Aula 14 - Testes de Hipóteses

- Hipóteses de interesse:

$$\begin{cases} H_0: & \text{hipótese nula} \\ H_a: & \text{hipótese alternativa} \end{cases}$$

$\begin{cases} H_0: \text{hip. nula} \\ H_1: \text{hip alternativa} \end{cases}$

- Tipos de erro:
  - Escolhendo H_0 podemos cometer o erro dessa hipótese ser falsa: erro tipo II.
  - Escolhendo H_a podemos cometer o erro dessa hipótese ser falsa: erro tipo I.

		hipótese verdadeira	
		H_0	H_a
sua	H_0	sem erro	erro tipo II (β)
escolha	H_a	erro tipo I (α)	sem erro

Exemplo: Hoje a noite você vai a uma festa. A previsão do tempo diz que há 80% de probabilidade de chuva. Você leva guarda-chuva?

$$\begin{cases} H_0: & \text{vai chover hoje a noite} \\ H_a: & \text{Não vai chover hoje a noite} \end{cases}$$

Erro tipo I: Você rejeita H_0, acredita que não vai chover, foi sem guarda-chuva e se molha.

Erro tipo II: Você não rejeita H_0, acredita que vai chover, leva guarda-chuva e passa a noite toda carregando um guarda-chuva sem usá-lo.

- **Regra da decisão:** Regra que estabelece com base nos dados obtidos, quando H_0 é rejeitada.
- **Nível de significância (α):** associado a uma regra de decisão. É a probabilidade de se cometer erro tipo I.

Teste para uma proporção populacional

Um produtor precisa decidir pela compra ou não de sementes fornecidas por um distribuidor, que afirma que a proporção de germinação de sementes é $\pi = 0,94$. Para tanto ele observou a proporção de germinação de uma amostra aleatória simples de 100 sementes e encontrou o valor $\hat{\pi} = 0,93$. Com base nesse resultado o produtor poderia discordar do distribuidor?

$\begin{cases} H_0: \pi = 0,94 \\ H_a: \pi < 0,94 \ (\text{ou } >) \end{cases}$ ou $\begin{cases} H_0: \pi = 0,94 \\ H_a: \pi \neq 0,94 \end{cases}$

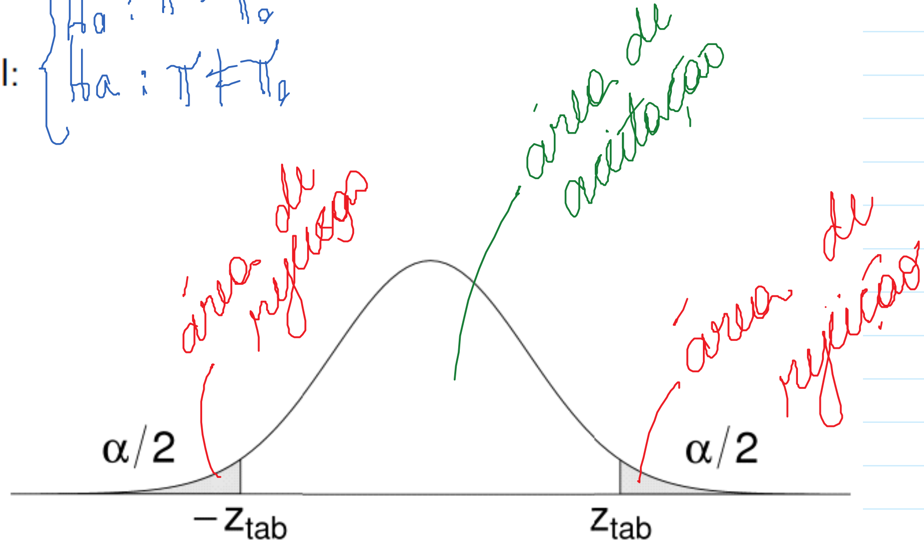
Procedimento para a realização de um teste de hipóteses:

1 Formular as hipóteses estatísticas: Hipótese nula (H_0) e Hipótese alternativa (H_1). ou H_a

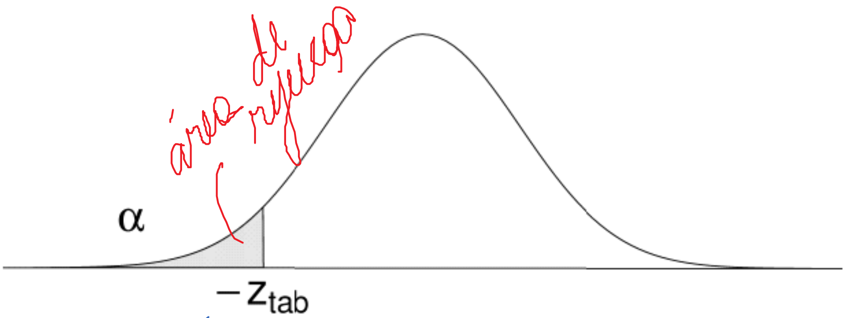
2 Definir da estatística adequada $\Rightarrow$ conhecer a distribuição amostral do estimador.

3 Fixar o nível de significância (α) para o teste $\Rightarrow$ limitar as regiões de rejeição e aceitação de $H_0 \Rightarrow z_{tab}$

- Teste Bilateral: $\begin{cases} H_0: \pi = \pi_0 \\ H_a: \pi \neq \pi_0 \end{cases}$



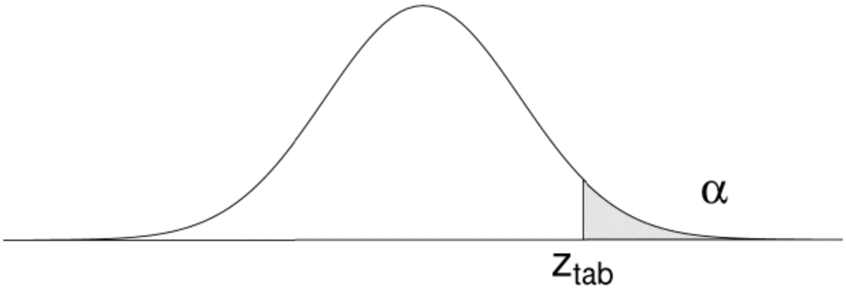
- Teste Unilateral (esquerda): $\begin{cases} H_0: \pi = \pi_0 \\ H_a: \pi < \pi_0 \end{cases}$



$\{ H_0: \pi = \pi_0$

$-z_{tab}$

- Teste Unilateral (direita): $\begin{cases} H_0: \pi = \pi_0 \\ H_a: \pi > \pi_0 \end{cases}$



Procedimento para a realização de um teste de hipóteses (continuação):

4 Com base na amostra, obter o valor z_{cal}, tal que

$$z_{cal} = \frac{\hat{\pi} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

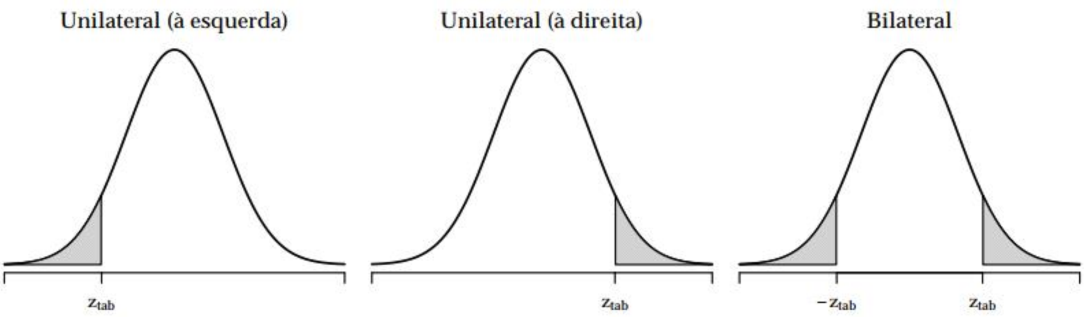
proporção amostral

5 Comparar z_{cal} com z_{tab}:

- Se z_{cal} pertencer a região de rejeição $\Rightarrow$ rejeita-se H_0 ao nível α de significância
- Se z_{cal} não pertencer a região de rejeição $\Rightarrow$ não se rejeita H_0 ao nível α de significância

- **Hipóteses simples**:
  H_0: $\theta = \theta_0$ versus H_a: $\theta = \theta_a$.

- **Hipóteses compostas**:
  - Unilateral (à esquerda): H_0: $\theta = \theta_0$ versus H_a: $\theta < \theta_0$;
  - Unilateral (à direita): H_0: $\theta = \theta_0$ versus H_a: $\theta > \theta_0$;
  - Bilateral: H_0: $\theta = \theta_0$ versus H_a: $\theta \neq \theta_0$.



Exemplo: Um produtor precisa decidir pela compra ou não de sementes fornecidas por um distribuidor, que afirma que a proporção de germinação de sementes é $\pi = 0,94$. Para tanto ele observou a proporção de germinação de uma amostra aleatória simples de 100 sementes e encontrou o valor $\hat{\pi} = 0,93$. Com base nesse resultado o produtor poderia discordar do distribuidor, considerando um nível de significância de 5%?

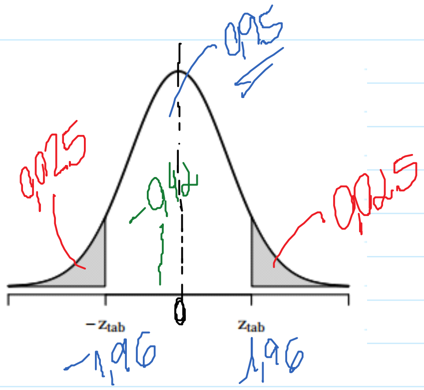
Situação 1: $\begin{cases} H_0: \pi = 0,94 \\ H_a: \pi \neq 0,94 \end{cases}$

$n = 100$ sementes

$\hat{\pi} = 0,93$

$$Z_{cal} = \frac{0,93 - 0,94}{\sqrt{\frac{0,94(0,06)}{100}}} = \boxed{-0,42}$$

$\alpha = 0,05$



$Z_{cal} > Z_{tab}$ não rejeita H_0

Não se rejeita Ho ao nível de significância de 5%, portanto, existem evidências de que o poder germinativo das sementes não difere de 94%.

Não se rejeita Ho ao nível de significância de 5%, portanto, existem evidências de que o poder germinativo das sementes não difere de 94%.

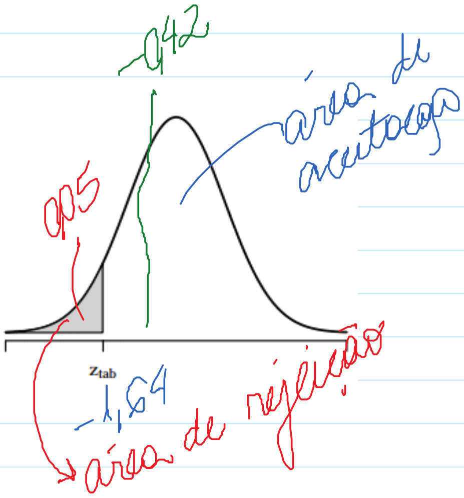
Situação 2: $\begin{cases} H_0: \pi = 0{,}94 \\ H_a: \pi < 0{,}94 \end{cases}$

$n = 100$ sementes

$\hat{\pi} = 0{,}93$

$$Z_{cal} = \frac{0{,}93 - 0{,}94}{\sqrt{\frac{0{,}94(0{,}06)}{100}}} = \boxed{-0{,}42}$$

$\alpha = 0{,}05$



Não se rejeita Ho ao nível de 5% de significância, portanto, existem evidências de que o poder germinativo das sementes não é menor do que 94%.

Exercício: Um biólogo, com base em conhecimentos teóricos e práticos, afirma que a proporção (π) de ferólitos sem bromélias, no estádio arbóreo pioneiro da Floresta Ombrófila, na Ilha de Santa Catarina, é igual a 0,47. Em uma amostra de 35 forófitos, 24 não apresentaram bromélias. Teste a afirmativa do biólogo ao nível de 5% de significância.

Situação 1: $\begin{cases} H_0: \pi = 0{,}47 \\ H_a: \pi \neq 0{,}47 \end{cases}$

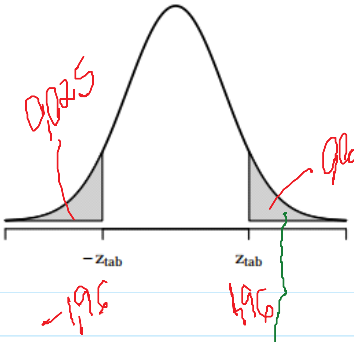
$n = 35$

$\hat{\pi} = \frac{24}{35} = 0{,}6857$

$\alpha = 0{,}05$

$$\pi = \frac{24}{35} = 0,6857$$

$$Z_{cal} = \frac{0,6857 - 0,47}{\sqrt{\frac{0,47(0,53)}{35}}} = 2,56$$



$Z_{cal} = 2,56$ → dentro da área de rejeição

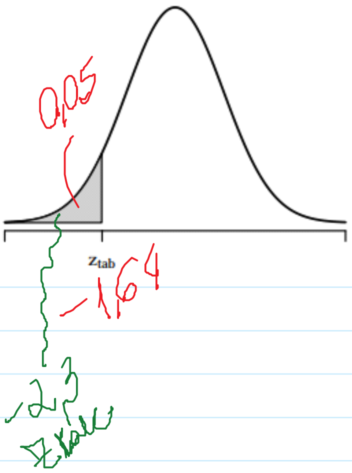
Rejeita-se a Ho ao nível de significância de 5%, portanto, existem evidências de que a proporção de ferólitos sem bromélias difere de 0,47.

Exercício: Uma estação de televisão afirma que 60% dos televisores estavam ligados no seu programa especial da última segunda-feira. Uma rede competidora deseja contestar essa afirmação e decide usar uma amostra de 200 famílias para um teste. Da pesquisa obteve-se 104 famílias que estavam assistindo ao programa. Verifique a informação da estação de televisão ao nível de 5% de significância.

$$\begin{cases} H_0: \pi = 0,60 \\ H_a: \pi < 0,60 \end{cases} \qquad \alpha = 0,05$$

$$\hat{\pi} = \frac{104}{200} = 0,52$$

$$Z_{cal} = \frac{0,52 - 0,60}{\sqrt{\frac{0,60(0,40)}{200}}} = -2,3$$



Rejeita-se Ho

Rejeita-se Ho ao nível de significância de 5%, portanto, há evidências de que a proporção de famílias assistindo ao programa foi menor que o afirmado pela estação.

Teste para médias populacionais

Existem basicamente três tipos de afirmações que podem ser feitas a respeito das médias populacionais:

1. a afirmação diz respeito a uma média populacional. Exemplo: a altura média da floresta nativa é igual a 20m.
2. a afirmação diz que duas médias populacionais são iguais. Exemplo: as produtividades médias de duas variedades de cana-de-açúcar são iguais.
3. a afirmação diz que as médias de duas ou mais populações são iguais. Exemplo: desejamos saber se há diferença entre três locais quanto ao número médio de micronúcleos por 5000 células sanguíneas de peixe do gênero Bagre.

$$H_0 : \mu = \mu_0,$$

em que μ_0 é um valor conhecido.

$$\begin{aligned} H_1 &: \mu \neq \mu_0 \quad \text{(teste bilateral)} \\ H_1 &: \mu > \mu_0 \quad \text{(teste unilateral à direita)} \\ H_1 &: \mu < \mu_0 \quad \text{(teste unilateral à esquerda)} \end{aligned}$$

Teste para uma média populacional com variância conhecida

Sabendo-se que, para uma população normal

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

Assim, quando a hipótese nula, $\mu = \mu_0$, for verdadeira, a estatística:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}},$$

$\bar{X} \rightarrow$ média amostral

segue a distribuição normal padrão, $N(0, 1)$.

Exemplo: Uma balança para encher pacotes de sementes automaticamente está programada para produzir pacotes com peso médio de 20 kg e desvio padrão de 0,20 kg. Periodicamente é feita uma inspeção para verificar se o peso médio está sob controle. Para este fim, foi selecionada uma amostra de 8 pacotes de sementes, cujos resultados foram:

20,3	19,8	20,3	19,7	19,8	19,7	19,8	19,8

Teste a hipótese que a a balança se desregulou e está produzindo um peso médio inferior a 20 kg. Use o nível de significância de 5%.

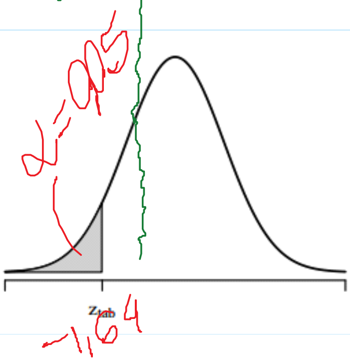
$$\begin{cases} H_0: \mu = 20\,kg \\ H_a: \mu < 20\,kg \end{cases}$$

$n = 8$ pacotes

$\bar{x} = 19{,}9$ kg

$\alpha = 0{,}05$

$Z_{cal} = -1{,}41$



$$Z_{cal} = \frac{19{,}9 - 20}{\frac{0{,}20}{\sqrt{8}}} = -1{,}41$$

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

Não se rejeita Ho

Não se rejeita Ho ao nível de significância de 5%, portanto, existem evidências de que o peso médio dos pacotes está sob controle e não é menor que 20kg.

Teste para uma média populacional com variância desconhecida

Quando não conhecemos a variância da população (mais comum), devemos estimá-la a partir da amostra. Nesse caso, a estatística apropriada para o teste da hipótese é dada por:

$$t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}},$$

a qual tem distribuição t de Student com $n - 1$ graus de liberdade, quando a hipótese nula for verdadeira.

Exemplo:Um pesquisador deseja verificar se o processamento térmico da soja reduz sua atividade ureática a níveis inferiores a 0,4 (nível máximo permitido para a utilização da soja na alimentação animal). Assim, retirou uma amostra de tamanho 10 e observou os resultados da tabela a seguir:

0,5	0,4	0,2	0,4	0,1
0,2	0,3	0,3	0,4	0,3

$n = 10$ $\quad \bar{x} = 0{,}31$ $\quad s = 0{,}1197$

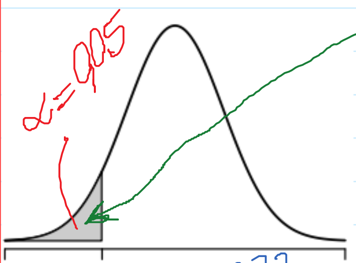
Tire uma conclusão supondo um nível de significância $\alpha = 0,05$.

$$\begin{cases} H_0: \mu = 0{,}4 \\ H_a: \mu < 0{,}4 \end{cases}$$

$$t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

na calculadora: $\bar{x}$ 1 $\quad \sqrt{x}$ 2 $\quad S_x$ 3 desvio padrão amostral

$$t_{calc} = \frac{0{,}31 - 0{,}4}{\frac{0{,}1197}{\sqrt{10}}} = -2{,}38$$



$t_T = -1{,}833$ $\quad \nu = 9\,gl$

Rejeita-se Ho

Rejeita-se Ho ao nível de 5% de significância e portanto, existem evidências de que o processo térmico da soja reduz sua atividade ureática a níveis inferiores a 0,4.

Exemplo: Um cientista deseja saber se o pH de um solo é ácido. Ele toma uma amostra de 10 unidades e obteve os valores de pH:

5,8	6,0	7,0	6,2	6,2
7,1	6,4	5,5	5,8	5,9

$n = 10$

Tire uma conclusão supondo nível de significância de 5%.

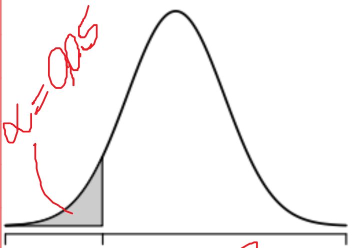
__Observação__: pH < 7,0 (acidez)
pH > 7,0 (alcalino)
pH = 7,0 (neutro)

$\bar{x} = 6,19$

$s = 0,519$

$$\begin{cases} H_0: \mu = 7,0 \\ H_a: \mu < 7,0 \end{cases}$$

$\alpha = 0,05$



$\alpha = 0,05$

$t_T = -1,833$

$\nu = 9\,gl$

$$t_{calc} = \frac{6,19 - 7,0}{\frac{0,519}{\sqrt{10}}} = -4,93$$

região de rejeição

Rejeita-se Ho ao nível de 5% de significância, e portanto, existem evidências de que o solo é ácido.

__Exercícios__:

1) Um agrônomo afirma que a produtividade média do feijão da safra das lavouras de agricultores familiares de um determinado ano é de 800 kg/ha. Para investigar a veracidade dessa afirmação selecionou-se uma amostra de nove lavouras onde obteve-se os seguintes valores de produtividade de feijão, em kg/ha:

Lavoura	1	2	3	4	5	6	7	8	9
Produtividade	767,8	764,1	716,8	750,2	756,0	692,5	736,1	746,1	731,4

a) Qual a conclusão ao nível de significância de 5%?

b) *Caso a afirmação do agrônomo não seja verdadeira, dê uma estimativa da média populacional, com grau de confiança de 95%.*

2) Foi retirada uma amostra de 10 bezerros da raça Nelore, aos 210 dias de idade, com o objetivo de verificar se o peso médio desses animais atingiu ou não 186kg. Os valores obtidos, em kg, foram os seguintes:

178	199	182	186	188	191	189	185	174	158

Teste as hipóteses:

$$\begin{cases} H_0: \mu = 186 \\ H_a: \mu < 186 \end{cases}$$

ao nível de significância de 5%.