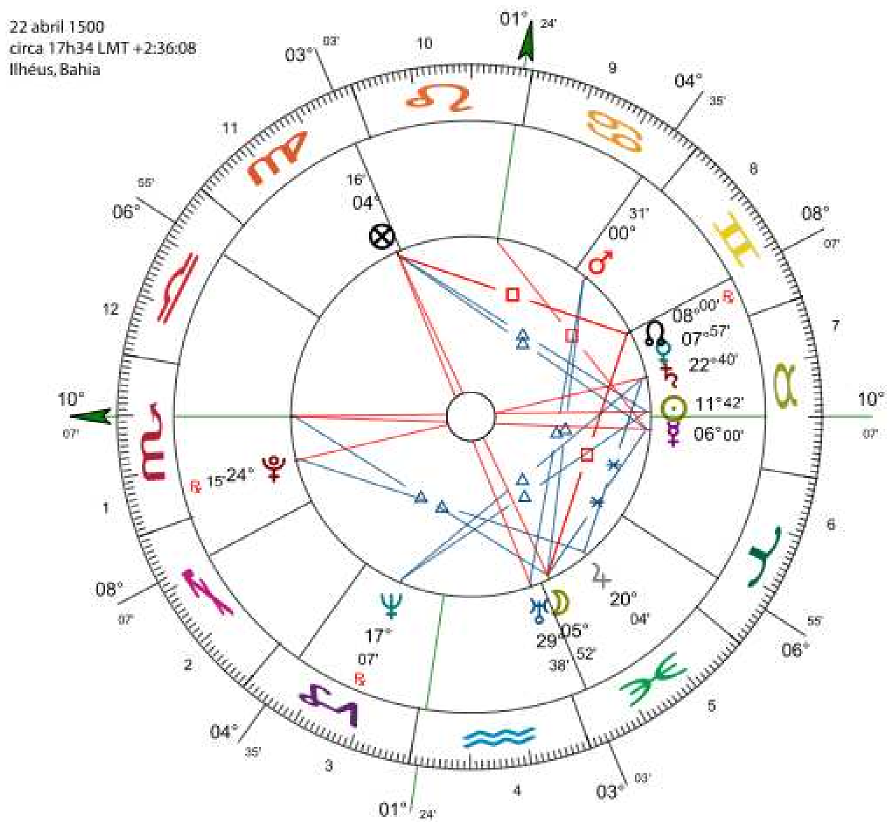
Slides Semana 9

Teste de Hipóteses: Introdução

Exemplo: Astrólogos conseguem predizer nossa personalidade com um mapa astral?



- Se você fornece data, hora e local de nascimento, um astrólogo monta o seu Mapa Astral.
- *De acordo com a astrologia, a posição dos astros no momento em que nascemos influencia nossa maneira de ser.* - Wikipedia
- *As configurações de um Mapa Astral se repetem apenas a cada 26.000 anos, portanto ele é quase como uma impressão digital - não existe um igual ao outro.* - Wikipedia
- Há comprovação científica de que seu mapa astral reflete sua personalidade?

Um teste foi feito da seguinte maneira: 116 pessoas selecionadas aleatoriamente forneceram data, hora e local de nascimento.

Um astrólogo preparou um mapa astral para essas 116 pessoas, usando apenas os dados fornecidos acima.

Cada voluntário também preencheu um questionário: “California Personality Index”.

Para **um outro astrólogo**, foram dados:

- data, hora, local, Mapa Astral de um dos voluntários, por exemplo, voluntário 3.
- questionário de personalidade preenchidos pelo voluntário 3.
- 2 questionários de personalidade, escolhidos ao acaso entre os 115 restantes, preenchidos por outros dois voluntários.
- Ao astrólogo, pediu-se então para identificar qual questionário havia sido preenchido pelo dono daquele Mapa Astral.
- Seja p a probabilidade de que o astrólogo identifique o questionário correto.
- Se de fato a informação do Mapa Astral não caracteriza a personalidade de uma pessoa e na verdade o astrólogo está apenas escolhendo um dos 3 questionários ao acaso, a probabilidade de acerto é $p = 1/3$.
- Astrólogos confiam em seus estudos e dizem que a probabilidade de acerto é maior do que $1/3$.
- Como testar se eles estão certos?
- Escolher ao acaso um astrólogo e fazer o teste com ele uma vez, é suficiente?

Exemplo: Mapa Astral

- Astrólogos foram selecionados aleatoriamente a partir de uma lista de astrólogos familiarizados com o "California Personality Index".
- A lista foi preparada pelo "National Council for Geocosmic Research".
- Vimos que podemos estudar fenômenos aleatórios definindo variáveis aleatórias e teoria da probabilidade.
- Um astrólogo pode associar corretamente o questionário ao mapa astral ou não.
- Para cada situação, há uma probabilidade associada. Portanto temos um evento aleatório.

Como definir a variável aleatória?

- X_i: astrólogo associa corretamente um questionário ao mapa astral i, ou seja,

$$X_i \sim \text{Bernoulli}(p)$$

- Podemos pensar em p como a proporção de acerto na população de astrólogos. Se astrólogos não têm a capacidade de predição, $p = 1/3$. Astrólogos alegam que são capazes: $p > 1/3$.
- Como usar dados para testar estes dois cenários?

Definindo hipóteses

- Objetivo em muitos estudos: checar se os dados apóiam certas afirmações que são feitas para uma população.
- Afirmações a serem testadas: **hipóteses**.
- Expressamos as hipóteses em termos dos parâmetros da população.
- Por exemplo: o parâmetro pode ser uma proporção populacional.

Hipótese: Usando o mapa astral de uma pessoa, a probabilidade p de um astrólogo predizer corretamente qual dos 3 questionários está associado àquele mapa astral é igual a $1/3$. Ou seja, os astrólogos apenas selecionam ao acaso um dos questionários.

Nesse caso, para saber se os astrólogos têm a capacidade de predizer a personalidade usando o mapa astral, usaríamos as seguintes **hipóteses**:

$$\begin{cases} H_0 : p = 1/3 & \text{(hipótese nula)} \\ H_A : p > 1/3 & \text{(hipótese alternativa)} \end{cases}$$

No experimento com os astrólogos, observar uma proporção alta de acertos pode ser uma evidência contra a hipótese de que $p = 1/3$?

Passos de um teste de hipótese

- **Passo 1: Suposições**

  O teste é válido sob algumas suposições. A mais importante assume que os dados do experimento foram produzidos através de um processo de aleatorização.

- **Passo 2: Hipóteses**

  O teste de hipótese tem sempre duas hipóteses sobre o parâmetro populacional de interesse. As hipóteses devem ser definidas **antes** de se realizar o experimento e coletar dados.

- **Hipótese Nula (H_0):** afirma que o parâmetro populacional assume um dado valor.

- **Hipótese Alternativa (H_A):** afirma que o parâmetro populacional assume outros valores, diferente do valor descrito na H_0.

No experimento dos astrólogos, $H_0: p = 1/3$ representa a hipótese de que **não há efeito**, no sentido de que os astrólogos não têm uma capacidade maior de predizer a personalidade usando o mapa astral.

A hipótese alternativa, $H_A: p > 1/3$, representa a hipótese de que **há efeito**, ou seja, os astrólogos têm uma capacidade de predizer a personalidade usando o mapa astral.

Em teste de hipóteses, mantém-se a favor de H_0 a menos que os dados tragam grande evidência contra.

A hipótese nula é conservadora: “o réu é inocente até que se prove o contrário”.

- **Passo 3: Estatística do teste**

Vimos que podemos usar uma estatística para estimar um parâmetro populacional. A **estatística do teste** descreve quão longe do parâmetro populacional usado na H_0 a estimativa está.

Por exemplo, se $H_0 : p = 1/3$, e se $\hat{p} = 40/116 = 0.345$, queremos uma estatística que quantifique quão longe está $\hat{p} = 0.345$ de $p = 1/3$.

- **Passo 4: valor-de-p**

Para interpretar uma estatística do teste, vamos usar uma probabilidade para resumir a evidência contra H_0. Esta probabilidade é chamada de **valor-de-p**.

- **Passo 5: Conclusão**

Baseado no valor-de-p, decidir se rejeita ou não a hipótese nula. Note que a conclusão é sempre em termos da hipótese nula: rejeitar ou não H_0.

Mas quão pequeno deve ser o valor-de-p para ser considerado forte evidência contra H_0?

Geralmente, fixamos o **nível de significância** do teste (α), e usamos a seguinte regra. É comum usarmos $\alpha = 0.05$.

- Se valor-de-p $\leq \alpha$: rejeitamos H_0, ou seja, os dados trazem forte evidência contra a hipóstese nula
- Se valor-de-p > α: não rejeitamos H_0, ou seja, não temos evidência nos dados contra a hipótese nula

Assumimos primeiro que H_0 é verdadeira.

Consideramos então todos os valores possíveis para a estatística do teste, de acordo com sua distribuição amostral.

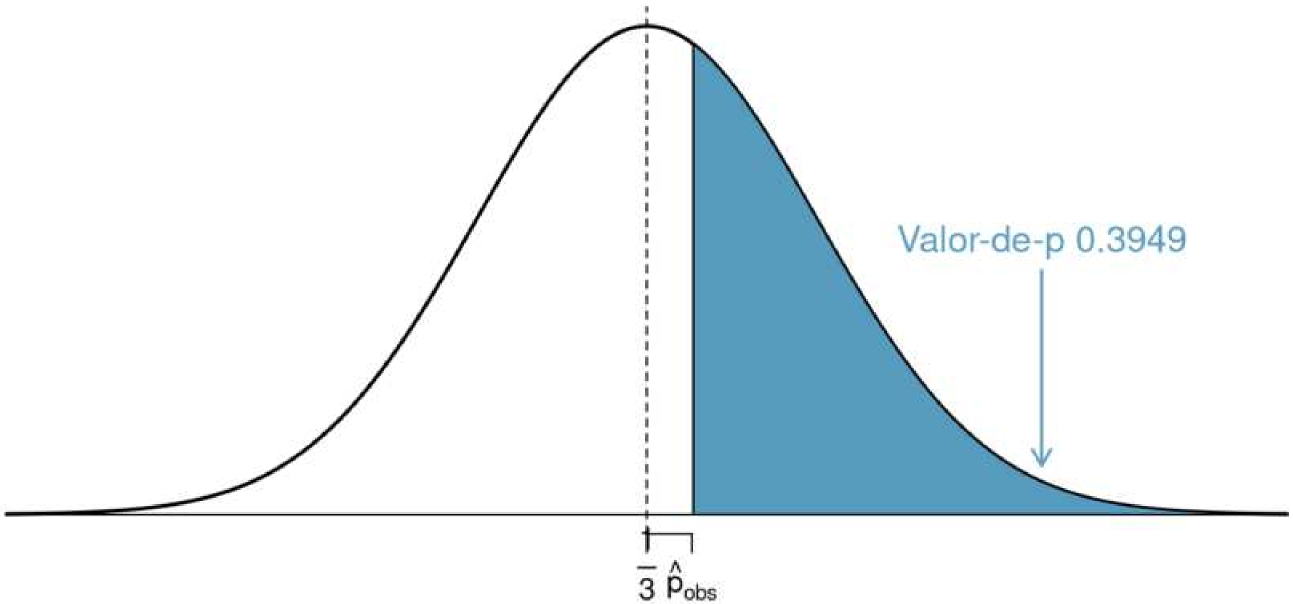
Calculamos a estatística do teste observada para o experimento realizado e verificamos onde, na distribuição amostral, ela se posiciona.

Calculamos a probabilidade de um valor igual ou mais extremo ao da estatística do teste observada (valor-de-p). Mais extremo: mais evidência contra H_0.

Se o valor-de-p obtido é bem pequeno, por exemplo, 0.01, isto quer dizer que se H_0 é verdadeira, então seria incomum obter uma amostra com os resultados como o observado. Um valor-de-p muito baixo traz fortes evidências contra H_0.

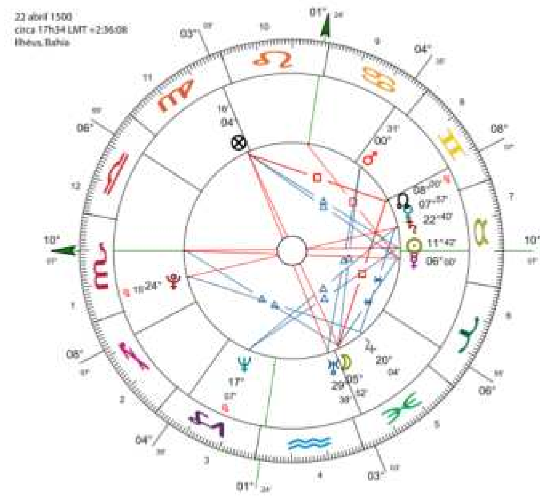
Exemplo: Mapa Astral

Distribuição amostral da proporção amostral $\hat{p}$ sob H_0.



valor-de-p (área em azul): probabilidade da proporção amostral assumir um valor igual ao observado, $\hat{p}_{obs}$, ou mais extremo, sob H_0.

Exemplo: Mapa Astral



Passo 1: Suposições

- A variável de interesse é binária. Sej, X_i: astrólogo i associa corretamente um questionário ao mapa astral, ou seja,

$$X_i \sim \text{Bernoulli}(p).$$

- Os dados foram obtidos usando processo de aleatorização: uma amostra aleatória de voluntários e astrólogos foi feita.
- Temos uma a.a. de tamanho 116. Portanto, a distribuição amostral da estimativa para p, $\hat{p}$, tem distribuição aproximadamente normal, pelo TCL.

Passo 2: Hipóteses

- H_0: $p = p_0 = 1/3$ *contra* H_A: $p > p_0 = 1/3$. Em outras palavras:
- H_0: Astrólogos *chutam* qual o questionário está associado ao mapa astral.
- H_A: Astrólogos predizem melhor do que um *chute* qual o questionário está associado ao mapa astral.

Passo 3: Estatística do teste

- Estatística do teste mede quão longe está a proporção amostral, $\hat{p}$, da proporção populacional, p, assumindo que H_0 seja verdadeira?
- Sabemos que:

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

- Se H_0 é verdadeira ($p = p_0$), então:

$$\hat{p} \sim \mathrm{N}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$$

Passo 3: Estatística do teste

- A estatística do teste é:

$$Z = \frac{\hat{p} - p_0}{EP_0(\hat{p})}$$

em que $EP_0(\hat{p})$ é o erro padrão de $\hat{p}$ sob H_0.

Portanto,

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{H_0}{\sim} \mathrm{N}(0, 1)$$

- A estatística do teste mede quão distante está $\hat{p}$ de p_0 em unidades de "erro padrão".
- No experimento dos astrólogos, dentre 116 mapas, 40 foram corretamente associados ao questionário de personalidade.

$$\hat{p} = 40/116 = 0.345$$

- A estatística do teste observada é:

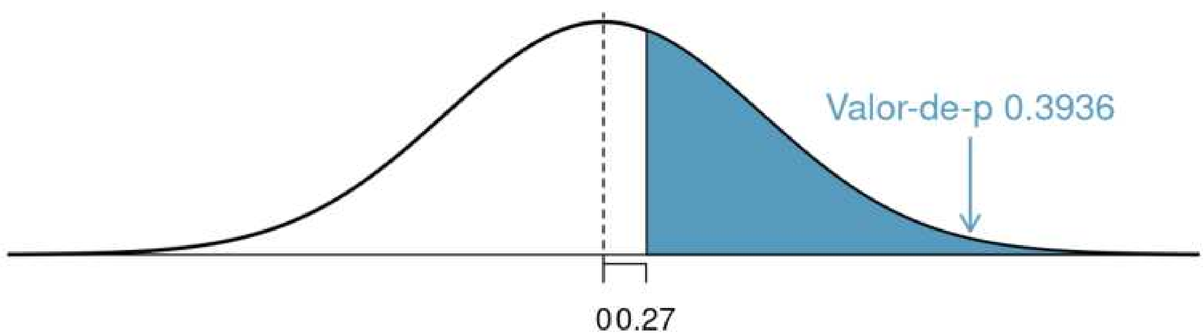
$$z_{obs} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.345 - 1/3}{\sqrt{\frac{1/3(1-1/3)}{116}}} = 0.27$$

- A proporção amostral está a 0.27 erro padrão de distância da proporção populacional, segundo H_0.

Passo 4: Valor-de-p

- Tendo observado $z_{obs} = 0.27$, isso traz evidência contra H_0 (a favor de H_A)?
- Quão improvável é $z_{obs} = 0.27$ se a proporção de acertos dos astrólogos é de fato $p = p_0 = 1/3$?
- valor-de-p: probabilidade de que uma estatística do teste assuma um valor igual ou mais extremo do que o observado, assumindo H_0 verdadeira.
- Mais extremo: neste caso, um valor maior que z_{obs}, pois equivale a um maior $\hat{p}$, maior proporção amostral de acertos (astrólogos alegam que $p > 1/3$).
- valor-de-p: $P(Z > z_{obs}) = P(Z > 0.27) = 0.3936$, onde $Z \sim \mathrm{N}(0, 1)$.

Distribuição amostral da estatística do teste Z sob H_0.



Passo 5: Conclusão

- O valor-de-p obtido no experimento foi 0.3936.
- O valor não é tão pequeno. Portanto, não temos evidências contra H_0.
- Não podemos concluir que astrólogos têm poderes preditivos especiais usando mapa-astral.


MENU

nature

International journal of science

Commentary | Published: 05 December 1985

A double-blind test of astrology

Shawn Carlson

Nature **318**, 419–425 (05 December 1985) | Download Citation

Two double-blind tests were made of the thesis that astrological 'natal charts' can be used to describe accurately personality traits of test subjects.

Detalhes da pesquisa podem ser encontrados no artigo da revista Nature: [A double-blind test of Astrology](#).

Resumo: Teste de Hipótese para uma proporção

Suponho que temos uma população e uma hipótese sobre a proporção p de indíviduos com certa característica.

Hipóteses:

$$\begin{aligned} H_0: p = p_0 \quad \text{vs} \quad H_A: & p \neq p_0 \text{ (bilateral)} \\ & p < p_0 \text{ (unilateral à esquerda)} \\ & p > p_0 \text{ (unilateral à direita)} \end{aligned}$$

Estatística do teste: Baseada na distribuição amostral de $\hat{p}$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \overset{H_0}{\sim} N(0,1)$$

Condição: $np_0 \geq 10$ e $n(1 - p_0) \geq 10$ para aproximação normal

Resumo: Teste de Hipótese para uma proporção

valor-de-p

- $H_A : p \neq p_0$ (bilateral): valor-de-p=$P(|Z| \geq |z_{obs}|)$
- $H_A : p < p_0$ (unilateral à esquerda): valor-de-p=$P(Z \leq z_{obs})$
- $H_A : p > p_0$ (unilateral à direita): valor-de-p=$P(Z \geq z_{obs})$

Conclusão

Para um nível de significância α:

- Se valor-de-p $\leq \alpha$: rejeitamos H_0
- Se valor-de-p > α: não rejeitamos H_0

Experimento da Coca vs Coca Zero

Em sala de aula, vários alunos disseram que conseguem distinguir entre Coca-Cola normal e Coca-Cola Zero.

Fizemos então o teste para comprovar se a afirmação é verdadeira.

Um dos alunos experimentou, em ordem aleatória, 20 amostras (ao acaso era Coca normal ou zero) e anotamos a quantidade de acertos.

Cada tentativa, X_i, é uma Bernoulli(p), em que p é a probabilidade de acerto.

Veja que $T = \sum_{i=1}^{20} X_i \sim Bin(20, p)$, onde T é o número de acertos.

Dos 20 testes, o aluno acertou 19! Temos então uma proporção amostral de acertos $\hat{p} = 19/20 = 0.95$. Isso mostra que o aluno realmente sabe a diferença?

Vamos testar o seguinte:

$$H_0 : p = 0.50 \quad \text{vs} \quad H_a : p > 0.50$$

Podemos testar essas hipóteses de duas maneiras:

- Usando a aproximação normal para a proporção de acertos, como vimos na última aula, já que as condições $np_0 \geq 10$ e $n(1-p_0) \geq 10$ são satisfeitas.
- Usando a distribuição exata do número total de acertos

Vamos revisar o que vimos na aula passada e também fazer o teste com a distribuição exata de T.

Usando a distribuição exata do número de acertos em 20 tentativas.

Hipóteses: $H_0 : p = 0.50 \quad \text{vs} \quad H_a : p > 0.50$

Hipóteses: $H_0 : Acertos = 10 \quad \text{vs} \quad H_a : Acertos > 10$

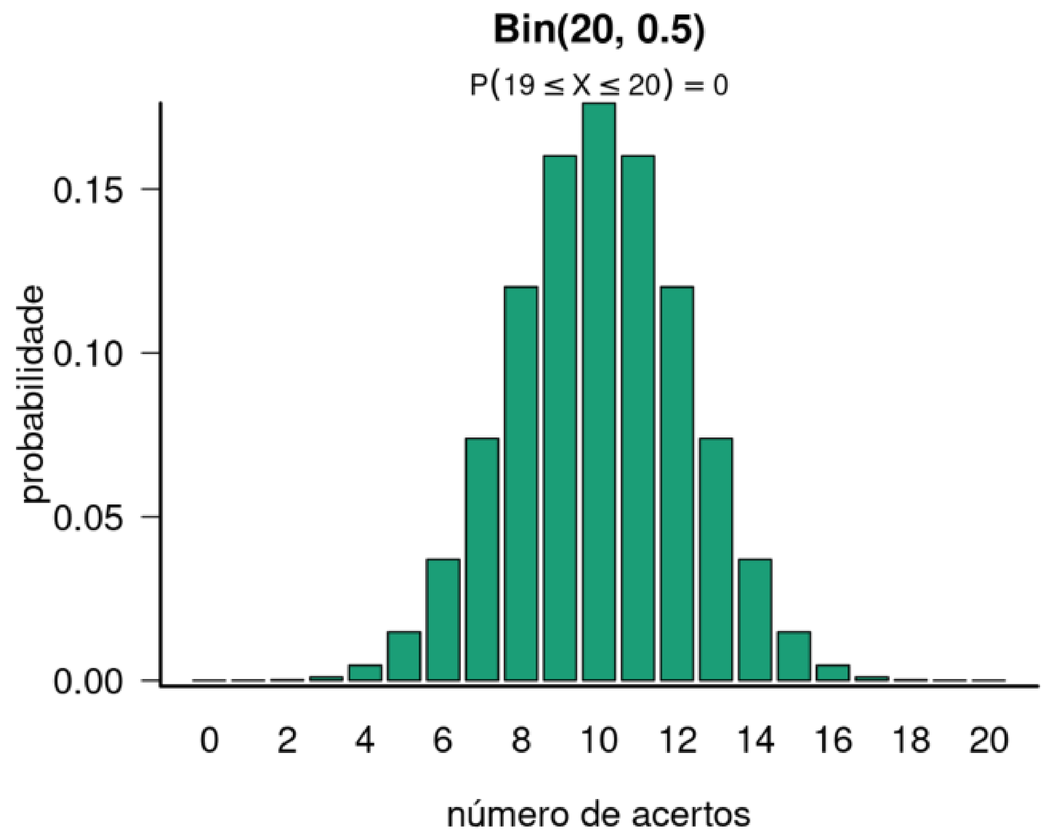
Estatística do teste: $T = \sum_{i=1}^{20} X_i \overset{H_0}{\sim} Bin(20, 0.5)$

O valor observado da estatística do teste é $t_{obs} = 19$, ou seja, o número total de acertos.

valor-de-p = $P(T \geq 19) = 0.00002$

Conclusão: Fixando $\alpha = 0.05$, rejeitamos a hipótese de que $p = 0.5$ e, portanto, acreditamos que a probabilidade de acertos é maior que 50%.

Experimento da Coca vs Coca Zero



Tipos de Erro

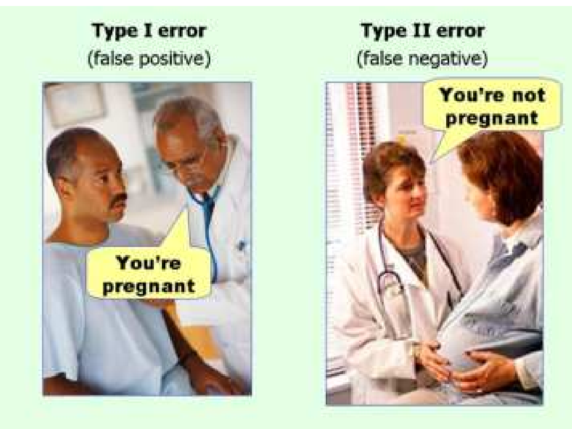
Quando realizamos um teste de hipóteses, podemos cometer 2 tipos de erros:

1- **Erro Tipo I:** Rejeitar a hipótese H_0, quando tal hipótese é verdadeira

2- **Erro Tipo II:** Não rejeitar a hipótese H_0, quando tal hipótese é falsa

Decisão	Ho	
	Verdadeira	**Falsa**
Rejeitar H_0	Erro Tipo I	OK ✓
Não Rejeitar H_0	OK ✓	Erro Tipo II

Erro Tipo I: erro mais grave



$$H_0 : \text{você não está grávida(o)}$$

$$H_A : \text{você está grávida(o)}$$

Tipos de Erro

Podemos calcular as probabilidades dos dois tipos de erro, chamadas de α e β:

$$\alpha = P(\text{Erro Tipo I}) = P(\text{Rejeitar } H_0 | H_0 \text{ verdadeira})$$

$$\beta = P(\text{Erro Tipo II}) = P(\text{Não Rejeitar } H_0 | H_0 \text{ falsa})$$

Na situação ideal, ambas as probabilidades de erro, α e β, seriam próximas de zero. Entretanto, à medida que diminuímos α, a probabilidade β tende a aumentar.

Levando isso em conta, em teste de hipóteses tentamos controlar a probabilidade do erro do tipo I, já que esse é o erro mais grave.

A probabilidade α é chamada de **nível de significância**, que geralmente fixamos em 5%.

No experimento da Coca-Cola tivemos 19 acertos em 20 tentativas e decidimos rejeitar H_0.

Mas e se tivéssemos observado 14 acertos? Ou 12?

Existe um valor, t_c, de maneira que se observarmos algo igual ou maior que ele decidimos rejeitar H_0?

Esse valor é chamado de **valor crítico** e vamos denotá-lo por t_c.

Tipos de Erro

No experimento da Coca-Cola: $H_0 : p = 0.5 \quad \text{vs} \quad H_a : p > 0.5$

Seja T o número de acertos em uma amostra de tamanho $n = 20$. Então $T \sim Bin(20, p)$.

Vamos considerar o seguinte valor crítico: $t_c = 12$.

Lembrando que T pode assumir os valores $0, 1, 2, \ldots, 20$.

O valor crítico t_c determina as probabilidades de cometer os erros tipo I e II. Considerando $t_c = 12$

$$\begin{aligned} P(\text{Erro Tipo I}) &= P(\text{Rejeitar } H_0 | H_0 \text{ verdadeira}) \\ &= P(T \geq t_c | p = 0.5) \\ &= \sum_{x=12}^{20} P(T = x | p = 0.5) \approx 0.25 \end{aligned}$$

$$\begin{aligned} P(\text{Erro Tipo II}) &= P(\text{Não Rejeitar } H_0 | H_0 \text{ falsa}) \\ &= P(T < t_c | p = 0.7) \\ &= \sum_{x=0}^{11} P(T = x | p = 0.7) \approx 0.11 \end{aligned}$$

Tipos de Erro

Observando a relação entre os erros tipo I e II, e t_c: $H_0 : p = 0.5 \text{ vs } H_a : p = 0.7$

t_c	P(Erro Tipo I)	P(Erro Tipo II)
12	0.25	0.11
13	0.13	0.23
14	0.06	0.39
15	0.02	0.58

Veja que à medida que $\alpha =$P(Erro Tipo I) diminui, $\beta =$P(Erro Tipo II) aumenta.

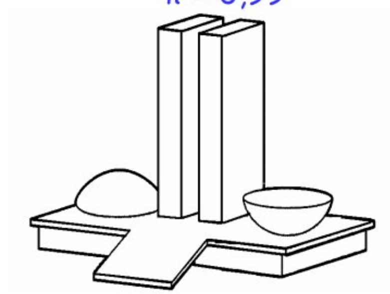
Então, optamos por controlar $\alpha =$P(Erro Tipo I), que é considerado o erro mais grave. Geralmente fixamos $\alpha = 0.05$ e rejeitamos H_0 se valor-de-p $< \alpha$.

Teste de hipóteses para média (σ conhecido)

Teste de hipóteses: proporção ou média

Proporção de políticos honestos

$\pi = 0{,}99$



Quantidade média de carboidratos

$\mu = 13g$

QUANTIDADE POR EMBALAGEM	
VALOR ENERGÉTICO	117kcal = 491kJ
CARBOIDRATOS	13g
PROTEINAS	10g
GORDURAS TOTAIS	2,8g
GORDURAS SATURADAS	1,9g

Exemplo: Café

Vamos voltar no problema da máquina que enche pacotes de café. Digamos que o peso nominal do pacote de café seja de 500g. Assume-se que o desvio padrão é conhecido ($\sigma = 10$).

Retiraram uma amostra de 25 pacotes e observaram um peso médio de 485g.

Isso nos traz evidência de que os pacotes têm menos de 500g?

Já calculamos o IC de 95% para esse problema:

$$IC(\mu, 0.95) = [481.08; 488.92]$$

Vamos agora testar as hipóteses:

$$H_0 : \mu = 500 \quad \text{vs} \quad H_a : \mu \neq 500$$

Suposições: Seja X_i o peso do i-ésimo pacote de café. Sabemos que $\mathbb{E}(X_i) = \mu$ e $Var(X_i) = \sigma^2$. Coletou-se uma amostra de tamanho $n = 25$. Pelo TCL:

$$\bar{X} \sim N(\mu, \sigma^2/n)$$

Hipóteses: $H_0 : \mu = \mu_0 = 500 \quad \text{vs} \quad H_a : \mu \neq \mu_0 = 500$

Estatística do teste:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \overset{H_0}{\sim} N(0,1)$$

Considerando a amostra obtida:

$$z_{obs} = \frac{485 - 500}{10/5} = -7.5$$

Como medir se -7.5 é evidência contra H_0?

O teste é bilateral, portanto o valor-de-p é calculado como:

Valor-de-p: $P(|Z| \geq 7.5) = 2P(Z \geq 7.5) \approx 0$

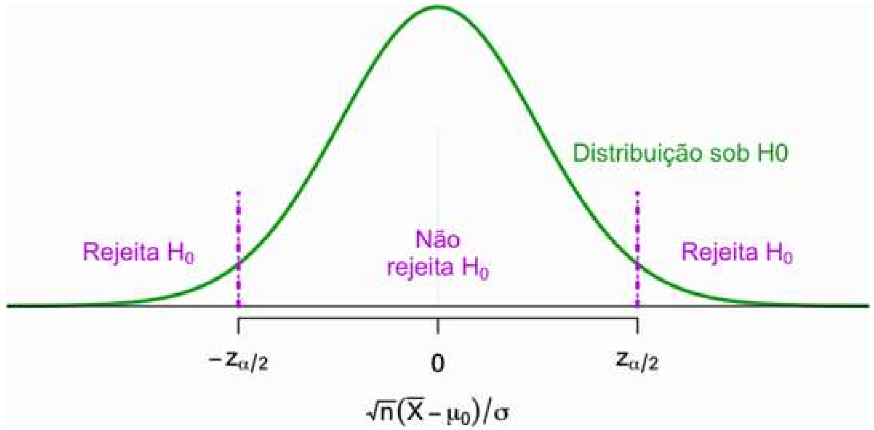
Conclusão: Como o valor-de-p é praticamente zero, rejeitamos H_0, ou seja, rejeitamos a hipótese de que a média é 500g.

Região Crítica (Região de Rejeição)

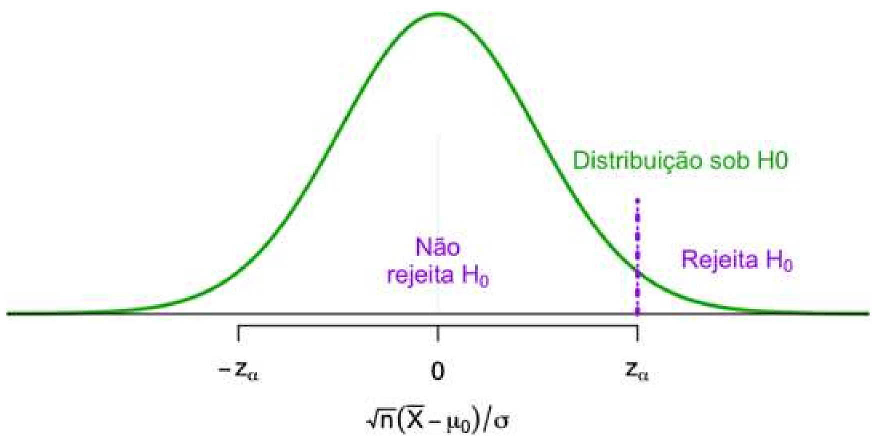
Outra forma de decidirmos se a evidência encontrada nos dados é forte o suficiente para rejeitar H_0 é determinando a **região crítica** ou **região de rejeição**.

Região Crítica: conjunto de valores da estatística do teste para os quais a hipótese nula é rejeitada.

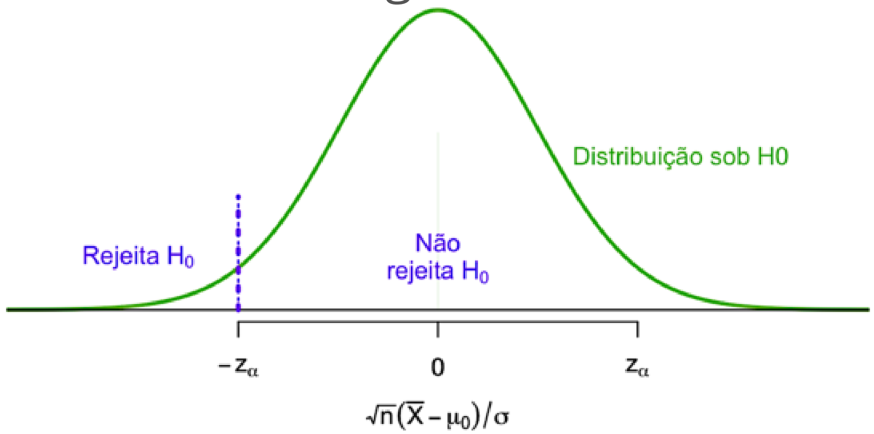
Região crítica: teste bilateral: $H_0 : \mu = \mu_0$ vs $H_a : \mu \neq \mu_0$ e um nível de significância α, definimos a região crítica do teste:



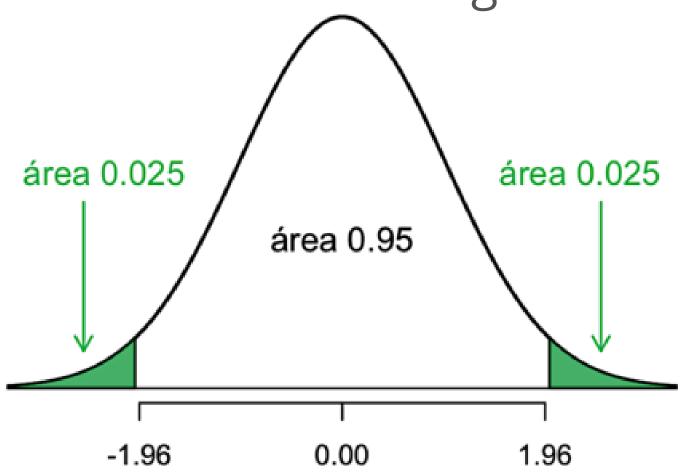
Região crítica: teste unilateral à direita: $H_0 : \mu = \mu_0$ vs $H_a : \mu > \mu_0$ e um nível de significância α, definimos a região crítica do teste:



Região crítica: teste unilateral à esquerda: $H_0 : \mu = \mu_0$ vs $H_a : \mu < \mu_0$ e um nível de significância α, definimos a região crítica do teste:

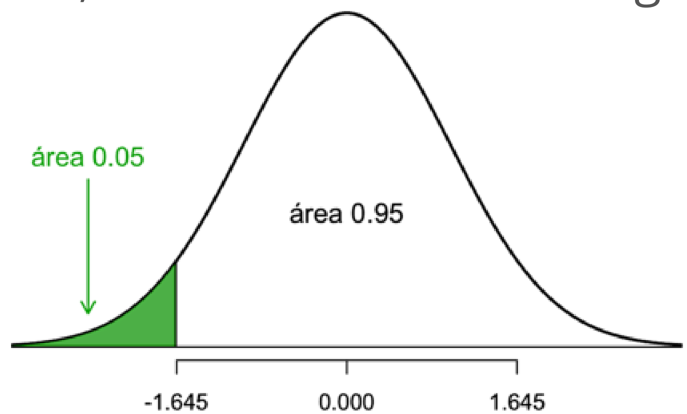


Quando o teste for bilateral: $H_0 : \mu = 500 \quad \text{vs} \quad H_a : \mu \neq 500$. A região critíca, para $\alpha = 0.05$, é a área em verde na figura abaixo:



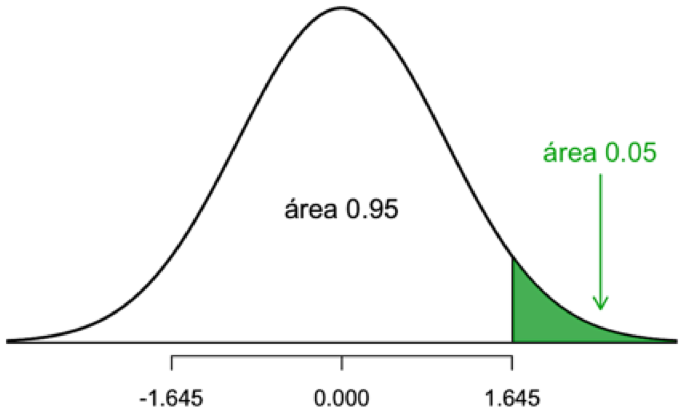
Decisão: Rejeitamos H_0 se $z_{obs} < -1.96$ ou $z_{obs} > 1.96$. No nosso exemplo, $z_{obs} = -7.5$. Portanto, rejeitamos H_0.

Quando o teste for unilateral à esquerda: $H_0 : \mu = \mu_0 \quad \text{vs} \quad H_a : \mu < \mu_0$. A região critíca, para $\alpha = 0.05$, é a área em verde na figura:



Decisão: Rejeitamos H_0 se $z_{obs} < -1.645$.

Região Crítica: teste unilateral à direita Quando o teste for unilateral à direita: $H_0 : \mu = \mu_0 \quad \text{vs} \quad H_a : \mu > \mu_0$. A região crítíca, para $\alpha = 0.05$, é a área em azul na figura:



Decisão: Rejeitamos H_0 se $z_{obs} > 1.645$

Teste de hipóteses para média (σ desconhecido)

No caso de testar

$$H_0 : \mu = \mu_0 \quad \text{vs} \quad H_a : \mu \neq \mu_0$$

quando σ é desconhecido e a amostra é pequena ($n < 30$) devemos utilizar a distribuição t.

Estatística do teste:

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} \overset{H_0}{\sim} t_{n-1}$$

valor-de-p: $P(|t_{n-1}| \geq |t_{obs}|) = 2P(t_{n-1} \geq |t_{obs}|)$

Para as hipóteses unilaterais, o raciocínio é semelhante ao que foi feito anteriormente quando σ é conhecido.

No nosso exemplo, suponha que não sabemos o valor de σ, mas o desvio padrão da amostra é 7.1g. Queremos testar

$$H_0 : \mu = 500 \quad \text{vs} \quad H_a : \mu \neq 500$$

Estatística do teste:

$$t_{obs} = \frac{\bar{x}_{obs} - \mu_0}{s/\sqrt{n}} = \frac{485 - 500}{7.1/5} = -10.56$$

valor-de-p: $P(|t_{24}| \geq 10.56) = 2P(t_{24} \geq 10.56) \approx 0$

Conclusão: Rejeitamos a hipótese de que a média é 500g.

valor crítico: para nível de significância $\alpha = 0.05$ e teste bilateral, t_{crit} é tal que $P(t_{24} > t_{crit}) = P(t_{24} < -t_{crit}) = 0.025$. De maneira que $t_{crit} = 2.06$. Portanto, se $|t_{obs}| > t_{crit}$, rejeita-se H_0.

Exemplo: Dieta LowCarb

- 41 pacientes obesos, selecionados aleatoriamente, foram submetidos a uma dieta com baixa quantidade de carboidratos.
- Pesquisadores responsáveis pelo estudo acreditam que essa dieta faz com que os pacientes apresentem uma redução de peso.
- Após 16 semanas, a diferença média de peso foi -9.7kg, com desvio padrão 3.4 kg.
- O que podemos concluir deste estudo?

Detalhes do estudo: [Effect of 6-month adherence to a very low carbohydrate diet program](#).

Suposições: X_i é a diferença entre peso inicial e final do i-ésimo obeso.

Sabemos que $\mathbb{E}(X_i) = \mu$ e $Var(X_i) = \sigma^2$.

Coletou-se uma amostra de tamanho $n = 41$.

Pelo TCL: $\bar{X} \sim N(\mu, \sigma^2/n)$

Hipóteses: $H_0 : \mu = 0 \quad \text{vs} \quad H_a : \mu < 0$

Ou seja, estamos testando se não há diferença no peso após a dieta versus a hipótese que há redução no peso após a dieta.

Estatística do teste: Como $n = 41$, podemos usar a aproximação normal

$$z_{obs} = \frac{\bar{x}_{obs} - \mu_0}{s/\sqrt{n}} = \frac{-9.7 - 0}{3.4/\sqrt{41}} = -18.3$$

Valor-de-p: Como o teste é unilateral à esquerda

$$\text{valor-de-p} = P(Z < -18.3) \approx 0$$

Conclusão: Como o valor-de-p é bem pequeno (<0.05) rejeitamos H_0, ou seja, rejeitamos a hipótese de que a dieta não produz diferença no peso.

Resumo: Teste de hipóteses para média

σ conhecido	**σ desconhecido**
$H_0: \mu = \mu_0$ vs	$H_0: \mu = \mu_0$ vs
$H_a: \mu \neq \mu_0$ ou	$H_a: \mu \neq \mu_0$ ou
$\mu > \mu_0$ ou $\mu < \mu_0$	$\mu > \mu_0$ ou $\mu < \mu_0$
Estatística do teste:	Estatística do teste:
$Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}} \overset{H_0}{\sim} N(0,1)$	$t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}} \overset{H_0}{\sim} t_{n-1}$
valor-de-p=	valor-de-p=
$P(\lvert Z\rvert \geq \lvert z_{obs}\rvert)$ se $H_a: \mu \neq \mu_0$	$P(\lvert t_{n-1}\rvert \geq \lvert t_{obs}\rvert)$ se $H_a: \mu \neq \mu_0$
$P(Z \geq z_{obs})$ se $H_a: \mu > \mu_0$	$P(t_{n-1} \geq t_{obs})$ se $H_a: \mu > \mu_0$
$P(Z \leq z_{obs})$ se $H_a: \mu < \mu_0$	$P(t_{n-1} \leq t_{obs})$ se $H_a: \mu < \mu_0$