Testes para dados categóricos

Teste de homogeneidade

Objetivo: testar se existe diferença entre frequências observadas (O_{ij}) e frequências esperadas (E_{ij}).

Dados amostrais: amostras aleatórias independentes de r populações classificadas em c categorias

Teste X^2-quadrado para r amostras e c categorias

populações	categoria 1	categoria 2	categoria 3	categoria c	combinada
1	O_{11}	O_{12}	O_{13}	O_{1c}	R_1
2	O_{21}	O_{22}	O_{23}	O_{2c}	R_2
3	O_{31}	O_{32}	O_{33}	O_{3c}	R_3
r	O_{r1}	O_{r2}	O_{r3}	O_{rc}	R_r
total	C_1	C_2	C_3	C_{kc}	N

$$C_c = O_{1c} + O_{2c} + O_{3c} + \ldots + O_{rc}$$

$$R_r = O_{r1} + O_{r2} + O_{r3} + \ldots + O_{rc}$$

Objetivo: testar se existe diferença entre frequências observadas (O_{ij}) e frequências esperadas (E_{ij}).

$$E_{ij} = \frac{R_i \, C_j}{N} \quad \text{(frequência esperada na linha i e coluna j)}$$

$$C_i = \text{total na categoria i}$$

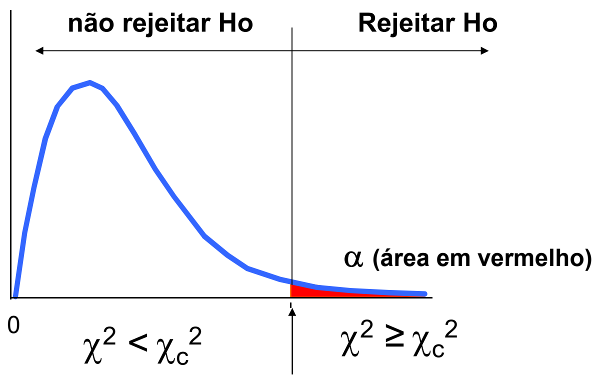
$$R_i = \text{total na população i}$$

Estatística de teste:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \sum_{\substack{\text{todas as} \\ \text{células}}} \frac{(O-E)^2}{E}$$

Valores críticos na tabela de χ^2 com **gl = (r-1)(c-1)** graus de liberdade.

- Fixar nível significância α.
- Encontrar na tabela da distribuição qui-quadrado com gl= (r-1)(c-1) o valor crítico χ_c^2 que determina a região de rejeição



Resumo do livro do Johnson

The χ^2 Test of Homogeneity in a Contingency Table

Null hypothesis

In each response category, the probabilities are equal for all the populations.

Test statistic

$$\chi^2 = \sum_{\text{cells}} \frac{(O - E)^2}{E} \qquad \begin{cases} O = \text{Observed cell frequency} \\ E = \dfrac{\text{Row total} \times \text{Column total}}{\text{Grand total}} \end{cases}$$

$$\text{d.f.} = (\text{No. of rows} - 1) \times (\text{No. of columns} - 1)$$

Rejection region

$$\chi^2 \geq \chi^2_\alpha$$

Exemplo (Johnson): incidência de alcoolismo em diferentes grupos de profissionais é a mesma.

Amostra:

variável	alcoólico	Não alcoólico
1. clero	32	268
2. educadores	51	199
3. executivos	67	233
4. comerciantes	83	267

Ho: p1=p2=p3=p4

H1: as proporções não são iguais

variável	alcoólico	Não alcoólico	No de elementos
clero	$O_{11} = 32$ $E_{11} = 58,25$	$O_{12} = 268$ $E_{12} = 241,75$	$R_1 = 32+268 = 300$
educadores	$O_{21} = 51$ $E_{21} = 48,54$	$O_{22} = 199$ $E_{22} = 201,46$	$R_2 = 51+199 = 250$
executivos	$O_{31} = 67$ $E_{31} = 58,25$	$O_{32} = 233$ $E_{32} = 241,75$	$R_3 = 67+233 = 300$
comerciantes	$O_{41} = 83$ $E_{41} = 67,96$	$O_{42} = 267$ $E_{42} = 282,04$	$R_4 = 83+267 = 350$
total	C1=233	C2= 967	N=1200

E_{11}= (300.233)/1200=58,25 E_{12}= (300.967)/1200=241,75

E_{21}= (250.233)/1200=48,54 E_{22}= (250.967)/1200=201,46

E_{31}= (300.233)/1200=58,25 E_{32}= (300.967)/1200=241,75

E_{31}= (350.233)/1200=67,96 E_{32}= (350.967)/1200 =282,04

Estatística teste: $\chi^2 = \sum_{\substack{\text{todas as} \\ \text{células}}} \frac{(O-E)^2}{E} = 20{,}59$

Ao nível de significância α=0.05 o valor crítico será para gl=(r-1)(c-1)=(4-1)(2-1)=3:

χ_c^2 = **7,81**

E temos χ^2 = **20,59** **>** χ_c^2 = **7,81**

Rejeitamos Ho

Há evidência amostral de que as proporções de alcoolismo não é a mesma nos diferentes grupos.

Teste de independência

Tabela de contingência com r linhas e c colunas

Dados amostrais: amostra aleatória de n elementos classificados em duas características classificadas em r (linhas) e c (colunas) categorias, respectivamente.

Lembrando:

Variáveis aleatórias X e Y independentes:

$P(X \cap Y) = P(X,Y) = P(X)P(Y)$

	x1	x2	x3	xc	Prob. marginal
y1	$P(x_1, y_1)$	$P(x_2, y_1)$	$P(x_3, y_1)$	$P(x_c, y_1)$	$P(y_1)$= soma sobre x
y2	$P(x_1, y_2)$	$P(x_2, y_2)$	$P(x_3, y_2)$	$P(x_c, y_2)$	$P(y_2)$= soma sobre x
y3	$P(x_1, y_3)$	$P(x_2, y_3)$	$P(x_3, y_3)$	$P(x_c, y_3)$	$P(y_3)$= soma sobre x
yr	$P(x_1, y_r)$	$P(x_2, y_r)$	$P(x_3, y_r)$	$P(x_c, y_r)$	$P(y_r)$= soma sobre x
Prob. marginal	$P(x_1)$= soma sobre y	$P(x_2)$= soma sobre y	$P(x_3)$= soma sobre y	$P(x_c)$= soma sobre y	N

Ho: As categorias são independentes: $P(x_1, y_1)=P(x_1,)P(y_1)$, $P(x_2, y_2)=P(x_2,)\ P(y_1)$, ..., $P(x_c, y_r)=P(x_c)P(y_r)$

H1: As categorias não são independentes

Em geral temos uma tabela com frequências observadas:

populações	x1	x2	x3	xc	total
y1	O_{11}	O_{12}	O_{13}	O_{1c}	R_1
y2	O_{21}	O_{22}	O_{23}	O_{2c}	R_2
y3	O_{31}	O_{32}	O_{33}	O_{3c}	R_3
r	O_{r1}	O_{r2}	O_{r3}	O_{rc}	R_r
total	C_1	C_2	C_3	C_{kc}	N

$$C_c = O_{1c}+ O_{2c} + O_{3c} + \dots + O_{rc}$$

$$R_r = O_{r1} + O_{r2} + O_{r3} + \dots + O_{rc}$$

As proporções observadas (PO) são dadas por:

$$PO_{ij} = \frac{O_{ij}}{N}$$

As proporções esperadas (PE) sob a hipótese nula (de independência) são dadas por:

$$PE_{ij} = P(x_i)\,P(y_j) = \left(\frac{C_i}{N}\right)\left(\frac{R_j}{N}\right)$$

e a frequencia esperada será :

$$E_{ij} = N.PE_{ij} = \frac{C_i R_j}{N}$$

- Estatística de teste:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}} = \sum_{\substack{\text{todas as}\\\text{células}}}\frac{(O\text{-}E)^2}{E}$$

- Valores críticos (χ_c^2): na tabela de χ^2 com **gl = (r-1)(c-1)** graus de liberdade.

- Rejeite Ho se $\chi^2 \geq \chi_c^2$

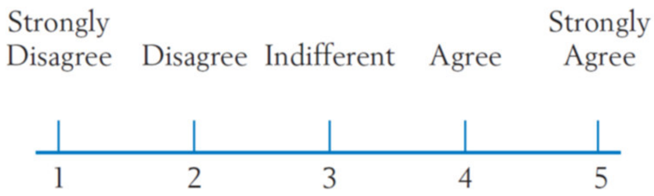
Os testes χ^2 são apropriados se as frequências esperadas nas células sejam grandes (≥ 5)

Alguns testes não paramétricos

Estatística não paramétrica: não requer população com distribuição normal ou outra distribuição específica.

Interessantes em situações com dados ordinais.

Escalas do tipo:



TESTE DE WILCOXON-MANN-WHITNEY
(Amostras independentes)

Usado para testar se dois conjuntos de dados foram obtidos de populações idênticas, ou seja, da mesma distribuição.

Exigência do Teste

- Nível de Mensuração em escala ordinal (pelo menos);
- **Amostras independentes**

Alternativa ao teste t que exige nível intervalar e condições sobre variâncias ou distribuição normal.

É um teste baseado em postos.

Se as duas amostras são retiradas de populações idênticas e aos valores são atribuídos postos considerando uma única coleção, então postos altos e baixos deveriam se distribuir equilibradamente entre as duas amostras.

Procedimento:

1 - ordene as amostras em ordem crescente.

2 - atribua postos aos dados considerando que vieram de uma única amostra combinada.

3 – Encontre a soma dos postos para uma das amostras

Postos empatados

5	**10**	**11**	**3**	**5**	**5**	**dados originais**
3	**5**	**5**	**5**	**10**	**11**	**dados ordenados**
1	**3**	**3**	**3**	**5**	**6**	**postos**

2 3 4 : posto médio = (2+3+4)/3=3

Encontre a média dos postos e atribua o posto médio para cada ítem empatado.

Notação (qualquer amostra pode ser usada como amostra 1)

n1 = número de casos da amostra 1
n2 = número de casos da amostra 2
N= n1+ n2
R1= soma dos postos da amostra 1
R2= soma dos postos da amostra 2
R = mesmo que R1

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}$$

média esperada da amostra R, esperada quando os valores das populações são idênticas

$$\sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

Desvio padrão dos valores da amostra R, esperado quando os valores das populações são idênticas

Para amostras grandes: **n_1>10 e n_2>10**

R tem distribuição aproximadamente normal com média μ_R e desvio σ_R

Estatística de teste: $$z = \frac{R - \mu_R}{\sigma_R}$$

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

n_1 = tamanho da amostra para a qual é encontrada a soma dos postos R

n_2 = tamanho da outra amostra

R = soma dos postos da amostra de tamanho n1

Amostras pequenas: $\mathbf{n_1 \leq 10}$ **e** $\mathbf{n_2 \leq 10}$

m = número de elementos da amostra de menor tamanho
n = número de elementos da amostra maior tamanho.

Estatística de teste:

$\mathbf{R_m}$= soma dos postos da amostra de menor tamanho.

Valores críticos: apêndice B-tabela 7

Ou use um software com o teste já programado.

Exemplo: Temos abaixo a média de provas de dois grupos de alunos. Determinar se os grupos tem mesma distribuição de notas. Considere α=0,05.

A	5,2	5,9	6,3	6,8	7,0	8,1	8,2	8,9	9,5	10,0	6,0

B	4,5	5,0	5,1	5,6	5,8	6,3	6,8	7,2	7,8	8,1	7,4	9,1

Amostras ordenadas e com postos considerando as duas juntas:

A	5,2	5,9	6,0	6,3	6,8	7,0	8,1	8,2	8,9	9,5	10,0
Posto	4	7	8	9,5	11,5	13	17,5	19	20	22	23

B	4,5	5,0	5,1	5,6	5,8	6,3	6,8	7,2	7,4	7,8	8,1	9,1
Posto	1	2	3	5	6	9,5	11,5	14	15	16	17,5	21

Ho : as populações são iguais

H1 : as populações não são iguais

Podemos aplicar **WILCOXON-MANN-WHITNEY**

Amostra A (1): n1=11>10 $A: \mu_R = n_1(n_1+n_2+1)/2 = 11(11+12+1)/2 = 132$

Amostra B (2): n2=12>10 $B: \mu_R = n_2(n_1+n_2+1)/2 = 12(11+12+1)/2 = 144$

podemos usar a estatítica de teste: $\mathbf{z} = \dfrac{\mathbf{R - \mu_R}}{\mathbf{\sigma_R}}$

Soma dos postos de 1: R= 4+7+8+9,5+11,5+13+17,5+19+20+22+23= 154,5

$$\mu_R = n_1(n_1 + n_2 + 1)/2 = 11(11 + 12 + 1)/2 = 132$$

$$\sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = \sqrt{\frac{11.12.24}{12}} = 16,2481$$

Temos:

$$\mathbf{z} = \frac{154,5 - 132}{16,2481} = 1,385$$

Para α=0,05 bilateral temos os valores críticos: $\mathbf{Z_c = \pm 1,96}$

Como Z_c=-1,96 < Z=1,385 < Z_c=1,96 não rejeitamos Ho.

Portanto, o resultado do teste indica que as populações tem mesma distribuição de notas.

Testes para amostras relacionadas
(amostras dependentes)

- O caso de duas amostras relacionadas pode ser entendido também como o caso de uma amostra com duas medidas ou replicações emparelhadas.
- A finalidade dos testes envolvendo amostras relacionadas é avaliar o efeito de algum "tratamento" numa variável de interesse.
- O "tratamento" pode consistir de uma diversidade de situações: treinamento, propaganda, modificação nas condições do ambiente de trabalho, introdução de um novo elemento na economia etc.

Teste de sinais

Nível de mensuração: ordinal

O Teste dos sinais é útil para uma pesquisa na qual a mensuração quantitativa não é possível, mas na qual é possível determinar, para cada par de observações, qual é o "maior“.

Exemplo:

Xi	10	2	6	8	9	10
Yi	9	5	8	8	7	4
Sinal da diferença (xi-yi)	**+**	**-**	**-**	**0**	**+**	**+**

Empates: *Xi* = *Yi* , ou seja , não há diferença entre os resultados do par i. **Os empates são excluídos da análise.**

Conceito básico do teste de sinais: analisar as frequências de sinais + e sinais – para determinar se existe diferença significante.

Considerações:
Amostra selecionada aleatoriamente.

Não há necessiadade que a amostra seja de uma população com uma distribuição particular, tal como distribuição normal.

Seja :
x = número de ocorrências do sinal menos frequente
n = $n_+ + n_-$ = número total combinado de sinais positivos e negativos

Ho: $P_+ = P_-$

Estatística de teste

n ≤ 25: distribuição binomial com p = q = ½

Para n > 25 e : $$z = \frac{x - \frac{n}{2}}{\sqrt{\frac{n}{4}}}$$

Exemplo

O gerente de um supermercado afirma que não há preferência entre duas marcas de chocolate. Foi realizada uma pesquisa com 12 pessoas para saber a preferência entre as duas marcas. A pesquisa consiste em atribuir nota 10 ao que achou melhor e nota 0 ao que achou pior. Temos o seguinte resultado.

X	10	10	0	10	0	0	0	10	10	10	0	10
Y	0	10	10	0	10	0	10	0	0	0	0	0
Sinal da diferença X-Y	+	0	-	+	-	0	-	+	+	+	0	+

Faça o teste com 5% de significância.

Será que preferência por x = preferência por y ?

Ho: Proporção de (+) = Proporção de (–)
H1: Proporção de sinais (+) ≠ Proporção de sinais (–)

n = $n_+ + n_-$ **=9**

x = 3 (número de sinais menos frequente)

nível de significância: α=0,05

estatística de teste: probabilidade dada pela distribuição binomial com n=9 e p=q=1/2

$p(y \leq 3)=p(0)+p(1)+p(2)+p(3) = 0,2539$

Teste bilateral: $2p=0,5078 > \alpha=0,05$ então não rejeitamos Ho

Portanto não rejeitamos a afirmação de que "não existe preferência por marca". Os resultados indicam que não há preferência por marca.

Teste de postos com sinal de wilcoxon

É um teste não-paramétrico que analisa as diferenças di=xi-yi entre as observações emparelhadas, levando em conta a magnitude das diferenças.

Utiliza os postos (rank) das diferenças de dados amostrais emparelhados.

Posto: é a posição ocupada (1º, 2º, 3º,...) quando os dados estão ordenados (menor para o maior, do pior para o melhor,..)

Exemplo:

5	**3**	**40**	**10**	**12**	**dados originais**
3	**5**	**10**	**12**	**40**	**dados ordenados**
↑	↑	↑	↑	↑	
1	**2**	**3**	**4**	**5**	**posto**

Postos empatados

5	10	11	3	5	5	**dados originais**
3	5	5	5	10	11	**dados ordenados**
↑	↑	↑	↑	↑	↑	
1	3	3	3	5	6	**postos**

2 3 4 : posto médio = (2+3+4)/3=3

Encontre a média dos postos e atribua o posto médio para cada ítem empatado.

Procedimento do teste de wilcoxon

PASSO 1: Para cada par de dados, determine a diferença $di=xi-yi$ subtraindo o segundo escore do primeiro. Mantenha os sinais, mas descarte os pares para os quais d=0.

PASSO 2: Ignore os sinais das diferenças, e ordene-as da menor para maior.Quando aparecerem diferenças com o mesmo valor numérico, atribua-lhes a média dos postos envolvidos no empate.

PASSO 3: Atribua a cada posto o sinal da diferença de onde proveio, isto é; insira os sinais que foram ignorados no PASSO 2.

PASSO 4: Ache a soma dos valores absolutos dos postos negativos, e a soma dos postos positivos.

PASSO 5: Seja T a menor das duas somas obtidas no PASSO 4. Poderíamos utilizar qualquer soma, mas para simplificar o processo, escolhemos a menor das duas.

PASSO 6: Seja *n* o número de pares de dados para os quais a diferença *d* não é 0.

•**PASSO 7:** Determine a estatística de teste e os valores críticos baseados no tamanho da amostra, conforme mostrado no quadro da Estatística de Teste.

Seja T a menor das duas somas seguintes:

1. A soma dos valores absolutos dos postos negativos

2. A soma dos valores dos postos positivos

Para $n \leq 30$: T

O valor crítico é obtido da tabela 8 do apendice B do Johnson ou da tabela de valores de T da pagina seguinte.

Para $n > 30$: $$z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

obtêm-se os valores críticos Zc na tabela da distribuição normal padronizada.

Ou use um software com o teste já programado.

Tabela: Valores Críticos de T para o Teste de Wilcoxon

n	Nível de Significância			
	0,005 Unilateral 0,01 bilateral	0,01 unilateral 0,02 bilateral	0,025 unilateral 0,05 bilateral	0,05 unilateral 0,10 bilateral
5	-	-	-	1
6	-	-	1	2
7	-	0	2	4
8	0	2	4	6
9	2	3	6	8
10	3	5	8	11
11	5	7	11	14
12	7	10	14	17
13	10	13	17	21
14	13	16	21	26
15	16	20	25	30
16	19	24	30	36
17	23	28	35	41
18	28	33	40	47
19	32	38	46	54
20	37	43	52	60
21	43	49	59	68
22	49	56	66	75
23	55	62	73	83
24	61	69	81	92
25	68	77	90	101
26	76	85	98	110
27	84	93	107	120
28	92	102	117	130
29	100	111	127	141
30	109	120	137	152

Tabela extraída e adaptada de TRIOLA (1999)

Notas:

1. - indica que não é possível obter um valor na região crítica.
2. Rejeitar a hipótese nula se a estatística de teste *T* é menor ou igual ao valor crítico da tabela. Não rejeitar a hipótese nula em caso contrário.

Note que se usar Tc dessa tabela:

Rejeite Ho se T ≤ Tc

Exemplo

Realizou-se um estudo para pesquisar a eficiência do hipnotismo na redução da dor. Temos o seguinte resultado, onde as medidas são em centímetros para representar uma escala de dor. Teste, com significância de 0,01 a afirmação de que o hipnotismo não tem efeito.

Antes da hipnose (X)	6,6	6,5	9,0	10,3	11,3	8,1	6,3	11,6
Depois da hipnose(Y)	6,8	2,4	7,4	8,5	8,1	6,1	3,4	2,0
Sinal da diferença di = Xi-Yi	-0,2	+4,1	+1,6	+1,8	+3,2	+2,0	+2,9	+9,6

H_o: dor antes = dor depois
H_1: dor antes ≠ dor depois

Ordenamento do |xi-yi|: 0,2; 1,6; 1,8; 2,0; 2,9; 3,2; 4,1; 9,6
posto: 1, 2, 3, 4, 5, 6, 7, 8

Antes X	6,6	6,5	9,0	10,3	11,3	8,1	6,3	11,6
Depois Y	6,8	2,4	7,4	8,5	8,1	6,1	3,4	2,0
Diferença di = Xi-Yi com sinais	-0,2	+4,1	+1,6	+1,8	+3,2	+2,0	+2,9	+9,6
Postos sem sinais	1	7	2	3	6	4	5	8
Posto com sinais originais	**-1**	**+7**	**+2**	**+3**	**+6**	**+4**	**+5**	**+8**

n = 8 (número de pares com diferença não nula)

1. A soma dos valores absolutos dos postos negativos: 1
2. A soma dos valores dos postos positivos: 2+3+4+5+6+7+8 = 35

Valor de T: o menor dentre 1 e 2

Portanto: T = 1

Estatística de teste:
como n = 8 < 25 a estatística de teste é T=1.
Valor crítico para $\alpha = 0{,}01$, bilateral e n=8 está na tabela: $T_{crítico} = 0$

Como $T = 1 > T_{crítico} = 0$ não rejeitamos Ho.

Não rejeitamos Ho. Ou seja, o resultado indica que a hipnose parece não ser eficaz.

Constatação banal, experimentada por todo pedagogo: *A primeira lição de leitura pega a criança ignorante e deve tratá-la como ignorante. Mas deve supor, ao mesmo tempo, não apenas um professor, que já saiba ler, mas também uma criança que já esteja em condições de tornar-se professor.*

Hegel, por François Châtelet, Jorge Zahar editor, 1995.

Fim