

**Estatística Computacional**  
(Licenciatura em Matemática)

Duração: 2h

Exame

29-06-10

NOME: \_\_\_\_\_

**Observação:** A resolução completa das perguntas inclui a justificação do raciocínio utilizado.

Uma equipa de biólogos dedicou-se ao estudo de certa espécie de ursos. Com este objectivo, seleccionou aleatoriamente 44 animais adultos daquela espécie, de entre os que vivem numa determinada reserva natural, e observou-os durante um ano. Os dados recolhidos foram sujeitos a um tratamento estatístico apoiado no SPSS, do qual se apresentam alguns excertos.

1. No que diz respeito à amostra das idades (expressas em meses) dos animais no início do estudo, foi obtido o seguinte *output*:

	Idade
Test Value <sup>a</sup>	142
Cases < Test Value	22
Cases >= Test Value	22
Total Cases	44
Number of Runs	23
Z	,000
Asymp. Sig. (2-tailed)	1,000
Exact Sig. (2-tailed)	1,000

a. Median

a) Indique as hipóteses em confronto. Que conclusão pode tirar, aos níveis de significância usuais?

b) Como são definidos os *runs*?

2. Um dos aspectos que interessou os investigadores foi a duração (expressa em dias) do período de hibernação dos animais. Nos quadros que se seguem figuram os resultados da análise descritiva das durações observadas:

**Descriptives**

		Statistic	Std. Error
Duração da hibernação (em dias)	Mean	173,45	2,733
	Median	175,00	
	Variance	328,672	
	Std. Deviation	18,129	
	Skewness	-,506	,357

**Extreme Values**

		Case Number	Value
Duração da hibernação (em dias)	Highest	1	22
		2	23
		3	21
		4	20
		5	43
	Lowest	1	24
		2	1
		3	26
		4	25
		5	27

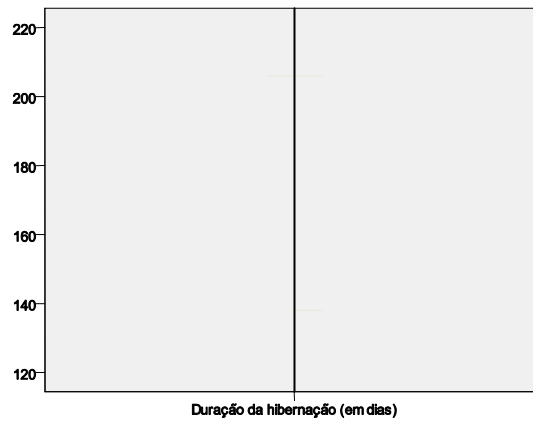
**Percentiles**

		Percentiles						
		5	10	25	50	75	90	95
Weighted Average (Definition 1)	Duração da hibernação (em dias)	140,00	146,50	162,75	175,00	186,50	196,50	204,00

a) Descreva a amostra no que diz respeito à localização central e à dispersão.

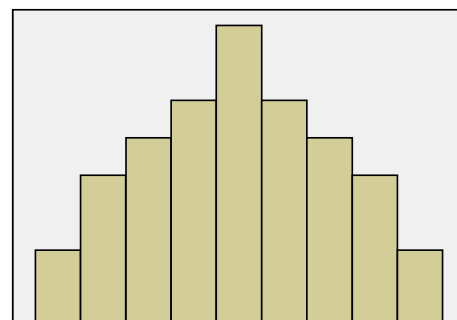
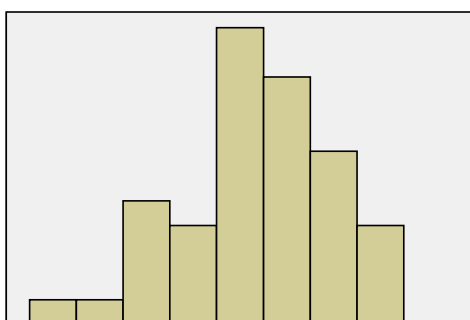
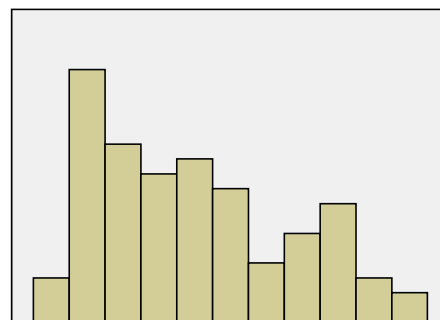
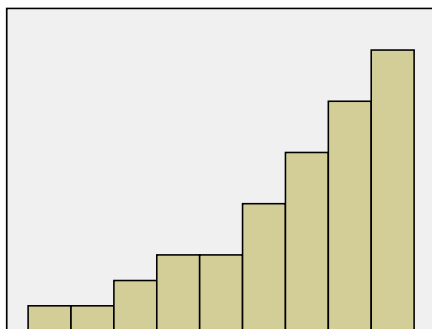
b) Indique o intervalo que contém os 90% valores mais centrais.

c) Desenhe a correspondente caixa de bigodes.



d) Que pode afirmar quanto à assimetria da amostra?

e) De entre os histogramas apresentados, seleccione aquele que poderá estar associado à amostra observada.



f) Estudos anteriores indicavam que a duração média da hibernação destes animais era inferior a 6 meses (aproximadamente 180 dias). Face aos resultados obtidos para a amostra observada, apresentados no quadro que se segue, para que níveis de significância se confirma tal afirmação?

One-Sample Test

	Test Value = 180					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Duração da hibernação (em dias)	-2,395	43	,021	-6,545	-12,06	-1,03

3. Pretende-se agora relacionar o sexo do animal com a duração do seu período de hibernação.

a) Qual é o coeficiente adequado para medir a intensidade da associação entre estes dois atributos?

b) Depois de se proceder à classificação da amostra dos tempos de hibernação dos animais, foi construída a seguinte tabela de contingência:

Sexo \* Duração da hibernação (em dias) (Binned) Crosstabulation

			Duração da hibernação (em dias) (Binned)			Total
			<= 150	151 - 180	181+	
Sexo	Macho	Count	1	12	10	23
		% within Sexo			a%	
		% within Duração da hibernação (em dias) (Binned)		b%		
		% of Total			c%	
Sexo	Fêmea	Count	5	10	6	21
		% within Sexo				
		% within Duração da hibernação (em dias) (Binned)				
		% of Total	d%			
Total		Count	6	22	16	44
		% within Sexo			e%	
		% within Duração da hibernação (em dias) (Binned)	f%			
		% of Total				

Determine os valores das constantes a, b, c, d, e e f e interprete os quatro primeiros.

4. Outro aspecto analisado foi a perda de peso dos animais durante o período de hibernação. Os estudos já referidos concluíram que é bastante elevada a proporção de animais que perdem, no período em causa, mais de 25% do seu peso. A fim de comprovar esta opinião, os biólogos determinaram a perda percentual de peso de cada animal neste

período e, usando a amostra assim obtida, efectuaram um teste binomial cujos resultados se encontram no próximo quadro:

		Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)	Exact Sig. (1-tailed)
Difperc	Group 1	<= 25	20	.5	.6	.036 <sup>a, b</sup>	.036
	Group 2	> 25	24	.5			
	Total		44	1,0			

a. Alternative hypothesis states that the proportion of cases in the first group < .6.

b. Based on Z Approximation.

a) Como foram definidos os dois grupos referidos no quadro anterior?

b) Indique as hipóteses em confronto. Que conclusão pode tirar sobre tal proporção? Justifique a resposta.

5. A quantidade de alimentos (expressa em kg) ingerida diariamente por cada animal foi também objecto de estudo.

a) Na pesquisa de um modelo adequado para representar esta característica em cada um dos sexos, utilizou-se o teste de Kolmogorov-Smirnov, com a correcção de Lilliefors, que forneceu o seguinte *output*.

Sexo		N	Mean	Std. Deviation	Std. Error Mean
Ingestdia	Macho	23	36,813	2,2642	,4721
	Fêmea	21	34,639	3,5817	,7816

Sexo		Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Ingestdia	Macho	,094	23	,200*	,972	23	,743
	Fêmea	,122	21	,200*	,976	21	,856

a. Lilliefors Significance Correction

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

(i) Que pode concluir, aos níveis de significância usuais?

(ii) Porque razão não é aconselhável efectuar este teste no menu *Nonparametric Tests*?

b) Admitindo que tais características são bem representadas por variáveis aleatórias independentes, proponha uma estimativa para a probabilidade de a quantidade de alimentos ingerida diariamente por um macho ser superior à ingerida por uma fêmea.

c) O quadro que se segue contém os resultados do teste que permite comparar as quantidades médias de alimentos ingeridas diariamente pelos machos e pelas fêmeas.

**Independent Samples Test**

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Ingestdia	Equal variances assumed	4,775	,035	2,428	42	,020	2,1736	,8951	,3673	3,9799
	Equal variances not assumed			2,380	33,235	,019	2,1736	,9131	,3164	4,0309

Utilizando o nível de significância 0.05, tire as conclusões que achar convenientes.

6. Ainda com base nos estudos atrás mencionados, foi possível postular uma lei de probabilidade para modelar o número de crias por ninhada. Para aferir a qualidade do modelo, a equipa observou, para cada fêmea da amostra, o número de crias nascidas no período em causa e realizou o teste de ajustamento do qui-quadrado. Obteve então o seguinte *output*:

	Observed N	Expected N	Residual
1	4	5,25	-1,25
2	7	6,3	0,7
3	8	6,3	1,7
4	2	3,15	-1,15
Total	21		

	Numcrias
Chi-Square	1,254*
df	3
Asymp. Sig.	,740

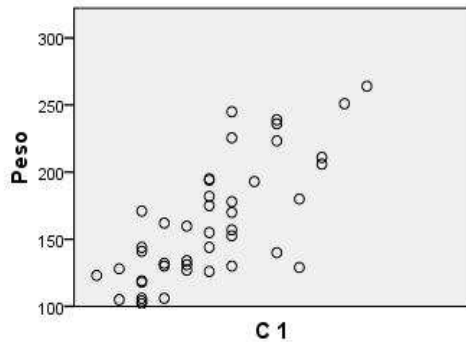
a. 1 cells (25,0%) have expected frequencies less than 5. The minimum expected cell frequency is 3,2.

a) Indique as hipóteses em confronto.

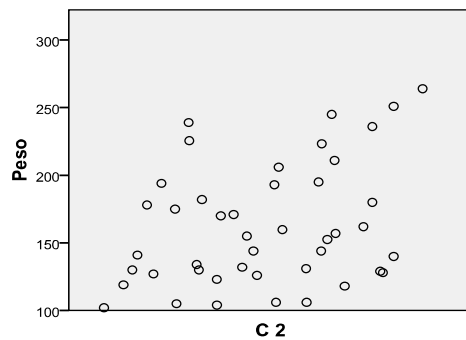
b) Como procederia para concluir o teste?

7. Os biólogos anestesiaram os animais por forma a conseguirem observar as suas características físicas. Deste modo, obtêm facilmente medidas do comprimento, do perímetro do tórax, do diâmetro da cabeça, etc. Contudo, é bastante difícil levantá-los e conseqüentemente pesá-los. Por esta razão, é importante relacionar funcionalmente o peso de um urso com algumas destas características (designadas, no que se segue, por C1, C2, C3 e C4).

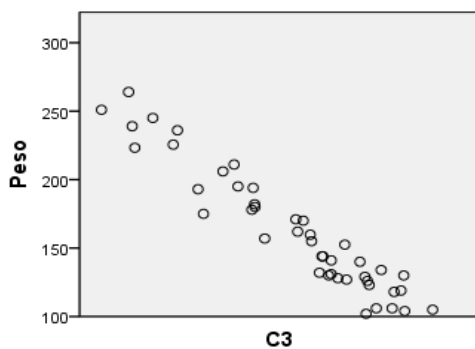
a) Com este objectivo, foram traçados os diagramas de dispersão que se seguem:



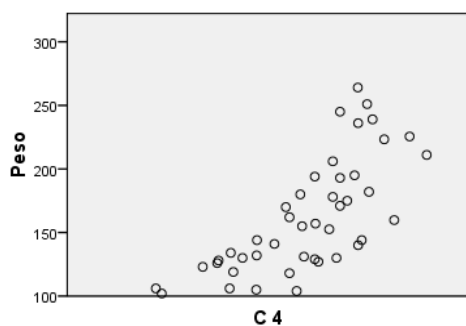
Simetric Measures		
		Value
Interval by Interval	Pearson's R	-,958
Ordinal by Ordinal	Spearman Correlation	-,952
N of Valid Cases		44



Symmetric Measures		
		Value
Interval by Interval	Pearson's R	,847
Ordinal by Ordinal	Spearman Correlation	,834
N of Valid Cases		44



Symmetric Measures		
		Value
Interval by Interval	Pearson's R	,702
Ordinal by Ordinal	Spearman Correlation	,769
N of Valid Cases		44



Symmetric Measures		
		Value
Interval by Interval	Pearson's R	,196
Ordinal by Ordinal	Spearman Correlation	,127
N of Valid Cases		44

Associe a cada diagrama o quadro respectivo.

b) Por forma a relacionar o peso de um urso ( $Y$ , em  $kg$ ) com o perímetro do seu tórax ( $x$ , em  $cm$ ), procedeu-se ao ajustamento de um modelo de regressão linear, tendo sido obtidos os seguintes resultados:

**Model Summary**

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,958 <sup>a</sup>	,919	,917	12,918

a. Predictors: (Constant), Perimtorax

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-52,155	10,044		a	,000
	Perimtorax	1,758	,081	,958	b	,000

a. Dependent Variable: Pesoant

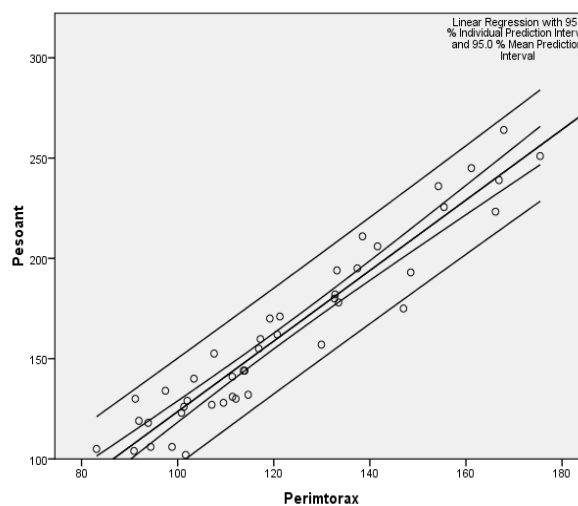
(i) Determine a e b.

(ii) Qual o significado do valor 12.918 que figura no primeiro quadro?

(iii) Sabendo que os valores do desvio padrão das duas amostras envolvidas são 44.75 e 24.4, determine a covariância da amostra bidimensional associada.

(iv) Que pode afirmar relativamente à qualidade do modelo ajustado?

(v) Identifique os elementos que constituem o gráfico seguidamente apresentado.



c) A comparação entre os valores observados do peso e os seus valores ajustados pelo modelo foi efectuada através do teste dos sinais, cujo *output* se apresenta de seguida.

Frequencies		N
Unstandardized Predicted Value - Pesoant	Negative Differences <sup>a</sup>	23
	Positive Differences <sup>b</sup>	21
	Ties <sup>c</sup>	0
	Total	44

a. Unstandardized Predicted Value < Pesoant

b. Unstandardized Predicted Value > Pesoant

c. Unstandardized Predicted Value = Pesoant

Test Statistics <sup>a</sup>	
	Unstandardized Predicted Value - Pesoant
Z	-,151
Asymp. Sig. (2-tailed)	,880
Exact Sig. (2-tailed)	,880
Exact Sig. (1-tailed)	,440

a. Sign Test

(i) Quais são as hipóteses do teste que corresponde ao valor de prova 0.44?

(ii) Indique uma expressão que permita calcular o valor referido na alínea anterior.

(iii) Que conclusão pode tirar, aos níveis de significância usuais?

(iv) Que outro teste não paramétrico poderia utilizar para o mesmo fim?