NOÇÕES DE TESTE DE HIPÓTESES

Estimação	Teste de Hipóteses
Qual é a probabilidade de "cara"no lançamento de uma moeda?	A moeda é honesta ou é desequilibrada?
Qual é a proporção de eleitores favoráveis ao candidato A?	O candidato A tem até 50% das intenções de voto ou tem mais?
Qual é o nível médio de colesterol para homens hipertensos e fumantes?	O nível médio de colesterol para homens hipertensos e fumantes é maior que 211mg/100ml?

Teste de hipóteses para a proporção populacional

Exemplo 1: **Queremos saber se uma moeda é honesta.**

Formulamos duas hipóteses:

H: a moeda é honesta (hipótese nula)

A: a moeda não é honesta (hipótese alternativa)

Essas hipóteses podem ser reescritas como:

H: $p = 0{,}5$
A: $p \neq 0{,}5$

sendo *p* a probabilidade de "cara".

Uma hipótese estatística é uma afirmação sobre um *parâmetro* da distribuição de uma variável aleatória.

Lançamos a moeda *n* vezes e observamos a proporção de caras $\hat{p} = \dfrac{X}{n}$ **(X é o nº de caras).**

Então, X ~ b (n; p).

Quando *n* é grande, $\hat{p}$ **tem distribuição aproximadamente normal, ou seja,** $\hat{p} \sim N\left(p, \dfrac{p(1-p)}{n}\right)$

Se observarmos 30 caras em 50 lançamentos, temos $\hat{p} = 0{,}60$, qual será nossa conclusão?

Para decidir, vamos realizar um teste de hipóteses.

Teste de uma hipótese estatística é uma regra que nos permite, com base na informação de uma amostra, decidir pela rejeição ou não de H.

Uma regra de decisão poderia ser, por exemplo, a seguinte:

- **se observamos $\hat{p} \leq 0{,}35$ ou $\hat{p} \geq 0{,}65$ rejeitamos H;**
- **caso contrário, não rejeitamos H.**

Nesse caso, o conjunto dos valores de $\hat{p}$ que levam à rejeição de H é $\{\ \hat{p};\ \hat{p} \leq 0{,}35$ ou $\hat{p} \geq 0{,}65\ \}$.

Região Crítica (Rc) ou região de rejeição é o conjunto de valores assumidos pela variável aleatória $\hat{p}$ para os quais a hipótese H é rejeitada.

Então,

$$Rc = \{\ \hat{p}\ ;\ \hat{p} \leq 0{,}35 \text{ ou } \hat{p} \geq 0{,}65\ \}: \textit{região crítica}$$

e

$$Rc^c = \{\ \hat{p}\ ;\ 0{,}35 < \hat{p} < 0{,}65\ \}: \textit{região de aceitação}.$$

Regra de decisão (teste):

$$\hat{p}_{obs} \in Rc \;\Rightarrow\; \text{rejeitamos H}$$
$$\hat{p}_{obs} \notin Rc \;\Rightarrow\; \text{não rejeitamos H}$$

No exemplo da moeda, observamos $\hat{p} = 0{,}60$ caras. $0{,}60 \notin Rc \Rightarrow$ não rejeitamos H (não temos evidência de que a moeda é desequilibrada).

Será que nossa conclusão está correta?

Ao decidir pela rejeição ou não da hipótese nula H, podemos cometer *dois tipos de erro*.

Erros

Erro tipo I: Rejeitar H quando H é verdadeira (afirmar que uma moeda não é honesta quando, na verdade, ela é).

Erro tipo II: Não rejeitar H quando H é falsa (afirmar que uma moeda é honesta quando, na verdade, é desequilibrada).

Probabilidades de erros

P(erro I) = P(rejeitar H | H é verdadeira) = α

α : nível de significância do teste

P(erro II) = P(não rejeitar H | H é falsa) = β

1 - β : poder do teste

Exemplo 2: Suponha que um medicamento existente no mercado produza o efeito desejado em 60% dos casos nos quais o mesmo é aplicado.

Laboratório: produz um novo medicamento e afirma que ele é melhor do que o existente.

Objetivo: Verificar estatisticamente se a afirmação do laboratório é verdadeira.

Hipóteses estatísticas:

$$\boxed{\begin{array}{l} H: p = 0{,}6 \\ A: p > 0{,}6 \end{array}}$$

sendo *p* a proporção de pacientes para os quais o novo medicamento é eficaz.

Aplicou-se o medicamento em n = 10 pacientes.

Seja,

X : nº de pacientes, dentre os 10, para os quais o novo medicamento produz o efeito desejado,

então X ~ b (10; p).

Simplificando, suponha que o laboratório afirma que o novo medicamento produz efeito desejado em 90% dos casos, ou seja,

H: p = 0,6
A: p = 0,9

Se H é verdadeira (p = 0,6), então E(X) = 6 pacientes.

Se H é falsa (p = 0,9), então E(X) = 9 pacientes.

Pergunta: ***Quando devemos rejeitar H*?**

$\hat{p}$

0,0 0,1 0,2 ... 0,55 0,6 0,65… k… 0,9 1

(H) ?

$$\mathbf{Rc = \{\ \hat{p} \geq k\ \}}$$

Por exemplo, Rc= {0,8; 0,9; 1} e Rc^C={0;0,1;0,2;..., 0,7}

Adotando Rc = {8, 9, 10},

α **= P(erro I)**

= P(rejeitar H | H é verdadeira)

= P($\hat{p} \in$ **Rc | p = 0,6)**

$= P(\hat{p} \geq 0{,}8 \mid p = 0{,}6)$

= P(X = 8 | p = 0,6) + P(X = 9 | p = 0,6) + P(X = 10 | p = 0,6)

= 0,1209 + 0,043 + 0,0060 = 0,1672

β **= P(erro II)**

= P(não rejeitar H | H é falsa)

= P($\hat{p} \notin$ **Rc | p = 0,9)**

$= P(\hat{p} \leq 0{,}7 \mid p = 0{,}9)$

= P(X = 0 | p=0,9) + ... + P(X = 7 | p=0,9)

= 0,0702

Binomial com
n = 10 e p = 0,60

x	P(X = x)
0	0,0001
1	0,0016
2	0,0106
3	0,0425
4	0,1115
5	0,2007
6	0,2508
7	0,2150
8	0,1209
9	0,0403
10	0,0060

Binomial com
n = 10 e p = 0,90

x	P(X = x)
3	0,0000
4	0,0001
5	0,0015
6	0,0112
7	0,0574
8	0,1937
9	0,3874
10	0,3487

Decisão	Verdadeiro valor de p	
	p = 0,6	p = 0,9
Não rejeitar H	Decisão correta 0,8328	Erro II β = 0,0702
Rejeitar H	Erro I α = 0,1672	Decisão correta 0,9298

P(não rejeitar H | H é verdadeira)
= 1 - P(rejeitar H | H é verdadeira) = 1 - α = 0,8328
(probabilidade de uma decisão correta)

P(rejeitar H | H é falsa)
= 1 - P(não rejeitar H | H é falsa) = 1 - β = 0,9298
(probabilidade de uma decisão correta)

Regiões críticas e probabilidades de erros

Rc	α	β
{8, 9, 10}	0,1672	0,0702
{9, 10}	0,0463	0,2639
{10}	0,0060	0,6513

Voltando ao caso real, temos

$$\boxed{\begin{array}{l} H: p = 0{,}6 \\ A: p > 0{,}6 \end{array}}$$

Se Rc = {8, 9, 10}, temos α = 0,1672.

$\beta(p) = P(\hat{q} \notin Rc \mid H \text{ é falsa})$

$= P(X=0) + \ldots + P(X=7), \quad p > 0{,}6$

$\beta(0{,}7) = P(X=0|p=0{,}7) + \ldots + P(X=7|p=0{,}7) = 0{,}6170$

$\beta(0{,}8) = P(X=0|p=0{,}8) + \ldots + P(X=7|p=0{,}8) = 0{,}3220$

$\beta(0{,}9) = P(X=0|p=0{,}9) + \ldots + P(X=7|p=0{,}9) = 0{,}0702$

Binomial com
n = 10 e p = 0,70

x	P(X = x)
0	0,0000
1	0,0001
2	0,0014
3	0,0090
4	0,0368
5	0,1029
6	0,2001
7	0,2668
8	0,2335
9	0,1211
10	0,0282

Binomial com
n = 10 e p = 0,80

x	P(X = x)
1	0,0000
2	0,0001
3	0,0008
4	0,0055
5	0,0264
6	0,0881
7	0,2013
8	0,3020
9	0,2684
10	0,1074

Binomial com
n = 10 e p = 0,90

x	P(X = x)
3	0,0000
4	0,0001
5	0,0015
6	0,0112
7	0,0574
8	0,1937
9	0,3874
10	0,3487

Até agora, o procedimento foi:

escolher Rc $\Rightarrow$ determinar α

Alternativamente, podemos

fixar α $\Rightarrow$ determinar Rc

Determinação da região crítica

Exemplo 3: Proporção de analfabetos em uma cidade.

H: $p = 0{,}15$
A: $p < 0{,}15$

X: nº de analfabetos em 60 cidadãos entrevistados

$$X \sim b(60; p)$$

Vamos fixar $\alpha = 0{,}05$.

Qual deve ser a Rc para este valor de α?

A região crítica deve ter a forma:

$$Rc = \{ \hat{q} \leq k \}$$

O valor de k deve ser tal que P(erro I) = α, ou seja,

$$P(\hat{q} \leq k \mid p=0{,}15) = 0{,}05.$$

$$P(\hat{p} \leq k \mid p = 0{,}5) = 0{,}05$$

$$P\left(Z \leq \frac{k - 0{,}15}{\sqrt{\frac{0{,}15 \times 0{,}85}{60}}}\right) = 0{,}05$$

$$\Rightarrow \frac{k - 0{,}15}{\sqrt{\frac{0{,}15 \times 0{,}85}{60}}} = -1{,}65 \Rightarrow k = 0{,}074$$

Binomial com n = 60 e p = 0,15

x	P(X = x)
0	0,0001
1	0,0006
2	0,0032
3	0,0110
4	0,0275
5	0,0544
6	0,0880
7	0,1199
8	0,1401
9	0,1429
10	0,1286
11	0,1031
12	0,0743
13	0,0484
14	0,0287
15	0,0155
16	0,0077
17	0,0035
18	0,0015
19	0,0006
20	0,0002
21	0,0001
22	0,0000

$$Rc = \{\ \hat{q} \leq 0{,}074\ \}$$

Se em 60 entrevistados observarmos 6 analfabetos, qual será a conclusão?

0,1 $\notin$ Rc $\Rightarrow$ não rejeitamos H
isto é,
não temos evidência para afirmar que a proporção de analfabetos é inferior a 15%.

Exemplo 4: Um industrial afirma que seu processo de fabricação produz 90% de peças dentro das especificações. O IPEM deseja investigar se este processo de fabricação ainda está sob controle.

Sendo p a proporção de peças dentro das especificações, as hipóteses de interesse são:

$$H: p = 0,9$$
$$A: p < 0,9$$

Selecionamos uma amostra aleatória de 15 itens e observamos o no.de itens satisfatórios (X), então

$$X \sim b(15, p).$$

proporção itens satisfatórios $\hat{p} = X/15$

Região crítica: $Rc = \{\hat{p} \leq k\}$ ou $\{X \leq 15k\}$

Logo, para α = 6% temos k = 11/15 e Rc = {X $\leq$ 11}.

Para α = 1% temos k = 9/15 e Rc = {X $\leq$ 9}.

Se observamos X = 10 peças satisfatórias, então

a) se α = 6% $\Rightarrow$ 10 $\in$ Rc

Rejeitamos H ao nível de significância de 6%.

b) se α = 1% $\Rightarrow$ 10 $\notin$ Rc

Não rejeitamos H ao nível de significância de 1%.

Crítica: arbitrariedade na escolha da Rc (nível de significância).

Binomial com n = 15 e p = 0,90

x	P (X = x)
0	0,0000
1	0,0000
2	0,0000
3	0,0000
4	0,0000
5	0,0000
6	0,0000
7	0,0000
8	0,0003
9	0,0019
10	0,0105
11	0,0428
12	0,1285
13	0,2669
14	0,3432
15	0,2059

Sugestão: determinar um nível de significância associado à evidência experimental.

NÍVEL DESCRITIVO: P (ou "**P**-value")

Nesse exemplo temos, $R_c = \{\hat{p} \leq k\}$ ou $\{X \leq 15k\}$

$X_{obs} = 10.$

Qual é o nível de significância k = 10/15?

$\mathbf{P} = P(\hat{p} \leq 10/15 \mid p = 0{,}9)$

$\quad = P(X \leq 10 \mid p = 0{,}9) = 0{,}0127$

Como concluir?

P é "pequeno" $\Rightarrow$ rejeitamos H

P é "grande" $\Rightarrow$ não rejeitamos H

Para um α fixado,

$P \leq \alpha \Rightarrow$ rejeitamos H

$P > \alpha \Rightarrow$ não rejeitamos H

Justificativa:

Se o nível de significância α é maior do que o nível descritivo *P*, então a região crítica de nível α contém o valor observado e, portanto, rejeitamos H.

No exemplo, adotando $\alpha = 0,05$

a) $P = 0,0127$

temos $P < \alpha$ e, portanto, rejeitamos H.

Se observarmos X = 12, *qual será o nível descritivo?*

P = P (X ≤ 12 | p = 0,9) = 0,1840

Conclusão: não rejeitamos H.

Nível Descritivo é o menor nível de significância para o qual o resultado observado é significante, ou seja, que conduz à rejeição da hipótese nula H.

ou é a probabilidade de observamos um valor tão extremo como o da amostra quando H é verdadeira.

Exemplo 5: A diretoria de uma escola acredita que neste ano a proporção p de alunos usuários da Internet é maior que os 70% encontrados no ano anterior. Se uma pesquisa com 30 alunos, escolhidos ao acaso, mostrou que 26 são usuários da Internet, podemos concluir que a afirmação da diretoria é verdadeira?

Hipóteses estatísticas:

H: $p = 0,7$
A: $p > 0,7$

X: nº de usuários em 30 entrevistados, então

$X \sim b(30, p)$.

Se observamos $X = 26$ usuários da Internet, então

$$\hat{p} = 26/30 = 0{,}87 \ ,$$

e o nível descritivo do teste é dado por:

$$P = P(\hat{p} \geq 0{,}87 \mid p = 0{,}7)$$
$$= P(X \geq 26 \mid p = 0{,}7) = 0{,}0301$$

Conclusão:

Se $\alpha = 5\%$, $P < \alpha \Rightarrow$ rejeitamos H, ou seja, os dados indicam que a proporção de usuários aumentou (é maior que 70%).

Binomial com n = 30 e p = 0,70

x	P (X = x)
10	0,0000
11	0,0001
12	0,0005
13	0,0015
14	0,0042
15	0,0106
16	0,0231
17	0,0444
18	0,0749
19	0,1103
20	0,1416
21	0,1573
22	0,1501
23	0,1219
24	0,0829
25	0,0464
26	0,0208
27	0,0072
28	0,0018
29	0,0003
30	0,0000

Exemplo 6: **Proporção de analfabetos em uma cidade**

Pelo Anuário IBGE/1988: 15% de analfabetos.

Duas pesquisas:

Em 1993, entre 200 entrevistados, 27 eram analfabetos.

Em 1998, entre 200 entrevistados, 19 eram analfabetos.

Esses dados suportam a tese de diminuição do analfabetismo na cidade de 1988 para 1993?
E de 1988 para 1998?

Em ambos os casos, o teste de interesse é das hipóteses:

$$\begin{array}{|l|}\hline H: p = 0{,}15 \\ A: p < 0{,}15 \\ \hline\end{array}$$

sendo p é a proporção populacional de analfabetos na cidade (no ano da pesquisa).

X: nº de analfabetos em 200 cidadãos entrevistados, então **X ~ b(200, p)**.

Sob H, $\hat{p} \sim N\left(0{,}15\,,\ \dfrac{0{,}15(1-0{,}15)}{200}\right)$ (aproximadamente)

Nível descritivo:

Em 1993, $\hat{p} = 0{,}135 \Rightarrow$ ***P*** **= P(** $\hat{p} \le$ **0,135 |** ***p*** **= 0,15)**

$$\cong P\left(Z \le \frac{0{,}135 \ \text{-}\ 0{,}15}{0{,}025} \right) = P\left(Z \le -0{,}60\right) = 0{,}27$$

Em 1998, $\hat{p} = 0{,}095 \Rightarrow$ ***P*** **= P(** $\hat{p} \le$ **0,095 |** ***p*** **= 0,15)**

$$\cong P\left(Z \le \frac{0{,}095\ \text{-}\ 0{,}15}{0{,}025} \right) = P\left(Z \le -2{,}20\right) = 0{,}014$$

Conclusão: **fixando α = 5%,**

em 93, *P* > α $\Rightarrow$ não rejeitamos H

em 98, *P* < α $\Rightarrow$ rejeitamos H