

TESTE DO “BOM AJUSTE” DO QUI QUADRADO



TABELA DE UMA ENTRADA

1	2	3	...	k
f_1	f_2	f_3	...	f_k

EXEMPLO



Em 2003, o número de AVCs masculinos no concelho de Braga foram os reportados na tabela, de acordo com a estação do ano.

Primavera	64
Verão	81
Outono	39
Inverno	28

Profª Ana Cristina Braga, DPS

3

TESTE DE HIPÓTESES



- Hipóteses $H_0 : p_1 = p_{1,0}; p_2 = p_{2,0}; \dots; p_k = p_{k,0}$

$$H_1 : p_i \neq p_{i,0}$$

- Estatística

$$Q = \sum_{i=1}^k \frac{(f_i - e_i)^2}{e_i}$$

- Região de Rejeição

$$Q > \chi_\alpha^2$$

Profª Ana Cristina Braga, DPS

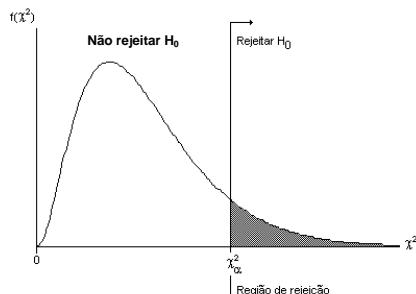
4



SOLUÇÃO

■ Região crítica

$$\chi^2 \geq \chi^2_{0,05;3} = 7,81$$



■ Estatística

$$Q = \sum_{i=1}^k \frac{(f_i - e_i)^2}{e_i} = \\ = \frac{(64-53)^2}{53} + \frac{(81-53)^2}{53} + \frac{(39-53)^2}{53} + \frac{(28-53)^2}{53} = 32,57$$

Profª Ana Cristina Braga, DPS

5



avcs_m.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

1 : fi 64

	estação	fi	var						
1	Primavera	64							
2	Verão	81							
3	Outono	39							
4	Inverno	28							
5									
6									
7									
8									
9									
10									
11									
12									
13									

Weight Cases dialog box:

- Do not weight cases
- Weight cases by
- Frequency Variable: Frequências obsevadas...

Current Status: Weight cases by fi

OK Paste Reset Cancel Help

Profª Ana Cristina Braga, DPS

6



avcs_m.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

Data View Variable View

IBM SPSS Statistics Processor is ready | Unicode ON | Weight On

Value Labels

Value Labels

Value: Spelling...
Label:
Add Change Remove
OK Cancel Help

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
estação	Numeric	8	0		(1, Primavera)	None	8	Right	Nominal	Input
fi	Numeric	8	0	Frequências o...	None	None	8	Right	Scale	Input

Profª Ana Cristina Braga, DPS

7



avcs_m.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Extensions Window Help

Data View Variable View

Visible: 2 of 2 Variables

Analyze menu expanded:

- Reports
- Descriptive Statistics
- Tables
- Compare Means
- General Linear Model
- Generalized Linear Models
- Mixed Models
- Correlate
- Regression
- Loglinear
- Neural Networks
- Classify
- Dimension Reduction
- Scale
- Nonparametric Tests
- Forecasting
- Survival
- Multiple Response
- Missing Value Analysis...
- Multiple Imputation
- Complex Samples
- Simulation
- Quality Control
- ROC Curve...
- Spatial and Temporal Modeling...

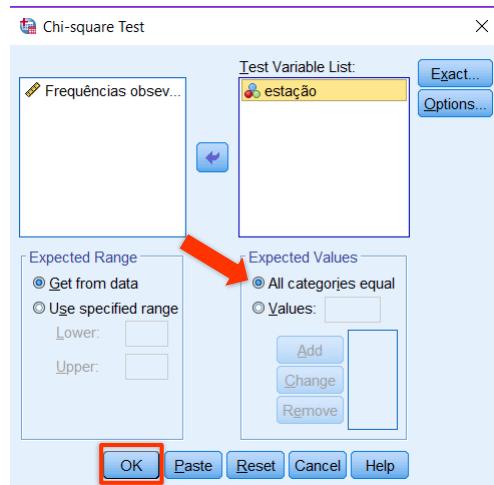
Legacy Dialogs menu expanded:

- One Sample...
- Independent Samples...
- Related Samples...
- Chi-square...
- Binomial...
- Runs...
- 1-Sample K-S...
- 2 Independent Samples...
- K Independent Samples...
- 2 Related Samples...
- K Related Samples...

IBM SPSS Statistics Processor is ready | Unicode ON | Weight On

Profª Ana Cristina Braga, DPS

8



Profª Ana Cristina Braga, DPS

9



NPar Tests

[DataSet1] C:\Users\ACB\OneDrive\Aulas2016_17\EE&IO_MEI\Avaliação\avcs_m.sav

Chi-Square Test

Frequencies

estação			
	Observed N	Expected N	Residual
Primavera	64	53,0	11,0
Verão	81	53,0	28,0
Outono	39	53,0	-14,0
Inverno	28	53,0	-25,0
Total	212		

Test Statistics

estação	
Chi-Square	32,566 ^a
df	3
Asymp. Sig.	,000

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell

Profª Ana Cristina Braga, DPS

10



Exemplo (fichas TP exercício 8)

Em cada dia, de segunda a sexta, um padeiro produz três grandes bolos de chocolate, e os que não são vendidos no mesmo dia são dados a um banco alimentar. Use os dados apresentados no quadro seguinte para testar, ao nível de significância de 0.05, se podem ser considerados como valores duma variável aleatória binomial.

Nº de bolos	0	1	2	3
Nº de dias	1	16	55	228

Profª Ana Cristina Braga, DPS

11

Formulação de hipóteses

X- v.a. que designa o nº de bolos vendidos

$$H_0: X \sim Bin(3, \pi)$$

probabilidades não estão completamente especificadas na hipótese nula, não conhecemos π , pelo que tem de ser estimado a partir dos dados da amostra através do seu estimador \bar{x} , pois sabemos que $\mu = n\pi \Leftrightarrow \pi = \frac{\mu}{n}$.

$$H_1: \bar{H}_0$$

Identificação da estatística de teste (ET): $Q \sim \chi^2$ e

$$Q = \sum_{i=1}^k \frac{(f_i - e_i)^2}{e_i} \text{ com } e_i = n \cdot p_i$$

Definição da regra de decisão

$$R.R: Q > c \quad \text{com } c = \chi^2_{gl, \alpha}$$

Profª Ana Cristina Braga, DPS

12



X_i - v.a representa o nº de bolos vendidos por dia

x_i	f_i	p_i	média = 2,7		$p = \text{média}/n = 0,9$
			e_i	e_i	
0	1	0,001	0,3		
1	16	0,027	8,1	8,4	8,8047619
2	55	0,243	72,9	72,9	4,3951989
3	228	0,729	218,7	218,7	0,3954733
somas	300	1			13,595434

Cálculos auxiliares:

$$\hat{\mu} = \bar{x} = \frac{\sum_{i=1}^k x_i f_i}{\sum_{i=1}^k f_i} = \frac{0 * 1 + 1 * 16 + 2 * 55 + 3 * 228}{300} = 2,7 \text{ e}$$

$$\hat{\pi} = p = \frac{\bar{x}}{n} = \frac{2,7}{3} = 0,9$$

Profª Ana Cristina Braga, DPS

13



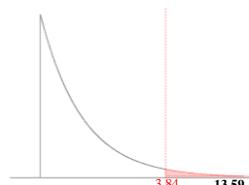
Nota: Quando $e_i < 5$ tem de se agrupar as classes até $e_{iag} > 5$ (condição para poder calcular Q)

$$Q = 13,595$$

$$g.l = \text{nº celas} - 1 - \text{nº parâmetros estimados} = 3 - 1 - 1 = 1$$

$$\alpha = 0,05$$

$$\text{tabela 7: } c = \chi^2_{1,0,05} = 3,84146$$



Decisão: Como $Q > c$, rejeitamos a hipótese nula para um nível de significância de 5%. Pelo que se conclui que a distribuição do número de bolos vendidos não segue uma distribuição binomial.

Profª Ana Cristina Braga, DPS

14