ACH2053 – Introdução à Estatística

Aula 12: Teste de Hipótese

Valdinei Freire

valdinei.freire@usp.br

http://www.each.usp.br/valdinei

Escola de Artes, Ciências e Humanidades - USP

2025

Uma senhora toma chá

Em um chá da tarde, uma convidada afirmou:

"Consigo perceber se o leite foi colocado antes ou depois do chá."

Ronald Fisher propôs então um experimento simples para testar essa afirmação:

- Preparar 8 xícaras de chá.
- Em 4 delas, colocar o leite antes do chá.
- Nas outras 4, colocar o leite depois.
- A convidada deveria identificar quais são quais.

O que é um teste de hipótese?

Um **teste de hipótese** é uma ferramenta estatística usada para tomar decisões com base em dados.

- **Hipótese nula ($\mathbf{H}_0$)**: A convidada está apenas chutando.
- **Hipótese alternativa ($\mathbf{H}_1$)**: A convidada realmente consegue distinguir.

Nosso objetivo: verificar se os dados observados fornecem evidência suficiente para rejeitar H_0.

Prova por contradição estocástica: argumento de que os dados não são verossímeis sob a hipótese nula.

Distribuição sob H_0

Se a convidada estiver apenas *chutando*, então ela tem $\binom{8}{4} = 70$ maneiras de escolher 4 xícaras.

Dessas 70 combinações possíveis:

- Apenas 1 corresponde à escolha perfeita (acertar as 4 certas).
- Há 16 maneiras de acertar exatamente 3.
- Há 36 maneiras de acertar 2.
- E assim por diante...

Se ela acertar todas as 4, qual a chance disso acontecer *por acaso*?

$$\Pr(\text{acertar 4 por sorte}) = \frac{1}{70} \approx 0{,}014$$

Rejeição da Hipótese Nula

Se a convidada acertar todas as 4 combinações, a chance disso acontecer por sorte é de apenas 1,4

Decisão

Se escolhermos um nível de significância de $\alpha = 5\%$, então:

- $P < \alpha \Rightarrow$ Rejeitamos H_0.
- Concluímos que há evidências de que a convidada distingue as xícaras.

Se ela acertasse apenas 3, o p-valor seria maior:

$$\Pr(acertos \geq 3) = \frac{1 + 16}{70} = \frac{17}{70} \approx 0,24 \Rightarrow \text{Não rejeitamos } H_0.$$

Estimadores vs Teste de Hipótese

Considere um exame com N questões dicotômicas e que se conhece o modelo paramétrico de cada questão i segundo o modelo TRI:

$$\Pr(X_i = 1|\theta = t) = c_i + (1 - c_i)\frac{e^{a_i(t-b_i)}}{1 + e^{a_i(t-b_i)}}.$$

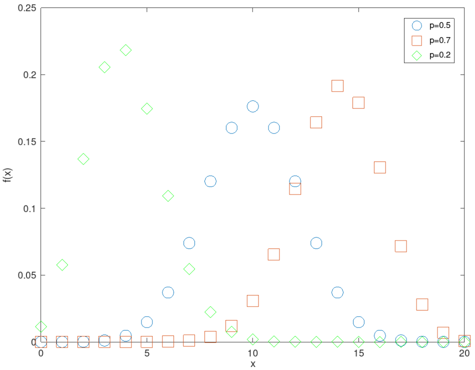
Dadas as respostas de uma aluna às N questões pode-se perguntar:

- **Estimador Pontual** Qual é a habilidade θ da aluna?
- **Teste de Hipótese** A habilidade θ da aluna é pelo menos θ_0?

Exemplo: Moedas Viciadas

Suponha que existem três tipos de moeda, cuja probabilidade de sair cara é, respectivamente, $p_0 = 0.5$, $p_1 = 0.7$, $p_2 = 0.2$. Dada uma amostra de 20 jogadas de uma moeda, crie um procedimento para determinar se a moeda jogada é a moeda com $p = 0.5$ ou não.

Densidade $f(x\|p)$			
x	$p = 0.5$	$p = 0.7$	$p = 0.2$
0	0.00	0.00	0.01
1	0.00	0.00	0.06
2	0.00	0.00	0.14
3	0.00	0.00	0.21
4	0.00	0.00	0.22
5	0.01	0.00	0.17
6	0.04	0.00	0.11
7	0.07	0.00	0.05
8	0.12	0.00	0.02
9	0.16	0.01	0.01
10	0.18	0.03	0.00
11	0.16	0.07	0.00
12	0.12	0.11	0.00
13	0.07	0.16	0.00
14	0.04	0.19	0.00
15	0.01	0.18	0.00
16	0.00	0.13	0.00
17	0.00	0.07	0.00
18	0.00	0.03	0.00
19	0.00	0.01	0.00
20	0.00	0.00	0.00



Teste de Hipótese

Considere um problema estatístico envolvendo um parâmetro $\theta \in \Omega$ com valor desconhecido.

Suponha que Ω pode ser particionado em dois subconjuntos disjuntos Ω_0 e Ω_1.

Denote por H_0 a hipótese de que $\theta \in \Omega_0$ e H_1 a hipótese de que $\theta \in \Omega_1$.

A hipótese H_0 é chamada de hipótese nula e a hipótese H_1 é chamada de hipótese alternativa.

Durante um teste, se decidimos que $\theta \in \Omega_1$, dizemos que rejeitamos H_0. Se decidimos que $\theta \in \Omega_0$, dizemos que não rejeitamos H_0.

Tipos de Erros

Erro do Tipo I A hipótese nula é verdadeira e decide-se rejeitar a hipótese nula.

Erro do Tipo II A hipótese nula é falsa e decide-se não rejeitar a hipótese nula.

Tipos de Erros		**Hipótese Nula (H_0)**	
		Verdadeira	**Falsa**
Decisão sobre a hipótese nula (H_0)	Não rejeitar	Inferência Correta (true negative) $\Pr = 1 - \alpha$	Erro do Tipo II (false negative) $\Pr = \beta$
	Rejeitar	Erro do Tipo I (false positive) $\Pr = \alpha$	Inferência Correta (true positive) $\Pr = 1 - \beta$

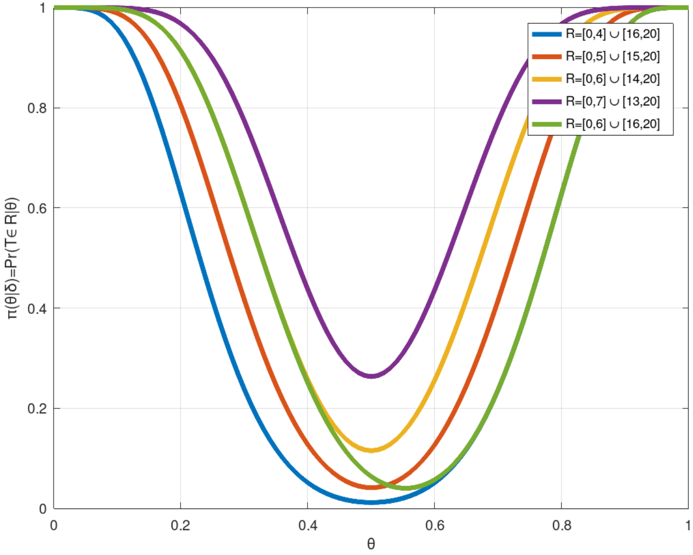
Procedimento de Teste

Região de Rejeição
Seja $\mathbf{X}$ uma amostra aleatória de uma distribuição que depende do parâmetro θ. Seja $T = r(\mathbf{X})$ uma estatística e seja R um subconjunto dos reais. Suponha que um procedimento de teste é da forma "rejeite H_0 se $T \in R$." Então T é uma estatística de teste e R é a região de rejeição do teste.

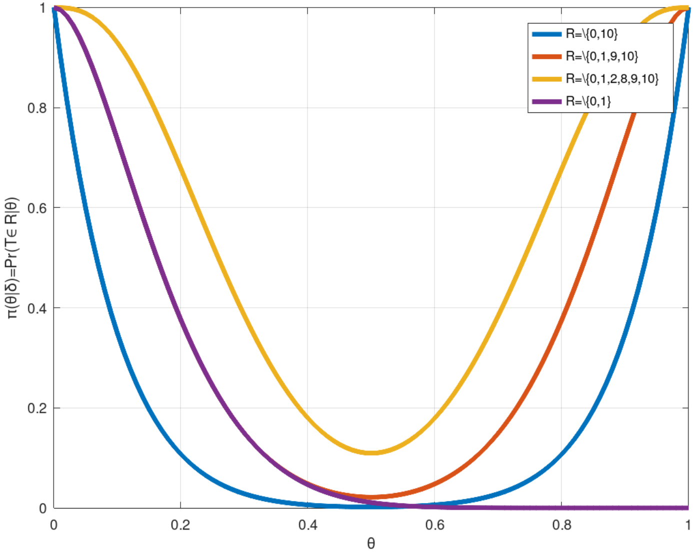
Função Poder do Teste
Seja δ um procedimento de teste. Se δ é descrita em termos de uma estatística de teste T e uma região de rejeição R, a função poder do teste é

$$\pi(\theta|\delta) = \Pr(T \in R|\theta).$$

Função Poder (n=20)



Função Poder (n=10)



Nível de Significância

Uma decisão errada que rejeita uma hipótese nula **verdadeira** é um erro do tipo I e (no pior caso) tem probabilidade $\alpha(\delta) = \max_{\theta \in \Omega_0} \pi(\theta|\delta)$.

Uma decisão errada que não rejeita uma hipótese nula **falsa** é um erro do tipo II e (no pior caso) tem probabilidade $\beta(\delta) = \max_{\theta \in \Omega_1} [1 - \pi(\theta|\delta)]$.

Suponha que um teste δ satisfaça a seguinte condição:

$$\pi(\theta|\delta) \leq \alpha_0, \text{ para todo } \theta \in \Omega_0,$$

então o teste tem nível de significância α_0.

Objetivo: construir teste δ_{α_0} com nível de significância α_0 e que minimize $\beta(\delta)$ (1 - $\beta(\delta)$ é o poder do teste), isto é,

$$\delta_{\alpha_0} = \arg \min_{\{\delta : \alpha(\delta) \leq \alpha_0\}} \beta(\delta) = \arg \max_{\{\delta : \alpha(\delta) \leq \alpha_0\}} 1 - \beta(\delta).$$

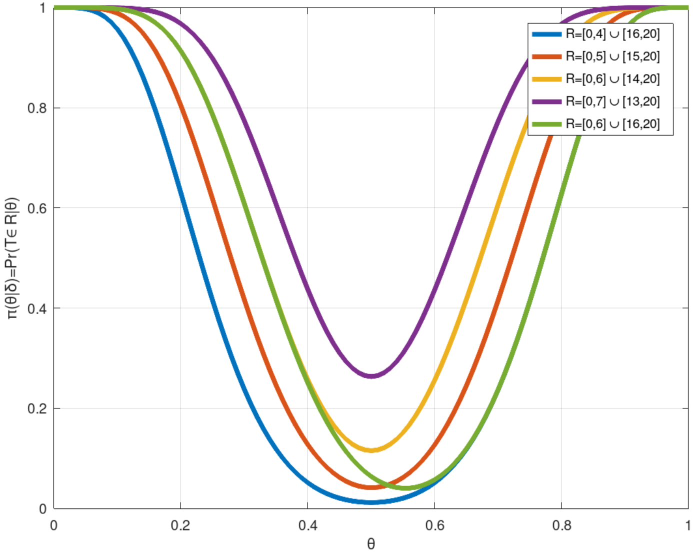
Moeda viciada

Considere que um experimento no qual uma moeda é jogada 20 vezes.

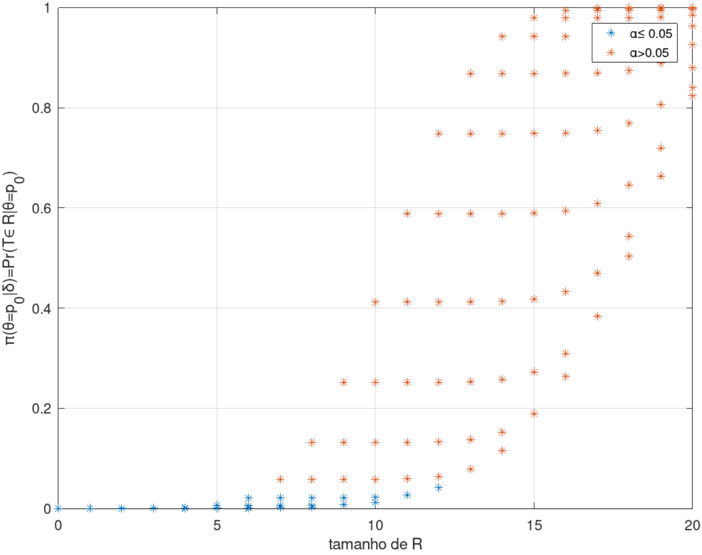
Construa um teste de hipótese para verificar se a moeda é viciada.

- escolha modelo
- escolha hipóteses H_0 e H_1
- escolha a estatística de teste T
- escolha um nível de significância α
- escolha uma região de rejeição R

Função Poder (n=20)

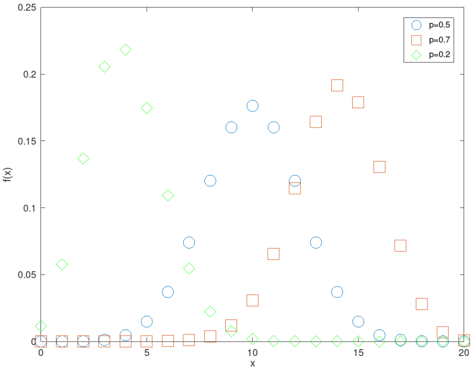


Função Poder (n=20)



Função Poder (n=20)

Densidade $f(x\|p)$			
x	$p = 0.5$	$p = 0.7$	$p = 0.2$
0	0.00	0.00	0.01
1	0.00	0.00	0.06
2	0.00	0.00	0.14
3	0.00	0.00	0.21
4	0.00	0.00	0.22
5	0.01	0.00	0.17
6	0.04	0.00	0.11
7	0.07	0.00	0.05
8	0.12	0.00	0.02
9	0.16	0.01	0.01
10	0.18	0.03	0.00
11	0.16	0.07	0.00
12	0.12	0.11	0.00
13	0.07	0.16	0.00
14	0.04	0.19	0.00
15	0.01	0.18	0.00
16	0.00	0.13	0.00
17	0.00	0.07	0.00
18	0.00	0.03	0.00
19	0.00	0.01	0.00
20	0.00	0.00	0.00



Valor-p

Suponha que o teste δ é do tipo "rejeite a hipótese nula se $T \geq c(\alpha)$", onde α é um nível de significância arbitrário. O valor-p é o menor nível de significância α_0 para o teste δ tal que, dadas as observações, a hipótese nula seria rejeitada.

Vacina

Considere um experimento realizado com a vacina CORONAVAC, no qual:

- 4653 tomaram a vacina e, dentre esses, 85 contraíram o Coronavirus; e
- 4599 tomaram um placebo e, dentre esses, 167 contraíram o Coronavirus.

Crie um teste de hipótese para determinar se a vacina CORONAVAC é diferente do placebo.

- escolha modelo
- escolha hipóteses H_0 e H_1
- escolha a estatística de teste T
- escolha um nível de significância α
- escolha uma região de rejeição R