



Instituto
Politécnico de Setúbal
Escola Superior de
Tecnologia de Setúbal



Licenciatura em Engenharia
Informática
Ano letivo 2021-2022

Unidade Curricular Métodos
Estatísticos

Docente José Palma

Ena Barão | 201400238

Nuno Reis | 202000753

Bernardo Teixeira | 201801954

Conteúdo

Resumo	4
Introdução	5
Dados fornecidos e tratamento dos dados	7
Variável year_born - análise estatística descritiva	8
Análise da distribuição year_born	11
Teste de ajustamento Qui-Quadrado year_born	11
Variável company_size - análise estatística descritiva	11
Regressão Linear	15
Diagrama de dispersão	15
Coefficiente de correlação linear de Pearson	16
Reta de regressão linear	16
Resíduos	17
Ajustamento de Kolmogorov-Smirnov	18
Relação family_member/year	18
Estatística descritiva family_member	18
Estatística descritiva year	21
Distribuição family_member/year	25
Teste de Independência Qui-Quadrado	25
Conclusões	26
Referências bibliográficas	28

Índice de gráficos

Gráfico 1- Histograma das frequências absolutas da variável year_born	9
Gráfico 2- Gráfico de barras das frequências absolutas da variável company_size	12
Gráfico 3-Gráfico circular das frequências relativas da variável company_size	13
Gráfico 4- Diagrama de dispersão com todo o conjunto de dados	15
Gráfico 5- Diagrama de dispersão com a reta da regressão linear	16
Gráfico 6 - Gráfico de resíduos	17
Gráfico 7 - Gráfico de barras das frequências absolutas por extensão de agregado familiar	19
Gráfico 8 - Gráfico circular das frequências relativas por extensão de agregado familiar ...	19
Gráfico 9 - Gráfico circular das frequências relativas do ano de estudo	22

Gráfico 10 - Gráfico de barras das frequências absolutas do ano de estudo.....	23
--	----

Índice de tabelas

Tabela 1- Tabela de frequências da variável year_born	8
Tabela 2 - Medidas de localização da variável year_born	9
Tabela 3 - Medidas de dispersão da variável year_born	10
Tabela 4- Tabela de frequências da variável company_size	12
Tabela 5- Medidas de dispersão da variável company_size	13
Tabela 6 - Medidas de localização da variável company_size	14
Tabela 7 - Tabela de frequências da variável family_member	18
Tabela 8 - Medidas de localização da variável family_member	20
Tabela 9 - Medidas de dispersão da variável family_member	20
Tabela 10 - Tabela de frequências da variável year	21
Tabela 11 - Medidas de dispersão da variável year	23
Tabela 12 - Medidas de localização da variável year	24

Índice de figuras

Figura 1- Caixa de bigodes da variável year_born	10
Figura 2-Caixa de bigodes variável company_size	14
Figura 3 - Caixa de bigodes da variável family_member	21
Figura 4 - Caixa de bigodes da variável year	24

Resumo

No âmbito da disciplina de métodos estatísticos, após o primeiro trabalho que se tratou da análise de um conjunto de dados na esfera da estatística descritiva, no segundo trabalho onde se abordou os modelos de regressão linear iremos dar seguimento com o mesmo conjunto de dados relativo a um estudo conduzido pela Coreia do Sul, de 2005 a 2018, que recolheu várias informações sobre os seus cidadãos, particularmente sobre o rendimento das famílias, para o terceiro trabalho.

Conforme solicitado no enunciado iremos testar as distribuições dos dados selecionados, realizar testes de ajustamento tal como testes de independência e daí tecer ilações e tomar decisões.

Durante a análise irá ser produzido um script de R e todos os materiais daí resultantes serão aqui devidamente apresentados.

Palavras-chave: Testes de ajustamento Qui-Quadrado; teste de ajustamento de Kolmogorov-Smirnov; Análise de resíduos;

Introdução

Este trabalho foi nos solicitado no âmbito da Unidade Curricular de Métodos Estatísticos e tem como principal objetivo estudar as distribuições teóricas dos dados selecionados, aplicar testes de ajustamento tal como testes de independência.

Para este efeito foram lançados diferentes desafios:

1. Estudar uma variável selecionada

Selecionamos a seguinte variável aleatória quantitativa:

- Ano de nascimento(no script de R denominada como year_born);

Para o estudo desta variável apresentamos:

- Estatística descritiva;
- Testamos a distribuição que parece adequada, justificando;
- Testes de ajustamento Qui-Quadrado considerando uma significância de 10%;
 - As hipóteses testadas;
 - Valor -p;
 - Região crítica;
 - Conclusões da tomada de decisão relativamente às hipóteses testadas

2. Estudar a relação de 2 variáveis quantitativas presentes no nosso conjunto de dados e no trabalho 2 representaram o exercício da regressão linear.

Selecionamos as seguintes variáveis aleatórias quantitativas:

- Tamanho da companhia (no script de R denominada como company_size);
- Ano de nascimento(no script de R denominada como year_born);

Para o estudo da relação das variáveis quantitativas apresentamos:

- Estatística descritiva de ambas as variáveis;
- Testamos a distribuição que parece adequada, justificando;
- Qual o modelo de regressão linear obtido com indicação da variável independente e dependente escolhidas;
- Caso o modelo seja uma regressão linear, quantificamos a força da correlação linear:
 - Através do diagrama de dispersão;
 - Coeficiente de correlação linear de Pearson;
- Através do teste de ajustamento de Kolmogorov-Smirnov, analisamos os resíduos obtidos pelo modelo;
 - As hipóteses testadas;
 - Valor -p;
 - Região crítica;

- Conclusões da tomada de decisão relativamente às hipóteses testadas
- Aplicamos independência Qui-Quadrado para verificar a associação entre as variáveis com um grau de significância de 1%
 - As hipóteses testadas;
 - Valor -p;
 - Região crítica;
 - Conclusões da tomada de decisão relativamente às hipóteses testadas

3. Estudar a relação de 2 variáveis quantitativas presentes no nosso conjunto de dados .

Selecionamos as seguintes variáveis aleatórias quantitativas:

- Agregado familiar (no script de R denominada como family_member);
- Ano em que foi realizado o estudo(no script de R denominada como year);

Para o estudo da relação das variáveis quantitativas apresentamos:

- Estatística descritiva de ambas as variáveis;
- Testamos a distribuição que parece adequada, justificando;
- Através do teste de independência de Qui-Quadrado, analisamos os resíduos obtidos;
 - As hipóteses testadas;
 - Valor -p;
 - Região crítica;
 - Conclusões da tomada de decisão relativamente às hipóteses testadas
- Aplicamos o teste de independência Qui-Quadrado para verificar a associação entre as variáveis com um grau de significância de 1%
 - As hipóteses testadas;
 - Valor -p;
 - Região crítica;
 - Conclusões da tomada de decisão relativamente às hipóteses testadas

Na conclusão deste processo será possível conhecer as distribuições dos dados, dos resíduos e mais importante tomar decisões sobre as hipóteses testadas no teste de ajustamento e no teste de independência.

Dados fornecidos e tratamento dos dados

O conjunto de dados Korea Income and Welfare apresentava as seguintes características:

1. Representa os dados que caracterizam o rendimento das famílias em determinada área geográfica e em determinado período de tempo;
2. Têm a dimensão de 92857 linhas (observações) e 14 colunas (variáveis aleatórias);
3. É composto pelas seguintes variáveis aleatórias:
 - a. id;
 - b. year;
 - c. wave;
 - d. region;
 - e. income;
 - f. family_member;
 - g. gender;
 - h. year_born;
 - i. education_level;
 - j. marriage;
 - k. religion;
 - l. occupation;
 - m. company_size;
 - n. reason_none_worker

Para este trabalho e apesar de trabalhar com todo o conjunto de dados vamos focar-nos no estudo da distribuição teórica da variável year_born num primeiro exercício . De seguida vamos incidir na análise da relação entre duas variáveis quantitativas são elas:

variável independente -> company_size = X

variável dependente -> year_born = Y

Apesar de se vir a demonstrar que têm uma fraca correlação linear, foi a melhor correlação encontrada de entre as variáveis aleatórias quantitativas constantes no conjunto de dados, para cumprir com os objetivos do trabalho fomos instruídos a perseguir este modelo.

No conjunto de dados Korea Income and Welfare, procedemos à remoção dos valores a Null, tal como, removemos os outliers conforme se pode observar no script de R constante na entrega do projeto.

Apresentamos de seguida a análise estatística descritiva das variáveis de maior foco neste estudo, apenas para conhecimento transversal do conjunto de dados.

Variável year_born - análise estatística descritiva

É uma variável quantitativa discreta, representa o ano do nascimento do indivíduo observado. Por apresentar tantos níveis(90), foi agrupada em classes.

Aplicando se a regra de Sturges consegui-o apurar 17 classes ($K=17$), as classes são fechadas à direita e a amplitude de cada classe é de 5.411765, pois $h= 5.411765$. Apresenta a seguinte tabela de frequências.

Apresenta se na seguinte tabela de frequências.

	classes	ni	fi	Ni	Fi
1	[1910,1915]	37	0.0004	37	0.0004
2	(1915,1921]	261	0.0028	298	0.0032
3	(1921,1926]	1766	0.0190	2064	0.0222
4	(1926,1932]	4647	0.0500	6711	0.0723
5	(1932,1937]	11895	0.1281	18606	0.2004
6	(1937,1942]	12397	0.1335	31003	0.3339
7	(1942,1948]	9058	0.0975	40061	0.4314
8	(1948,1953]	8541	0.0920	48602	0.5234
9	(1953,1959]	7807	0.0841	56409	0.6075
10	(1959,1964]	10657	0.1148	67066	0.7223
11	(1964,1970]	7972	0.0859	75038	0.8081
12	(1970,1975]	8218	0.0885	83256	0.8966
13	(1975,1980]	6225	0.0670	89481	0.9636
14	(1980,1986]	2236	0.0241	91717	0.9877
15	(1986,1991]	873	0.0094	92590	0.9971
16	(1991,1997]	244	0.0026	92834	0.9998
17	(1997,2002]	23	0.0002	92857	1.0000

Tabela 1- Tabela de frequências da variável year_born

No histograma do gráfico 11 podemos visualizar a distribuição das observações desta variável.

Histograma do ano de nascimento

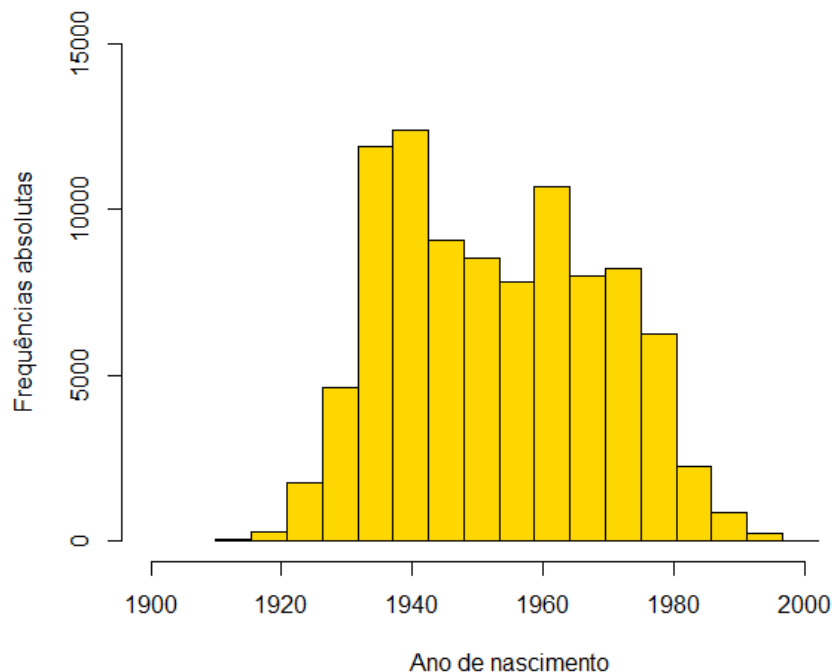


Gráfico 1- Histograma das frequências absolutas da variável year_born

Podemos visualizar como se comportam as medidas de localização e de dispersão respetivamente nas tabelas 18 e 19.

Medidas de Localização							
Moda	1942	Quartis		Decis			
Média	1952.957						
Mediana	1952						
	25%	1939					
	50%	1952					
	75%	1966					
						10%	1933
						20%	1937
						30%	1941
				40%	1946		
				50%	1952		
				60%	1958		
				70%	1963		
				80%	1969		
				90%	1975		

Tabela 2 - Medidas de localização da variável year_born

Medidas de dispersão	
Variância	256.1941
Desvio Padrão	16.00607
Amplitude Total	92
Amplitude Interquartil	27

Tabela 3 - Medidas de dispersão da variável year_born

Na caracterização da distribuição das frequências verificamos o valor da assimetria é de ($b_1 = 0.01768795$). Podemos agora afirmar que a assimetria é positiva pois, $b_1 > 0$.

Relativamente aos valores de curtose, podemos afirmar que a variável apresenta uma curva platicúrtica ou achatada dado que $b_2 < 0$, ($b_2 = -0.9478836$).

Podemos verificar que os quartis da caixa de bigodes tem uma concentração de dados muito uniforme.

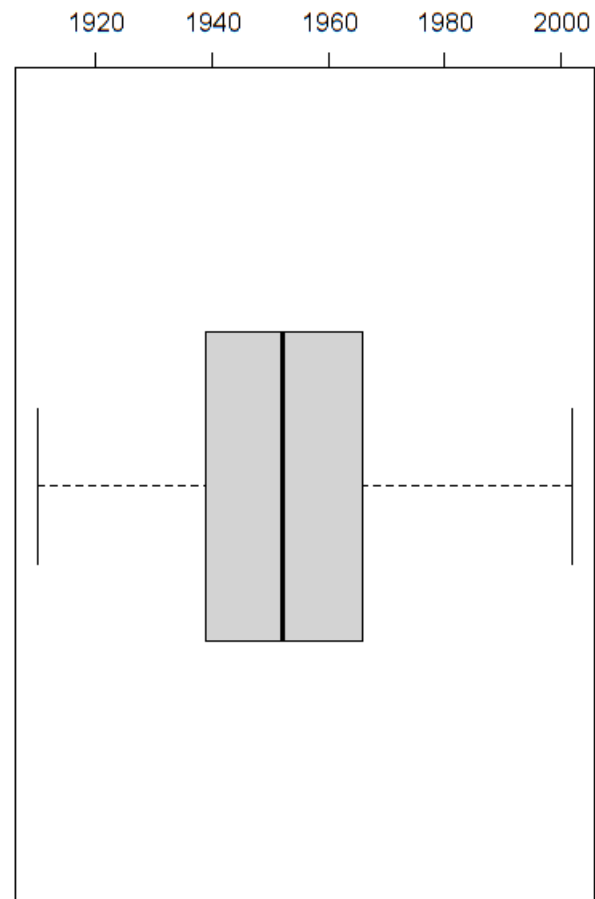


Figura 1- Caixa de bigodes da variável year_born

Analise da distribuição year_born

Através da observação do histograma conseguimos prever que a distribuição da variável é aproximadamente Normal.

O teste de ajustamento do Qui-Quadrado é válido para distribuições discretas e a amostra segue todas as regras recomendadas (a dimensão da amostra deve ser maior que 30, todas as frequências esperadas devem ser maiores ou iguais a 1, não há mais de 20% das frequências esperadas inferiores a 5, para tal foi necessário agrupar dados), por isso faz sentido usar este teste para confirmar ou anular esta previsão.

Teste de ajustamento Qui-Quadrado year_born

Neste teste de ajustamento temos como hipótese nula (H_0) - os dados provêm de uma população que segue uma distribuição Normal e hipótese alternativa (H_1) - os dados provêm de uma população que não segue uma distribuição Normal. Vamos considerar um nível de significância de 10%.

Foram agrupados os dados e registadas as frequências absolutas, calculadas e registadas as frequências esperadas e posteriormente calculado o valor-p e o valor observado da estatística de teste sob a hipótese H_0 .

Visto que o valor observado da estatística de teste sob a hipótese H_0 não está dentro da região crítica e que o valor-p é maior que o nível de significância, não rejeitamos a hipótese H_0 , logo conseguimos concluir que os dados provêm de uma população que segue uma distribuição Normal.

Variável company_size - análise estatística descritiva

É uma variável quantitativa discreta, representa o tamanho de uma companhia através do número de funcionários. O estudo incidiu em empresas com um intervalo do número de funcionários de 1 a 99.

Apresenta a seguinte tabela de frequências.

Licenciatura em Engenharia Informática - Unidade Curricular Métodos Estatísticos
Ano Letivo 2021-2022

i	x_i	n_i	f_i	N_i	F_i
1	1	28319	0.478304930	28319	0.4783049
2	2	5612	0.094786089	33931	0.5730910
3	3	6497	0.109733646	40428	0.6828247
4	4	2669	0.045079129	43097	0.7279038
5	5	1860	0.031415204	44957	0.7593190
6	6	1346	0.022733798	46303	0.7820528
7	7	3478	0.058743054	49781	0.8407959
8	8	1031	0.017413482	50812	0.8582093
9	9	1097	0.018528215	51909	0.8767375
10	10	6905	0.116624723	58814	0.9933623
11	11	393	0.006637729	59207	1.0000000

Tabela 4- Tabela de frequências da variável *company_size*

Podemos visualizar as frequências quer relativas quer absolutas da variável nos gráficos 14 e 15.

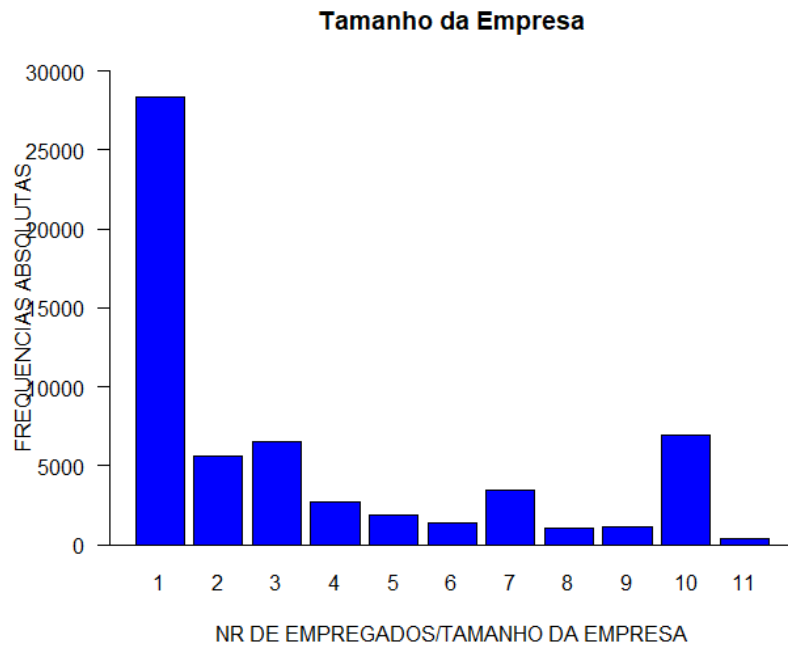


Gráfico 2- Gráfico de barras das frequências absolutas da variável *company_size*

Gráfico Circular: TAMANHO DA EMPRESA

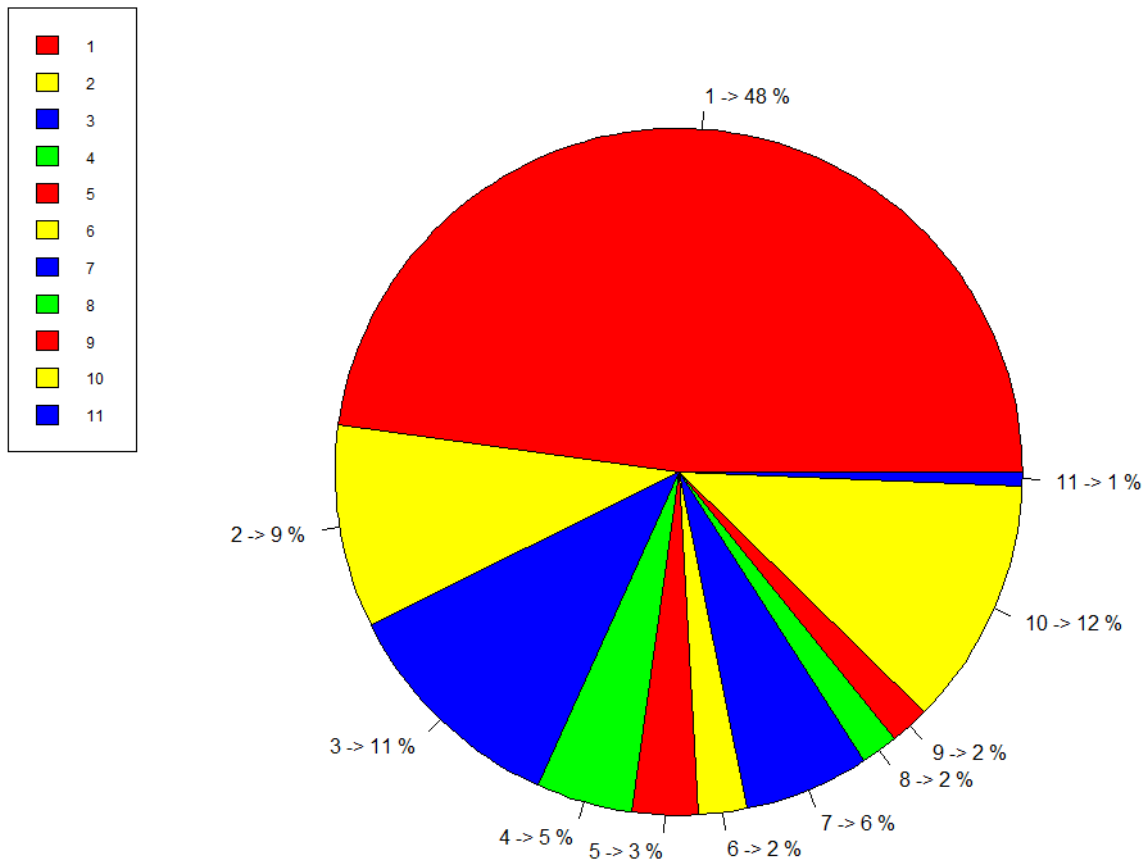


Gráfico 3-Gráfico circular das frequências relativas da variável company_size

Podemos visualizar como se comportam as medidas de dispersão e de localização respetivamente nas tabelas 24 e 25.

Medidas de dispersão	
Variância	10.38252
Desvio Padrão	3.222192
Amplitude Total	10
Amplitude Interquartil	4

Tabela 5- Medidas de dispersão da variável company_size

Medidas de Localização							
Moda	1	Quartis		Decis			
Média	3.427399						
Mediana	2						
		25%	1				
		50%	2				
		75%	5				
						10%	1
						20%	1
						30%	1
				40%	1		
				50%	2		
				60%	3		
				70%	4		
				80%	7		
				90%	10		

Tabela 6 - Medidas de localização da variável *company_size*

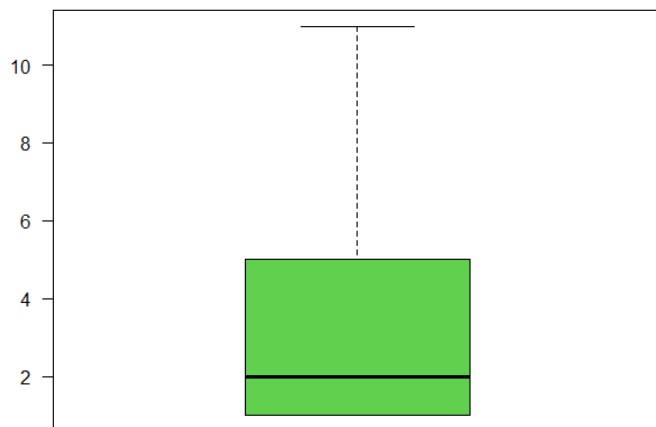


Figura 2-Caixa de bigodes variável *company_size*

Na caracterização da distribuição das frequências verificamos o valor da assimetria é de ($b_1 = 1.104122$). Podemos assim afirmar que a assimetria é positiva pois, $b_1 > 0$.

Relativamente aos valores de curtose verificamos que a variável apresenta uma curva platicúrtica ou achatada dado que $b_2 < 0$, ($b_2 = -0.284578$).

Na figura 6 podemos visualizar a caixa de bigodes da variável *company_size* e verificamos o estudo focou se em empresas de pequena dimensão pois a mediana situa se em empresas com 2 funcionários .

Regressão Linear

Para atingir este objetivo vamos investigar a presença ou ausência de relação linear entre as duas variáveis com todo o conjunto de dados.

Após diversos ensaios, onde foram exploradas as combinações entre 5 variáveis quantitativas tomadas de 2 a 2, para verificar qual seria a melhor combinação para uma correlação linear mais forte, chegou se há conclusão de que as melhores variáveis aleatórias quantitativas em estudo são:

variável independente -> company_size = X

variável dependente -> year_born = Y

Esta relação estuda se o do tamanho da companhia está a ser alterada pelo ano de nascimento dos funcionários.

Diagrama de dispersão

Pela análise do diagrama de dispersão não se vê uma relação linear entre as variáveis, pois não é possível imaginar uma reta nem com declive negativo nem com declive positivo a passar pela nuvem de pontos

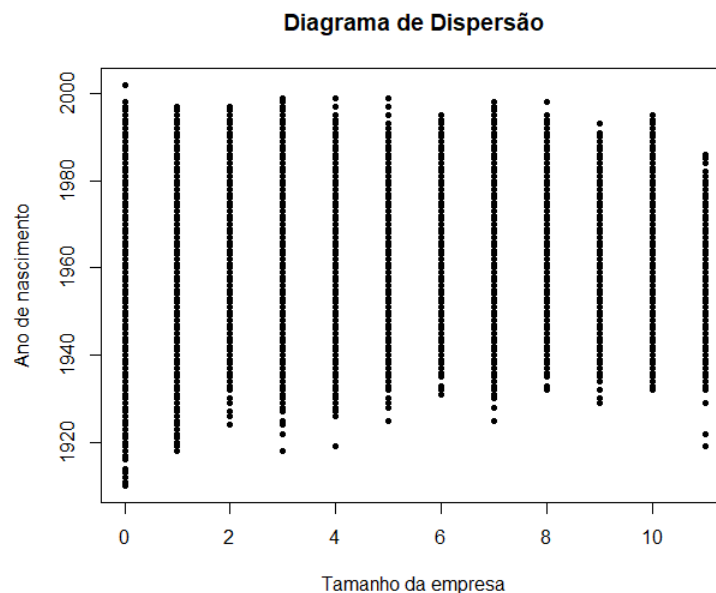


Gráfico 4- Diagrama de dispersão com todo o conjunto de dados

Coeficiente de correlação linear de Pearson

Coeficiente confirma o que vimos no diagrama de dispersão, a correlação linear é muito fraca.

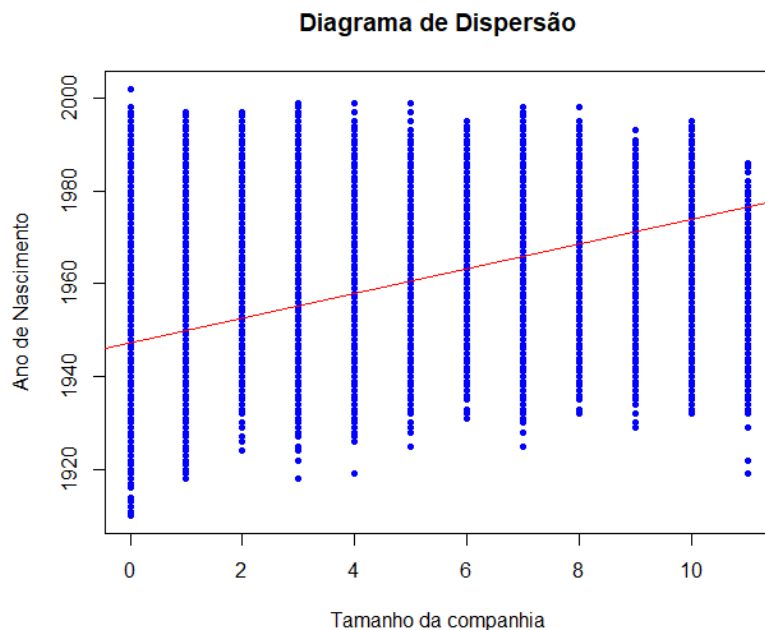
O $r_{xy} = 0.5082532$ não se encontra entre $-1 < r_{xy} < -0.8$ nem entre $0.8 < r_{xy} < 1$, onde poderíamos considerar um coeficiente de correlação linear muito forte.

Sabemos que seria a melhor opção abandonar este modelo, contudo no âmbito deste trabalho vamos mantê-lo e considera-lo como válido para cumprir os objetivos que nos foram propostos.

Reta de regressão linear

Quando a correlação linear é forte, podemos inferir o valor de uma se conhecermos a outra. A reta que atravessa a nuvem de pontos conforme podemos verificar no gráfico **X** divide o diagrama de dispersão em dois grupos idênticos.

A reta de regressão passa pelo ponto cujas coordenadas são, respetivamente, as médias das variáveis em estudo, ou seja, o centro de gravidade da nuvem de pontos (ponto de coordenadas (x, y)).



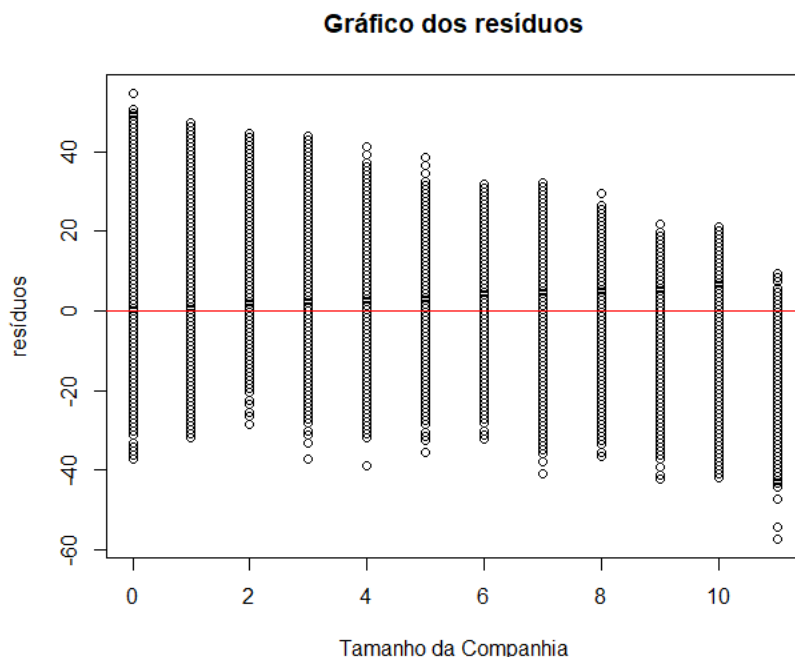
No gráfico 10 podemos visualizar a reta regressão linear que apresenta os valores de interceção para a variável independente, $company_size \Leftrightarrow X=2.663$ e para a variável dependente, $year_born \Leftrightarrow Y=1947.138$.

Gráfico 5- Diagrama de dispersão com a reta da regressão linear

Resíduos

Ao analisarmos os resíduos podemos concluir a qualidade do nosso modelo. Vamos analisar a diferença entre os valores observados e os valores ajustados.

Começamos por analisar o diagrama de dispersão dos resíduos no gráfico 8.



Os resíduos são muito grandes (entre mais 40 e -40), têm um padrão bem definido e constante e isto é sintoma que o modelo ajustado não é bom. Este padrão indica que os resíduos não são independentes.

Gráfico 6 - Gráfico de resíduos

Isto apenas vêm reforçar o que já havíamos concluído anteriormente, que este modelo não têm uma correlação linear forte e apenas não abandonamos e escolhemos outro modelo mais ajustado para correta conclusão dos objetivos propostos neste trabalho.

Ajustamento de Kolmogorov-Smirnov

Relação family_member/year

Estatística descritiva family_member

É uma variável quantitativa discreta, representa o número de elementos do agregado familiar a que foi efetuado o estudo. O estudo incidiu em famílias com 1 elemento até famílias de grandes dimensões com 9 elementos.

Apresenta a seguinte tabela de frequências.

i	x_i	n_i	f_i	N_i	F_i
1	1	25086	2.701573e-01	25086	0.2701573
2	2	28668	3.087328e-01	53754	0.5788901
3	3	16030	1.726310e-01	69784	0.7515212
4	4	16857	1.815372e-01	86641	0.9330584
5	5	4845	5.217700e-02	91486	0.9852354
6	6	1123	1.209386e-02	92609	0.9973292
7	7	211	2.272311e-03	92820	0.9996015
8	8	28	3.015389e-04	92848	0.9999031
9	9	9	9.692323e-05	92857	1.0000000

Tabela 7 - Tabela de frequências da variável family_member

Podemos visualizar as frequências quer relativas quer absolutas da variável nos gráficos 7,8.

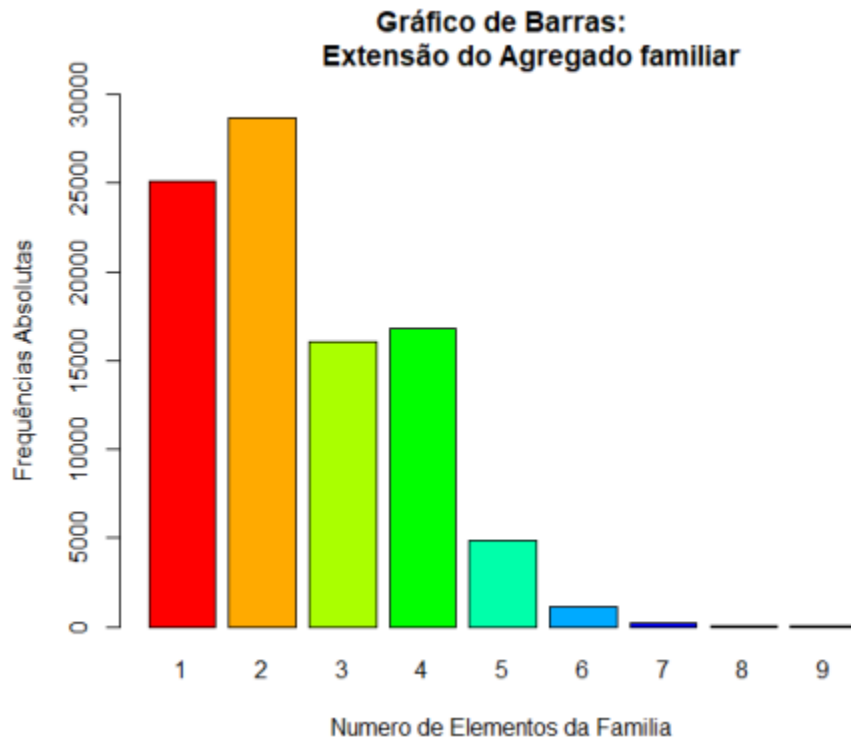


Gráfico 7 - Gráfico de barras das frequências absolutas por extensão de agregado familiar

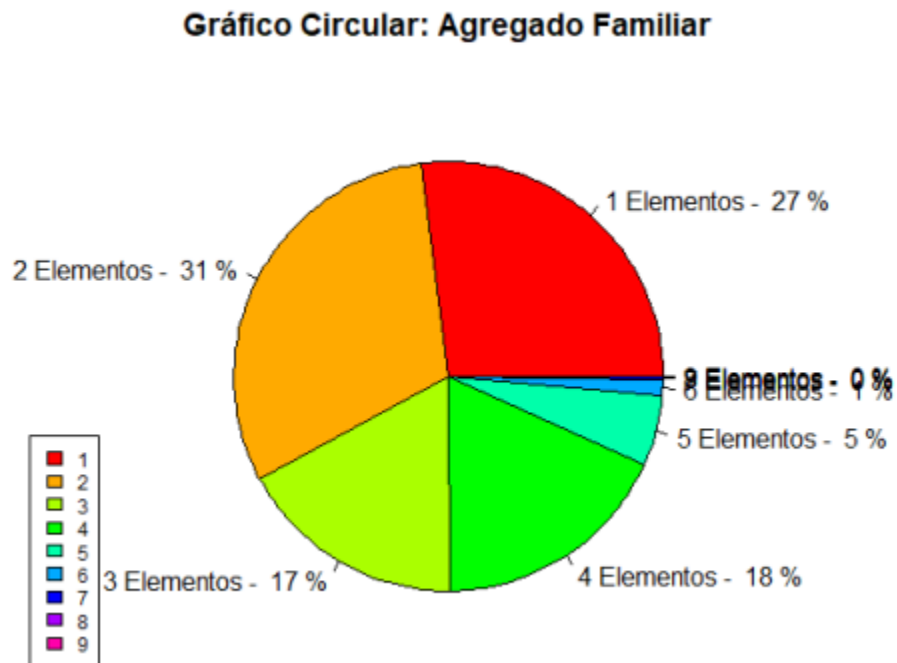


Gráfico 8 - Gráfico circular das frequências relativas por extensão de agregado familiar

Podemos visualizar como se comportam as medidas de localização e de dispersão respetivamente nas tabelas 12 e 13.

Medidas de Localização			
Moda	2	Quartis	
Média	2.484304		
Mediana	2		
		25%	1
		50%	2
		75%	3
		Decis	
		10%	1
		20%	1
		30%	2
		40%	2
		50%	2
		60%	3
		70%	3
		80%	4
		90%	4

Tabela 8 - Medidas de localização da variável *family_member*

Medidas de dispersão	
Variância	1.669912
Desvio Padrão	1.292251
Amplitude Total	8
Amplitude Interquartil	2

Tabela 9 - Medidas de dispersão da variável *family_member*

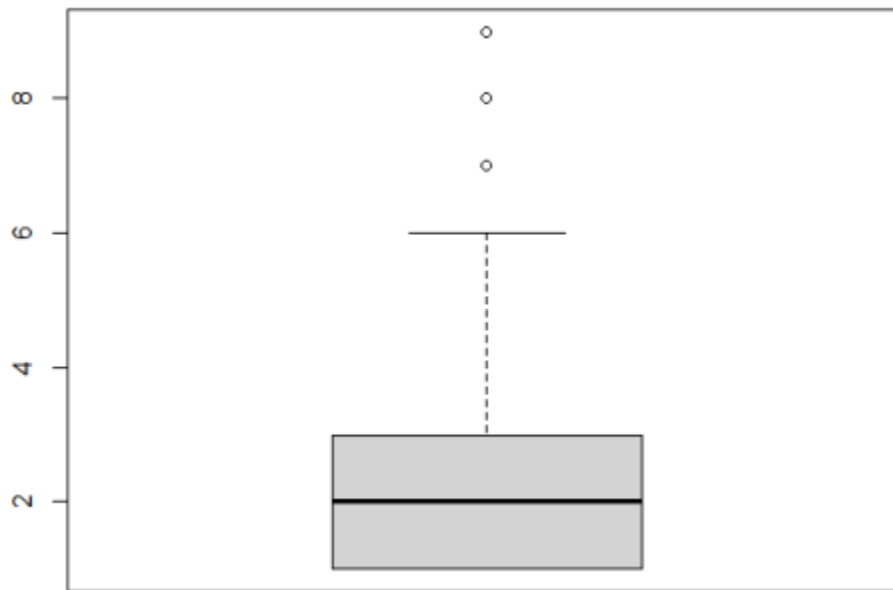


Figura 3 - Caixa de bigodes da variável family_member

Na caixa de bigodes podemos verificar que o estudo abrangeu maioritariamente famílias com 1 e 2 elementos, pois 50% dos dados encontram-se até ao 2ª Quartil.

Mais esporadicamente foram estudadas famílias com 7, 8 e 9 elementos, podemos verificar que são outliers, sem expressão.

Na caracterização da distribuição das frequências verificamos o valor da assimetria é de ($b_1 = 0.6395767$).

Podemos agora afirmar que a assimetria é positiva pois, $b_1 > 0$. Relativamente aos valores de curtose, podemos afirmar que a variável apresenta uma curva platicúrtica ou achatada dado que b_2

Estatística descritiva year

É uma variável quantitativa discreta, representa o ano em que o estudo foi efetuado. E organiza-se na seguinte tabela de frequências.

i	x _i	n _i	f _i	N _i	F _i
1	2005	7072	0.07616012	7072	0.07616012
2	2006	6580	0.07086165	13652	0.14702176
3	2007	6314	0.06799703	19966	0.21501879
4	2008	6207	0.06684472	26173	0.28186351
5	2009	6034	0.06498164	32207	0.34684515
6	2010	5735	0.06176163	37942	0.40860678
7	2011	7532	0.08111397	45474	0.48972075
8	2012	7312	0.07874474	52786	0.56846549
9	2013	7048	0.07590166	59834	0.64436715
10	2014	6914	0.07445858	66748	0.71882572
11	2015	6723	0.07240165	73471	0.79122737
12	2016	6581	0.07087242	80052	0.86209979
13	2017	6474	0.06972011	86526	0.93181990
14	2018	6331	0.06818010	92857	1.00000000

Tabela 10 - Tabela de frequências da variável year

Podemos visualizar as frequências quer relativas quer absolutas da variável nos gráficos 1 e 2.

Gráfico Circular: Ano do Estudo

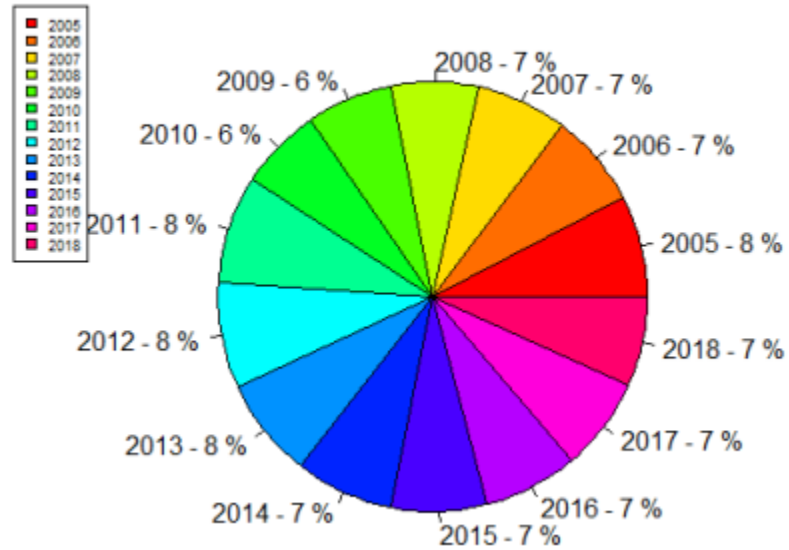


Gráfico 9 - Gráfico circular das frequências relativas do ano de estudo

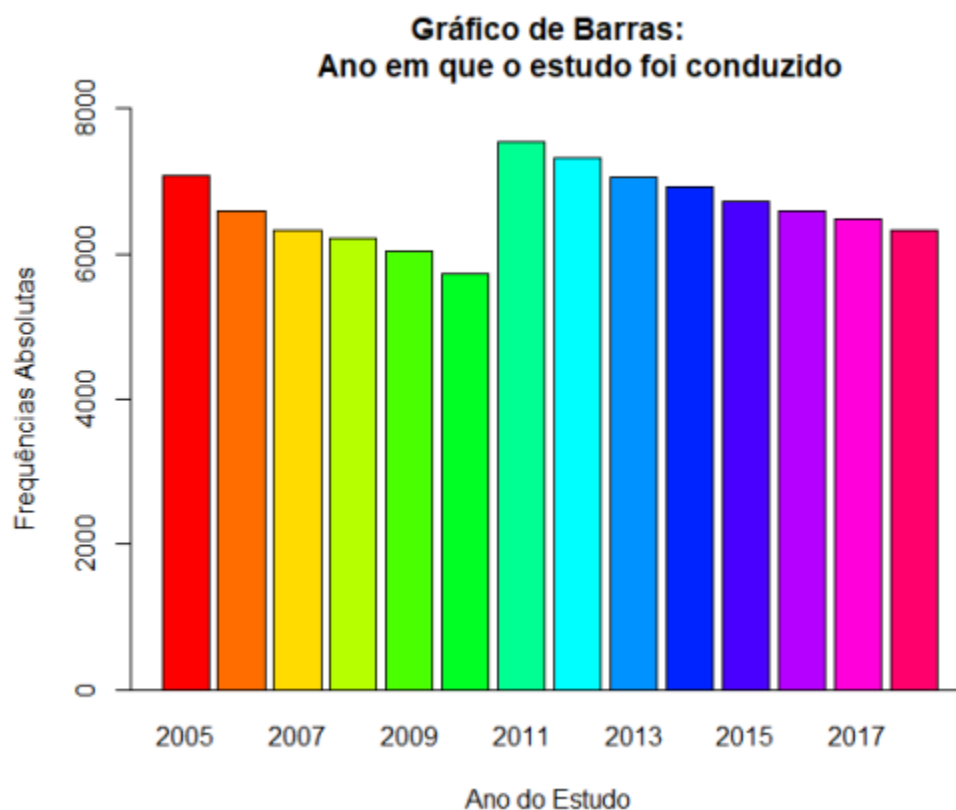


Gráfico 10 - Gráfico de barras das frequências absolutas do ano de estudo

Podemos visualizar como se comportam as medidas de dispersão e de localização respetivamente nas tabelas 3 e 4.

Medidas de dispersão	
Variância	16.08991
Desvio Padrão	4.011222
Amplitude Total	13
Amplitude Interquartil	7

Tabela 11 - Medidas de dispersão da variável year

Medidas de Localização							
Moda	2011	Quartis		Decis			
Média	2011.518						
Mediana	2012						
		25%	2008				
		50%	2012				
		75%	2015				
						10%	2006
						20%	2007
						30%	2009
				40%	2010		
				50%	2012		
				60%	2013		
				70%	2014		
				80%	2016		
				90%	2017		

Tabela 12 - Medidas de localização da variável year

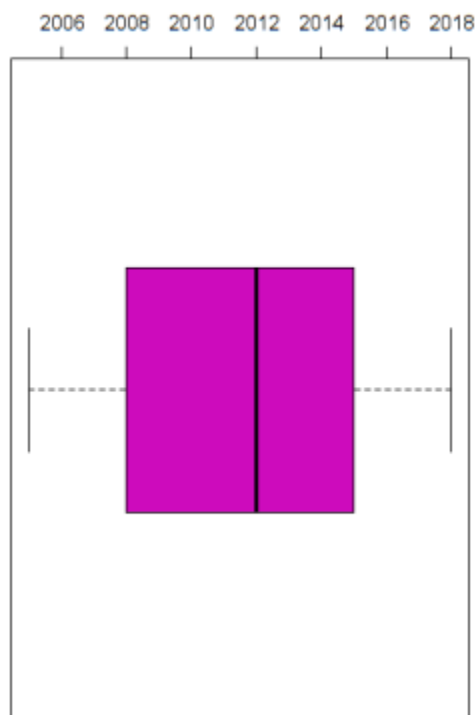


Figura 4 - Caixa de bigodes da variável year

Na caracterização da distribuição das frequências verificamos o valor da assimetria é de ($b_1 = -0.03954851$). Podemos assim afirmar que a assimetria é negativa pois, $b_1 < 0$.

Relativamente aos valores de curtose verificamos que a variável apresenta uma curva platicúrtica ou achatada dado que $b_2 < 0$, ($b_2 = -1.182807$).

Na figura 1 podemos visualizar a caixa de bigodes da variável year.

Distribuição family_member/year

Teste de Independência Qui-Quadrado

Conclusões

Neste projeto abordamos o conjunto de dados intitulado '*Korea Income and Welfare*', com o objetivo de estudar as distribuições teóricas dos dados selecionados, aplicar testes de ajustamento tal como testes de independência.

Com o auxílio do RStudio e através da linguagem de R produzimos um script que nos permitiu analisar os dados.

Apenas cumprimos um dos objetivos a que nos tínhamos proposto, objetivo esse que consistia em testar a distribuição da população á qual pertence a amostra da variável ano de nascimento a que nos pareceu mais adequada, considerando um nível de significância de 10%. Infelizmente não conseguimos concluir os outros dois objetivos.

Este projeto teve uma importância valiosa na aquisição de conhecimentos na esfera dos testes de ajustamento e independência e da linguagem de R, pois obrigou todos os elementos deste grupo a pesquisar e analisar e aperfeiçoar técnicas fundamentais nesta área.

Licenciatura em Engenharia Informática - Unidade Curricular Métodos Estatísticos
Ano Letivo 2021-2022

Referências bibliográficas

- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 1 – Estatística Descritiva. 2021-2022. Materiais de apoio. Disponível em: <<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em: 10/04/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 2 – Regressão Linear Simples. 2021-2022. Materiais de apoio. Disponível em: <<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em: 07/05/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 3 – Distribuições Teóricas. 2021-2022. Materiais de apoio. Disponível em: <<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em: 07/05/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 4 – Elementos da Teoria da Amostragem. 2021-2022. Materiais de apoio. Disponível em: <<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em: 07/05/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 5 – Elementos da Teoria da Estimação. 2021-2022. Materiais de apoio. Disponível em: <<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em: 07/05/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 6 – Testes de Hipótese Paramétricos. 2021-2022. Materiais de apoio. Disponível em: <<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em: 07/05/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 7 – Testes de Hipótese Não Paramétricos. 2021-2022. Materiais de apoio. Disponível em: <<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em: 07/05/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 7 – Testes de Hipótese Não Paramétricos-Parte 1 -Testes de Ajustamento. 2021-2022. Materiais de apoio. Disponível em:

<<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em:
07/05/2022

- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 7 – Testes de Hipótese Não Paramétricos- Parte 2 -Testes de Independência. 2021-2022. Materiais de apoio. Disponível em:
<<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em:
07/05/2022
- Departamento de Matemática Escola Superior de Tecnologia de Setúbal. Capítulo 7 – Testes de Hipótese Não Paramétricos- Parte 3 -Testes à igualdade de duas distribuições. 2021-2022. Materiais de apoio. Disponível em:
<<https://moodle.ips.pt/2122/mod/resource/view.php?id=3386>>. Acesso em:
07/05/2022