



Departamento de Matemática

BIOESTATÍSTICA

EXERCÍCIOS – Folha 7

ANO LECTIVO: 2007/2008

ANOVA

1. A anemia é uma doença que afecta muitas pessoas e que pode ter diversas origens. Pretendendo-se avaliar possíveis diferenças entre diferentes tratamentos de estados anémicos, planeou-se uma experiência com 120 indivíduos anémicos, divididos aleatoriamente em três grupos de 40, aos quais se atribuiu cada um dos tratamentos. O primeiro tratamento era constituído apenas por uma dieta rica em ferro. O segundo tratamento combinava um suplemento de ferro com a dieta do primeiro tratamento e o último acrescentava um complexo vitamínico. No sentido de avaliar possíveis diferenças entre os tratamentos, efectuou-se uma ANOVA com base nos valores de hemoglobina dos 120 indivíduos após um período de 3 meses de tratamento. Os resultados da ANOVA encontram-se na tabela seguinte. Complete a tabela nos espaços apropriados.

ANOVA

Hemoglobina

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1,522	_____	_____	_____	,483
Within Groups	121,403	_____	_____		
Total	122,925	_____			

2. Um treinador pretende saber qual o número óptimo de dias semanais de treino para os seus atletas. Para tal mediu a performance de três grupos de atletas separados consoante o número de dias de treino: um, dois e três dias. Os dados encontram-se no ficheiro Atletas2.sav. Teste através de uma ANOVA paramétrica (assumindo que os grupos têm distribuição Normal com variância constante), se existem diferenças entre as performances dos 3 grupos ($\alpha=0.05$).
3. Com o intuito de estudar a influência da temperatura ambiente na capacidade de trabalho das pessoas, 80 estudantes universitários foram divididos em 4 grupos equilibrados que foram sujeitos a 4 temperaturas seleccionadas ao acaso. Cada aluno teve de responder a um conjunto de 150 questões simples tendo sido contabilizado o número de respostas correctas. Os resultados encontram-se no ficheiro Respostas.sav.
 - a. Construa caixas de bigodes comparativas para visualizar a posição relativa dos grupos.

- b. Considerando que são válidos os pressupostos de realização de uma ANOVA paramétrica, conduza uma ANOVA de efeitos aleatórios e retire conclusões ($\alpha=0.05$).
 - i. Escreva as hipóteses em causa;
 - ii. Registe o valor da estatística de teste;
 - iii. Registe o valor do p-value do teste;
 - iv. Conclua.
 - c. No caso de concluir que a temperatura influencia a capacidade de trabalho, forneça uma estimativa da variância, σ_{τ}^2 .
4. Considere o ficheiro Factura.sav que contém informação relativa a uma factura telefónica associada a um telefone da rede fixa da PT de um cliente que não possui nenhum telemóvel. Sabendo que de acordo com o plano tarifário deste cliente, por um lado, as chamadas de curta distância têm as tarifas mais baixas e que as chamadas para redes móveis têm as tarifas mais altas, e por outro, à noite as tarifas são mais reduzidas do que durante o dia, é de esperar encontrar diferenças entre a hora a que se efectuam as chamadas consoante o tipo de chamada. Assumindo que são válidos os pressupostos de realização de uma ANOVA paramétrica:
- a. Obtenha a tabela de ANOVA que permite analisar se a hora a que se realizam as chamadas (variável hora) depende do tipo de chamada (variável tipochamada), $\alpha=0.05$.
 - i. Escreva as hipóteses em causa;
 - ii. Registe o valor da estatística de teste;
 - iii. Registe o valor do p-value do teste;
 - iv. Conclua.
 - b. Efectue uma análise de comparações múltiplas utilizando os métodos de Bonferroni e de Tuckey para $\alpha=0.05$. O que pode concluir?
5. Considere o ficheiro Rolhas.sav que contém os valores de várias variáveis medidas num conjunto de 150 rolhas. Uma dessas variáveis é a classe a que pertence a rolha e que tem que ver com a qualidade geral da rolha (Super, Normal, Pobre). Outra das variáveis é o número de defeitos observados em cada rolha. Pretende-se estudar se o número de defeitos está relacionado com a qualidade da rolha.
- a. Construa caixas de bigodes comparativas para visualizar a posição relativa dos grupos.
 - b. Construa QQ-plots dos diferentes grupos e efectue testes de ajustamento para avaliar a possibilidade de efectuar uma ANOVA paramétrica aos dados ($\alpha=0.05$).
 - c. Efectue o teste de homogeneidade de variâncias para $\alpha=0.05$.
 - d. Efectue uma ANOVA paramétrica aos dados.
 - i. Escreva as hipóteses em causa;
 - ii. Registe o valor da estatística de teste;
 - iii. Registe o valor do p-value do teste;
 - iv. Conclua.
 - e. Efectue uma análise de comparações múltiplas utilizando os métodos de Bonferroni e de Tuckey para $\alpha=0.05$. O que pode concluir?

- f. Seguidamente considerou-se a variável Área Média dos defeitos e obtiveram-se as seguintes tabelas de ANOVA e de comparações múltiplas:

ANOVA

Área média dos defeitos

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	277,867	2	138,933	174,243	,000
Within Groups	117,211	147	,797		
Total	395,078	149			

Multiple Comparisons

Dependent Variable: Área média dos defeitos

	(I) Classe	(J) Classe	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Super	Normal	-,918000*	,178589	,000	-1,34084	-,49516
		Pobre	-3,234600*	,178589	,000	-3,65744	-2,81176
	Normal	Super	,918000*	,178589	,000	,49516	1,34084
		Pobre	-2,316600*	,178589	,000	-2,73944	-1,89376
	Pobre	Super	3,234600*	,178589	,000	2,81176	3,65744
		Normal	2,316600*	,178589	,000	1,89376	2,73944
Bonferroni	Super	Normal	-,918000*	,178589	,000	-1,35049	-,48551
		Pobre	-3,234600*	,178589	,000	-3,66709	-2,80211
	Normal	Super	,918000*	,178589	,000	,48551	1,35049
		Pobre	-2,316600*	,178589	,000	-2,74909	-1,88411
	Pobre	Super	3,234600*	,178589	,000	2,80211	3,66709
		Normal	2,316600*	,178589	,000	1,88411	2,74909

*. The mean difference is significant at the .05 level.

O que pode concluir a partir destas duas tabelas assumindo que são válidos os pressupostos de realização da ANOVA?

6. Foi realizada uma experiência para testar o efeito da alimentação na aprendizagem. Para o efeito separou-se uma turma em quatro grupos: um grupo de controlo ao qual foi fornecida bebida e comida à discrição, um outro grupo ao qual se forneceu apenas comida, um grupo que apenas ingeriu bebidas e um quarto grupo que não teve direito nem a comer nem a beber. Em seguida, os alunos assistiram a uma aula sobre a ANOVA no SPSS. No final, cada aluno realizou um teste cujos resultados se encontram no ficheiro Aprend.sav.
- Construa caixas de bigodes comparativas e observe as diferenças entre os grupos.
 - Efectue uma ANOVA paramétrica para averiguar se existem diferenças significativas entre os grupos, ao nível de significância de 5%.
 - Escreva as hipóteses em causa;
 - Registe o valor da estatística de teste;
 - Registe o valor do p-value do teste;
 - Conclua.

- c. Efectue uma análise de comparações múltiplas utilizando o teste de Dunnett para comparar o grupo de controle com cada um dos restantes grupos, $\alpha=0.05$.
7. Recolheram-se os valores do índice se Apgar medido ao fim de 1 e de 5 minutos num conjunto de 227 recém-nascidos em 3 hospitais portugueses: Hospital de S. João do Porto, Hospital de S. António do Porto e Hospital da Universidades de Coimbra. O índice de Apgar toma valores (inteiros) de 0 a 10 que resultam de uma classificação feita pelos médicos aos recém-nascidos. Esta classificação é distribuída por 5 atributos: frequência cardíaca, esforço respiratório, tónus muscular, irritabilidade reflexa e cor. Cada atributo é classificado com 0, 1 ou 2 e o índice resulta da soma das 5 classificações. Um bebé normal deve ter índice superior ou igual a 8. Os resultados encontram-se no ficheiro APGAR.sav
- a. Tendo em conta a natureza (ordinal) da variável em causa, efectue uma ANOVA não paramétrica (teste de Kruskal-Wallis) para averiguar se existem diferenças significativas entre os três hospitais ($\alpha=0.01$), no que respeita ao índice de Apgar ao fim de 1 minuto.
- Escreva as hipóteses em causa;
 - Registe o valor da estatística de teste;
 - Registe o valor do p-value do teste;
 - Conclua.
- b. Seguidamente foi feita uma ANOVA não paramétrica à variável Índice de APGAR ao fim de 5 minutos e obteve-se os seguintes resultados.

Ranks

	Hospital	N	Mean Rank
Apgar 5	HUC	26	134,94
	HGSA	48	128,07
	HSJ	150	103,63
	Total	224	

Test Statistics^{a,b}

	Apgar 5
Chi-Square	14,091
df	2
Asymp. Sig.	,001

a. Kruskal Wallis Test

b. Grouping Variable: Hospital

- c. Ao nível de significância de 1% o que pode concluir?
- Escreva as hipóteses em causa;
 - Registe o valor da estatística de teste;
 - Registe o valor do p-value do teste;
 - Conclua.
8. Considere novamente o ficheiro Factura.sav do exercício 4. Efectue uma ANOVA factorial (dupla) para averiguar se o custo das chamadas depende dos factores tipo de chamada e tipo de tarifa. (Menu Analyze / General Linear Model / Univariate, ambos os factores de efeitos fixos).

- a. O que pode concluir a partir da tabela de ANOVA?
 - b. Utilizando o botão Save do menu utilizado, guarde os resíduos (unstandardized) e efectue um QQ-plot e teste de ajustamento à Normal (Menu Analyze / Descriptive Statistics / Explore) para averiguar se os erros se podem considerar Normais.
9. Cinco conjuntos de quatro porquinhos da índia (cada conjunto pertence a uma mesma linhagem) foram sujeitos a uma experiência para averiguar possíveis diferenças entre 4 dietas alimentares distintas. Foi feito um planeamento por blocos aleatorizados tendo sido atribuída cada uma das 4 dietas a cada porquinho da mesma linhagem. Obtiveram-se assim 5 blocos com 4 observações cada. A atribuição das dietas dentro de cada bloco foi aleatória e a escolha dos blocos também. Os dados encontram-se no ficheiro DietaPorquinhos.sav. Efectue uma ANOVA com blocos aleatorizados para averiguar se existem diferenças significativas entre as dietas, ao nível de significância de 5%. Utilize o menu General Linear Model / Univariate colocando a variável peso na janela da variável dependente, a variável dieta como efeito fixo e a variável bloco como efeito aleatório. Seguidamente seleccione o botão Model e seleccione Main effects no centro da janela (Build Terms) passando para a janela da direita as duas variáveis (bloco e dieta). Seleccione o botão Save e guarde os resíduos (unstandardized).
- i. Escreva as hipóteses em causa;
 - ii. Registe o valor da estatística de teste F relativa às dietas;
 - iii. Registe o valor do p-value do teste relativo às dietas;
 - iv. Registe o valor da estatística de teste F relativa aos blocos;
 - v. Registe o valor do p-value do teste relativo aos blocos;
 - vi. Conclua.
 - vii. Efectue uma análise dos resíduos para averiguar se os erros se podem considerar Normais (Menu Analyze / Descriptive Statistics / Explore).
10. Foi efectuado um estudo sobre a apanha do casulo (*Diopatra neapolitana*) na ria de Aveiro. Os apanhadores de casulo foram observados durante um ano inteiro e na análise dos resultados foram considerados dois factores que poderiam influenciar a apanha: a estação do ano e o tipo de maré (viva ou morta). Consideraram-se duas variáveis de interesse: o esforço de captura (mariscador * min / dia) e a taxa de captura (quilos capturados / (mariscador * min). Foi feita uma ANOVA a dois factores sobre cada uma das variáveis de interesse tendo-se obtido as seguintes tabelas de resultados:

Esforço de captura

Fonte de variação	Graus de liberdade	MS	F	p-value
Maré	1	5151180	2,0823	>0.1
Estação	3	1430122	0,5781	>0.5
Interacção	3	118744	0,4482	>0.5
Erro	45	2473847		

Taxa de captura

Fonte de variação	Graus de liberdade	MS	F	p-value
Maré	1	,001403	5,1871	<0.05
Estação	3	,000995	3,6760	<0.05
Interacção	3	,000290	1,0716	>0.3
Erro	21	,000271		

- a. O que pode concluir ao nível de significância 0.05 no que respeita a variável Esforço de captura?
- b. O que pode concluir ao nível de significância 0.05 no que respeita a variável Taxa de captura?