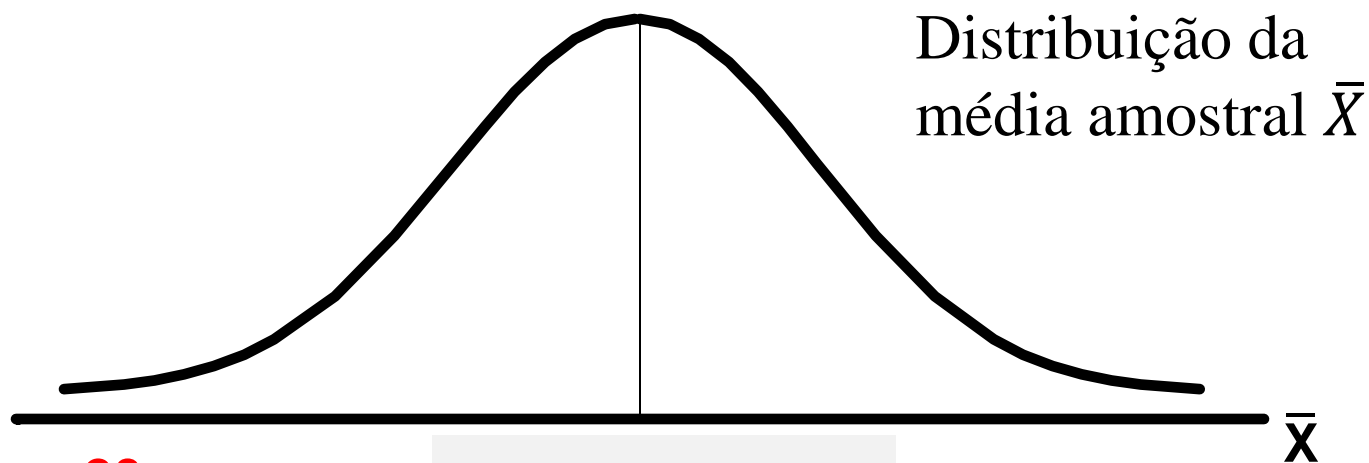




Procedimento de um Teste

- Suponha que a idade média da amostra seja $\bar{X} = 20$.
- Este valor é menor do que a afirmação de que a idade média da população é de 50, mas é um valor de uma amostra e amostras variam...
- Esta idade média na amostra é « suficientemente menor » para considerarmos que a hipótese nula é falsa?
- O que significa ser « suficientemente menor »?

Procedimento de um Teste



20



Quando for pouco provável encontrar uma amostra com este valor de média ou menos...

$\mu = 50$
Se H_0 é verdadeira



... se de fato esta fosse a média populacional...

... Então nós rejeitamos a hipótese nula de que $\mu = 50$.



Procedimento de um Teste

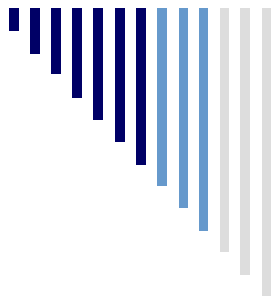
- Começamos o teste assumindo que a hipótese nula é verdadeira.
- Se este for o caso, a probabilidade de encontrarmos uma média amostral tão longe seria muito pequena, então rejeitamos a hipótese nula.
- Em outras palavras, uma média amostral de 20 é tão pouco provável se a verdadeira média da população fosse 50, que nós concluimos que a média da população não deve ser 50.



Procedimento de um teste

- 2 abordagens alternativas são possíveis:
 - Valor crítico
 - P-valor

- As duas abordagens levam a mesma conclusão.



ABORDAGEM POR VALOR CRÍTICO



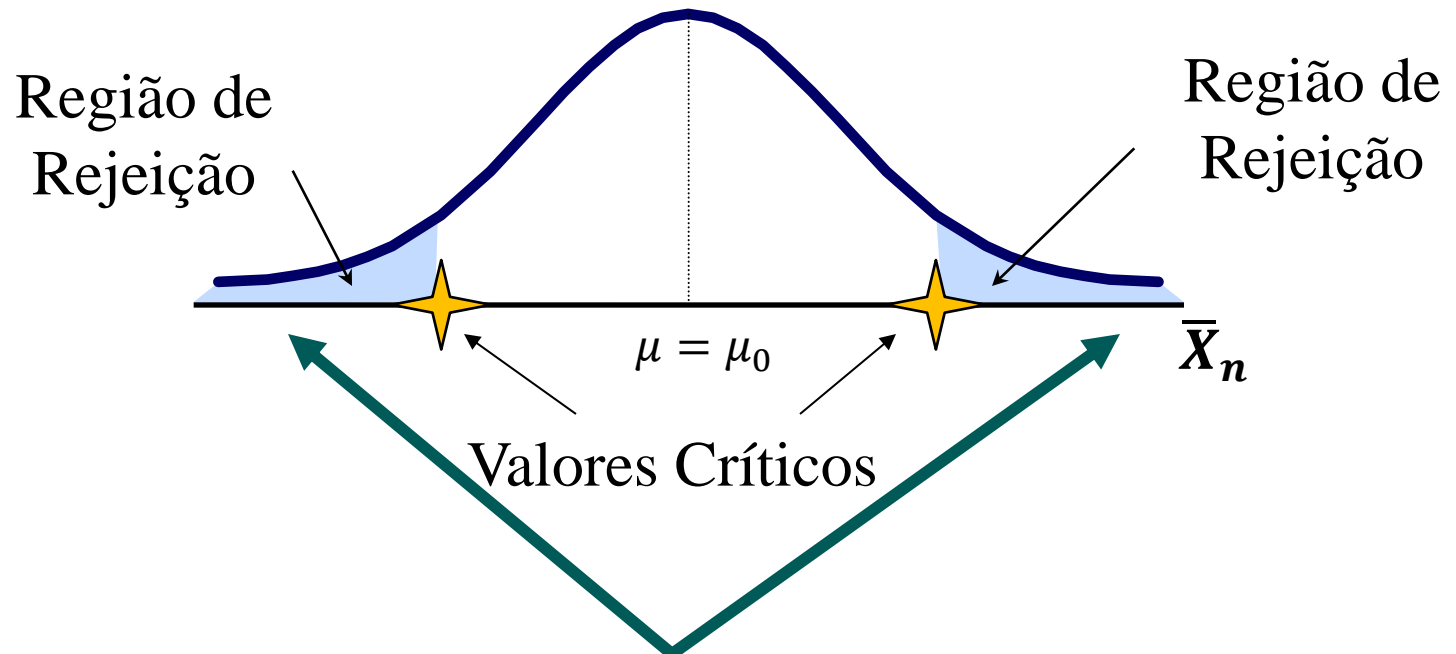
Estatística de Teste e Valores Críticos

Procedimento por valor crítico:

- Se a **média amostral**, \bar{X}_n , é **próxima da média populacional**, μ , **suposta na Hipótese nula**, então a hipótese nula **não é rejeitada**.
- Se a **média amostral**, \bar{X}_n , estiver longe **da média populacional**, μ , **suposta na Hipótese nula**, então a hipótese nula **é rejeitada**.
- **O quão longe é “longe o suficiente” para rejeitar H_0 ?**
- **O valor crítico** de uma estatística de teste cria uma “linha limítrofe” para a tomada de decisão – ela responde a pergunta de o quão longe é longe o suficiente.

Estatística de Teste e Valores Críticos

Distribuição da média amostral, supondo que a hipótese nula é verdadeira, $\mu = \mu_0$:



Região de Rejeição: \bar{X}_n é “Longe Demais” da Média da Distribuição Amostral (μ_0) se H_0 fosse verdadeira.



Estatística de Teste e Valores Críticos

- Suponha que encontramos um valor de estatística de teste na amostra que está na região de rejeição
 - Ex: média amostral igual a 20, quando H_0 supõe que a média populacional é igual a 50
- Existem duas possibilidades para encontrarmos um valor tão baixo:
 - A verdadeira média realmente é 50, nós que pegamos uma amostra muito atípica
 - A verdadeira média não é 50.
- Quando decidimos rejeitar a hipótese nula, podemos ter cometido um erro: de fato a amostra que era atípica. **Nunca saberemos qual é a verdade!**



Erros em Testes de Hipóteses

Possíveis resultados de um julgamento		
	Situação Real (desconhecida)	
Decisão	Réu é inocente	Réu é culpado
Não condenar Réu	Decisão certa	Erro: Libertar um culpado.
Condenar Réu	Erro: Mandar inocente para a cadeia!	Decisão certa



Erros em Testes de Hipóteses

Possíveis resultados de um julgamento		
	Situação Real	
Decisão	Réu é inocente	Réu é culpado
Não condenar Réu	Decisão Certa	Erro Tipo II: Libertar um culpado.
Condenar Réu	Erro Tipo I: Mandar inocente para a cadeia!	Decisão Certa

Erros em Testes de Hipóteses



Para julgamento:

H_0 : Réu é inocente, H_1 : Réu não é inocente

Erro Tipo I

- Definido como: **Rejeitar uma hipótese nula verdadeira**
- Ex: mandar um inocente para a cadeia
- Considerado um tipo de erro muito grave
- Probabilidade do Erro Tipo I é α
 - Chamado de nível de significância de um teste
 - Escolhido pelo pesquisador antes do teste

Definição!

Erros em Testes de Hipóteses



Para julgamento:

H_0 : Réu é inocente, H_1 : Réu não é inocente

▪ Erro Tipo II

- Definido como: Não rejeitar uma hipótese nula falsa
- Ex: não mandar um culpado para a cadeia
- A probabilidade de Erro Tipo II é β
- Não é controlada diretamente pelo pesquisador, pois depende do verdadeiro valor do parâmetro.

Definição!



Erros em Testes de Hipóteses

Possíveis Resultados de um Teste de hipóteses		
	Situação Real	
Decisão	H_0 é Verdadeira	H_0 é Falsa
Não Rejeitar H_0	Decisão Certa Pr: $1 - \alpha$	Erro Tipo II Pr: β
Rejeitar H_0	Erro Tipo I Pr: α	Decisão Certa Pr: $1 - \beta$



Erros em Testes de hipóteses

- O **coeficiente de confiança** $(1-\alpha)$ é a probabilidade de não rejeitar H_0 quando ela é verdadeira.

Definição!

- O nível de confiança de um teste de hipóteses é de $(1-\alpha)*100\%$, em que α é o nível de significância ou a probabilidade de Erro tipo I.

- O **poder estatístico de um teste** $(1-\beta)$ é a probabilidade de rejeitar H_0 quando ela é falsa. Um teste com baixo poder é um teste ruim.

Definição!



Relação entre Erros Tipo I e Tipo II

- Os Erros Tipo I e II não podem acontecer ao mesmo tempo:
 - O Erro Tipo I só acontece quando H_0 é verdadeira
 - O Erro Tipo II só acontece quando H_0 é falsa
- As probabilidades de erro tipo I e II estão relacionadas entre si e dependem do grau de exigência que usamos para o procedimento do teste!
- Se a probabilidade de Erro Tipo I (α) **diminui**, então, a probabilidade de Erro Tipo II (β) **aumenta**.



Relação entre Erros Tipo I e Tipo II

- Se o teste exige provas muito fortes para rejeitar H_0 (menor erro tipo I), então necessariamente fica mais fácil não rejeitarmos H_0 quando deveríamos (erro tipo II);
 - É muito fácil construir um teste com erro tipo I baixo: basta nunca rejeitar H_0 (não prender ninguém)! Porém, este teste comete bastante erro tipo II (vários criminosos serão soltos).
- **Testes de hipóteses controlam diretamente a probabilidade de erro tipo I!**
- **Por isso, devemos escrever as hipóteses de forma que o erro mais grave seja o erro tipo I.**

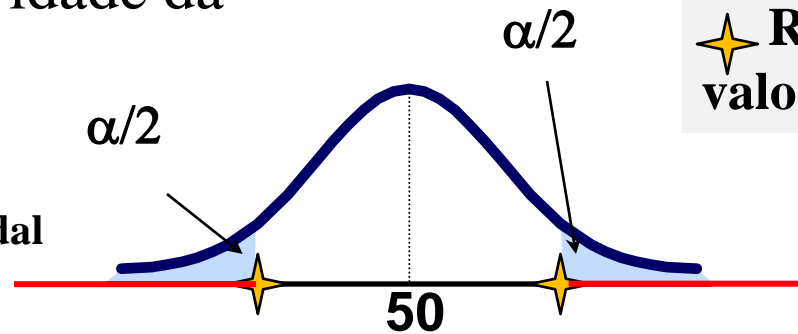
Nível de Significância, α

Afirmção: A média de idade da população é 50.

$$H_0: \mu = 50$$

$$H_1: \mu \neq 50$$

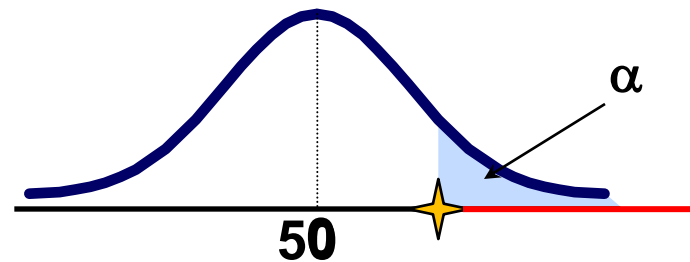
Teste bi-caudal



$$H_0: \mu \leq 50$$

$$H_1: \mu > 50$$

Teste de cauda Superior

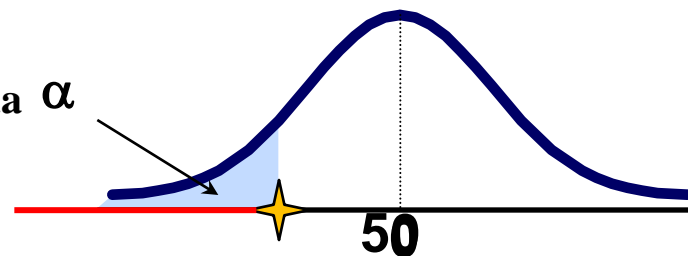


A região de rejeição está em vermelho

$$H_0: \mu \geq 50$$

$$H_1: \mu < 50$$

Teste de cauda Inferior



Teste Z: σ conhecido

Teste de hipóteses
para μ

σ conhecido
(teste Z)

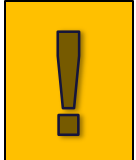
σ desconhecido
(teste t)

A estatística de teste é

$$Z_{amostra} = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0, 1)$$

sob certas condições. Quais?

Teste Z: σ conhecido



Para um teste bi-caudal para a média com σ conhecido se a população tem distribuição normal, ou o tamanho da amostra é grande:

- Converta a estatística amostral (\bar{X}_n) para a estatística de teste:

$$Z_{amostra} = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}$$

- Determine os valores críticos para distribuição normal padronizada a partir da tabela para um dado nível de significância α .
- **Decisão:** Se a estatística de teste cair na região de rejeição, rejeite H_0 ; caso contrário não rejeite H_0 .

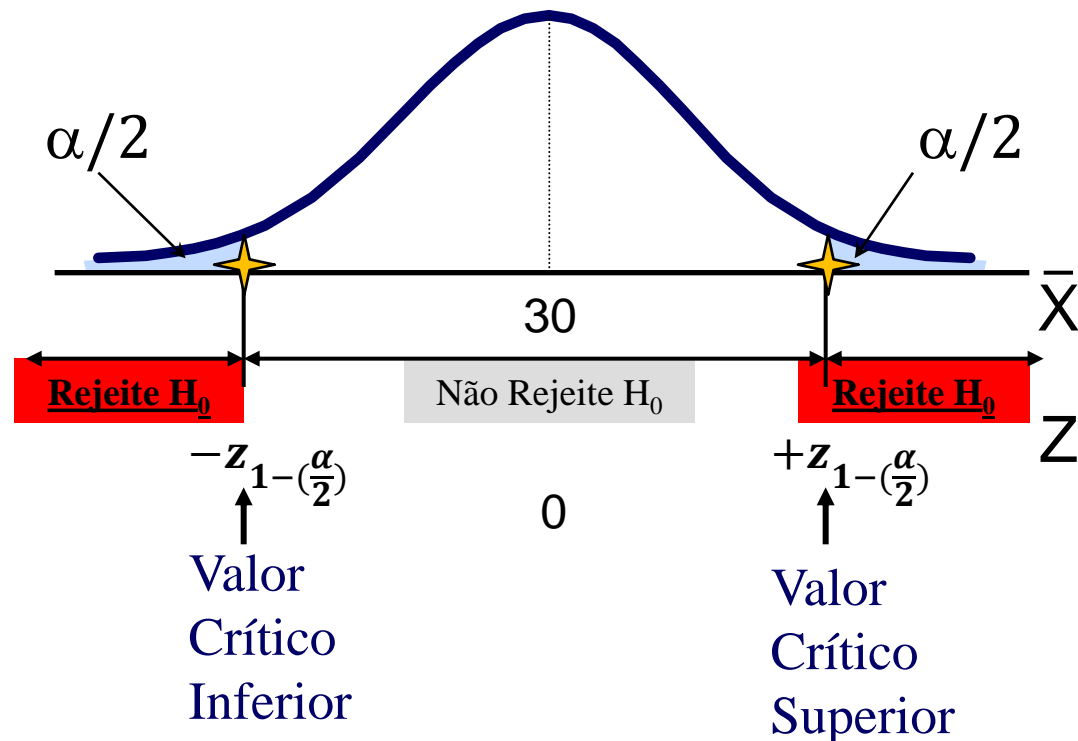
Teste Z: σ conhecido

Teste bicaudal:

- Para o teste bicaudal, existem dois valores limiares (valores críticos), que definem a região de rejeição.

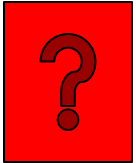
$$H_0: \mu = 30$$

$$H_1: \mu \neq 30$$





Teste Z: σ conhecido



Exercício: Teste a afirmação de que o verdadeiro diâmetro médio de um parafuso produzido em uma fábrica é de 30mm a um nível de significância de 5%.

Assuma que $\sigma = 0.8$ e que, para uma amostra de 100 parafusos observou-se uma média de 29.84 mm.



Teste Z: σ conhecido

Solução:

- 1) Estabeleça as hipóteses nula e alternativa apropriadas. Como o teste é para ver se a média é igual a 30 mm, temos que considerar que o diâmetro « real » pode tanto ser maior quanto ser menor do que 30mm. Assim:
 $H_0: \mu = 30$ $H_1: \mu \neq 30$ (Teste bicaudal)
- 2) O nível de significância $\alpha = 0.05$ foi indicado para o teste
- 3) A amostra utilizada para o teste tem $n = 100$ parafusos.



Teste Z: σ conhecido

- 4) Como $\sigma = 0.8$ é conhecido historicamente pela fábrica, usamos um teste Z
- 5) Assim, para $\alpha = 0.05$ os valores críticos de uma $N(0,1)$ são ± 1.96
- 6) Use os dados para computar a estatística de teste:
A amostra ($n=100$) resultou em $\bar{x}_n = 29.84$
Então a estatística de teste é:

$$Z_{amostra} = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{29.84 - 30}{\frac{0.8}{\sqrt{100}}} = -2.0$$

Teste Z: σ conhecido

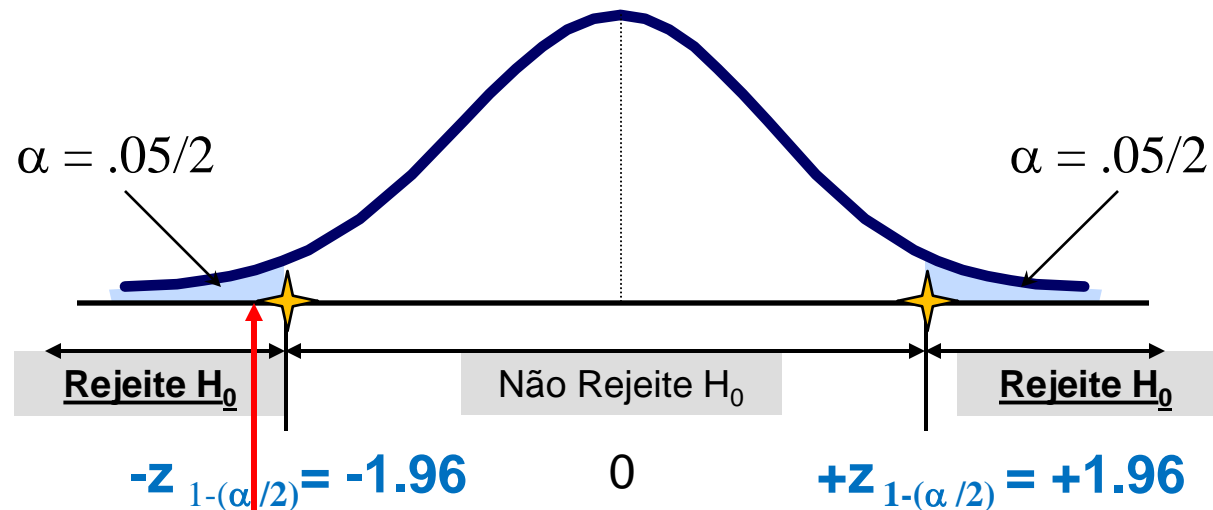
7) A estatística de teste pertence a região de rejeição?

Rejeite H_0 se:

$Z_{amostra} < -1.96$ ou

$Z_{amostra} > 1.96$;

Caso contrário não rejeite H_0 .



Aqui, $z_{amostra} = -2.0 < -1.96$, tal que a estatística de teste cai na região de rejeição.



Teste Z: σ conhecido

- 8) Chegue a uma conclusão e interprete o resultado:

Visto que a estatística de teste cai na região de rejeição do teste ($z_{amostra} = -2.0 < -1.96$), podemos rejeitar a hipótese nula e concluir que existe evidência suficiente de que o diâmetro médio de parafusos produzidos pela fábrica não é igual a 30 mm.

Teste Z: σ conhecido



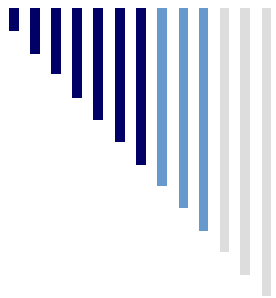
7 Passos de um Teste de hipóteses:

1. Estabeleça a hipótese nula, H_0 , e a hipótese alternativa, H_1 ;
2. Escolha o nível de significância, α , e o tamanho da amostra, n ;
3. Determine a estatística de teste apropriada e a sua distribuição amostral;
4. Determine os valores críticos que separam as regiões de rejeição e de não-rejeição;

Teste Z: σ conhecido



5. Colete os dados e compute o valor da estatística de teste;
6. Tome uma decisão estatística:
 - Se a estatística de teste cair na região de não-rejeição, então não é possível rejeitar a hipótese nula, H_0 .
 - Se a estatística de teste cair na região de rejeição, então rejeite a hipótese nula em favor da hipótese alternativa, H_1 .
7. Expresse a conclusão resultante da decisão no contexto do problema.



ABORDAGEM POR P-VALOR



Abordagem por p-valor

- Uma outra forma de realizar um teste de hipóteses é através do p-valor (em vez de usar valores críticos)
- **p-valor:** Probabilidade de obter uma estatística de teste igual a ou ainda mais extrema do que a observada na amostra supondo que H_0 é verdadeira
 - O p-valor também é chamado de nível de significância observado na amostra, pois:
 - é o menor valor de α para o qual H_0 não pode ser rejeitada.

Definição!

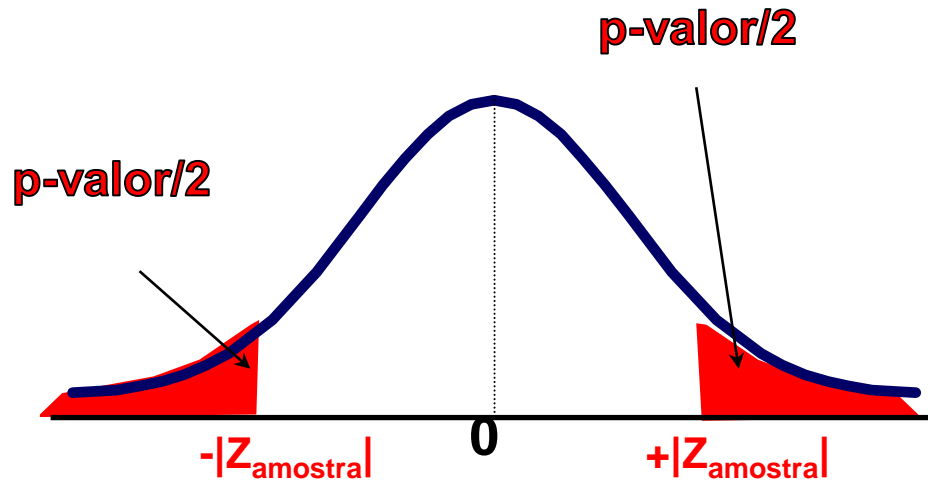
Abordagem por p-valor

Afirmação: A média de idade da população é igual a 50.

Teste bi-caudal:

$$H_0: \mu = 50$$

$$H_1: \mu \neq 50$$



Para o teste bicaudal, o p-valor é a soma das áreas hachuradas para a cauda superior e inferior!

Abordagem por p-valor



Procedimento de teste usando o p-valor:

- 1) Converta a estatística amostral (\bar{X}_n) na estatística de teste ($Z_{amostra}$)
- 2) Obtenha o **p-valor** de uma tabela ou usando um programa
- 3) Compare o **p-valor** a α
 - Se **p-valor** $< \alpha$, rejeite H_0
 - Se **p-valor** $\geq \alpha$, não rejeite H_0

Portanto, se o p-valor é suficientemente pequeno (menor do que α), o que significa que a média da amostra está muito longe da suposta média da população (μ_0), então H_0 pode ser rejeitada.



Teste Z: σ conhecido



Exercício: Teste a afirmação, usando a abordagem por p-valor, de que o verdadeiro diâmetro médio de um parafuso produzido em uma fábrica é de 30mm a um nível de significância de 5%.

Assuma que $\sigma = 0.8$ e que, para uma amostra de 100 parafusos observou-se uma média de 29.84 mm.

Abordagem por p-valor: σ conhecido

- **Cálculo do p-valor:** Qual a probabilidade de "encontrarmos uma média amostral igual a 29.84 ou algum valor ainda mais afastado da média $\mu = 30$ nas "duas direções" se H_0 é verdadeira?

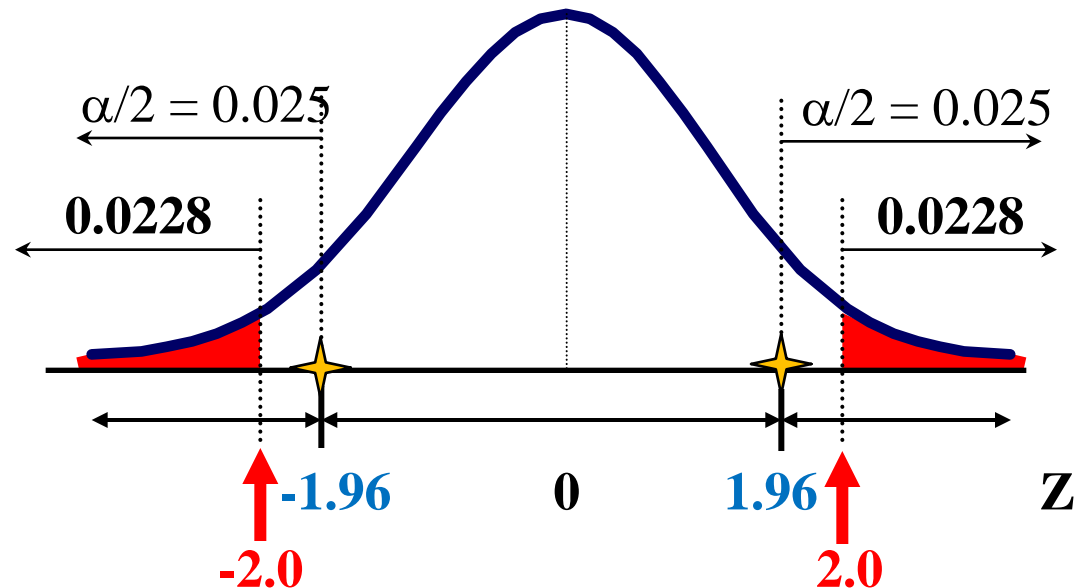
1) $\bar{x}_n = 29.84$ é transformado em escore Z:
 $Z_{amostra} = -2.0$

2) Cálculo do p-valor:

$$P(Z < -2.0) = 0.0228$$

$$P(Z > +2.0) = 0.0228$$

$$\begin{aligned} \text{p-valor} &= 0.0228 + 0.0228 \\ &= 0.0456 \end{aligned}$$



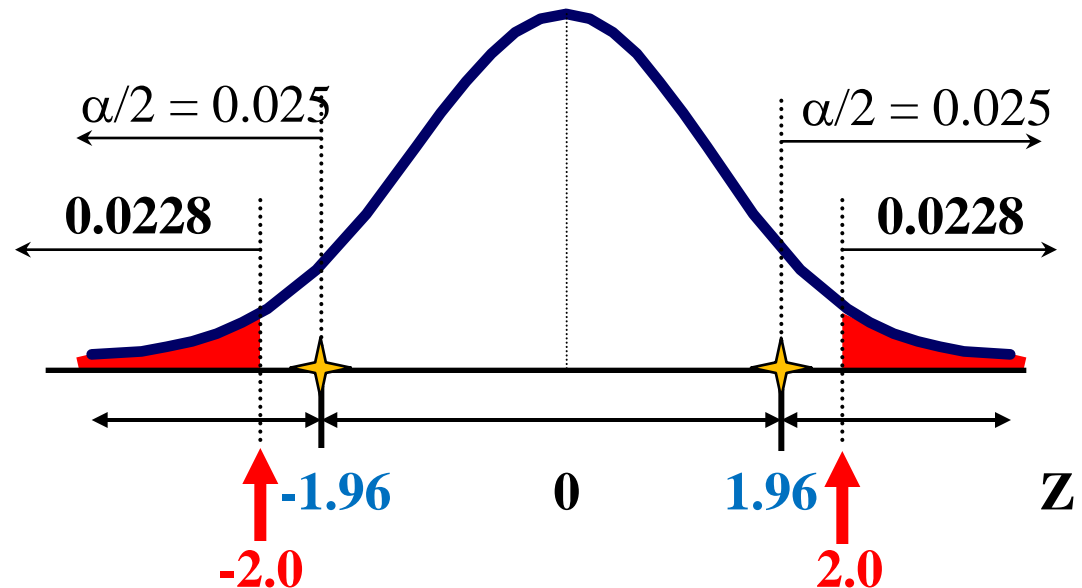
Abordagem por p-valor: σ conhecido

3) Compare o p-valor a α

- Se o *p – valor* $< \alpha$, rejeite H_0
- Se o *p – valor* $\geq \alpha$, não rejeite H_0

Aqui: p-valor=0.0456 e
 $\alpha = 0.05$

**Já que $0.0456 < 0.05$,
rejeitamos a hipótese
nula**





Teste de hipóteses: Relação com Intervalo de Confiança

- Para $\bar{x}_n = 29.84$, $\sigma = 0.8$ e $n = 100$, o intervalo de confiança de 95% para a média é:

$$\left(29.84 - (1.96) \frac{0.8}{\sqrt{100}}, 29.84 + (1.96) \frac{0.8}{\sqrt{100}} \right)$$

$$29.6832 \leq \mu \leq 29.9968$$

- Uma vez que este intervalo não contém a média idealizada (30), rejeitamos a hipótese nula a $\alpha = 0.05$



Teste Z: σ conhecido e teste unicaudal

- Em muitos casos, a hipótese alternativa foca em uma direção particular.

$$H_0: \mu \geq 30$$

$$H_1: \mu < 30$$



Este é um teste de *cauda inferior* já que a hipótese alternativa estabelece valores *abaixo* da média de 30

$$H_0: \mu \leq 30$$

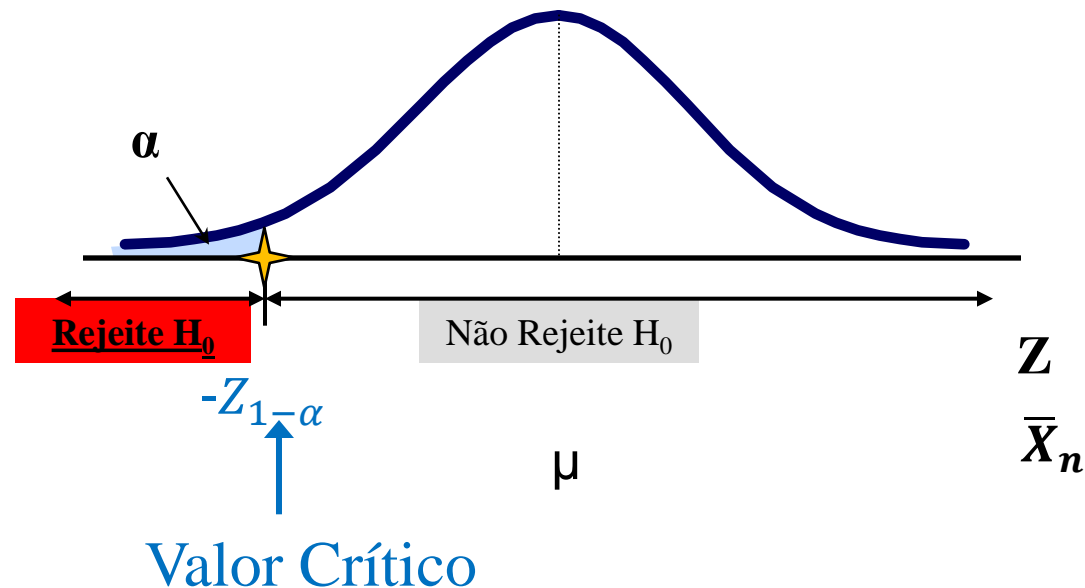
$$H_1: \mu > 30$$



Este é um teste de *cauda superior*, já que a hipótese alternativa estabelece valores *acima* da média de 30

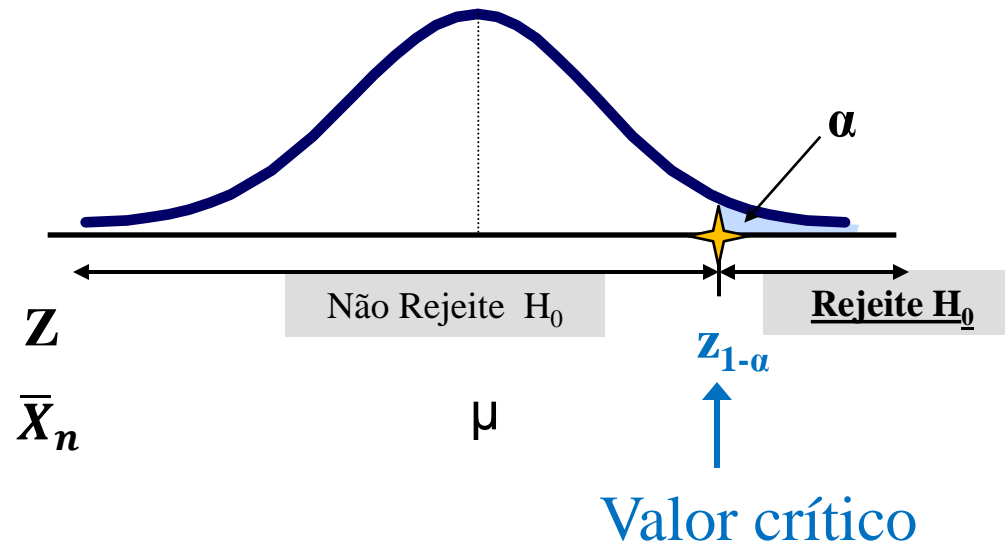
Teste Z: σ conhecido e teste de cauda inferior

- Existe apenas um único valor crítico, uma vez que a região de rejeição possui apenas uma cauda.



Teste Z: σ conhecido e teste de cauda superior

- Existe apenas um único valor crítico, uma vez que a região de rejeição possui apenas uma cauda.



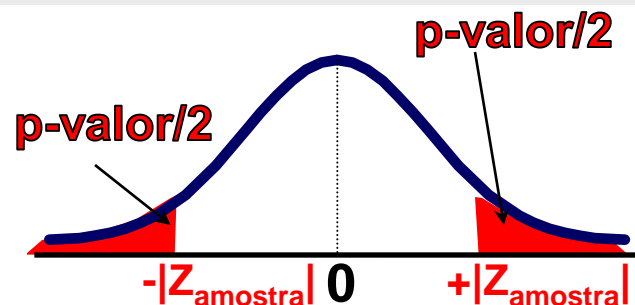
Abordagem por p-valor

Com σ conhecido:

$$H_0: \mu = 50$$

$$H_1: \mu \neq 50$$

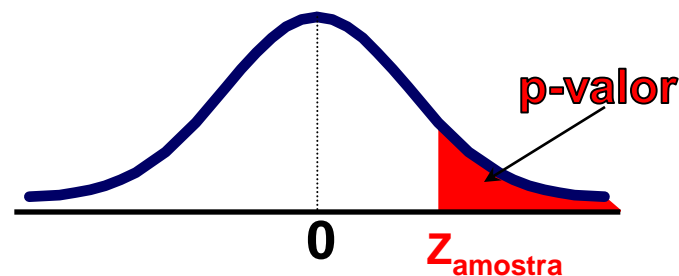
Teste bi-caudal



$$H_0: \mu \leq 50$$

$$H_1: \mu > 50$$

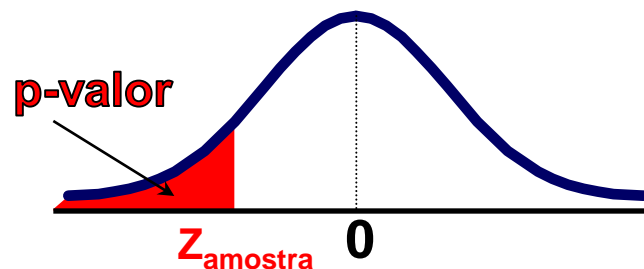
Teste de cauda Superior



$$H_0: \mu \geq 50$$

$$H_1: \mu < 50$$

Teste de cauda Inferior



O p-valor é a área hachurada!



Teste Z: σ conhecido e teste de cauda superior



Exercício: um gerente da indústria de telecomunicações acha que a conta mensal média de celular **aumentou** e passou a ser **maior** do que R\$52 por mês. O gerente deseja testar esta afirmação. Registros passados da companhia indicam que o desvio-padrão de todas as contas é de aproximadamente R\$10.

Suponha que a amostra obtida leva aos resultados:

$$n = 64, \bar{x}_{64} = 53.1$$

Use um nível de significância de 10%.



Teste Z: σ conhecido e teste de cauda superior

Solução:

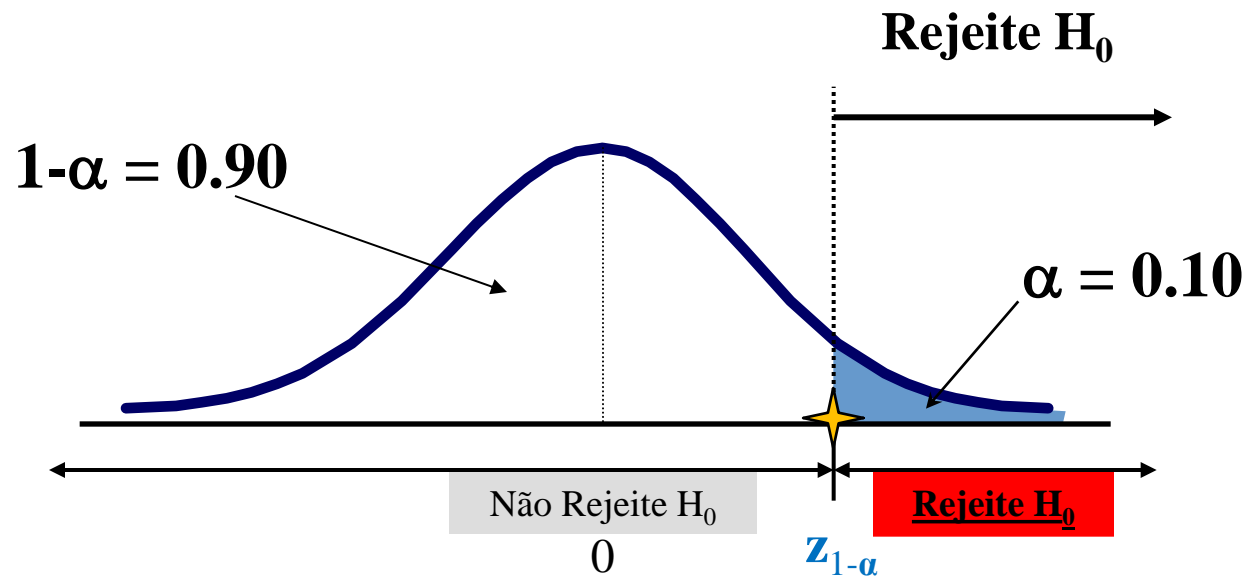
1) Formule as hipóteses do teste (lembre-se que a nova teoria que desafia o senso comum deve ser colocada na hipótese alternativa)

$H_0: \mu \leq 52$ a média é menor ou igual a R\$52 por mês

$H_1: \mu > 52$ a média é maior do que R\$52 por mês.
(i.e., existe evidência suficiente para suportar a afirmação do gerente.)

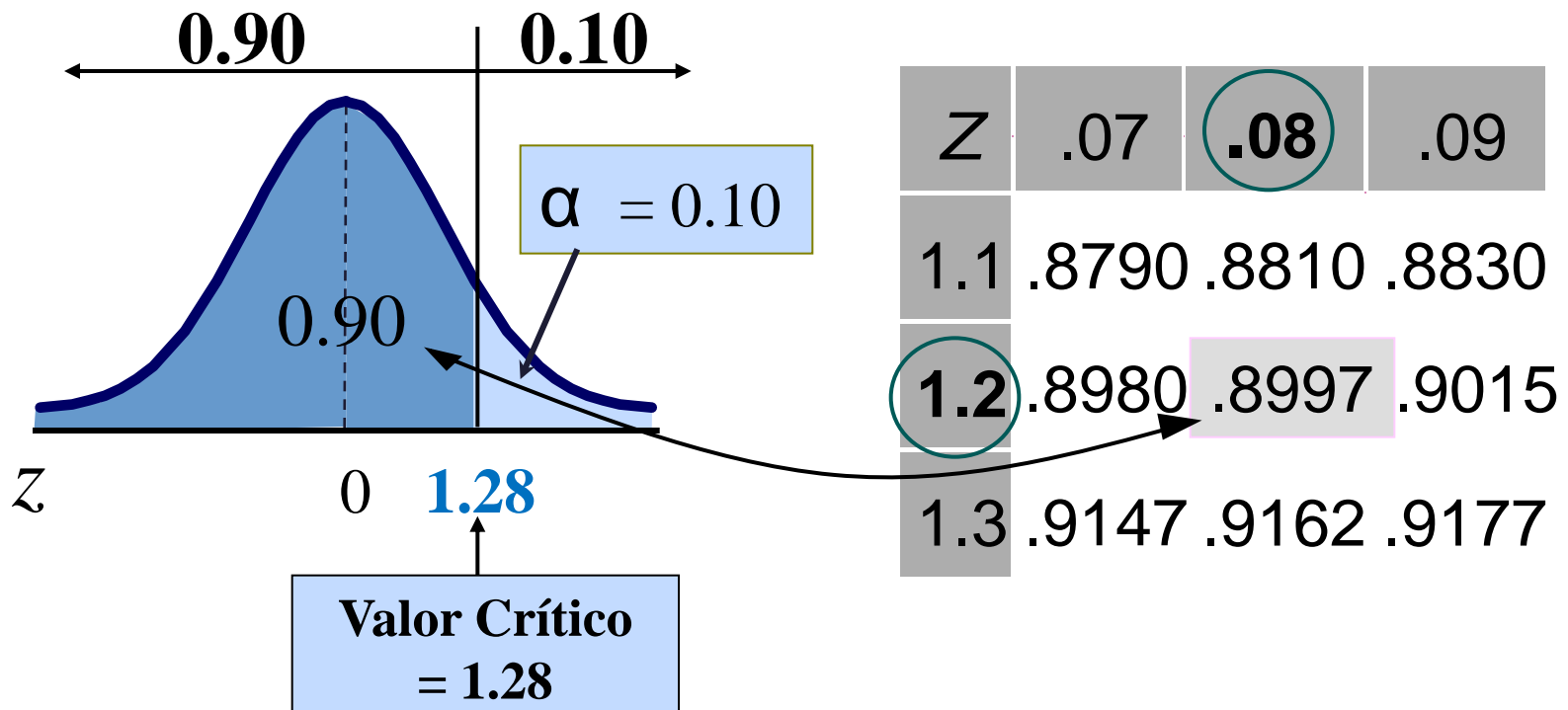
Teste Z: σ conhecido e teste de cauda superior

2) Encontre a região de rejeição para o teste de cauda superior:



Teste Z: σ conhecido e teste de cauda superior

Qual o valor crítico $z_{1-\alpha}$ dado $\alpha = 0.10$?





Teste Z: σ conhecido e teste de cauda superior

3) Obtenha uma amostra e compute a estatística de teste.

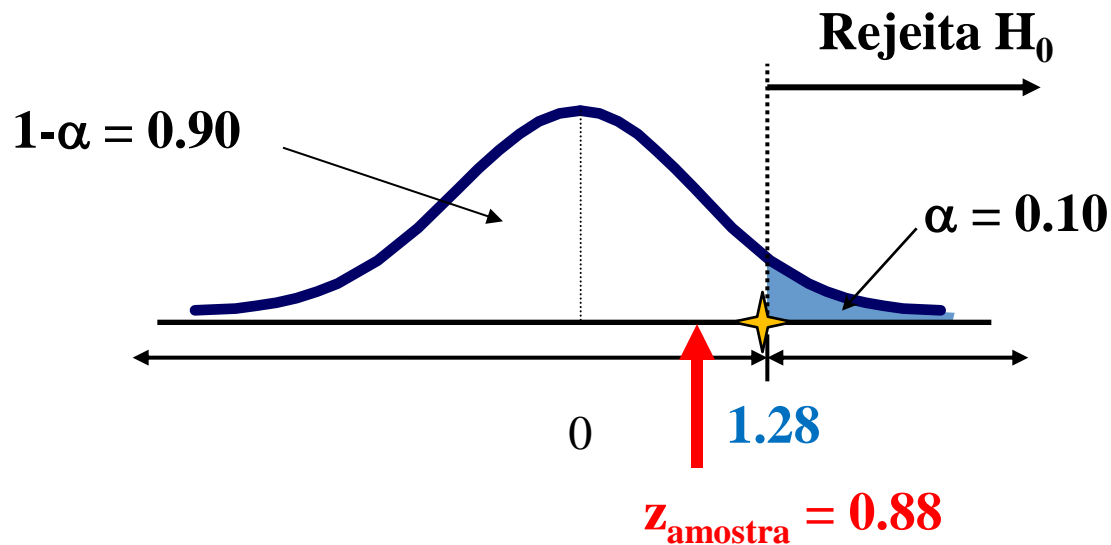
Suponha que a amostra obtida leva aos resultados: $n = 64$, $\bar{x}_{64} = 53.1$ ($\sigma=10$ é conhecido a partir de registros prévios da companhia)

- A estatística de teste é:

$$Z_{amostra} = \frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{53.1 - 52}{\frac{10}{\sqrt{64}}} = 0.88$$

Teste Z: σ conhecido e teste de cauda superior

4) Tome uma decisão e interprete o resultado:

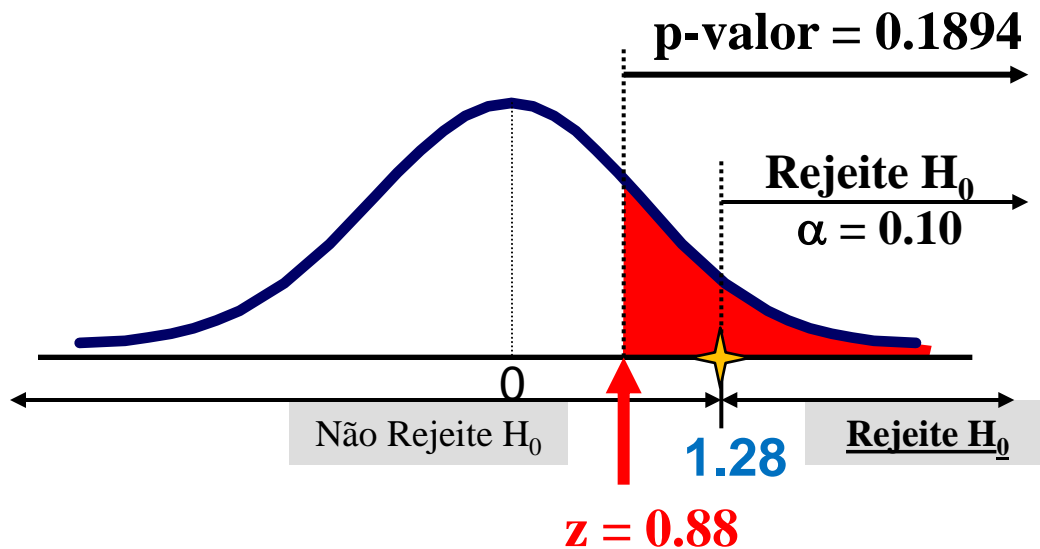


Não rejeitamos H_0 uma vez que $z_{\text{amostra}} = 0.88 \leq 1.28$

i.e.: não existe evidência o suficiente para corroborar a afirmação que a conta média mensal é maior do que R\$52.

Teste Z: σ conhecido e teste de cauda superior

Alternativamente, calcule o p-valor e compare a α :



$$\begin{aligned} P(\bar{X}_n \geq 53.1) &= \\ P\left(Z \geq \frac{53.1 - 52}{\frac{10}{\sqrt{64}}}\right) &= \\ P(Z \geq 0.88) &= \\ 1 - 0.8106 &= \\ 0.1894 & \end{aligned}$$

Não rejeitamos H_0 uma vez que $p\text{-valor} = 0.1894 > \alpha = 0.10$



Teste Z: σ conhecido e teste de cauda inferior



Exercício: um gerente da indústria de telecomunicações acha que a conta mensal média de celular **diminuiu** e passou a ser **menor** do que R\$52 por mês. O gerente deseja testar esta afirmação. Registros passados da companhia indicam que o desvio-padrão das contas é de aproximadamente R\$10.

Suponha que uma amostra de 100 contas resultou em uma média de R\$ 51.8.

- a) Faça o teste de hipóteses pela abordagem do valor crítico usando um nível de significância de 5%.
- b) Calcule o p-valor e refaça o teste usando um nível de significância de 10%.



Conhecemos o verdadeiro σ ?

Provavelmente não!

- Em praticamente todas as situações reais, σ é desconhecido.
- Se existe uma situação em que σ é conhecido, então geralmente μ também é conhecido (uma vez que para calcular σ precisamos saber μ .)
- Se você realmente conhece a média populacional μ , não é necessário pegar uma amostra para estimá-la.



Teste de hipóteses: σ desconhecido

- Desvio-padrão populacional é desconhecido: substituir pelo desvio amostral, S .
- Então, usamos distribuição T de Student em vez da distribuição Z (normal padrão), para o escore-Z.
- Para usar a distribuição T
 - a distribuição da população que estamos amostrando é normal;
 - ou o tamanho da amostra é suficientemente grande para usar o TLC.



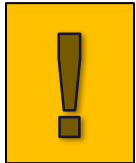
Teste de hipóteses: σ desconhecido

Todos os outros passos, conceitos e conclusões são os mesmos!

- O teste pode ser bicaudal, de cauda superior ou de cauda inferior.
- Podemos usar abordagem por valor crítico ou por p-valor.



Teste de hipóteses: σ desconhecido



Se: a população X tem distribuição normal ou a amostra é grande o suficiente; se o desvio-padrão populacional é σ e média igual a μ ,

Então:

a estatística de teste, $T_{amostra}$, tem distribuição T com $n-1$ graus de liberdade:

$$T_{amostra} = \frac{\bar{X}_n - \mu}{\frac{S}{\sqrt{n}}} \sim T_{n-1}$$

Teste Z: σ desconhecido



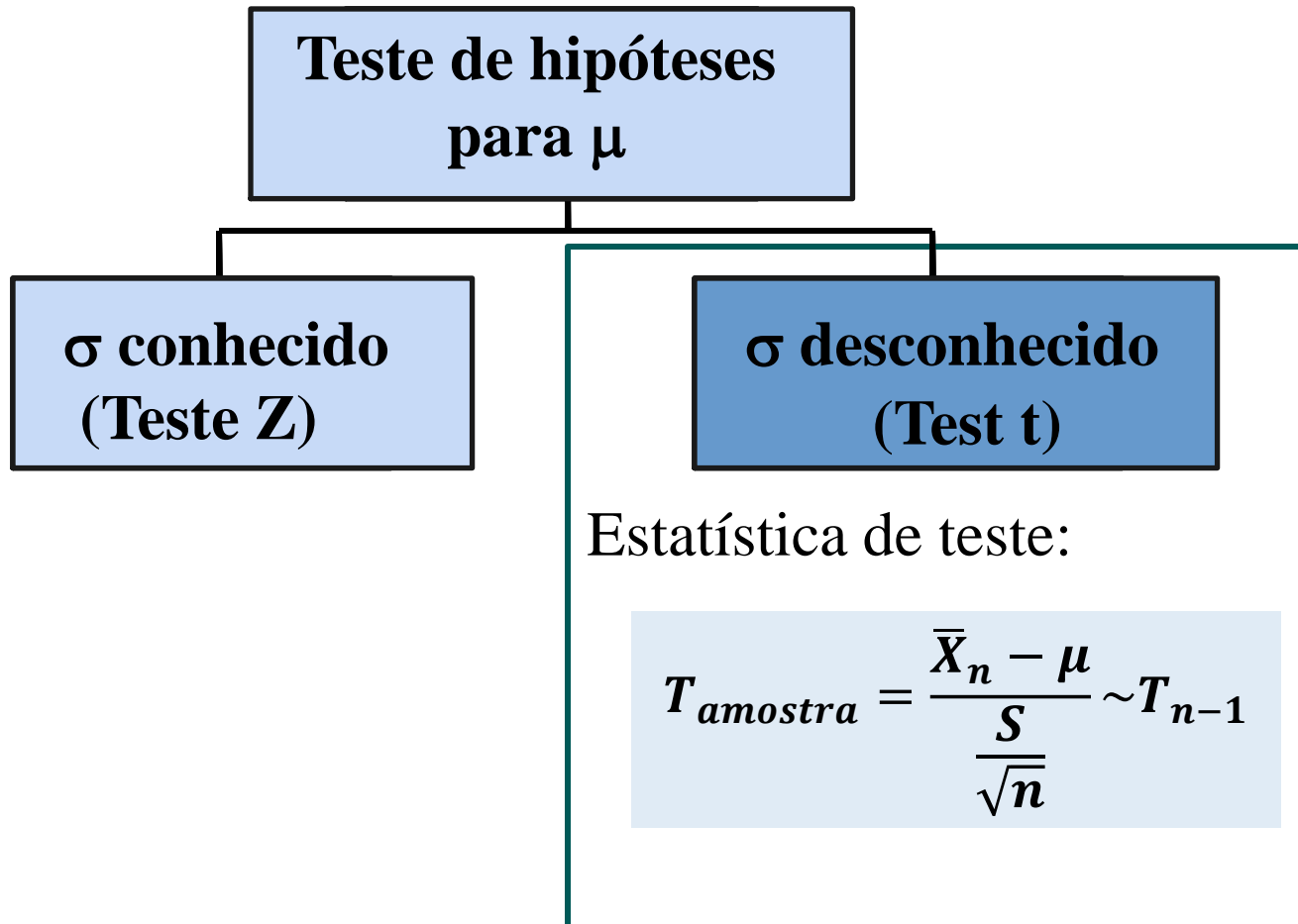
Para um teste para a média com σ **desconhecido** e população normal ou n grande:

- Converta a estatística amostral (\bar{X}_n) para a estatística de teste:

$$T_{amostra} = \frac{\bar{X}_n - \mu}{\frac{S}{\sqrt{n}}}$$

- Determine o(s) valor(es) crítico(s) para uma t de Student com $n - 1$ graus de liberdade a partir da tabela para um dado nível de significância α .
- **Decisão:** Se a estatística de teste cair na região de rejeição, rejeite H_0 ; caso contrário não rejeite H_0 .

Teste t: σ desconhecido





Teste t: σ desconhecido



Exercício: Diz-se que o custo médio de um quarto de hotel em NY é de \$168 por noite. Uma amostra aleatória de 25 hotéis resultou em $\bar{x} = \$172.50$ e $s = \$15.40$. Teste a afirmação a $\alpha = 0.05$.

(Um histograma dos dados indica que o preço médio tem distribuição aproximadamente normal)

$$H_0: \mu = 168$$

$$H_1: \mu \neq 168$$

Teste t: σ desconhecido

$$H_0: \mu = 168$$

$$H_1: \mu \neq 168$$

$$\alpha = 0.05$$

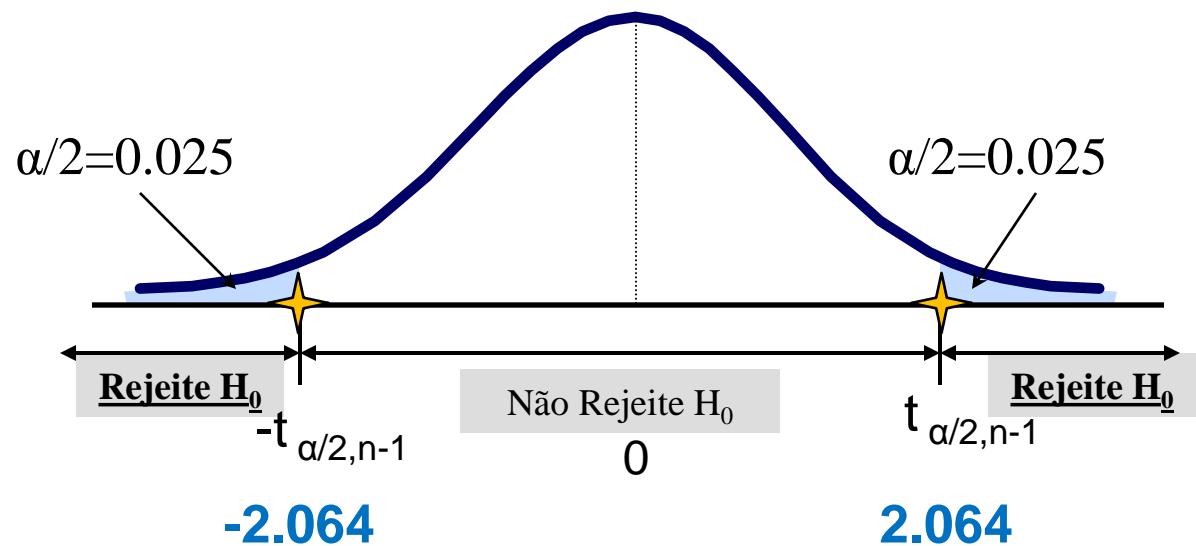
$$n = 25$$

σ é desconhecido,
então use
estatística T

Valores críticos:

$$t_{24,0.025} = \pm 2.064$$

Determine a região de rejeição:



Distribuição t de Student

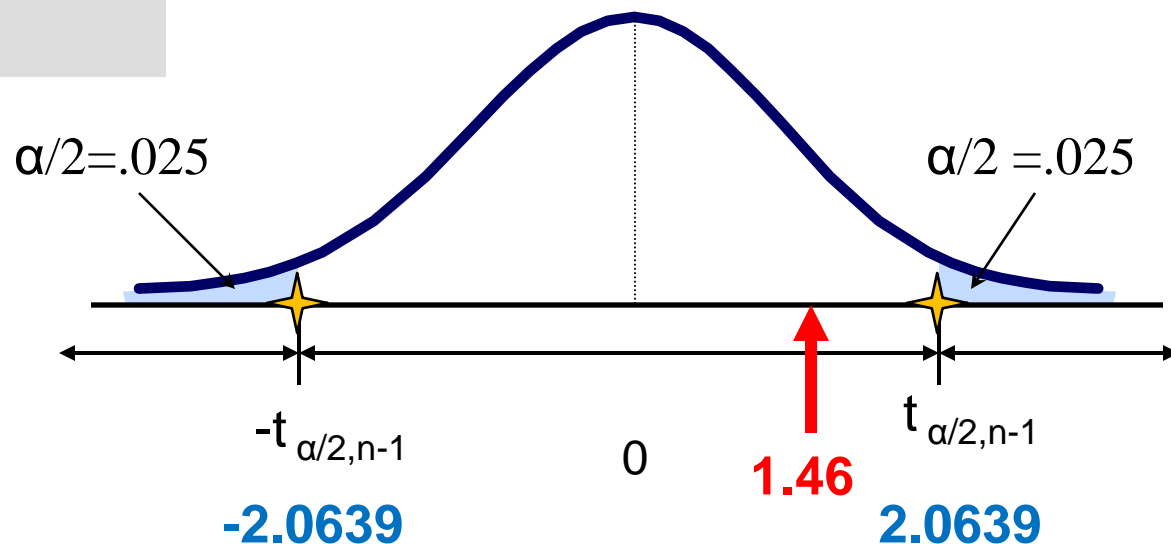


gl	0,25	0,10	0,05	0,025	0,01	0,0083	0,005
1	1,000	3,078	6,314	12,706	31,821	38,343	63,656
2	0,816	1,886	2,920	4,303	6,965	7,664	9,925
3	0,765	1,638	2,353	3,182	4,541	4,864	5,841
4	0,741	1,533	2,132	2,776	3,747	3,966	4,604
5	0,727	1,476	2,015	2,571	3,365	3,538	4,032
6	0,718	1,440	1,943	2,447	3,143	3,291	3,707
7	0,711	1,415	1,895	2,365	2,998	3,130	3,499
8	0,706	1,397	1,860	2,306	2,896	3,018	3,355
9	0,703	1,383	1,833	2,262	2,821	2,936	3,250
10	0,700	1,372	1,812	2,228	2,764	2,872	3,169
11	0,697	1,363	1,796	2,201	2,718	2,822	3,106
12	0,695	1,356	1,782	2,179	2,681	2,782	3,055
13	0,694	1,350	1,771	2,160	2,650	2,748	3,012
14	0,692	1,345	1,761	2,145	2,624	2,720	2,977
15	0,691	1,341	1,753	2,131	2,602	2,696	2,947
16	0,690	1,337	1,746	2,120	2,583	2,675	2,921
17	0,689	1,333	1,740	2,110	2,567	2,657	2,898
18	0,688	1,330	1,734	2,101	2,552	2,641	2,878
19	0,688	1,328	1,729	2,093	2,539	2,627	2,861
20	0,687	1,325	1,725	2,086	2,528	2,614	2,845
21	0,686	1,323	1,721	2,080	2,518	2,603	2,831
22	0,686	1,321	1,717	2,074	2,508	2,593	2,819
23	0,685	1,319	1,714	2,069	2,500	2,584	2,807
24	0,685	1,318	1,711	2,064	2,492	2,575	2,797
25	0,684	1,316	1,708	2,060	2,485	2,568	2,787
26	0,684	1,315	1,706	2,056	2,479	2,561	2,779
27	0,684	1,314	1,703	2,052	2,473	2,554	2,771
28	0,683	1,313	1,701	2,048	2,467	2,548	2,763
29	0,683	1,311	1,699	2,045	2,462	2,543	2,756
30	0,683	1,310	1,697	2,042	2,457	2,537	2,750
35	0,682	1,306	1,690	2,030	2,438	2,516	2,724
40	0,681	1,303	1,684	2,021	2,423	2,501	2,704
45	0,680	1,301	1,679	2,014	2,412	2,488	2,690
50	0,679	1,299	1,676	2,009	2,403	2,479	2,678
60	0,679	1,296	1,671	2,000	2,390	2,465	2,660
70	0,678	1,294	1,667	1,994	2,381	2,454	2,648
80	0,678	1,292	1,664	1,990	2,374	2,447	2,639
90	0,677	1,291	1,662	1,987	2,368	2,441	2,632
100	0,677	1,290	1,660	1,984	2,364	2,436	2,626
110	0,677	1,289	1,659	1,982	2,361	2,433	2,621
120	0,677	1,289	1,658	1,980	2,358	2,430	2,617
130	0,676	1,288	1,657	1,978	2,355	2,427	2,614

Tabela que fornece valores t tais que $P(t_n > t) = p$, onde n é o número de graus de liberdade

Teste t: σ desconhecido

$$t_{\text{amostra}} = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{172.50 - 168}{\frac{15.40}{\sqrt{25}}} = 1.46$$



Não rejeitamos H_0 : não temos evidência o suficiente de que o verdadeiro custo médio é diferente de \$168

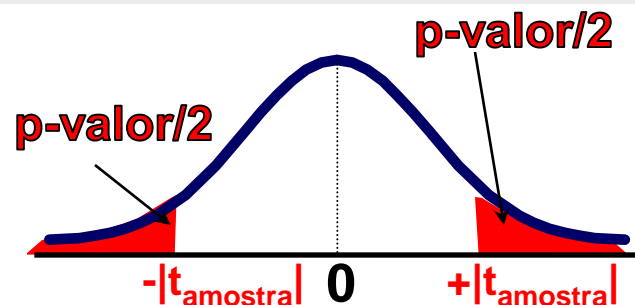
Abordagem por p-valor

Com σ desconhecido:

$$H_0: \mu = 50$$

$$H_1: \mu \neq 50$$

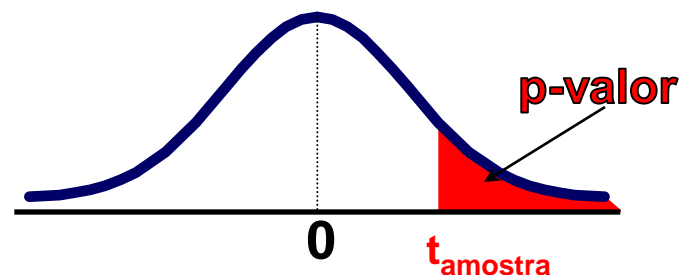
Teste bi-caudal



$$H_0: \mu \leq 50$$

$$H_1: \mu > 50$$

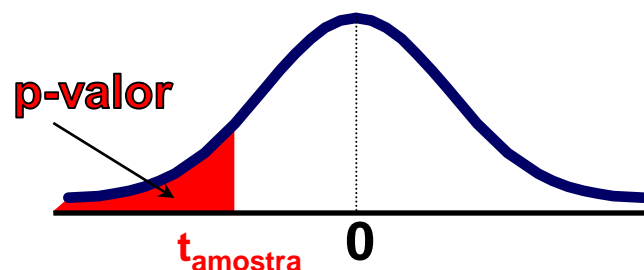
Teste de cauda Superior



$$H_0: \mu \geq 50$$

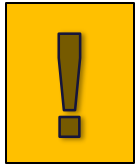
$$H_1: \mu < 50$$

Teste de cauda Inferior



O p-valor é a área hachurada, mas da distribuição T!

Abordagem por p-valor



Procedimento de teste usando o p-valor:

- 1) Converta a estatística amostral (\bar{X}_n) na estatística de teste ($T_{amostra}$)
- 2) Obtenha o **p-valor** de uma tabela ou usando um programa
- 3) Compare o **p-valor** a α
 - Se **p-valor** $< \alpha$, rejeite H_0
 - Se **p-valor** $\geq \alpha$, não rejeite H_0

Portanto, se o p-valor é suficientemente pequeno (menor do que α) então H_0 pode ser rejeitada.



Teste de hipóteses: Relação com Intervalo de Confiança

- Para $\bar{x} = 172.5$, $s = 15.40$ e $n = 25$, o intervalo de confiança de 95% para a média é:

$$IC(\mu, 95\%) = \left(172.5 - (2.064) \frac{15.4}{\sqrt{25}}, 172.5 + (2.064) \frac{15.4}{\sqrt{25}} \right)$$

$$166.14 \leq \mu \leq 178.86$$

- Uma vez que o intervalo contém a média hipotetizada (168), não rejeitamos a hipótese nula a $\alpha = 0.05$



Teste t: σ desconhecido

- Lembre-se que assumimos que a estatística amostral vem de uma amostra aleatória com distribuição normal.
- Se a amostra é pequena (< 30), deve-se verificar graficamente se podemos assumir a hipótese de normalidade.
- Se a amostra é grande, o teorema do limite central se aplica e a distribuição da média amostral é aproximadamente normal.



Teste de hipóteses: proporção

- Envolve variáveis categóricas (qualitativas)
- Dois resultados possíveis
 - “Sucesso” (possui uma certa característica)
 - “Fracasso” (não possui a característica)
- Fração ou **proporção de “sucessos”** na população é denotado por p



Teste de hipóteses: proporção

- A proporção amostral de sucessos é denotada por \hat{p} :

$$\hat{p} = \frac{n^{\circ} \text{ de sucessos na amostra}}{n}$$

- Quando tanto np quanto $n(1 - p)$ são de pelo menos 5, a distribuição de p pode ser aproximada por uma distribuição normal com média e desvio-padrão:

$$\hat{p} \sim N \left(p, \sqrt{\frac{p(1 - p)}{n}} \right)$$

Teste de hipóteses: proporção



A distribuição amostral de \hat{p} é aproximadamente normal, então a estatística de teste é uma estatística Z:

$$Z_{amostra} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

Teste Z: σ conhecido



Para um teste bi-caudal para a média com σ conhecido:

- Converta a estatística amostral (\bar{X}_n) para a estatística de teste:

$$Z_{amostra} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0, 1)$$

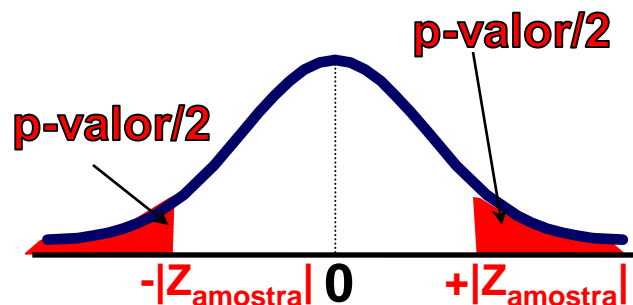
- Determine os valores críticos para distribuição normal padronizada a partir da tabela para um dado nível de significância α .
- **Decisão:** Se a estatística de teste cair na região de rejeição, rejeite H_0 ; caso contrário não rejeite H_0 .

Abordagem por p-valor

$$H_0: p = 0.68$$

$$H_1: p \neq 0.68$$

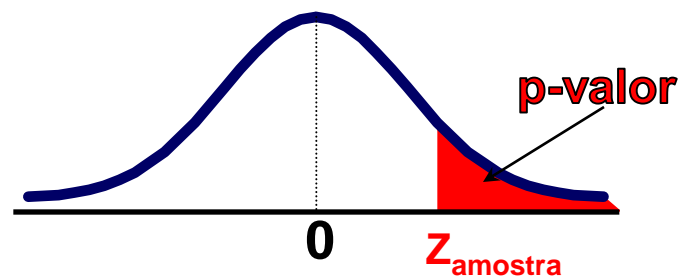
Teste bi-caudal



$$H_0: p \leq 0.68$$

$$H_1: p > 0.68$$

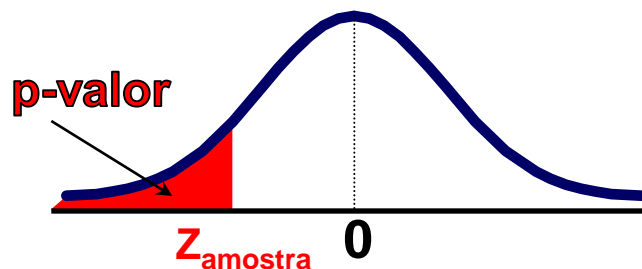
Teste de cauda Superior



$$H_0: p \geq 0.68$$

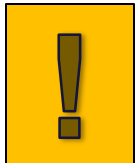
$$H_1: p < 0.68$$

Teste de cauda Inferior



O p-valor é a área hachurada!

Abordagem por p-valor



Procedimento de teste usando o p-valor:

- 1) Converta a estatística amostral (\bar{X}_n) na estatística de teste ($Z_{amostra}$)
- 2) Obtenha o **p-valor** de uma tabela ou usando um programa
- 3) Compare o **p-valor** a α
 - Se **p-valor** $< \alpha$, rejeite H_0
 - Se **p-valor** $\geq \alpha$, não rejeite H_0

Portanto, se o p-valor é suficientemente pequeno (menor do que α) então H_0 pode ser rejeitada.

Teste de hipóteses: proporção



Exercício: Uma empresa de marketing afirma receber 8% de respostas para a sua lista de email. Para testar esta afirmação, uma amostra aleatória de 500 emails foram verificados, sendo encontradas 30 respostas. Use $\alpha = 0.05$ como nível de significância.

Primeiro checamos:

$$\begin{aligned}np &= (500)(0.08) = 40 \\n(1 - p) &= (500)(0.92) = 460\end{aligned}$$

Teste de hipóteses: proporção

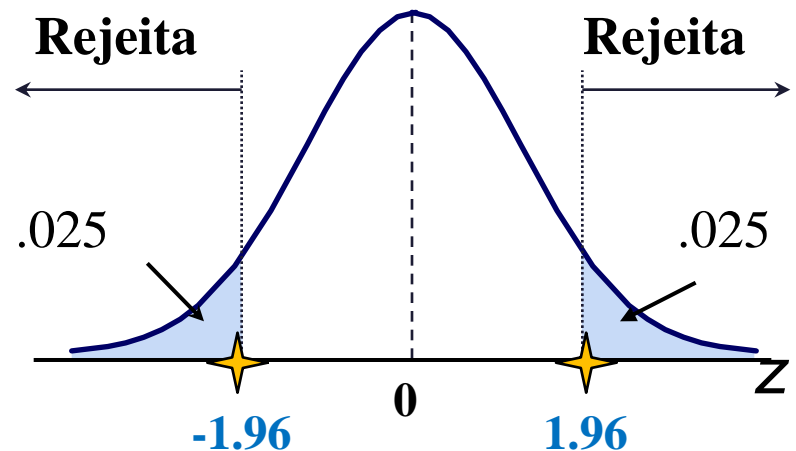
$$H_0: p = 0.08 \quad H_1: p \neq 0.08$$

$$\alpha = 0.05$$

$$n = 500, \quad \hat{p} = 0.06$$

Valores Críticos: ± 1.96

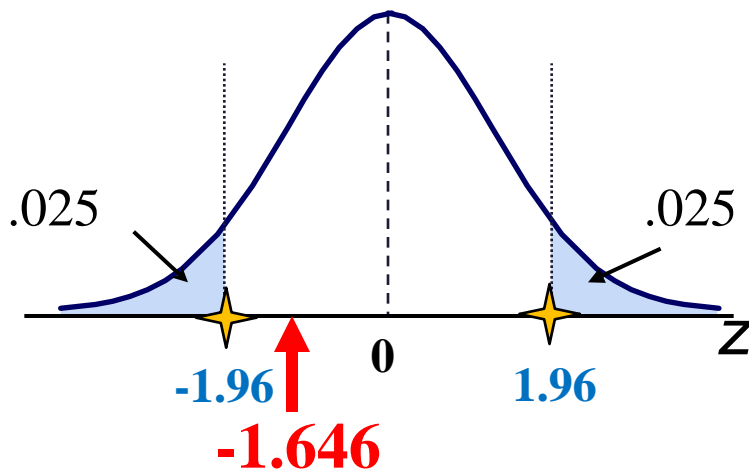
1) Determine a região de rejeição



Teste de hipóteses: proporção

Estadística de teste:

$$Z_{amostra} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.06 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{500}}} = -1.648$$



Decisão:

Não rejeitamos H_0 a $\alpha = 0.05$

Conclusão:

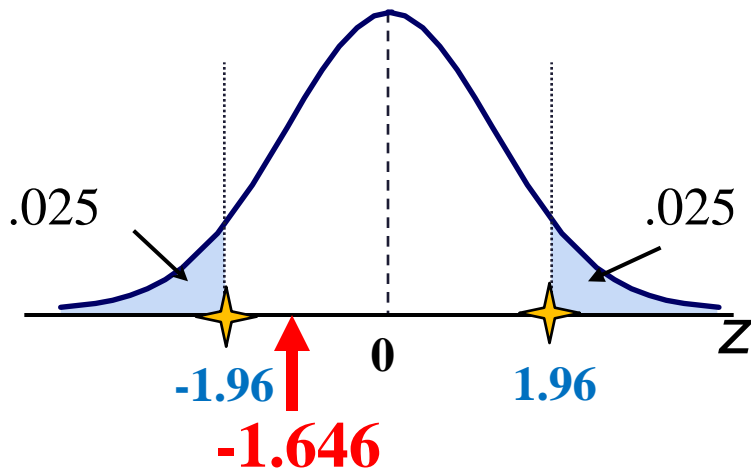
Não temos evidência o suficiente para rejeitar a tese da empresa de receber uma taxa de 8% de respostas.

Teste de hipóteses: proporção

Estatística de teste:

$$Z_{amostra} = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} = \frac{0.06 - 0.08}{\sqrt{\frac{0.08(1-0.08)}{500}}} = -1.648$$

$$\begin{aligned} p - \text{valor} &= 2 * P(Z \leq -1.648) \\ &= 2 * 0.0497 = 0.0993 \end{aligned}$$



$p - \text{valor} > \alpha = 0.05$: Não rejeitamos H_0

Não temos evidência o suficiente para rejeitar a tese da empresa de receber uma taxa de 8% de respostas.



Questões Éticas

- Use dados coletados aleatoriamente para evitar tendências de seleção
- Escolha o nível de significância, α , antes da coleta de dados
- Não use técnicas de “vai-e-vem” para escolher entre um teste uni-caudal ou bi-caudal ou para determinar o nível de significância
- Não use “limpeza de dados” para esconder observações que não corroboram a hipótese considerada
- Apresente todos os resultados pertinentes



Resumo

Nesta aula, vimos:

- A metodologia de teste de hipóteses
- O teste Z para a média populacional (σ conhecido)
- Discutimos as abordagens por valor crítico e por p-valor
- Realizamos testes bi-caudais e uni-caudais
- O teste t para a média (σ desconhecido)
- Realizamos o teste Z para a proporção populacional
- Discutimos questões éticas