



Departamento de
Matemática

Introdução à Probabilidade e Estatística 2015/2016 - 2º Semestre

Ficha N.º 8: Testes de Hipóteses Não-Paramétricos

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 de é de admitir que tenha sido proveniente de uma 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. Suponha que recolheu uma amostra aleatória referente a um determinado item de interesse. Antes de utilizar a amostra para os fins a que foi recolhida resolve testar a distribuição da população da qual essa amostra foi recolhida.

Diga como procederá para efectuar este teste. (Não se esqueça de enunciar qual o tipo de teste bem como o nome do teste escolhido, quais as respectivas hipóteses nula e alternativa do teste e, após o teste realizado, qual o critério de rejeição da hipótese nula.)

3. Considere duas amostras, ambas de dimensão 20, correspondentes a medições físicas em dois grupos de pacientes. A um dos grupos foi administrado um placebo e ao outro administrado um tratamento efectivo. 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 seguinte *output* abaixo, teste de é de admitir que ambas as amostras sejam provenientes de distribuições normais.

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.

4. Um novo de casino consiste em jogar 3 vezes um dado. O vencedor é aquele que obter o maior número de pintas 6. Suponha que um jogador joga 100 vezes com 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 avaliar a falsidade, ou não, do dado. O que conclui?

5. Efectuou-se um estudo para determinar se a opinião pública era favorável à construção de uma barragem eléctrica. 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 indiferentes e 33 contrários à construção.

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

Opinião				Test Statistics	
	Observed N	Expected N	Residual		Opinião
1	42	60,0	-18,0	Chi-Square	11,456 ^a
2	61	45,0	16,0	df	3
3	33	30,0	3,0	Asymp. Sig.	P
4	14	15,0	-1,0		
Total	150				

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

6. 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?

7. 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.

- a) Determine os valores de **a, b, c, d e e**.
b) 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.

8. 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		
Total	300				

a. 0 cells (,0%) have expected frequencies less than 5. The minimum expected cell frequency is 60,0.

9. Com o objectivo de participarem numa dada 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 $\alpha=5\%$.

10. Em 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 seguintes questões

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.

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

11. 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.

12. 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 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?