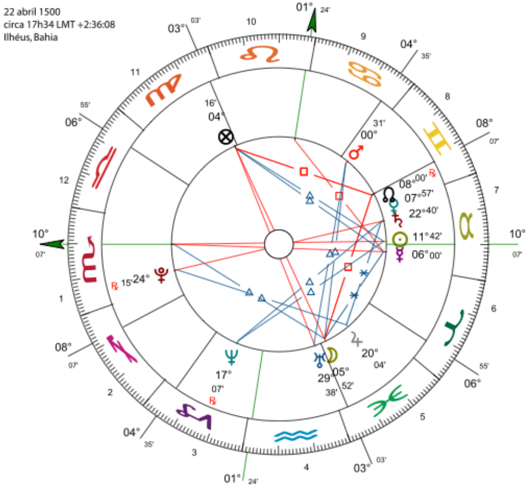
ME951 - Estatística e Probabilidade I

Parte 17

Notas de aula de ME414 produzidas pelos professores **Samara Kiihl**, **Tatiana Benaglia** e **Benilton Carvalho** modificadas e alteradas pela Profa. **Larissa Avila Matos**

Teste de Hipóteses: Introdução

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



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?

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

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".

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

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.

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

28 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?

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

X_i: astrólogo associa corretamente um questionário ao mapa astral i.

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

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?

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

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.

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

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**:

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

No experimento com os 28 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.

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

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".

Passos de um teste de hipótese

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 o que chamamos de **valor-de-p**.

Passos de um teste de hipótese

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 $> \alpha$: não rejeitamos H_0, ou seja, não temos evidência nos dados contra a hipótese nula

Passos de um teste de hipótese

Assumimos primeiro que H_0 é verdadeira.

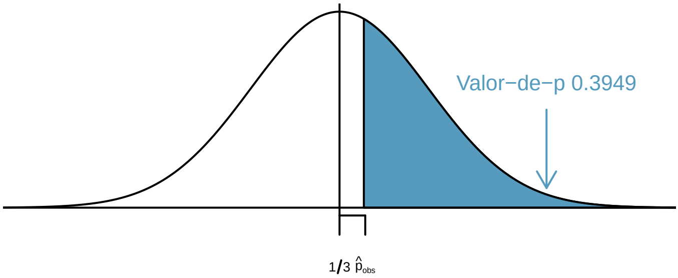
Consideramos então todos os valores possíveis para a estatística do teste, de acordo com sua distribuição amostral.

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: Astrólogos conseguem predizer nossa personalidade com um mapa astral?



Distribuição amostral da proporção amostral $\hat{p}$ sob H_0. A área em azul representa uma probabilidade, que chamamos de valor-de-p: probabilidade de proporção amostral assumir um valor igual ao observado, $\hat{p}_{obs}$, ou mais extremo, sob H_0.

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

Passo 1: suposições.

A variável de interesse é binária.

X_i: astrólogo i associa corretamente um questionário ao mapa astral.

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

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.

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

Passo 2: hipóteses.

H_0: $p = p_0 = 1/3$. Astrólogos *chutam* qual o questionário está associado ao mapa astral.

H_a: $p > p_0 = 1/3$. Astrólogos predizem melhor do que um *chute* qual o questionário está associado ao mapa astral.

Exemplo: Astrólogos conseguem predizer nossa personalidade com um 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?

$\hat{p} \sim \text{Normal}\left(p, \frac{p(1-p)}{n}\right)$.

Se H_0 é verdadeira: $\hat{p} \sim \text{Normal}\left(p_0, \frac{p_0(1-p_0)}{n}\right)$.

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

Passo 3: estatística do teste.

A estatística do teste é:

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

onde $EP_0(\hat{p})$ é o erro padrão de $\hat{p}$ sob H_0: $\sqrt{p_0(1-p_0)/n}$.

Então,

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

A estatística do teste mede quão distante está $\hat{p}$ de p em unidades de "erro-padrão".

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

Passo 3: estatística do teste.

No experimento dos astrólogos, dentre 116 mapas, 40 foram corretamente associados ao questionário de personalidade.
$\hat{p} = 40/116 = 0.345$.

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

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

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

Passo 4: valor-de-p.

$z_{obs} = 0.27$ traz evidência contra H_0 a favor de H_a?

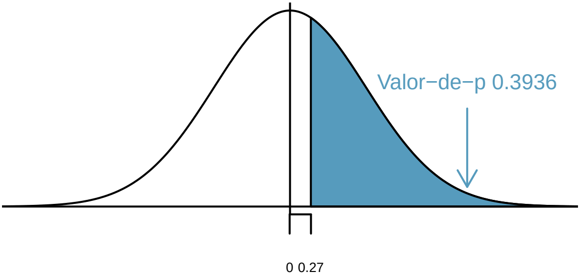
Quão improvável é $z_{obs} = 0.27$ se a proporção de acertos dos astrólogos é de fato $p = p_0 = 1/3$?

Valor-de-p: probabilidade de que uma estatística do teste assuma um valor igual ou mais extremo, assumindo que $p = p_0 = 1/3$.

Mais extremo: neste caso, é um maior valor de 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 $Z \sim \text{Normal}(0, 1)$.

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



Distribuição amostral da estatística do teste Z sob H_0. A área em azul representa a probabilidade de valores mais extremos que z_{obs} ocorrerem, que chamamos de valor-de-p.

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

Passo 5: conclusão.

O valor-de-p obtido no experimento foi 0.3936.

O valor não é tão pequeno.

Não encontramos evidências contra H_0.

Não podemos concluir que astrólogos têm poderes preditivos especiais usando mapa-astral.

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 $> \alpha$: não rejeitamos H_0

Exemplo

Uma indústria farmacêutica diz que menos de 20% dos pacientes que estão usando um certo medicamento terão efeitos colaterais.

Realizou-se então um ensaio clínico com 400 pacientes e verificou-se que 68 pacientes apresentaram efeitos colaterais.

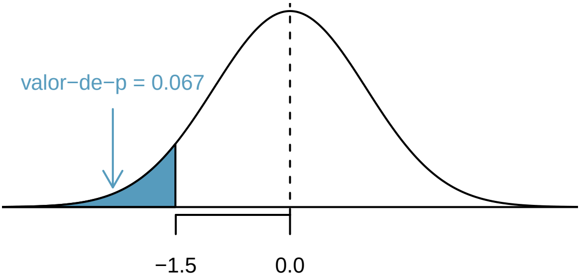
Hipóteses: $H_0 : p = 0.20 \quad \text{vs} \quad H_a : p < 0.20$

Estatística do teste: Da amostra temos que $\hat{p} = 68/400 = 0.17$

$$z_{obs} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.17 - 0.20}{\sqrt{\frac{0.20(1-0.20)}{400}}} = -1.5$$

Exemplo (continuação)

$$\text{valor-de-p} = P(Z \leq -1.5) = 0.067$$

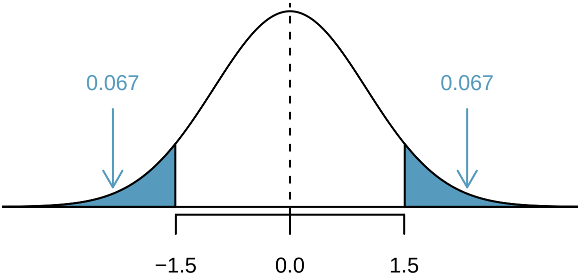


Conclusão: Para $\alpha = 0.05$, como o valor-de-p é maior que 0.05, não temos evidências nos dados para rejeitar a hipótese de que $p = 0.20$. Na verdade, a evidência está na direção que a indústria farmacêutica queria, mas não é o suficiente para rejeitar H_0.

Exemplo (continuação)

E se estivéssemos testando: $H_0 : p = 0.20 \quad \text{vs} \quad H_a : p \neq 0.20$

$$\begin{aligned} \text{valor-de-p} &= P(|Z| \geq 1.5) = P(Z \leq -1.5) + P(Z \geq 1.5) \\ &= 2P(Z \leq -1.5) = 2 \times 0.067 = 0.134 \end{aligned}$$



Conclusão: Para $\alpha = 0.05$, como o valor-de-p é maior que 0.05, não temos evidências nos dados para rejeitar a hipótese de que $p = 0.20$.

Coca vs Coca Zero - você consegue distinguir?

Experimento da Coca vs Coca Zero

Algumas pessoas afirmam que conseguem distinguir o sabor da coca-cola normal da coca zero.

Faremos então um teste para comprovar se a afirmação é verdadeira.

Experimento:

- Sorteia-se, sem a pessoa saber, coca ou coca zero, usando um dado (se sair par, recebe uma coca-cola normal, se sair ímpar, uma coca zero.
- A bebida sorteada é então dada à pessoa sendo testada, que deve então dizer qual ela acredita que é.
- Repetimos isso 20 vezes.
- Anota-se o total de acertos.

Experimento da Coca vs Coca Zero

Suposições:

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

Estamos interessados no total de acertos em 20 tentativas:
$T = \sum_{i=1}^{20} X_i \sim \text{Binomial}(20, p)$.

Podemos usar a aproximação pela Normal caso as condições sejam satisfeitas.

Hipóteses:

$H_0 : p = 1/2$, indicando que a pessoa não consegue diferenciar as duas bebidas.

$H_a : p > 1/2$.

Experimento da Coca vs Coca Zero

Estatística do teste:

Usamos $T = \sum_{i=1}^{20} X_i \sim \text{Binomial}(20, p)$.

Valor-de-p:

Evidência contra H_0. Calculamos a probabilidade, sob H_0, de ocorrer um valor igual ou mais extremo ao valor observado no experimento.

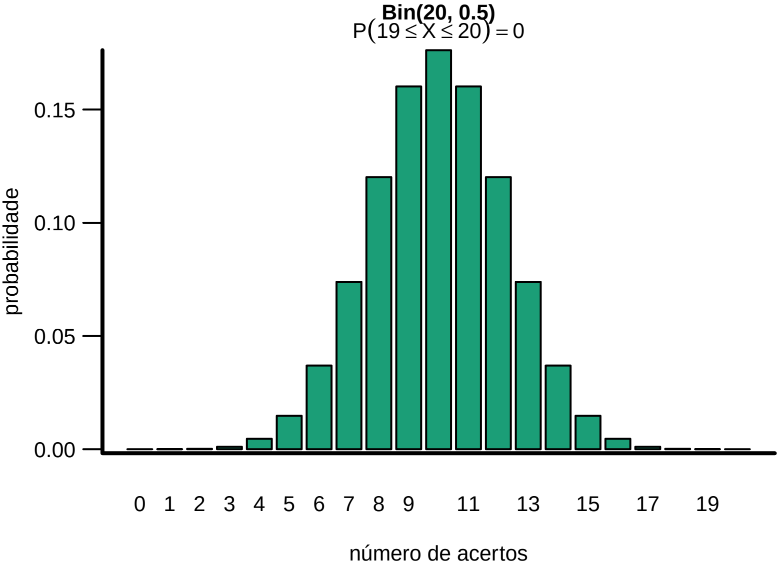
Experimento da Coca vs Coca Zero

Resultado do experimento:

Seja $t_{obs} = 19$ o número de acertos, o valor-de-p foi então:
$P(T \geq 19) = 0$, onde $T \overset{H_0}{\sim}$ Binomial$(20, 1/2)$.

Conclusão: Decidimos rejeitar H_0.

Experimento da Coca vs Coca Zero



Experimento da Coca vs Coca Zero

Utilizando a aproximação pela Normal:

Temos que $T \sim Bin(20, p)$, onde T é o número de acertos.

A proporção amostral de acertos $\hat{p} = \frac{T}{20} = 19/20 = 0.95$.

Vamos testar o seguinte: $H_0 : p = 0.50 \quad \text{vs} \quad H_a : p > 0.50$.

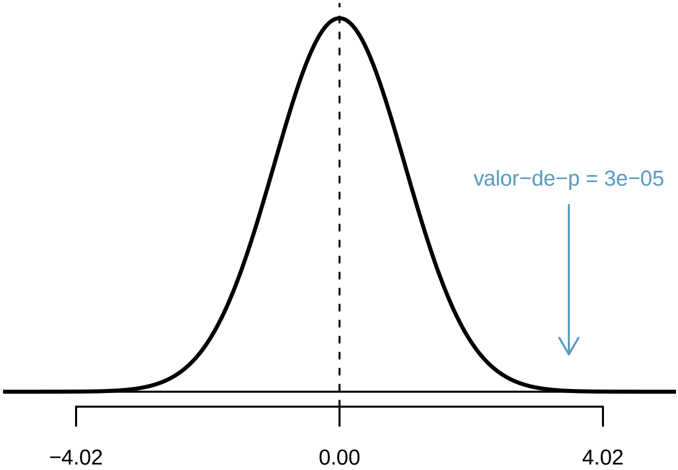
Estatística do teste:

$$z_{obs} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.95 - 0.5}{\sqrt{\frac{(0.50)(0.50)}{20}}} = 4.02$$

valor-de-p $= P(Z \geq 4.02) = 0$

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

Experimento da Coca vs Coca Zero



Leituras

- Ross: capítulo 9.
- OpenIntro: seções 4.3 e 6.1.
- Magalhães: capítulo 8.