Estatística aplicada a ensaios clínicos

Luís Vicente Garcia
Disciplina de Anestesiologia

Faculdade de Medicina de Ribeirão Preto
Universidade de São Paulo

Aula 18

Luís Vicente Garcia
lvgarcia@fmrp.usp.br

Faculdade de Medicina de Ribeirão Preto
Universidade de São Paulo

Estatística aplicada a ensaios clínicos

testes para dados categorizados

Aula reproduzida a partir do capítulo 9 do livro "Bioestatística Teórica e Computacional" de Hector Gustavo Arango Editora Guanabara Koogan

dados categorizados

contagem de frequência de uma variável classificada ou subdividida em categorias

- serve para variáveis qualitativas
- serve para variáveis quantitativas também

grupo	masculino	feminino
controle	80 %	20 %
tratamento	75 %	25%

Peso (Kg)	masculino	feminino
até 100 Kg	50 %	50 %
acima de l00 Kg	80 %	20%

tabela de contingência

fator discriminante x (variável independente)

fator discriminado Y (variável dependente)

	discriminado 1	discriminado 2	discriminado 3	total X
discriminante 1	x1,y1	x1,y2	x1,y3	X1
discriminante 2	x2,y1	x2,y2	x2,y3	X2
discriminante 3	x3,y1	x3,y2	x3,y3	X3
discriminante 4	x4,y1	x4,y2	x4,y3	X4
total Y	Y1	Y2	Y3	total geral

tabela de contingência

fator discriminante X (variável independente)

fator discriminado Y (variável dependente)

	discriminado 1	discriminado 2	discriminado 3	total X
discriminante 1	$O_{1,1}$	$O_{1,2}$	$O_{1,3}$	X1
discriminante 2	$O_{2,1}$	$O_{2,2}$	$O_{,2,3}$	X2
discriminante 3	$O_{3,1}$	$O_{,3,2}$	$O_{3,3}$	X3
discriminante 4	$O_{4,1}$	$O_{,4,2}$	$O_{4,3}$	X4
total Y	Y1	Y2	Y3	total geral

O = observado

Testes p/ dados categorizados

determinar se o fator discriminante exerce alguma influência sobre o fator discriminado

H0: as categorias de X exercem a mesma influência sobre as categorias de Y

H1: pelo menos uma categoria de X exerce influência diferente sobre as categorias de Y

Tipos de tabela de contingência

2 x 2

2 x 3

2 x 4

...

...

...

...

...

Testes p/ dados categorizados

☼ **Qui-quadrado clássico**
☼ **Exato de Fisher**
☼ **McNemar**
☼ **Mantel-Haenszel**

qui-quadrado clássico

1. construir a matriz dos valores esperados, de dimensão r x s (r = linha, s = coluna)

2. valores esperados são calculados conforme abaixo:

$$E_{ij} = \frac{\sum_{j=1}^{s} O_{ij} \cdot \sum_{i=1}^{r} O_{ij}}{\sum_{i=1}^{s} \sum_{j=1}^{r} O_{ij}} = \frac{A_i \cdot B_j}{T}$$

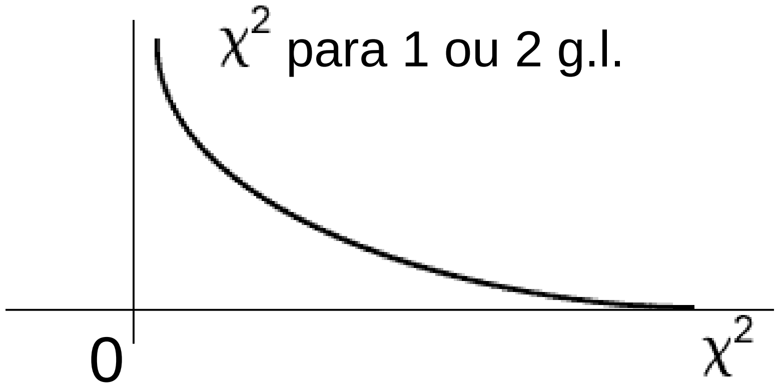
qui-quadrado clássico

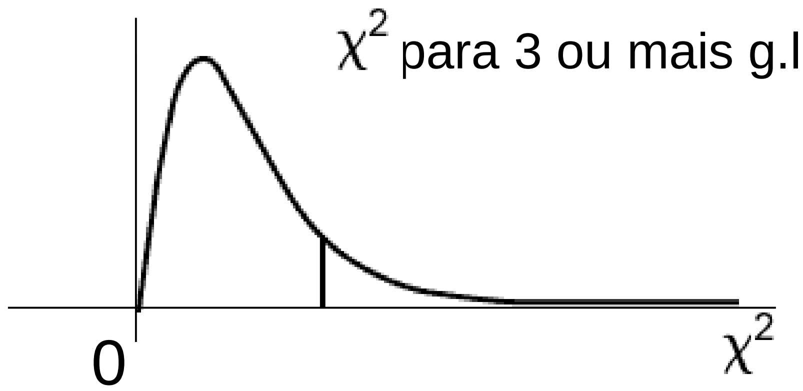
3. a estatística teste é:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Distribuição qui-quadrado

Alguns testes usam a distribuição de probabilidade conhecida como qui-quadrado (χ^2)





é uma família de distribuições. O gráfico depende do número de graus de liberdade (número de opções disponíveis) em um experimento estatístico.

As distribuições são anti-simétricas à direita e o valor de qui-quadrado é igual ou superior a 0.

exemplo

	cefaleia		
Grupo	sim	não	total
vicentex	6	57	63
placebo	30	61	91
Total	36	118	154

H_0: vicentex = placebo

H_1: vicentex ≠ placebo

qui-quadrado clássico

1. construir a matriz dos valores esperados, de dimensão r x s (r = linha, s = coluna)

2. valores esperados são calculados conforme abaixo:

$$E_{ij} = \frac{\sum_{j=1}^{s} O_{ij} \cdot \sum_{i=1}^{r} O_{ij}}{\sum_{i=1}^{s} \sum_{j=1}^{r} O_{ij}} = \frac{A_i \cdot B_j}{T}$$

	cefaleia		
Grupo	sim	não	total
vicentex	6	57	63
placebo	30	61	91
Total	36	118	154

	cefaleia		
Grupo	sim	não	total
vicentex	E1,1	E1,2	63
placebo	E2,1	E2,2	91
Total	36	118	154

construção da matriz E (esperado)

E1,1/63 = 36/154 ➡ E1,1 = 14,73

E1,2/63 = 118/154 ➡ E1,2 = 48,27

E2,1/91 = 36/154 ➡ E2,1 = 21,27

E2,2/91 = 118/154 ➡ E2,2 = 69,73

	cefaleia		
Grupo	sim	não	total
vicentex	6	57	63
placebo	30	61	91
Total	36	118	154

	cefaleia		
Grupo	sim	não	total
vicentex	14,73	48,27	63
placebo	21,27	69,73	91
Total	36	118	154

construção da matriz E (esperado)

E1,1/63 = 36/154 ⟹ E1,1 = 14,73

E1,2/63 = 118/154 ⟹ E1,2 = 48,27

E2,1/91 = 36/154 ⟹ E2,1 = 21,27

E2,2/91 = 118/154 ⟹ E2,2 = 69,73

qui-quadrado clássico

3. a estatística teste é:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

	cefaleia		
Grupo	sim	não	total
vicentex	6	57	63
placebo	30	61	91
Total	36	118	154

	cefaleia		
Grupo	sim	não	total
vicentex	14,73	48,27	63
placebo	21,27	69,73	91
Total	36	118	154

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

	cefaleia		
Grupo	sim	não	total
vicentex	6	57	63
placebo	30	61	91
Total	36	118	154

	cefaleia		
Grupo	sim	não	total
vicentex	14,73	48,27	63
placebo	21,27	69,73	91
Total	36	118	154

$$\chi^2 = \frac{(6 - 14{,}73)^2}{(14{,}73)} + \frac{(57 - 48{,}27)^2}{(48{,}27)} + \frac{(30 - 21{,}27)^2}{(21{,}27)} + \frac{(61 - 69{,}73)^2}{(14{,}73)}$$

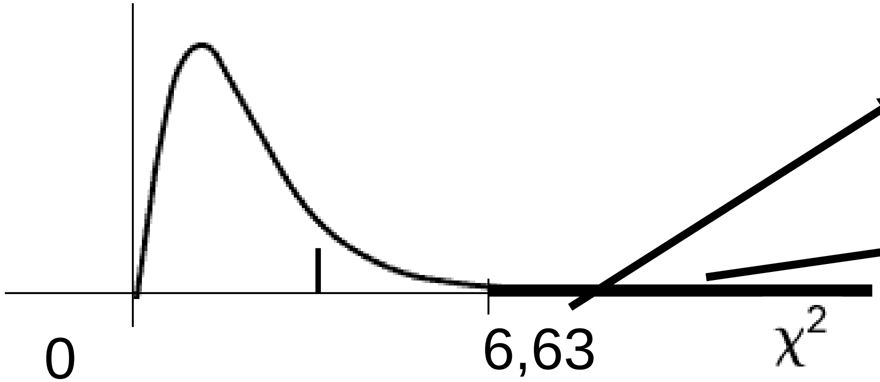
$$\chi^2 = \mathbf{11{,}4290}$$

tabela 2 x 2 possui 1 grau de liberdade

$\chi^2 =$ **11,4290**

α	qui-quadrado
10%	2,70
5%	3,84
1%	6,63

Uma distribuição qui-quadrado com 2 – 1 = 1 g.l.



Determine o valor crítico.

Determine a região de rejeição.

tabela 2 x 2 possui 1 grau de liberdade

$\chi^2 = 11{,}4290$

α	qui-quadrado
10%	2,70
5%	3,84
1%	6,63

H_0: vicentex = placebo

H_1: vicentex ≠ placebo

rejeito H0

observado

	VIVO	MORTO	TOTAL
AZT	144	1	145
PLACEBO	121	16	137
TOTAL	265	17	282

observado

	VIVO	MORTO	TOTAL
AZT	144	1	145
PLACEBO	121	16	137
TOTAL	265	17	282

$E1,1/145 = 265/282 \Rightarrow E1,1 = 136,26$

$E1,2/145 = 17/282 \Rightarrow E1,2 = 8,74$

$E2,1/137 = 265/282 \Rightarrow E2,1 = 128,74$

$E2,2/137 = 17/282 \Rightarrow E2,2 = 8,26$

esperado

	VIVO	MORTO	TOTAL
AZT	136,26	8,74	145
PLACEBO	128,74	8,26	137
TOTAL	265	17	282

CÁLCULO DO X^2

i	observado	esperado	O - E	$(O - E)^2$	$(O - E)^2: E$
1	144	136,26	7,74	59,91	0,44
2	121	128,74	-7,74	59,91	0,47
3	1	8,84	-7,74	59,91	6,85
4	16	8,26	7,74	59,91	7,25
total	282	282	0	239,64	15,01

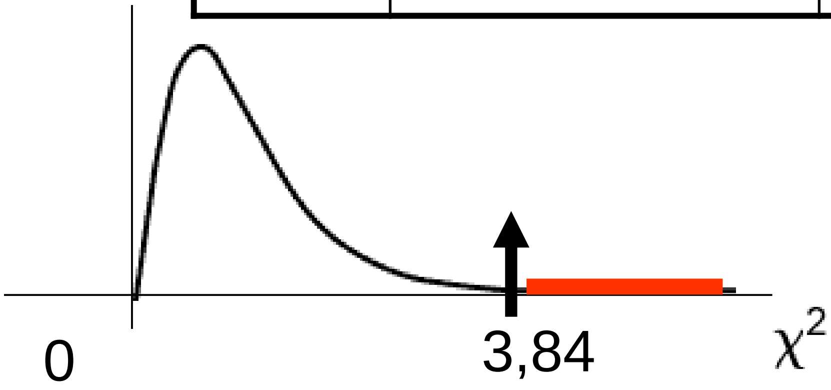
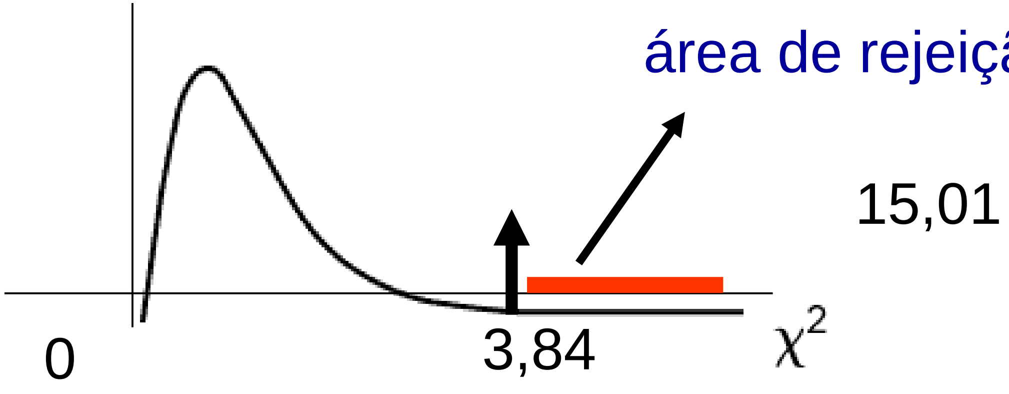


TABELA DE PROBABILIDADES

CÁLCULO DO X^2



Conclusão: o AZT é bom.

condições para uso do qui-quadrado clássico

1. **se n > 40**

2. **se n entre 20 e 40 e o valor esperado das células > 5, então utiliza-se a correção de Yates**

$$\chi^2 = \sum \frac{(|O - E| - 0,5)^2}{E}$$

n < 20 – Teste exato de Fisher

1. se n < 20

2. tabela 2 x 2

3. constrói-se duas outras tabelas denominadas matrizes extremas (X1 e X2)

4. a partir da matriz original e das matrizes extremas é calculada uma estatística para cada matriz, por meio da fórmula que será mostrada adiante

n < 20 – Teste exato de Fisher

1. se n < 20

2. tabela 2 x 2

3. constrói-se duas outras tabelas denominadas matrizes extremas (X1 e X2)

4. a partir da matriz original e das matrizes extremas é calculada uma estatística para cada matriz, por meio da fórmula que será mostrada adiante

observado (sobrevida 15 anos)

	VIVO	MORTO	TOTAL
sadios	5	2	7
diabéticos	1	8	9
TOTAL	6	10	16

matriz extrema I

	VIVO	MORTO	TOTAL
sadios	0	7	7
diabéticos	6	3	9
TOTAL	6	10	16

todos os mortos pertencem ao grupo sadio

matriz extrema II

	VIVO	MORTO	TOTAL
sadios	6	1	7
diabéticos	0	9	9
TOTAL	6	10	16

todos os vivos pertencem ao grupo sadio

teste exato

$$F = \frac{A1!A2!B1!B2!}{\left[\prod_{i=1}^{r}\prod_{j=1}^{s} . O_{i,j}\right] T!}$$

matriz original

	VIVO	MORTO	TOTAL
sadios	5	2	7
diabéticos	1	8	9
TOTAL	6	10	16

$$F = \frac{7!9!10!6!}{2!5!8!1!16!} = 0{,}02360$$

$$F = \frac{A1!A2!B1!B2!}{\left[\prod_{i=1}^{r}\prod_{j=1}^{s} . O_{i,j}\right] T!}$$

matriz extrema X1

	VIVO	MORTO	TOTAL
sadios	0	7	7
diabéticos	6	3	9
TOTAL	6	10	16

$$F = \frac{7!9!10!6!}{0!7!6!3!16!} = 0{,}0105$$

teste exato

$$F = \frac{A1!A2!B1!B2!}{\left[\prod_{i=1}^{r}\prod_{j=1}^{s} . O_{i,j}\right] T!}$$

matriz extrema X2

	VIVO	MORTO	TOTAL
sadios	6	1	7
diabéticos	0	9	9
TOTAL	6	10	16

$$F = \frac{7!9!10!6!}{6!1!0!9!16!} = 0,0009$$

resultado final

- p = Foriginal + FX1 + FX2 = 0,0236 + 0,0105 + 0,009 = 0,035

- P = 3,5%

- a afirmação de que a sobrevida de 15 anos de diabéticos é diferente dos sadios envolve uma probabilidade de erro de 3,5%, ou seja, ao nível de 3,5% de significância rejeita-se a hipótese de nulidade. Diabéticos e sadios não têm a mesma sobrevida

Testes p/ dados categorizados

- ☼ **Qui-quadrado clássico**
- ☼ **Exato de Fisher**
- ☼ **McNemar**
- ☼ **Mantel-Haenszel**

Teste de McNemar

1. O fator discriminante é categorizado segundo duas situações, mas refere-se ao mesmo grupo

2. A estatística é a seguinte:

$$Mc = \frac{4.\left(\left| O_{1,2} - \frac{O_{1,2} + O_{2,1}}{2} \right| - 0,5 \right)^2}{O_{1,2} + O_{2,1}}$$

3. O valor Mc tem distribuição qui-quadrado com 1 grau de liberdade (tabela 2 x 2)

exemplo

40 idosos fizeram um teste de montagem de um quebra-cabeça num determinado tempo. Depois tomaram um medicamento e refizeram o mesmo teste. Abaixo está o resultado obtido, antes e depois do uso do medicamento que tinha a promessa de melhorar o desempenho cognitivo dos indivíduos.

	montagem de um quebra-cabeça		
Grupo	**Correto**	**Incorreto**	**total**
antes	**9**	**31**	**40**
depois	**22**	**18**	**40**

A tabela acima não indica os resultados obtidos por cada indivíduo. antes e depois do medicamento. Alguns indivíduos podem ter piorado o desempenho e estejam sendo compensados por indivíduos que melhoraram o desempenho. A tabela deve ser arranjada, tal qual fazemos num teste pareado, para representar o desempenho de cada par.

exemplo

Tabela original

	montagem de um quebra-cabeça		
Grupo	Correto	Incorreto	total
antes	9	31	40
depois	22	18	40

Tabela arranjada

		depois		
antes	Grupo	Correto	Incorreto	total
	correto	8	1	9
	incorreto	14	17	31
	Total	22	18	40

Teste de McNemar

$$Mc = \frac{4.\left(\left| O_{1,2} - \frac{O_{1,2} + O_{2,1}}{2} \right| - 0,5\right)^2}{O_{1,2} + O_{2,1}}$$

$$Mc = \frac{4.\left(\left| 1 - \frac{1 + 14}{2} \right| - 0,5\right)^2}{1 + 14}$$

Mc = 9,6

qui-quadrado crítico = 6,64

rejeito H0

Testes p/ dados categorizados

- ☼ Qui-quadrado clássico
- ☼ Exato de Fisher
- ☼ McNemar
- ☼ Mantel-Haenszel

Teste de Mantel-Haenszel

1. igual ao qui-quadrado
2. pode existir algum fator associado ao fator discriminante que esteja também exercendo influência sobre o fator determinado
3. Poderia ser interessante entender o efeito isolado do fator discriminante sobre o fator discriminado

$$\chi^2 = \frac{(|SO - SE| - 0{,}5)^2}{SV_{\text{(soma das variâncias)}}}$$

so = soma dos valores observados que relacionam positivamente o fator discriminante e o fator discriminado

se = soma dos valores esperados que relacionam positivamente o fator discriminante e o fator discriminado

exemplo

efeito do fumo sobre a irritação da mucosa estomacal. fumantes x não fumantes

Grupos da variável interferência: beber café. grupo que ingere menos de uma xícara por dia e grupo que ingere um pouco mais de café, mas nada exagerado.

CAFÉ

	irritação estomacal		
Grupo	**sim**	**não**	**total**
fumantes	**12**	**45**	**57**
não fumantes	**6**	**92**	**98**
total	**18**	**137**	**155**

não CAFÉ

	irritação estomacal		
Grupo	**sim**	**não**	**total**
fumantes	**16**	**89**	**105**
não fumantes	**10**	**115**	**125**
total	**26**	**204**	**230**

exemplo

efeito do fumo sobre a irritação da mucosa estomacal. fumantes x não fumantes

Grupos da variável interferência: beber café. grupo que ingere menos de uma xícara por dia e grupo que ingere um pouco mais de café, mas nada exagerado.

CAFÉ	irritação estomacal		
Grupo	sim	não	total
fumantes	$\boxed{12}$	45	57
não fumantes	6	92	98
total	18	137	155

não CAFÉ	irritação estomacal		
Grupo	sim	não	total
fumantes	$\boxed{16}$	89	105
não fumantes	10	115	125
total	26	204	230

valores esperados para situação fumante e irritação estomacal

E(I)1,1/18 = 57/155 → **E1,1 = 6,62 (tabela I)**

E(II)1,2/26 = 105/230 → **E1,2 = 11,87 (tabela II)**

SE = 6,62 + 11,87 = 18,49

exemplo

efeito do fumo sobre a irritação da mucosa estomacal. fumantes x não fumantes

Grupos da variável interferência: beber café. grupo que ingere menos de uma xícara por dia e grupo que ingere um pouco mais de café, mas nada exagerado.

CAFÉ	irritação estomacal		
Grupo	sim	não	total
fumantes	$\boxed{12}$	45	57
não fumantes	6	92	98
total	18	137	155

não CAFÉ	irritação estomacal		
Grupo	sim	não	total
fumantes	$\boxed{16}$	89	105
não fumantes	10	115	125
total	26	204	230

valores esperados para situação fumante e irritação estomacal

SO = 12 + 16 = 28

exemplo

efeito do fumo sobre a irritação da mucosa estomacal. fumantes x não fumantes

Grupos da variável interferência: beber café. grupo que ingere menos de uma xícara por dia e grupo que ingere um pouco mais de café, mas nada exagerado.

CAFÉ	irritação estomacal		
Grupo	sim	não	total
fumantes	$\boxed{12}$	45	57
não fumantes	6	92	98
total	18	137	155

não CAFÉ	irritação estomacal		
Grupo	sim	não	total
fumantes	$\boxed{16}$	89	105
não fumantes	10	115	125
total	26	204	230

$$VI = \frac{57.98.18.137}{155^2 \,.\, (155 - 1)}$$

$$VII = \frac{105.125.26.204}{230^2 \,.\, (230 - 1)}$$

$$V = \frac{A1.A2.B1.B2}{T^2 \,.\, (T - 1)}$$

VI + VII = 3,72 + 5,74 = 9,47

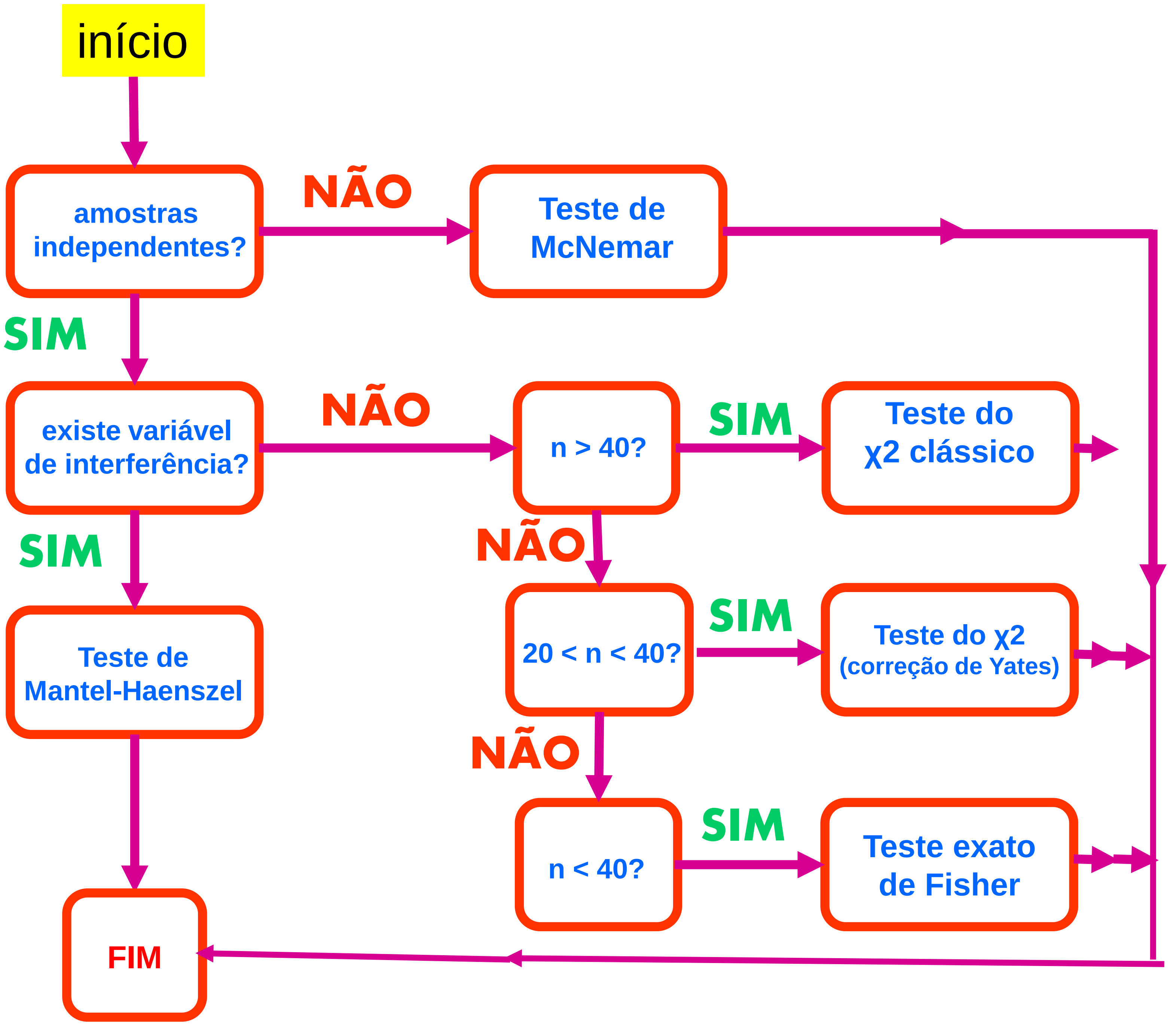
Teste de Mantel-Haenszel

$$\chi^2 = \frac{(\left| SO - SE \right| - 0{,}5)^2}{SV \text{ (soma das variâncias)}}$$

$$\chi^2 = \frac{(\left| 28 - 18{,}49 \right| - 0{,}5)^2}{9{,}47}$$

$$\chi^2 = 8{,}57$$

rejeito H0



início
amostras independentes?
NÃO
Teste de McNemar
SIM
existe variável de interferência?
NÃO
n > 40?
SIM
Teste do χ2 clássico
SIM
NÃO
Teste de Mantel-Haenszel
20 < n < 40?
SIM
Teste do χ2 (correção de Yates)
NÃO
n < 40?
SIM
Teste exato de Fisher
FIM