

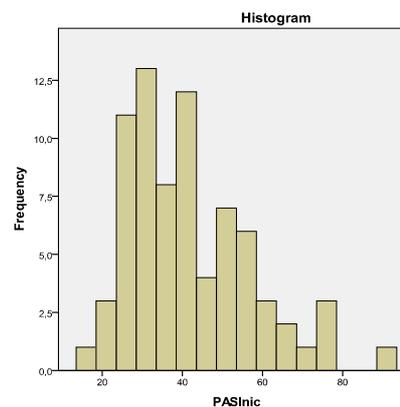
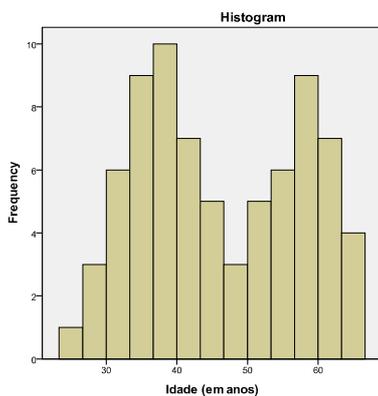
NOME: _____

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

A psoríase é uma doença crónica da pele que pode afectar pessoas de todas as idades. Embora as suas causas exactas sejam desconhecidas, sabe-se que os antecedentes familiares, o stress e factores metabólicos como a diabetes contribuem para o seu aparecimento. Uma forma usual de representar numericamente o estado dos pacientes com esta doença é através do índice PASI que traduz, na escala 0-100, a extensão da área afectada e a gravidade das lesões, de acordo com a avaliação de um dermatologista.

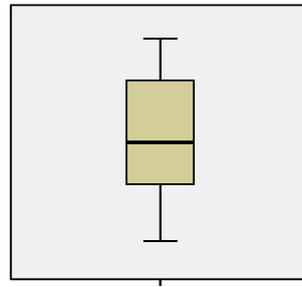
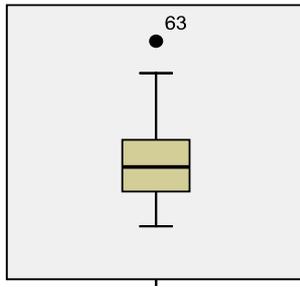
A fim de analisar o efeito de um novo tratamento para a psoríase, foi realizada uma experiência envolvendo uma amostra aleatória de 75 pacientes de ambos os sexos e várias faixas etárias. Para cada paciente, registou-se o sexo (0-Feminino, 1-Masculino), a idade actual (em anos), a idade em que lhe foi diagnosticada a doença (em anos), o valor do índice PASI, a existência de antecedentes familiares da doença (0-Não, 1-Sim) e o estado psicológico do paciente face ao tratamento (0-Pessimista, 1-Indiferente, 2-Optimista).

1. Apresentam-se seguidamente os histogramas relativos às observações das variáveis Idade e PASInic (valor do índice PASI no início do tratamento).



- a) Que informações fornecem estes gráficos sobre as respectivas amostras?

- b) Associe a cada histograma o correspondente diagrama de extremos-e-quartis.



2. Por forma a modelar a lei de probabilidade do índice PASInic na população dos pacientes com psoríase, considerou-se a variável LnPASInic cujos valores são os logaritmos neperianos dos valores de PASInic.

a) A análise descritiva desta nova amostra conduziu aos seguintes resultados:

Descriptives				Statistics		
		Statistic	Std. Error	LnPASInic		
LnPASInic	Mean	3,6649	,04096	N	Valid	75
	Variance	,126			Missing	0
	Std. Deviation	,35468		Percentiles	25	3,4012
	Minimum	2,77			50	3,6889
	Maximum	4,51			75	3,9318
	Skewness	,114	,277			

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

(ii) Que pode afirmar quanto à assimetria da amostra?

(iii) Averigúe se existem outliers.

b) O próximo quadro constitui o *output* do teste de Kolmogorov-Smirnov, aplicado a tal amostra:

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
LnPASInic	,076	75	,200*	,989	75	,763

a. Lilliefors Significance Correction

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

(i) Indique as hipóteses em confronto.

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

c) (i) Sabe-se que as formas graves de psoríase estão associadas a valores do índice PASI superiores a 80.
Tendo em conta a conclusão obtida na alínea anterior, determine uma estimativa para a probabilidade de um paciente apresentar uma forma grave da doença.

(ii) Utilizando o menu *Compare Means*, foi construído um intervalo de confiança para a esperança da variável LnPASInic, com grau de confiança 0.95.

One-Sample Test

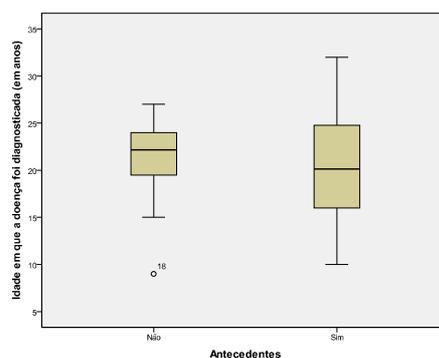
	Test Value = 0					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
LnPASInic	a	b	.000	c	3,5833	d

Indique os valores das constantes a, b, c e d.

(iii) Como poderá utilizar o intervalo de confiança obtido na alínea anterior para estimar a localização da variável PASInic? Justifique convenientemente a resposta.

3. Seguidamente, os investigadores procuraram relacionar a idade em que foi diagnosticada a doença com o sexo e os antecedentes familiares dos pacientes.

a) Com base nos seguintes *box-plots* paralelos, compare as idades em que foi diagnosticada a doença nos pacientes com e sem antecedentes familiares.



b) No que diz respeito ao sexo, optou-se por um teste t de comparação de médias, depois de verificados os devidos pressupostos. O *output* obtido é apresentado nos quadros que se seguem.

Group Statistics

	Sexo	N	Mean	Std. Deviation	Std. Error Mean
Idade em que a doença foi diagnosticada (em anos)	Feminino	43	20,00	5,476	,835
	Masculino	32	22,27	3,991	,705

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
Idade em que a doença foi diagnosticada (em anos)	Equal variances assumed	3,587	,062	-1,984	73	,051	-2,270	1,144	-4,550	,010
	Equal variances not assumed			-2,076	72,980	,041	-2,270	1,093	-4,448	-,091

(i) Quais são as hipóteses do teste de Levene?

(ii) Indique uma expressão que permita calcular o valor 0.051 que figura na 1ª linha da tabela.

(iii) Para que níveis de significância poderá afirmar que, em média, o aparecimento da doença é mais precoce nos indivíduos do sexo feminino?

4. Os pacientes foram sujeitos ao referido tratamento durante 6 meses e, após este período, foram examinados novamente, tendo-lhes sido atribuído um novo valor do índice PASI. A comparação dos valores inicial e final deste índice permitiu, então, avaliar a eficácia do tratamento.

a) Os resultados que se seguem dizem respeito ao teste de Wilcoxon.

Ranks

	N	Mean Rank	Sum of Ranks	
PASIPos -PASInic	Negative Ranks	52 ^a	41,32	2148,50
	Positive Ranks	21 ^b	26,31	552,50
Ties	0 ^c			
Total	73			

- a. PASIPos < PASInic
- b. PASIPos > PASInic
- c. PASIPos = PASInic

Test Statistics^a

	PASIPos - PASInic
Z	-4,388 ^a
Asymp. Sig. (2-tailed)	,000
Exact Sig. (2-tailed)	,000
Exact Sig. (1-tailed)	,000
Point Probability	,000

- a. Based on positive ranks.
- b. Wilcoxon Signed Ranks Test

(i) Apresente uma justificação para o facto de terem sido consideradas apenas 73 observações.

(ii) Indique as hipóteses em confronto.

(iii) Que pode concluir, ao nível de significância 0.05?

b) Qual é o teste paramétrico alternativo ao teste anterior?

5. Outros aspectos importantes que contribuíram para avaliar o sucesso do tratamento foram estudados com base na diferença entre os valores inicial e final do índice PASI definida, no ficheiro dos dados, pela variável DifPASI.

Considera-se que o estado de saúde de um paciente melhorou se lhe corresponde uma diferença positiva.

a) O primeiro aspecto diz respeito à proporção de pacientes que melhoram o seu estado de saúde com o tratamento em causa. Para inferir sobre esta proporção, efectuou-se o teste binomial, que conduziu aos seguintes resultados:

	Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)
DifPASI	Group 1	22	,3	,4	,037 ^a ^b
	Group 2	53	,7		
Total		75	1,0		

a. Alternative hypothesis states that the proportion of cases in the first group < .4.

b. Based on Z Approximation.

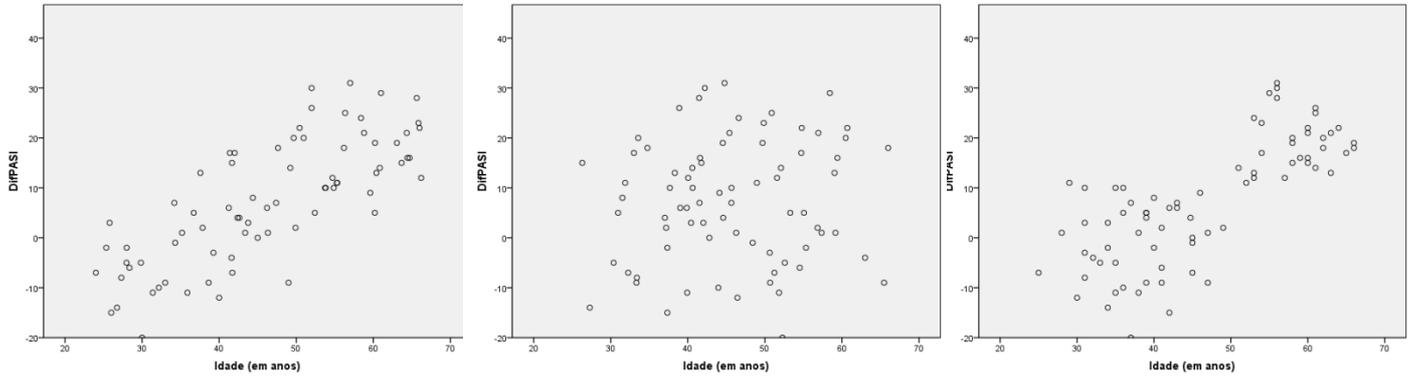
(i) Qual o significado da expressão *Based on Z Approximation*?

(ii) Que conclusão pode tirar deste *output* sobre a proporção em causa?

b) O segundo prende-se com a influência da idade na redução do índice PASI. Por forma a analisar tal influência, foram calculados os coeficientes de correlação da amostra bidimensional relativa às idades dos pacientes e às correspondentes reduções do índice PASI, cujos valores figuram no próximo quadro:

	Value
Interval by Interval Pearson's R	,757
Ordinal by Ordinal Spearman Correlation	,752
N of Valid Cases	75

- (i) Dos diagramas de dispersão seguidamente apresentados, qual poderá corresponder a esta amostra? Justifique a resposta.



- (ii) Como é calculado o coeficiente de correlação de Spearman?

6. Procurou-se também relacionar o estado psicológico do paciente face ao tratamento com a redução do seu índice PASI. Para tal, procedeu-se à classificação da variável DifPASI e considerou-se a amostra bidimensional correspondente.

- a) Esta amostra foi resumida na tabela de contingência que se segue.

DifPASI (Binned) * Estado psicológico do paciente Crosstabulation

Count		Estado psicológico do paciente			Total
		Pessimista	Indiferente	Optimista	
DifPASI (Binned)	<= 0	12	6	4	22
	1 - 20	9	21	11	41
	21+	1	2	9	12
Total		22	29	24	75

- (i) De entre os pacientes que se encontravam otimistas, qual a percentagem dos que se apresentaram melhoras?

- (ii) Que tipo de gráficos de barras poderia utilizar para representar graficamente a amostra anterior?

b) Com base na mesma amostra, foi ainda efectuado o teste do qui-quadrado, que forneceu o *output* apresentado no início da página seguinte

- (i) Quais são as hipóteses em confronto?

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	19,834 ^a	4	,001
Likelihood Ratio	18,337	4	,001
N of Valid Cases	75		

a. 3 cells (33,3%) have expected count less than 5. The minimum expected count is 3,52.

(ii) Explique o significado do comentário fornecido pelo SPSS. Como procederia para efectuar o teste?

7. Por fim, foi ajustado um modelo de regressão linear para exprimir o valor do índice PASI no fim do tratamento (PASIPos) à custa do seu valor no início (PASInic). Os resultados obtidos constam das tabelas que se seguem.

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	15945,878	1	15945,878	103,565	,000 ^a
	Residual	11239,802	73	153,970		
	Total	27185,680	74			

a. Predictors: (Constant), Índice da severidade da doença

b. Dependent Variable: PASIPos

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-5,882	4,195		-1,402	,165
	Índice da severidade da doença	,965	,095	,766	10,177	,000

a. Dependent Variable: PASIPos

a) Indique as estimativas dos coeficientes do modelo.

b) Que informação fornecem os valores de prova apresentados na 2ª tabela?

c) Determine o coeficiente de determinação e comente o valor obtido.

d) Quais são as hipóteses do modelo? Como poderia validá-las?