

Introdução à Probabilidade e Estatística

Ficha Nº6: Testes de Hipóteses Não Paramétricos

Para as licenciaturas em: Eng. Civil, Eng. das Energias Renováveis, Eng. Geológica, Eng. Informática e Eng. Mecatrónica

2º Semestre 2014/15 - 2h Teóricas + 2h Práticas

Docentes: Patrícia Filipe e Ana Isabel Santos

1. Considere a seguinte amostra:

1.26	0.34	0.7	1.75	5.57	1.55	0.8	0.42	0.51	3.2
0.15	0.49	0.95	0.24	1.37	0.17	6.98	0.13	0.94	0.38

Com base no seguinte *output*, teste se é de admitir que a amostra seja proveniente de uma população com distribuição normal.

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Ex1	,416	20	,000	,324	20	,000

a. Lilliefors Significance Correction

2. Considere duas amostras, ambas de dimensão 20, correspondentes a medições realizadas em dois grupos de pacientes. A um dos grupos foi administrado um placebo e ao outro administrado um tratamento específico. Os resultados registados foram os seguintes:

Dados de Placebo									
1.26	0.34	0.7	1.75	50.57	1.55	0.8	0.42	0.51	3.2
0.15	0.49	0.95	0.24	1.37	0.17	6.98	0.13	0.94	0.38
Dados do Tratamento									
3,59	6,72	37,2	0,69	6,94	14,81	9,39	-10,95	13,28	12,84
4,65	18,59	30,54	2,28	1,54	1,41	-2,64	11,99	22,95	-16,03

Com base no *output* abaixo, teste se podemos admitir que ambas as amostras sejam provenientes de populações com distribuição normal.

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Placebo	0,416	20	0,000	0,324	20	0,000
Tratamento	0,119	20	0,200*	0,971	20	0,773

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

3. Um novo jogo de casino consiste em lançar 3 vezes um dado. O vencedor é aquele que obtiver o maior número de pintas 6. Um jogador, selecionado ao acaso, jogou 100 vezes e obteve os seguintes resultados:

Nº pintas 6	ni
0	47
1	35
2	15
3	3

O jogador não aceita estes resultados e suspeita da honestidade do casino. Para tal, vai realizar um teste de hipóteses para avaliar a falsidade, ou não, do dado. O que conclui?

4. Foi realizado um estudo para determinar se a opinião pública era favorável à construção de uma barragem hidroeléctrica. Os resultados foram: 40% a favor da construção, 30% são indiferentes, 20% opõem-se e os restantes disseram não terem pensado no assunto. Uma amostra aleatória de 150 indivíduos da região afectada revelou que 42 eram a favor, 61 eram indiferentes e 33 eram contrários à construção.

- Com base nos *outputs* seguintes, estarão estes dados de acordo com os resultados obtidos no referido estudo?
- Determine o valor do *p-value* associado a este teste de hipóteses.

Opinião			
	Observed N	Expected N	Residual
1	42	60,0	-18,0
2	61	45,0	16,0
3	33	30,0	3,0
4	14	15,0	-1,0
Total	150		

Test Statistics	
	Opinião
Chi-Square	11,456 ^a
df	3
Asymp. Sig.	P

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 15,0.

5. A procura diária de um certo produto foi, em 40 dias escolhidos ao acaso, a seguinte:

Número de unidades	Número de dias
0	6
1	14
2	10
3	7
4	2
5	1

Será que se pode admitir que tais observações foram extraídas de uma população com distribuição Poisson, isto é, será de admitir que a procura diária segue uma distribuição de Poisson?

6. Numa dada sala de cinema da região de Évora realizou-se um inquérito a 400 estudantes, escolhidos aleatoriamente, da Universidade de Évora, relativamente à sua preferência sobre 4 tipos de filmes: A, B C e D. Os resultados obtidos foram:

Filme				Test Statistics	
	Observed N	Expected N	Residual		Filme
A	a	100,0	30,0	Chi-Square	14,000 ^a
B	b	100,0	-10,0	df	3
C	c	100,0	-20,0	Asymp. Sig.	,003
D	d	100,0	,0		
Total	e				

a. 0 cells (.0%) have expected frequencies less than 5. The minimum expected cell frequency is 100,0.

- Determine os valores de **a**, **b**, **c**, **d** e **e**.
- Ao nível de significância de 1%, poderá afirmar-se que não existe preferência por nenhum dos tipos de filmes? Efectue o teste estatístico que considere mais adequado.

7. Foi registado o número de nascimentos num hospital durante os quatro períodos do ano: Jan-Mar, Abr-Jun, Jul-Set, Out-Dez. Diz-se que durante o período de Jan-Mar nascem duas vezes mais crianças do que nos outros períodos. Verifique se os resultados obtidos na experiência contradizem a afirmação.

Trimestre				Test Statistics	
	Observed N	Expected N	Residual		Trimestre
Jan-Mar	110	120,0	-10,0	Chi-square	8,467 ^a
Abr-Jun	57	60,0	-3,0	df	3
Jul-Set	53	60,0	-7,0	Asymp. Sig.	,037
Out-Dez	80	60,0	20,0	a. 0 cells (,0%) have expected frequencies less than 5. The minimum expected cell frequency is 60,0.	
Total	300				

8. Com o objectivo de participarem numa determinada actividade social, os estudantes de uma escola foram submetidos a dois testes: um psicotécnico e um sobre regras de conduta. Obtiveram-se os seguintes resultados:

		Psicotécnico	
		Aprovado	Reprovado
Regras de conduta	Aprovado	54	73
	Reprovado	47	167

Existe relação entre os resultados obtidos nos dois testes? Considere $\alpha=5\%$.

9. Numa certa região do país foi feito um inquérito às preferências clubistas dos adeptos de futebol relativamente aos clubes de futebol Porto, Benfica e Sporting. Os inquiridos foram classificados por faixa etária em 35 anos ou menos e mais de 35 anos. De acordo com o seguinte *output* do SPSS responda às questões apresentadas.

Faixa etária * Clube Crosstabulation

			Clube			Total
			Porto	Benfica	Sporting	
Faixa etária <=35 anos	Count		75	75	d	200
	Expected Count		a	80,0	60,0	200,0
>35 anos	Count		75	125	100	300
	Expected Count		90,0	120,0	90,0	b
Total	Count		150	c	150	500
	Expected Count		150,0	200,0	150,0	500,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9,549 ^a	2	,008
Likelihood Ratio	9,488	2	,009
Linear-by-Linear Association	8,663	1	,003
N of Valid Cases	500		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 60,00.

- a) Determine os valores de **a**, **b**, **c** e **d**.
- b) É de concluir que as preferências se podem considerar independentes da idade?

10. Foi conduzida uma experiência no âmbito da qual se procurou testar se existe alguma relação entre a qualidade da secagem de máquinas de lavar roupa de um certo tipo e a velocidade de rotação a que se eleva o tambor da roupa na fase de

centrifugação. Os resultados desta experiência, efectuada com base no comportamento de 90 máquinas, estão apresentados na seguinte tabela:

		Qualidade da secagem			
		Medíocre	Suficiente	Boa	Muito Boa
Velocidade de Rotação (rpm)	600	12	8	7	3
	900	9	10	7	4
	1200	2	9	8	11

Teste, ao nível de significância de 5%, se as duas variáveis são ou não independentes.

11. Os dados seguintes são o resultado de um inquérito de opinião efetuado a residentes no litoral e no interior acerca do desenvolvimento da sua região. Para averiguar se a opinião acerca do desenvolvimento local é condicionada pela região de residência, procedeu-se a um teste adequado tendo-se obtido o seguinte quadro de resultados:

Regiao * Opiniao Crosstabulation

			Opinião			Total
			Elevado	Suficiente	Insuficiente	
Regiao	Interior	Count	10	B	140	200
		Expected Count	A	90,0	80,0	C
	Litoral	Count	D	130	20	200
		Expected Count	30,0	90,0	F	200,0
Total			60	E	160	G

Chi-Square Tests

	Value	df
Pearson Chi-Square	152,222 ^a	2

- Diga qual o teste realizado, formule as hipóteses correspondentes e determine os valores de **A** a **G**.
- O que pode concluir, para 10% de significância?