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.

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.

<table>
<tr><td colspan="2" rowspan="2">Tipos de Erros</td><td colspan="2">Hipótese Nula (H_0)</td></tr>
<tr><td>Verdadeira</td><td>Falsa</td></tr>
<tr><td rowspan="2">Decisão sobre a hipótese nula (H_0)</td><td>Não rejeitar</td><td>Inferência Correta
(true negative)
$\Pr = 1 - \alpha$</td><td>Erro do Tipo II
(false negative)
$\Pr = \beta$</td></tr>
<tr><td>Rejeitar</td><td>Erro do Tipo I
(false positive)
$\Pr = \alpha$</td><td>Inferência Correta
(true positive)
$\Pr = 1 - \beta$</td></tr>
</table>

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

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

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

Moeda viciada

Densidade $f(t\|p)$			
t	$p = 0.5$	$p = 0.7$	$p = 0.2$
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

- escolha modelo

$$\Pr(X = 1) = p \quad \text{e} \quad \Pr(X = 0) = 1 - p$$

- escolha hipóteses H_0 e H_1

$$\begin{aligned} H_0 &: p = 0.5 \\ H_1 &: p \in \{0.2, 0.7\} \end{aligned}$$

- escolha a estatística de teste T

$$T(\mathbf{x}) = \sum_{i=1}^{n} x_i$$

- escolha um nível de significância α

$$\alpha = 0.05$$

- escolha uma região de rejeição

$$R = \{0, 1, 2, 3, 4, 5, 15, 16, 17, 18, 19, 20\}$$

$$\alpha(\delta) = 0.02 \leq 0.05 \quad \text{e} \quad \beta(\delta) = \max\{0.19, 0.57\} = 0.57$$

Moeda viciada

Densidade $f(t\|p)$			
t	$p = 0.5$	$p = 0.7$	$p = 0.2$
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

- escolha modelo

$$\Pr(X = 1) = p \quad \text{e} \quad \Pr(X = 0) = 1 - p$$

- escolha hipóteses H_0 e H_1

$$\begin{aligned} H_0 &: p = 0.5 \\ H_1 &: p \neq 0.5 \end{aligned}$$

- escolha a estatística de teste T

$$T(\mathbf{x}) = \sum_{i=1}^{n} x_i$$

- escolha um nível de significância α

$$\alpha = 0.05$$

- escolha uma região de rejeição

$$R = \{0, 1, 2, 3, 4, 5, 15, 16, 17, 18, 19, 20\}$$

$$\alpha(\delta) = 0.02 \leq 0.05 \quad \text{e} \quad \beta(\delta) = 0.98$$

Moeda viciada

Densidade $f(t|p)$			
t	$p = 0.5$	$p = 0.7$	$p = 0.2$
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

- escolha modelo

$$\Pr(X = 1) = p \quad \text{e} \quad \Pr(X = 0) = 1 - p$$

- escolha hipóteses H_0 e H_1

$$H_0 : p \leq 0.5$$
$$H_1 : p > 0.5$$

- escolha a estatística de teste T

$$T(\mathbf{x}) = \sum_{i=1}^{n} x_i$$

- escolha um nível de significância α

$$\alpha = 0.05$$

- escolha uma região de rejeição

$$R = \{14, 15, 16, 17, 18, 19, 20\}$$

$$\alpha(\delta) = 0.05 \leq 0.05 \quad \text{e} \quad \beta(\delta) = 0.95$$

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.

Suponha que foi observado $t = 13$, encontre o valor-p nos seguintes casos:

- $H_0 : p = 0.5$
- $H_0 : p \leq 0.5$

Distribuição Normal (σ conhecido)

- escolha modelo

$$X \sim \mathcal{N}(\mu, \sigma)$$

- escolha hipóteses H_0 e H_1

$$H_0 : \mu = 0$$
$$H_1 : \mu \neq 0$$

- escolha a estatística de teste T

$$T(\mathbf{x}) = \frac{1}{\sqrt{n}\sigma} \sum_{i=1}^{n} x_i$$

- escolha um nível de significância α

$$\alpha = 0.05$$

- escolha uma região de rejeição

$$R = \{x : |x| > 1.96\}$$

Distribuição Normal (σ conhecido)

- escolha modelo

$$X \sim \mathcal{N}(\mu, \sigma)$$

- escolha hipóteses H_0 e H_1

$$H_0 : \mu \leq 0$$
$$H_1 : \mu > 0$$

- escolha a estatística de teste T

$$T(\mathbf{x}) = \frac{1}{\sqrt{n}\sigma} \sum_{i=1}^{n} x_i$$

- escolha um nível de significância α

$$\alpha = 0.05$$

- escolha uma região de rejeição

$$R = \{x : x > 1.645\}$$

Distribuição Normal - duas amostras (σ conhecido)

Observa-se $x_1, \dots, x_m$ e $y_1, \dots, y_n$.

- escolha modelo

$$X \sim \mathcal{N}(\mu, \sigma)$$

- escolha hipóteses H_0 e H_1

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

- escolha a estatística de teste T

$$T(\mathbf{x}, \mathbf{y}) = \frac{1}{\sigma\sqrt{\frac{1}{m} + \frac{1}{n}}}\left(\frac{1}{m}\sum_{i=1}^{m} x_i - \frac{1}{n}\sum_{i=1}^{n} y_i\right) = \frac{\bar{X} - \bar{Y}}{\sigma\sqrt{\frac{1}{m} + \frac{1}{n}}}$$

- escolha um nível de significância α

$$\alpha = 0.05$$

- escolha uma região de rejeição

$$R = \{x : |x| > 1.96\}$$

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

Distribuição Bernoulli - duas amostras

Observa-se $x_1, \ldots, x_m$ e $y_1, \ldots, y_n$.

- escolha modelo
$$\Pr(X = 1) = p_1 \quad \text{e} \quad \Pr(Y = 1) = p_2$$
- escolha hipóteses H_0 e H_1
$$H_0 : p_1 = p_2$$
$$H_1 : p_1 \neq p_2$$
- escolha a estatística de teste T
$$T(\mathbf{x}, \mathbf{y}) = \frac{\bar{X} - \bar{Y}}{\hat{p}(1-\hat{p})\sqrt{\left(\frac{1}{m} + \frac{1}{n}\right)}}$$
- escolha um nível de significância α
$$\alpha = 0.05$$
- escolha uma região de rejeição
$$R = \{x : |x| > c\} \quad \text{e} \quad \max_p \sum_{x \in R} \Pr(T = x | p_1 = p_2 = p) \leq \alpha$$

Teste de Fisher

A New Test for 2 × 2 Tables

If an experiment yields results in the form of a 2×2 table:

	P	not-P	Total
A	a	c	m
B	b	d	n
Total	r	s	N

where m and n have been fixed in advance, a test for deciding whether there is evidence of association between the attributes A and B and the mutually exclusive and exhaustive attributes P and not-P has been given by Fisher[1]. This test, on the null hypothesis of no association, associates the results with a probability $m!n!r!s!/N!a!b!c!d!$.

It has been usual in the past to apply this test, or some approximation to it[2,3], in cases where we can assume that the probability p_1 that A has P and the probability p_2 that B has P are both constant. The hypothesis tested then becomes $H(p) \equiv p_1 = p_2 = p$.

Teste de Barnard

It is, however, possible to construct a more powerful test of the hypothesis $H(p)$ on the data given. The table above can be represented geometrically by the point (a,b) in a plane lattice diagram of points with integer co-ordinates. Since m and n are fixed in advance, all possible results will then be represented by points of the lattice lying in a rectangle bounded by the x and y axes and the lines $x = m$ and $y = n$. The hypothesis $H(p)$ assigns a 'weight'

$$W(a,b,p) = (m!n!/a!b!c!d!)p^r(1-p)^s$$

to the point (a,b). To obtain a valid test of $H(p)$ on significance level α, we have only to choose a region R in the rectangle such that

$$\underset{0 \leqslant p \leqslant 1}{\text{Max}} \sum_R W(a,b,p) \leqslant \alpha.$$

This validity condition does not determine R uniquely. To obtain a reasonable test, we must require that R should consist of as many points as possible, and should lie away from that diagonal of the rectangle which passes through the origin. Formulated mathematically, these latter requirements mean that the complement of R must in a certain sense be convex, symmetrical and minimal.

For example, when $m = n = 3$, and $\alpha = 1/32$, R consists of the two points (3,0), (0,3). The corresponding level of significance with Fisher's test is 1/10. Thus the new test is more powerful than Fisher's.