Universidade do Minho
Escola de Engenharia
Dep. de Produção e Sistemas

# ESTATÍSTICA APLICADA

2018/2019

Profª Ana Cristina Braga

12-09-2018

# ÍNDICE


FICHA Nº 1 - DESCRITIVA....................................................................3

FICHA Nº 2 - PROBABILIDADES.............................................................7

FICHA Nº 3 – DISTRIBUIÇÕES DE PROBABILIDADE ...............................9

FICHA Nº 4 – ESPERANÇA MATEMÁTICA..............................................12

FICHA Nº 5 – FAMÍLIAS DE DISTRIBUIÇÕES ........................................13

FICHA Nº 6 - ESTIMADORES PONTUAIS..............................................17

FICHA Nº 7 - DISTRIBUIÇÕES AMOSTRAIS..........................................18

FICHA Nº 8 - INTERVALOS DE CONFIANÇA ..........................................19

FICHA Nº 9 - TESTES DE HIPÓTESES...................................................22

FICHA Nº 10 - ANÁLISE DA VARIÂNCIA...............................................27

FICHA Nº 11 – QUI-QUADRADO – "BOM AJUSTE" ................................29

FICHA Nº 12 – REGRESSÃO E CORRELAÇÃO (SPSS) .............................31

# FICHA Nº 1 - DESCRITIVA

**1.** Pediu-se a 36 pessoas para classificarem o Sistema de Saúde em Portugal de acordo com a seguinte escala: 1 (péssimo), 2 (mau), 3 (pouco razoável), 4 (razoável), 5 (muito razoável), 6 (bom), 7 (muito bom), 8 (excelente). As classificações obtidas foram:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 2 | 7 | 6 | 3 | 7 | 8 | 3 | 2 | 6 | 3 | 6 |
| 3 | 7 | 5 | 3 | 6 | 7 | 3 | 7 | 6 | 4 | 3 | 5 |
| 8 | 6 | 5 | 4 | 3 | 6 | 6 | 5 | 7 | 8 | 4 | 3 |

A organização dos dados conduziu à seguinte tabela:

| $x_i$ | $f_i$ | $f_{ri}(\%)$ | $F_i$ | $F_{ri}\ (\%)$ |
|---|---|---|---|---|
| 2 | 2 | 5,6 | 2 | 5,6 |
| 3 | 9 | 25,0 | 11 | 30,6 |
| 4 | 3 | 8,3 | 14 | 38,9 |
| 5 | 5 | 13,9 | 19 | 52,8 |
| 6 | 8 | 22,2 | 27 | 75,0 |
| 7 | 6 | 16,7 | 33 | 91,7 |
| 8 | 3 | 8,3 | 36 | 100,0 |
| Total | 36 | 100,0 | | |

**a)** Identifique o tipo de variável apresentada.
**b)** Calcule a média, o desvio padrão, a mediana e a moda;
**c)** Desenhe o gráfico mais adequado de frequências absolutas;
**d)** Será a distribuição de frequências unimodal? Justifique.
**e)** Que pode concluir sobre a distribuição de opiniões?
**f)** Calcule a percentagem de pessoas que têm opinião:
 i) Desfavorável;
 ii) Favorável.

**2.** Numa empresa, a fabricação de peças é feita em série. Retirou-se uma amostra aleatória simples de 45 lotes, cada um com 50 peças e, registou-se o número de peças defeituosas em cada lote, tendo-se obtido os seguintes resultados:

| $x_i$ | $f_i$ | $f_{ri}(\%)$ | $F_i$ | $F_{ri}(\%)$ |
|---|---|---|---|---|
| 1 | 2 | 4,4 | 2 | 4,4 |
| 2 | 5 | 11,1 | 7 | 15,6 |
| 3 | 5 | 11,1 | 12 | 26,7 |
| 4 | 10 | 22,2 | 22 | 48,9 |
| 5 | 9 | 20,0 | 31 | 68,9 |
| 6 | 5 | 11,1 | 36 | 80,0 |
| 7 | 3 | 6,7 | 39 | 86,7 |
| 8 | 3 | 6,7 | 42 | 93,3 |
| 9 | 2 | 4,4 | 44 | 97,8 |
| 10 | 1 | 2,2 | 45 | 100,0 |
| Total | 45 | 100,0 | | |

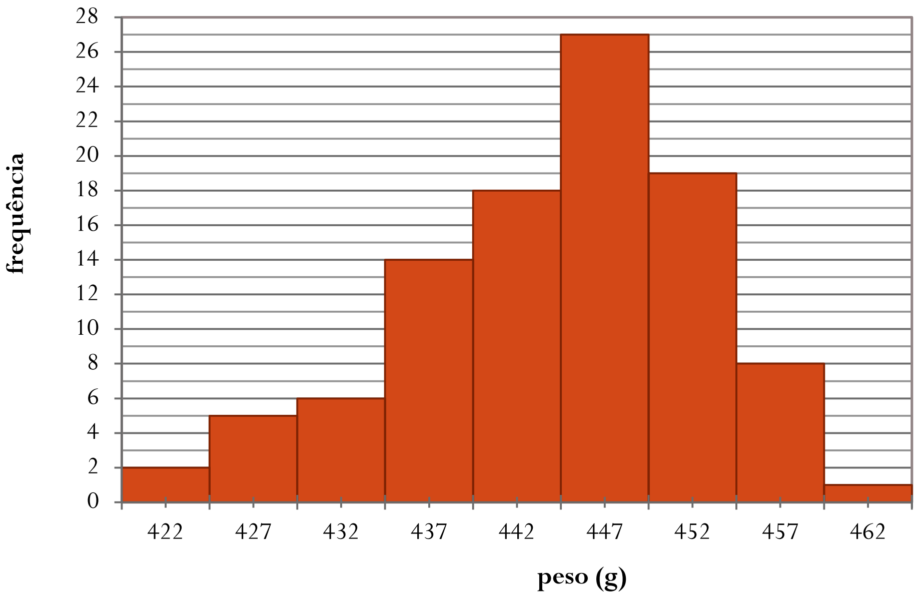
**a)** Identifique o tipo de variável apresentada;
**b)** Construa o gráfico de frequências absolutas;
**c)** Calcule a média, a variância, a mediana e a moda da amostra.

**3.** O vencimento/hora de 100 operários é dado pela tabela

| Vencimento | Nº operários |
|---|---|
| [120 ;125[ | 10 |
| [125 ;130[ | 20 |
| [130 ;135[ | 38 |
| [135 ;140[ | 25 |
| [140 ;145[ | 7 |

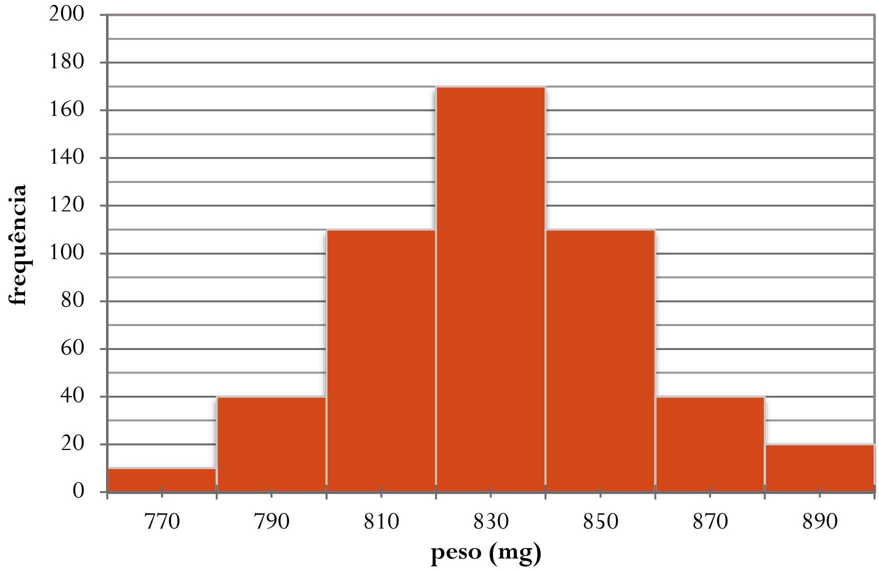
**a)** Construa o histograma de frequências absolutas;
**b)** Calcule o vencimento/hora médio, o desvio padrão, a mediana e a moda;
**c)** Determine o n.º de operários com vencimentos compreendidos entre:
(i) $\overline{x}-s$ e $\overline{x}+s$
(ii) $\overline{x}-2s$ e $\overline{x}+2s$

**4.** Considere-se uma amostra constituída por 100 latas de leite em pó, cujo rótulo indica um peso médio de 450 gramas




**a)** Apresente a tabela das frequências relativas;
**b)** Calcule a média da amostra e o desvio padrão das observações;
**c)** Qual a proporção de latas com peso superior a 450 g?
**d)** O que pode concluir sobre a forma da distribuição?

**5.** A distribuição dos pesos (mg) de 500 cigarros de uma determinada marca está representada no seguinte histograma




**a)** Calcule a média, a mediana, a moda e a variância da amostra;
**b)** Qual a proporção de cigarros com menos de 820 mg?
**c)** Qual a proporção de cigarros cujo peso está compreendido entre 800 e 880 mg?

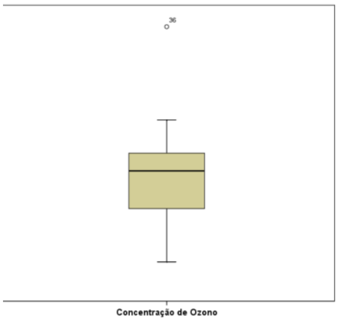
**6.** Considere os dados referentes a doentes com cancro do colon retal e respetivo estadio da doença **(opcional)**

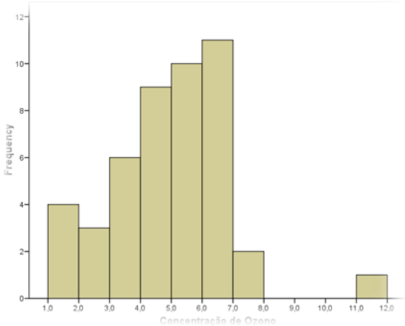
| Doente | Idade | estadio | Doente | Idade | estadio | Doente | Idade | estadio | Doente | Idade | estadio | Doente | Idade | estadio | Doente | Idade | estadio |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 51 | 3 | 21 | 71 | 2 | 41 | 60 | 3 | 61 | 83 | 2 | 81 | 70 | 2 | 101 | 69 | 3 |
| 2 | 56 | 4 | 22 | 83 | 3 | 42 | 60 | 2 | 62 | 63 | 2 | 82 | 73 | 2 | 102 | 70 | 4 |
| 3 | 81 | 3 | 23 | 58 | 4 | 43 | 61 | 2 | 63 | 65 | 3 | 83 | 82 | 3 | 103 | 71 | 2 |
| 4 | 64 | 3 | 24 | 70 | 2 | 44 | 63 | 4 | 64 | 75 | 3 | 84 | 87 | 3 | 104 | 73 | 4 |
| 5 | 82 | * | 25 | 83 | 2 | 45 | 67 | 2 | 65 | 89 | * | 85 | 86 | 3 | 105 | 61 | 3 |
| 6 | 88 | 2 | 26 | 88 | 3 | 46 | 67 | 3 | 66 | 71 | 2 | 86 | 51 | 3 | 106 | 65 | 3 |
| 7 | 58 | 1 | 27 | 73 | 3 | 47 | 70 | 3 | 67 | 43 | 3 | 87 | 58 | 3 | 107 | 74 | 3 |
| 8 | 56 | 3 | 28 | 79 | 2 | 48 | 71 | 3 | 68 | 73 | 4 | 88 | 75 | 4 | 108 | 65 | 3 |
| 9 | 61 | 1 | 29 | 62 | 3 | 49 | 76 | 3 | 69 | 28 | 3 | 89 | 47 | 4 | 109 | 70 | 2 |
| 10 | 64 | * | 30 | 73 | 2 | 50 | 91 | 2 | 70 | 47 | 3 | 90 | 81 | 2 | 110 | 63 | 3 |
| 11 | 68 | 3 | 31 | 75 | 2 | 51 | 73 | 3 | 71 | 50 | 4 | 91 | 73 | 3 | 111 | 77 | 3 |
| 12 | 45 | 2 | 32 | 90 | 3 | 52 | 77 | 2 | 72 | 71 | 3 | 92 | 63 | 3 | 112 | 86 | 3 |
| 13 | 70 | 2 | 33 | 76 | 4 | 53 | 84 | 2 | 73 | 71 | 3 | 93 | 63 | 3 | 113 | 90 | 3 |
| 14 | 58 | 2 | 34 | 85 | 4 | 54 | 73 | 3 | 74 | 72 | 3 | 94 | 80 | 3 | 114 | 58 | 3 |
| 15 | 54 | 3 | 35 | 53 | 3 | 55 | 80 | 3 | 75 | 79 | 4 | 95 | 84 | 4 | 115 | 80 | 3 |
| 16 | 52 | 2 | 36 | 72 | 3 | 56 | 85 | 3 | 76 | 93 | 3 | 96 | 73 | 3 | 116 | 75 | 3 |
| 17 | 60 | 3 | 37 | 72 | 4 | 57 | 76 | 4 | 77 | 76 | 1 | 97 | 80 | 4 | 117 | 61 | 3 |
| 18 | 63 | 3 | 38 | 77 | 4 | 58 | 75 | 4 | 78 | 82 | 3 | 98 | 58 | 4 | 118 | 82 | 4 |
| 19 | 74 | 3 | 39 | 80 | 3 | 59 | 77 | 3 | 79 | 83 | 3 | 99 | 59 | 3 | | | |
| 20 | 73 | 2 | 40 | 46 | 2 | 60 | 65 | 2 | 80 | 64 | 4 | 100 | 59 | 3 | | | |

**a)** Introduza os dados em EXCEL
**b)** Classifique as variáveis idade e estadio quanto ao tipo e à escala de medida.
**c)** Utilizando o SPSS efetue a análise descritiva de acordo com a natureza das variáveis em estudo.
**d)** Utilizando o comando *Explore* avalie qual a idade média, mediana, desvio padrão, máxima e mínima dos pacientes em cada estadio da doença e produza os gráficos de caixa de bigodes para a idade em cada estadio.

**e)** Comente os resultados obtidos

**7.** Um dos melhores indicadores da poluição atmosférica nas grandes cidades é a concentração de ozono no ar. Durante o Verão de 2006 foram registadas 46 medições desta concentração numa certa cidade obtendo-se os resultados apresentados na tabela e gráfico seguintes:




**Statistics**

Concentração de Ozono

| | | |
|---|---|---|
| N | Valid | 46 |
| | Missing | 0 |
| Mean | | 4,878 |
| Median | | 5,200 |
| Mode | | 4,7[a] |
| Std. Deviation | | 1,9528 |
| Variance | | 3,813 |
| Range | | 10,6 |
| Minimum | | 1,1 |
| Maximum | | 11,7 |
| Percentiles | 10 | 2,190 |
| | 25 | 3,500 |
| | 75 | 6,050 |
| | 90 | 6,800 |

a. Multiple modes exist. The smallest value is shown

**a)** Classifique a variável (tipo e escala) e comente o quadro das estatísticas resumo.

**b)** Comente as representações gráficas (a) e (b). Que pode afirmar quanto à simetria da distribuição? Justifique.

**c)** Qual a proporção de medições que apresentam uma concentração de ozono inferior a 4,0?

# FICHA Nº 2 - PROBABILIDADES

**1.** Tira-se uma carta de um baralho de 52. Considere os acontecimentos relativos à experiência: $A_1$= {extração de um ás do baralho}, $A_2$= {extração de uma carta de espadas do baralho}

**a)** A1 e A2 são independentes?
**b)** A1 e A2 são mutuamente exclusivos?
**c)** Calcule a probabilidade da extração de um ás ou de uma carta de espadas.

**2.** No lançamento de um dado determine a probabilidade de que saia:

**a)** Face par, ou número primo
**b)** Face par e múltiplo de 3

**3.** Considere a experiência aleatória que consiste no lançamento simultâneo de dois dados perfeitos e os seguintes acontecimentos
A= "soma dos resultados igual a 7"
B= "ambos os resultados ímpares"
C= "produto dos resultados igual a 12"
Determine: a) $P(A \cup C)$, b) $P(A \cup B)$

**4.** Dois indivíduos *A* e *B* estão afetados por uma doença incurável. Atendendo ao estado de evolução da doença em cada um dos indivíduos estimaram-se as probabilidades de falecimento, ao fim de cinco anos, respetivamente em $P(A) = \frac{2}{3}$ e $P(B) = \frac{1}{2}$. Calcule a probabilidade ao fim de cinco anos,

**a)** apenas A tenha falecido;
**b)** apenas B tenha falecido;
**c)** ambos tenham falecido;
**d)** pelo menos um tenha falecido

**5.** Durante a travessia do Canal da Mancha, um velejador tem 2/3 de probabilidades de ser atingido pelo mau tempo, e, independentemente disso, 1/4 de probabilidades de ter uma colisão com um petroleiro. Definindo os seguintes acontecimentos: M={ser atingido pelo mau tempo} e C={colisão com um petroleiro}, calcule

**a)** $P(M \cap C)$, **b)** $P(M \cap \overline{C})$, **c)** $P(\overline{M} \cap C)$, **d)** $P(M \cup C)$

**6.** Seja A e B acontecimentos independentes e P(A)=1/6 e P(B)=1/4. Determine

**a)** $P(A \cap B)$, **b)** $P(A \cup B)$, **c)** $P(\overline{A} \cap \overline{B})$, **d)** $P(A \cap \overline{B})$

**7.** Sejam A e B dois acontecimentos tais que P(A)=1/4, P(B)=2/3 e $P(A \cap B) = 1/6$. Determine

**a)** $P(A \cup B)$, **b)** $P(\overline{A})$, **c)** $P(\overline{B})$, **d)** $P(A | B)$, **e)** $P(B | A)$, **f)** $P(\overline{A} \cap \overline{B})$, **g)** $P(\overline{A} | \overline{B})$,

**h)** $P(\overline{B} | \overline{A})$

**8.** Sejam $M_1$, $M_2$ acontecimentos independentes, tais que $P(M_1 \cup M_2) = 0.8$ e $P(M_1 | M_2) = 0.2$. Calcule $P(M_2)$.

**9.** Lançaram-se alguns dados ao ar e sabe-se que a soma de pontos obtidos foi igual a 2. Qual a probabilidade de que se tivessem lançado 2 dados?

**10.** Tira-se uma carta de um baralho de 52. Sabendo que a carta extraída é de espadas, qual a probabilidade de ser um ás? (ver exercício 1)

**11.** Um acidente pode ser devido a falha humana, falha de travões ou rebentamento de pneu, sendo a 1ª causa 2 vezes mais provável do que cada uma das outras.

**a)** Determine a probabilidade de um acidente se dever a cada uma destas causas.

**b)** A probabilidade de que um acidente seja corretamente atribuído a falha humana é de 80% e erradamente atribuído a essa causa é de 4%. Calcule e probabilidade de que um acidente atribuído a falha humana tenha tido essa causa.

**12.** Num hospital ingressam 50% de indivíduos com a doença K, 30% com a doença L e 20% com a doença M. A probabilidade de cura da doença K é 0.7; para as doenças L e M, a probabilidade é de respetivamente 0.8 e 0.9. A um doente internado foi dada alta. Calcule a probabilidade de que esse indivíduo tenha sofrido da doença K.

**13.** Num laboratório um investigador fez uma preparação com 3 classes de bactérias A, B e C, na proporção de 10%, 30% e 60% de cada classe, respetivamente. As bactérias da classe A reagem com sulfato em 80% dos casos, as da classe B em 60% e as da classe C em 40%.

**a)** Qual a probabilidade de uma bactéria escolhida ao acaso da preparação reaja com sulfato?

**b)** O investigador colheu uma bactéria da preparação e ela reagiu com o sulfato. Concluiu então que ela pertencia à classe C. Concorda com o investigador?

# FICHA Nº 3 – DISTRIBUIÇÕES DE PROBABILIDADE

**1.** Determinar se os valores dados podem ser usados como os valores de uma função de probabilidade de uma variável aleatória na gama de valores de x=1,2,3,4.

**a)** f(1)=0.25 f(2)=0.75 f(3)=0.25 f(4)=-0.25

**b)** f(1)=0.15 f(2)=0.27 f(3)=0.29 f(4)=0.29

**c)** f(1)=1/19 f(2)=10/19 f(3)=2/19 f(4)=5/19

**2.** Determine se as funções dadas podem servir como funções de probabilidade na gama de valores dada.

**a)** $f(x)=\frac{x-2}{5} \qquad x=1,2,3,4,5$

**b)** $f(x)=\frac{x^2}{30} \qquad x=0,1,2,4$

**c)** $f(x)=\frac{1}{5} \qquad x=0,1,2,3,4,5$

**d)** $f(x)=c\left(\frac{1}{4}\right)^x \qquad x=1,2,3,\ldots$

**3.** Determinar se os valores dados podem ser usados como os valores de uma função de distribuição acumulada de uma variável aleatória na gama de valores de x=1,2,3,4.

**a)** F(1)=0.3 F(2)=0.5 F(3)=0.8 F(4)=1.2

**b)** F(1)=0.5 F(2)=0.4 F(3)=0.7 F(4)=1.0

**c)** F(1)=0.25 F(2)=0.65 F(3)=0.83 F(4)=1.0

**4.** Encontre a função de distribuição acumulada da variável aleatória que tem a distribuição de probabilidade

$$f(x)=\frac{x}{15} \qquad x=1,2,3,4,5$$

Apresente o respetivo gráfico.

**5.** A variável aleatória X tem a função de distribuição acumulada

$$F(x)=\begin{cases} 0 & x<1 \\ 1/3 & 1\le x<4 \\ 1/2 & 4\le x<6 \\ 5/6 & 6\le x<10 \\ 1 & x\ge 10 \end{cases}$$

Apresente o respetivo gráfico. Calcule

**a)** P(2<X≤6);

**b)** P(X=4);

**c)** f(x).

**6.** A variável aleatória X tem a função de distribuição acumulada

$$F(x) = \begin{cases} 0 & x < -1 \\ 1/4 & -1 \le x < 1 \\ 1/2 & 1 \le x < 3 \\ 3/4 & 3 \le x < 5 \\ 1 & x \ge 5 \end{cases}$$

Apresente o respetivo gráfico. Calcule: P(X≤3); P(X=3); P(X<3); P(-0.4<X<4); P(X=5); f(x).

**7.** Determine *c* para que a função possa servir como f.p. na gama de valores dada.

**a)** $f(x) = cx \qquad x = 1, 2, 3, 4, 5$

**b)** $f(x) = cx^2 \qquad x = 1, 2, 3, \ldots, k$

**8.** A função densidade de probabilidade da variável aleatória contínua X é dada por

$$f(x) = \begin{cases} \frac{1}{5} & 2 < x < 7 \\ 0 & \text{outros valores} \end{cases}$$

**a)** Desenhe o gráfico e verifique que a área total debaixo da curva é igual a 1.

**b)** Calcule P(3<X<7).

**c)** Determine a função de distribuição acumulada e use-a para calcular a alínea b).

**9.** A função densidade de probabilidade da variável aleatória contínua X é dada por

$$f(x) = \begin{cases} \frac{1}{8}(x+1) & 2 < x < 4 \\ 0 & \text{outros valores} \end{cases}$$

**a)** Calcule P(X<3.2) e P(2.9<X<3.2).

**b)** Determine a função de distribuição acumulada e use-a para calcular a alínea a).

**10.** A função densidade de probabilidade da variável aleatória contínua X é dada por

$$f(x) = \begin{cases} \frac{c}{\sqrt{x}} & 0 < x < 4 \\ 0 & \text{outros valores} \end{cases}$$

**a)** Determine o valor de c.

**b)** Calcule P(X<1/4) e P(X>1).

**c)** Determine a função de distribuição acumulada e use-a para calcular a alínea b).

**11.** A função densidade de probabilidade da variável aleatória contínua X é dada por

$$f(x) = \begin{cases} kxe^{-x^2} & x > 0 \\ 0 & \text{outros valores} \end{cases}$$

Encontre o valor de k e desenhe o gráfico da função.

**12.** A função densidade de probabilidade da variável aleatória contínua X é dada por

$$f(x) = \begin{cases} kx(1-x) & 0 < x < 1 \\ 0 & \text{outros valores} \end{cases}$$

**a)** Calcule k.

**b)** Calcule P(X<1/4) e P(X>1/2)

**c)** Determine a função de distribuição acumulada e use-a para calcular a alínea b).

**13.** A função de distribuição acumulada da variável aleatória contínua X é dada por

$$F(x) = \begin{cases} 0 & x < -1 \\ (x+1)/2 & -1 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

Calcule P(-1/2<X<1/2) e P(2<X<3). Determine a função densidade de probabilidade e use-a para calcular as duas probabilidades.

**14.** A função de distribuição acumulada da variável aleatória contínua X é dada por

$$F(x) = \begin{cases} 1 - \dfrac{9}{x^2} & x > 3 \\ 0 & \text{outros valores} \end{cases}$$

Calcule P(X≤5) e P(X>8). Determine a função densidade de probabilidade e use-a para calcular as duas probabilidades. Desenhe os gráficos das funções de densidade de probabilidade e de distribuição acumulada.

# FICHA Nº 4 – ESPERANÇA MATEMÁTICA

**1.** Encontre o valor esperado e a variância da variável aleatória discreta X com a seguinte distribuição de probabilidade

$$f(x) = \frac{|x-2|}{7} \qquad x = -1, 0, 1, 3$$

**2.** Encontre o valor esperado e a variância da variável aleatória X com a seguinte a seguinte função densidade de probabilidade

$$f(x) = \begin{cases} \frac{1}{8}(x+1) & 2 < x < 4 \\ 0 & \text{outros valores} \end{cases}$$

**3.** Encontre o valor esperado e a variância da variável aleatória X com a seguinte a seguinte função densidade de probabilidade

$$f(x) = \begin{cases} x & 0 < x < 1 \\ 2 - x & 1 \le x < 2 \\ 0 & \text{outros valores} \end{cases}$$

**4.** Considere a seguinte função de probabilidade

$$f(x) = \frac{x}{15} \qquad x = 1, 2, 3, 4, 5$$

Encontre:

**a)** $E[X]$, $E\left[X^2\right]$ e $V[X]$

**b)** $E\left[(3X+2)^2\right]$

**5.** Considere a seguinte função densidade de probabilidade

$$f(x) = \begin{cases} \frac{1}{x(\ln 3)} & 1 < x < 3 \\ 0 & \text{outros valores} \end{cases}$$

Encontre:

**a)** $E[X]$, $E\left[X^2\right]$, $E\left[X^3\right]$ e $V[X]$;

**b)** $E\left[X^3 + 2X^2 - 3X + 1\right]$.

# FICHA Nº 5 – FAMÍLIAS DE DISTRIBUIÇÕES

## *DISTRIBUIÇÕES DE PROBABILIDADE DISCRETAS*

**1.** Dez por cento da população tem sangue do tipo B. Numa amostra aleatória de 20 pessoas encontre a probabilidade de encontrar com o tipo B:
   - **a)** Exatamente três pessoas?
   - **b)** Mais de cinco pessoas?
   - **c)** Menos de duas pessoas?

**2.** Para admissão a um concurso para uma vaga de secretária exige-se uma prova de conhecimentos que consiste em 16 questões. Cada questão tem cinco escolhas, uma correta e quatro erradas. Uma das candidatas questiona-se acerca das probabilidades se responder à sorte a cada uma das questões colocadas.
   - **a)** Qual a probabilidade de obter três respostas corretas?
   - **b)** Qual a probabilidade de obter duas ou mais questões corretas?
   - **c)** Se 50 candidatas fizessem o exame e se todas respondessem à sorte, qual seria a média de respostas certas?

**3.** Quando uma determinada máquina funciona devidamente apenas 1% das peças produzidas são defeituosas. Assuma o funcionamento correto da máquina.
   - **a)** Se forem examinadas, duas peças qual a probabilidade de 1 ser defeituosa?
   - **b)** Se forem examinadas 5 peças, qual a probabilidade de nenhuma ser defeituosa?
   - **c)** Qual o número esperado de peças defeituosas numa produção de 200?
   - **d)** Qual o desvio padrão das peças defeituosas numa amostra de 200?

**4.** Os sistemas de deteção de mísseis e radares militares permitem avisar contra ataques inimigos. Uma questão importante está relacionada com a capacidade do sistema em identificar e avisar corretamente um ataque. Assuma que um sistema particular de deteção tem 90% de probabilidades de detetar um ataque de míssil.
   - **a)** Qual a probabilidade de que um único sistema detete o ataque?
   - **b)** Se na mesma área forem instalados dois sistemas de deteção com funcionamento independente, qual a probabilidade de pelo menos um deles detetar o ataque?
   - **c)** Se forem instalados três sistemas, qual a probabilidade de pelo menos um deles detetar o ataque?

**5.** Sabe-se que com um determinado tratamento administrado a doentes em condições bem definidas se consegue 70% de curas. Se esse tratamento for aplicado a 20 doentes nas mesmas condições, qual a probabilidade de obter:
   - **a)** Máximo 15 curas?
   - **b)** 12 ou mais curas?
   - **c)** Entre 12 e 15 curas, inclusive?

**6.** Um determinado restaurante tem reputação de boa comida. O gerente registou que no sábado à noite os grupos de clientes chegam a uma média de 15 grupos cada meia hora.
   - **a)** Qual a probabilidade de que passem 5 minutos sem chegar nenhum cliente?
   - **b)** Qual a probabilidade de que oito grupos de clientes cheguem em 10 minutos?

c) Qual a probabilidade de que mais de 5 grupos cheguem num período de 10 minutos?

7. Os passageiros chegam aleatoriamente e independentemente a um grande aeroporto internacional, a uma média de 10 passageiros por minuto
   a) Qual a probabilidade de não chegar nenhum passageiro durante um minuto?
   b) Qual a probabilidade de chegarem 3 ou mais passageiros durante um minuto?
   c) Qual a probabilidade de não chegar nenhum passageiro durante 15 segundos?
   d) Qual a probabilidade de pelo menos 1 passageiro chegar num período de 15 segundos?
8. Numa empresa Têxtil existem numerosos teares de um certo tipo. A experiência mostra que, o número de teares que se avaria em cada mês é uma variável aleatória X que segue uma distribuição de Poisson com média igual a 3. Calcule:
   a) A probabilidade de que, durante um mês, se avariem 7 ou mais teares?
   b) A capacidade mínima que deve ter a oficina de reparação, de modo que, a probabilidade de não haver teares a aguardar reparação seja pelo menos de 90%.
9. Uma empresa de contabilidade prevê erros em 1% dos balanços das suas contas de clientes. Uma amostra de 150 contas foi selecionada para auditoria.
   a) Qual a probabilidade de que nenhuma das contas selecionadas tenha erros?
   b) Qual a probabilidade de 4 ou mais das contas conterem erros?
   c) Qual a probabilidade de exatamente 2 contas conterem erros?

10. Para um determinado modelo de calculadora de bolso, o fabricante sabe que 3% das calculadoras irão falhar nos primeiros 30 dias de operação e serão devolvidas para reparação. Assuma que tem um lote de 120 calculadoras:
   a) Qual o número esperado de calculadoras a falhar nos primeiros 30 dias de operação?
   b) Qual a probabilidade de pelo menos 2 falhem?
   c) Qual a probabilidade de que falhem exatamente 3?

## *DISTRIBUIÇÕES DE PROBABILIDADE CONTÍNUAS*

11. Em certas experiências, o erro obtido na determinação da densidade de uma substância é uma variável aleatória uniforme com a=-0.015 e b=0.015. Encontre as probabilidades de tal erro de determinação:
   a) estar entre –0.002 e 0.003;
   b) exceder 0.005 em valor absoluto.

12. O departamento de *I&D* de uma fábrica de aço acredita que uma das máquinas produz folhas de aço com espessura variada. A espessura é uma variável aleatória uniforme com valores entre 150 e 200 milímetros. Qualquer folha de aço com espessura inferior a 160 milímetros é rejeitada, pois são inaceitáveis para os clientes. Calcule a fração de folhas de aço produzidas por esta máquina que são rejeitadas?

**13.** A duração, em milhares de horas de um componente de um tipo de aparelhos de radar é uma variável aleatória X cuja f.d.p. é

$$f(x) = \begin{cases} 0.1e^{-0.1x} & \text{para } x > 0 \\ 0 & \text{para } x \leq 0 \end{cases}$$

**a)** Menos de 4 milhares de horas?
**b)** Entre 5 e 10 milhares de horas?

**14.** A quilometragem (em milhares de quilómetros) que um dono de um carro realiza com um certo tipo de pneus é uma variável aleatória com uma distribuição exponencial com $\theta$=40. Encontre as probabilidades de um desses pneus dure:
**a)** no mínimo 20 000 quilómetros;
**b)** no máximo 30 000 quilómetros.

**15.** O tempo que um relógio de pêndulo trabalha sem necessidade de dar corda é uma variável aleatória com uma distribuição exponencial com $\theta$=120 dias. Encontre as probabilidades para tal relógio de:
**a)** necessitar de corda em menos de 24 dias;
**b)** não necessitar de corda, no mínimo, durante 180 dias.

**16.** O intervalo de tempo que um ferry demora a fazer a travessia entre duas ilhas é normalmente distribuído com média de 2 horas e desvio padrão de 12 minutos. Nas últimas viagens, qual a proporção de vezes que o ferry fez a travessia em:
**a)** Menos de 1 hora e 45 minutos?
**b)** Mais de 2 horas e 5 minutos?
**c)** Entre 1 hora e 50 minutos e 2 horas e 20 minutos?

**17.** Alguns fabricantes automóveis desenvolvem os sensores de emissão de forma a estes serem substituídos depois de 100 000 milhas. Um desses fabricantes determinou que o tempo de serviço (em meses), desses sensores, segue uma distribuição normal com média 48 meses e desvio padrão de 9 meses.
**a)** O fabricante decidiu dar uma garantia aos sensores de 3 anos. Que percentagem de sensores não satisfaz a garantia?
**b)** O fabricante decidiu substituir apenas 1% de todos os sensores. Qual deverá ser a duração da garantia (em meses)?

**18.** Seja X o número de minutos depois das 11:00 que um autocarro deixa a estação. Assuma que a distribuição do tempo é aproximadamente normal com média 15 e desvio padrão de 4 minutos.
**a)** Se uma pessoa chegar à estação às 11:10, qual a probabilidade dessa pessoa ter perdido o autocarro?
**b)** Se a pessoa estiver disposta correr um risco de 20% de perder o autocarro, qual o número máximo de minutos depois das 11:00 que poderá chegar à estação?
**c)** A que horas deverá chegar à estação para ter uma probabilidade de 50% de apanhar o autocarro?

**19.** As classificações de um exame de admissão a um colégio seguem uma distribuição normal de média 500 e desvio padrão 100. Determine a probabilidade de um estudante ter classificação:

**a)** superior a 650;

**b)** inferior a 250;

**c)** entre 325 e 675.

**20.** Num processo fotográfico, o tempo de processamento da imagem pode ser visto como uma variável aleatória que segue uma distribuição normal com $\mu$=15.40 segundos e $\sigma$=0.48 segundos. Encontre as probabilidades de o tempo de processamento demorar:

**a)** no mínimo 16.00 segundos;

**b)** no máximo 14.20 segundos;

**c)** entre 15.00 e 15.80 segundos.

**21.** Uma fábrica de sapatos sabe que a medida dos pés dos clientes (senhoras) segue a lei normal, cuja média é 36cm e o desvio padrão é 1.5cm. O responsável da produção pretende programar a produção de um novo modelo pelo que necessita da distribuição por medidas.

**a)** Qual a percentagem prevista de pares com as medidas 32-34; 34-36; 36-38; 38-40; 40-42?

**b)** Numa produção total de 3000 pares, quantos devem ter medidas inferiores a 32cm ou superiores a 42cm?

**22.** Um operador de telecomunicações recebe um carregamento de 800 telemóveis. O fabricante garante um máximo de 1% dos telemóveis defeituosos. Se a afirmação for verdadeira, encontre a probabilidade do carregamento conter 15 ou mais telemóveis defeituosos.

**23.** Trinta por cento dos estudantes de uma determinada universidade frequentaram colégios particulares. Assuma uma amostra aleatória de 50 estudantes.

**a)** Qual a probabilidade de exatamente 10 dos estudantes selecionados terem frequentado um colégio particular?

**b)** Qual a probabilidade de 20 ou mais dos estudantes selecionados terem frequentado um colégio particular?

**c)** Qual a probabilidade de o número de estudantes provenientes de colégios particulares estar entre 10 e 20 inclusive?

**24.** De um questionário conduzido há 5 anos, concluiu-se que 30% dos adultos de uma cidade bebiam regularmente álcool. Se esta for ainda a percentagem, qual a probabilidade de, numa amostra aleatória de 1000 adultos, o número de pessoas que bebe álcool, ser:

**a)** Menor que 280?

**b)** Maior ou igual a 316?

# FICHA Nº 6 - ESTIMADORES PONTUAIS

**1.** Dada a f.d.p. da variável aleatória X:

$$f(x;\theta)=\begin{cases}\dfrac{x}{\theta^2}\exp\left(-\dfrac{x^2}{2\theta^2}\right) & , \quad para \quad x>0 \quad e \quad \theta\geq 0\\ 0 & , \quad nos \quad outros \quad casos\end{cases}$$

Considere o estimador $T=\left(\sum_{i=1}^{n}x_i^2\right)\Big/2n$ para o parâmetro $\theta^2$ da distribuição. Calcule a sua tendência.

**2.** Baseando-se numa amostra de tamanho 3 considere três estimativas possíveis para $\mu$ (média da distribuição):

$$W_1=\frac{1}{4}(X_1+X_2)+\frac{1}{2}X_3 \qquad W_2=\frac{1}{4}(X_1+X_2+X_3) \qquad W_3=0.3(X_1+X_2)+0.4X_3$$

**a)** Quais são os estimadores não tendenciosos? Justifique.
**b)** Quais as variâncias destes estimadores?
**c)** Determine a eficiência relativa dos estimadores não tendenciosos.

**3.** Um engenheiro pretende estimar a média do resultado de um processo químico, baseada em três medições, $X_1$, $X_2$ e $X_3$, resultantes de 3 experiências. Considere os seguintes estimadores para a média $\mu$:

$$T_1=\frac{X_1+X_2+X_3}{3} \text{ e } T_2=\frac{X_1+2X_2+X_3}{4}$$

Qual dos dois estimadores deve preferir? Justifique.

**4.** Considere os seguintes estimadores construídos com base na amostra aleatória $(X_1,X_2,\cdots,X_{n_1},\cdots X_n)$ proveniente de uma população $X$ com média $\mu$ e variância $\sigma^2$.

$$T_1=\frac{1}{n_1}\sum_{i=1}^{n_1}X_i,\quad T_2=\frac{1}{n-n_1}\sum_{i=n_1+1}^{n}X_i \text{ e } T=\frac{T_1+T_2}{2}$$

**a)** Averigue se $T_1$ e T são estimadores centrados de $\mu$.
**b)** Determine, justificando, o valor de n1 a partir do qual $T_1$ é mais eficiente do que T.

**5.** Suponha que $(X_1,X_2,\cdots,X_{n_1},\cdots X_n)$ é uma amostra aleatória de uma população cuja função densidade de probabilidade é dada por:

$$f(x)=\begin{cases}(\theta+1)x & 0<x<1, com\ \theta>1\\ 0 & caso\ contrário\end{cases}$$

em que θ é um parâmetro desconhecido.

**a)** Determine E [X] .

**b)** Considere os seguintes estimadores para θ:

$$T_1=\bar{X} \text{ e } T_2=-2+\frac{1-\bar{X}}{2}$$

Verifique se algum destes estimadores é um estimador centrado paraθ.

# FICHA Nº 7 - DISTRIBUIÇÕES AMOSTRAIS

**1.** Uma população tem uma média de 325 e variância 144. Considere uma amostra de tamanho 36. Determine

**a)** A média amostral

**b)** O desvio padrão amostral

**c)** $P(320 \leq \bar{x} \leq 322)$

**d)** $P(321 < \bar{x} < 327)$

**e)** $P(\bar{x} < 323)$

**f)** $P(\bar{x} > 328)$

**2.** Numa escola os registos indicam que os exames finais têm uma classificação com média de 510 e um desvio padrão de 90. Sabendo que 100 estudantes fazem o teste, qual a probabilidade da sua classificação média ser de:

**a)** Mais de 530?

**b)** Menos de 500?

**c)** Entre 495 e 515?

**3.** Um catálogo de um fabricante indica para um determinado produto uma vida média de 1200 horas. Assuma o desvio padrão igual a 120 horas. Um cliente decide selecionar aleatoriamente 35 itens do referido produto e rejeitar a amostra se $\bar{x} < 1160$ horas. Se a indicação do fabricante for verdadeira, qual a probabilidade de rejeitar a amostra?

**4.** Uma fábrica de sapatos tem uma máquina que corta peças de borracha comprimida para serem usadas em solas. A espessura dessas solas é uma variável aleatória normalmente distribuída com desvio padrão igual a 2 mm, com valor médio $\mu$. Para se tentar corrigir estas medidas, reajustando a máquina, é conveniente verificar a qualidade do produto, medindo espessura das solas de uma amostra aleatória retirada periodicamente da máquina. De uma amostra de 5 elementos foram registadas as espessuras respetivas e calculada a média aritmética. Se $\bar{x} < 24.8$ ou $\bar{x} > 25.2$ diz-se que a máquina não está controlada, pelo que é parada e reajustada.

**a)** Com a média $\mu = 25mm$, qual a probabilidade de a amostra indicar que a máquina não está controlada?

**b)** Se a média mudar para $\mu = 25.3mm$, qual a probabilidade de a amostra indicar que a máquina não está controlada?

**5.** Duas amostras aleatórias independentes foram retiradas de uma população normal com média 150 e variância 28.6. As amostras têm respetivamente tamanhos 10 e 25 e médias aritméticas de $\bar{x}_1$ e $\bar{x}_2$ . Determine:

**a)** $Var[\bar{X}_1 - \bar{X}_2]$;

**b)** $P(|\bar{X}_1 - \bar{X}_2| > 4)$.

# FICHA Nº 8 - INTERVALOS DE CONFIANÇA

**1.** Se uma amostra aleatória de tamanho n=20 duma população normal com variância $\sigma^2$=225 tem média $\bar{x} = 64.3$:

**a)** Construa o intervalo de confiança de 95% para a média da população $\mu$.

**b)** Construa também o intervalo de 90% de confiança.

**2.** Um inspetor alimentar, ao examinar 12 frascos de compota, obteve as seguintes percentagens de impureza: 2.3, 1.9, 2.1, 2.8, 2.3, 3.6, 1.4, 1.8, 2.1, 3.2, 2.0 e 1.9. Assumindo que estas determinações são distribuídas normalmente:

**a)** Construa o intervalo de 99% de confiança para o teor médio de impurezas nesta marca de compotas.

**b)** Construa também o intervalo de 90% e de 95% de confiança.

**3.** Uma mostra de n=100 empregados de uma companhia foi selecionada, e, o salário mensal foi registado. A média e o desvio padrão dos seus salários foram respetivamente $\bar{x} = 177500$ e s=9000. Construa o intervalo de confiança de 95% para o salário médio da população $\mu$.

**4.** Uma amostra aleatória de tamanho n é retirada duma população com média $\mu$ e desvio padrão $\sigma$. A média e o desvio padrão da amostra são $\bar{x}$ =45 e s=5.8. Calcule o intervalo de confiança de 95% para $\mu$ usando os seguintes tamanhos da amostra:

**a)** n=30 **b)** n=60 **c)** n=90

Compare os tamanhos dos três intervalos.

**5.** Sejam $\bar{X}_1$ e $\bar{X}_2$ as médias aritméticas de duas amostras aleatórias e independentes de tamanho n, tiradas respetivamente das distribuições $N(\mu_1, \sigma^2)$, $N(\mu_2, \sigma^2)$.

Determine $n$ de modo que $P\left[\bar{X}_1 - \bar{X}_2 - \frac{\sigma}{5} < \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + \frac{\sigma}{5}\right] = 0.90$

**6.** Cinco pessoas selecionadas aleatoriamente foram usadas num teste para medir as suas capacidades em termos de volume de ar inspirado, antes e depois de um tratamento. Se $\mu_x$ for a capacidade média da população antes do tratamento e $\mu_y$ for a capacidade média da população depois do tratamento, construa um intervalo de confiança que tenha 90% de probabilidade de conter $\mu_y - \mu_x$.

| | **Volume de ar inspirado** | |
|---|---|---|
| **pessoas** | **antes (X)** | **depois (Y)** |
| **A** | 2750 | 2850 |
| **B** | 2360 | 2380 |
| **C** | 2950 | 2930 |
| **D** | 2830 | 2860 |
| **E** | 2250 | 2320 |

**7.** As capacidades caloríficas do carvão de duas minas são (em milhões de calorias por tonelada):

Mina A: 8500 8330 8480 7960 8030

Mina B: 7710 7890 7920 8270 7860

Assumindo que os dados constituem amostras independentes de populações normais com variâncias iguais:

**a)** Construa o intervalo de 99% para a diferença entre as verdadeiras médias das capacidades caloríficas do carvão das duas minas.

**b)** Construa também o intervalo de 90%. O que pode concluir?

**8.** Duas amostras independentes de tamanho $n_1$ e $n_2$ foram retiradas de duas populações com médias $\mu_1$ e $\mu_2$ e desvios padrão $\sigma_1^2$ e $\sigma_2^2$ respetivamente. A seguinte informação amostral é conhecida:

Amostra 1: $n_1 = 50$, $\bar{x}_1 = 91.1$, $s_1 = 5.4$

Amostra 2: $n_2 = 50$, $\bar{x}_2 = 92.3$, $s_2 = 7.6$

**a)** Estime a diferença e construa o intervalo de 95% de confiança.

**b)** Qual o limite de 95% de confiança para o erro na estimação?

**9.** Um clube de compras por correio oferece mensalmente produtos que podem ser adquiridos pelos sócios. É feito um teste de aceitação do produto A enviando-o a 250 sócios, escolhidos aleatoriamente dentre os 9000 membros. Baseada nesta amostra, somente 70 sócios decidiram comprar o produto A.

**a)** Dê uma estimativa pontual da proporção de sócios que se espera comprem o produto.

**b)** Calcule, com 95.4% de certeza, um limite do erro cometido.

**10.** Quarenta e uma pessoas, de uma amostra aleatória de 500 trabalhadores, estão desempregadas. Calcule um intervalo de confiança que tenha 95% de probabilidade de conter a percentagem de desempregados no país.

**11.** Numa pesquisa de mercado, 30 famílias, de uma amostra aleatória de 150, afirmaram que tencionavam comprar um carro novo no próximo ano. Construa um intervalo de confiança com 95% de probabilidade de conter a proporção de todas as famílias que tencionam comprar um carro novo no próximo ano.

**12.** Uma amostra aleatória de tamanho n=60 é retirada duma população binomial com parâmetro p, a proporção de sucessos na população. A amostra produz $x=35$ sucessos.

**a)** Estime $p$.

**b)** Construa um intervalo de 95% de confiança indicando o erro da estimativa.

**13.** Um estudo está a ser conduzido para estimar a proporção de votantes numa grande comunidade que apoiam a construção duma central nuclear. De 400 votantes selecionados aleatoriamente, só 140 apoiam o projeto.

**a)** Construa um intervalo de 90% de confiança para a proporção de todos os votantes nesta comunidade que apoiam o projeto.

**b)** Construa também os intervalos de 95% e de 98% de confiança.

**14.** Na freguesia A, 132 votantes de 400 apoiam um candidato à presidência, enquanto que na freguesia B, 90 votantes de 150 apoiam o mesmo candidato à presidência. Encontre o intervalo de 99% de confiança para o intervalo ($p_1$-$p_2$), a diferença entre a proporção atual de votantes das duas freguesias que apoiam o candidato.

**15.** Um produtor de extintores de moscas quer comparar duas novas formulações, 1 e 2. Dois quartos de igual tamanho, cada um contendo 1000 moscas, são usados na experiência, um tratado com o extintor 1 e o outro tratado com igual quantidade do extintor 2. Um total de 825 e 760 moscas sucumbem aos extintores 1 e 2 respetivamente. Estime a diferença na taxa de mortalidade dos dois extintores, quando usados no ambiente de teste.

**16.** Um relojoeiro pretende conhecer as variações do produto que fabrica. Para construir um intervalo de confiança para $\sigma$, baseou-se numa amostra aleatória de 10 relógios escolhidos dentre os relógios que passaram o último teste de qualidade. Os valores dos desvios dos 10 relógios, em relação a um relógio *padrão* foram registados ao fim de um mês. Considere $\bar{x} = 7$ seg. e $s = 4$ seg. Supondo que a distribuição dessas medidas pode ser aproximada por uma distribuição normal, determine o intervalo de confiança que tenha 90% de probabilidade de conter $\sigma$.

**17.** Um controlador de qualidade numa fábrica de refrigerantes sabe que a quantidade exata de cada lata variará, uma vez que existem fatores incontroláveis que afetam o enchimento. A quantidade média é importante, mas também é a variação dessa quantidade. Se é grande, algumas latas conterão pouco líquido, enquanto outras terão muito líquido. Para estimar a variação do enchimento, o supervisor seleciona aleatoriamente 10 latas e determina o volume do conteúdo de cada uma delas. Foram obtidos os seguintes resultados: $\bar{x} = 32.98$ cl. e $s = 0.04$ cl. Construa um intervalo de confiança de 90% para a verdadeira variação do enchimento.

**18.** Uma empresa está a experimentar dois arranjos físicos diferentes para a sua linha de montagem. Ambos os arranjos produzem aproximadamente o mesmo número de peças acabadas por dia. Para obter um maior controlo sobre o processo, o arranjo com menor variância no número de peças acabadas deve ser mantido. Duas amostras independentes produziram os seguintes resultados.

**Linha 1:** $n_1 = 21$, $s_1^2 = 1432$ **Linha 2:** $n_2 = 25$, $s_2^2 = 3761$

Construa um intervalo de confiança de 95% para $\sigma_1^2 / \sigma_2^2$, a razão das variâncias do número de peças acabadas em cada uma das linhas de montagem. Qual dos dois arranjos deve ser usado?

# FICHA Nº 9 - TESTES DE HIPÓTESES

**1.** Suponha que um tratamento pretende aumentar um resultado. Considere a hipótese nula de que o aumento da média é $\mu = 0$. Vamos fazer um estudo que vai usar uma amostra de tamanho n, para determinar os efeitos do tratamento. Um efeito é significativo se $\bar{x} > k$. O teste terá um nível de significância de 0.05 e uma probabilidade de 80% de detetar um aumento na média de duas unidades. Determine os valores do tamanho da amostra a usar, n e do valor k da região de rejeição, uma vez que o desvio padrão é $\sigma = 5$.

**2.** Deverá decidir quais das duas distribuições discretas descreve o comportamento de uma variável aleatória X. Chamaremos às distribuições $p_0(x)$ e $p_1(x)$. As probabilidades associadas a cada valor de X=x são as seguintes nos dois modelos:

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $p_0$ | 0,1 | 0,1 | 0,1 | 0,1 | 0,2 | 0,1 | 0,3 |
| $p_1$ | 0,2 | 0,1 | 0,1 | 0,2 | 0,2 | 0,1 | 0,1 |

Observe a variável X uma única vez e formule:

$H_0$: $p_0$ é a distribuição correta

$H_1$: $p_1$ é a distribuição correta

Um procedimento possível de decisão consiste em não rejeitar $H_0$ se X=4 ou X=6 e rejeitar $H_0$ nos outros casos.

**a)** Determine a probabilidade de cometer um erro do tipo I;

**b)** Determine a probabilidade de cometer um erro do tipo II.

**3.** Um experimentador testou diferenças de atitude quanto ao fumar, antes e depois de um filme sobre cancro do pulmão ter sido visionado. Foi encontrada uma diferença significativa entre os níveis de $\alpha = 0.05$ e $\alpha = 0.02$

**a)** Qual é a hipótese nula assumida?

**b)** Qual o nível que indica o maior grau de significância, 0.05 ou 0.02? Justifique.

**c)** Para $\alpha = 0.05$, rejeitar-se-á $H_1$? Justifique. Será $H_1$ rejeitada para $\alpha = 0.02$? Justifique.

**d)** Ao escolher $\alpha = 0.02$ em vez de $\alpha = 0.05$, aumenta-se o risco de um dos dois tipos de erro? Qual? Justifique.

**4.** O espaço amostral da "estatística" de um teste é formado por cinco valores {a,b,c,d,e}. Considere o teste sobre a função de probabilidade da variável X, de $f_0(x)$ e $f_1(x)$, definidas da seguinte maneira:

| X | a | b | c | d | e |
|---|---|---|---|---|---|
| $f_0(x)$ | 0 | 0,1 | 0,2 | 0,3 | 0,4 |
| $f_1(x)$ | 0,3 | 0 | 0,2 | 0,4 | 0,1 |

**a)** Calcule as funções α e β, probabilidades associadas com os erros do tipo I e II, respetivamente, se for definida como região de rejeição o conjunto

i) $C_1$={b,c}

ii) $C_2$={d}

**b)** Face aos resultados de a) qual o melhor teste (associado à região de rejeição $C_1$ ou $C_2$)? Justifique.

**5.** De uma amostra casual de 100 horas, uma máquina produziu em média 678 artigos por hora com um desvio padrão de 25. Depois de ter sido introduzido um esquema de controlo, a máquina passou a produzir em média 674 artigos com desvio padrão de 5, tirada de uma amostra aleatória de 500 horas. O administrador da empresa afirma que o esquema de controlo reduziu a produção. O sindicato, no entanto, afirma que os 4 artigos a menos na média calculada, são devidos a flutuações estatísticas.

**a)** Calcule a função potência, quando a hipótese nula é verdadeira.

**b)** Se o nível de significância for 1%, considerar-se-á a afirmação da administração ou do sindicato?

**6.** Até agora, a percentagem de empregados de uma firma que usavam transporte público para se deslocarem para o emprego e do emprego para casa, era de 20%. Foi feita uma campanha para a utilização dos transportes públicos. Pretende-se saber se a campanha foi eficaz. Para isso, considerou-se uma amostra aleatória de 25 empregados o número de empregados que passou a utilizar os transportes públicos é dado por X.

**a)** Formule a hipótese nula, em termos de p, proporção da população de empregados da firma que utiliza os transportes públicos.

**b)** Qual seria a região de rejeição, se o nível de significância do teste deve ser controlado para um valor menor que 0.1?

**7.** Quando o resultado de um processo de produção é estável a um nível aceitável, diz-se que está controlado. Suponha que o processo tem estado controlado desde há algum tempo e que a proporção de produtos defeituosos é de 5%. Para automatizar o processo, o chefe de produção decide considerar o processo não controlado se forem encontrados mais de dois produtos defeituosos numa amostra aleatória de 15 produtos.

**a)** Determine a probabilidade de aparecer não controlado, quando p=0.05

**b)** Determine o gráfico da curva de potência para este esquema de controlo quando p=0.05, 0.1, 0.3 e 0.4.

**8.** Uma única observação vai ser usada para testar a hipótese nula de que o tempo médio de espera entre tremores de terra numa estação sismológica (a média de uma distribuição exponencial) é $\theta = 10$ horas, contra a alternativa de que $\theta \neq 10$ horas. A hipótese nula é rejeitada, se e só se, o valor observado for menor que 8 ou maior que 12.

**a)** Encontre a probabilidade de um erro do tipo I;

**b)** Encontre a probabilidade de erros do tipo II, quando $\theta$=2, 4, 6, 8, 12, 16 e 20

**c)** Apresente o gráfico da função potência para este teste.

**9.** Uma amostra aleatória de tamanho n=64 vai ser usada para testar a hipótese nula de que, para um certo grupo etário, a nota média num teste (a média duma população normal com $\sigma^2 = 256$) é menor que ou igual a 40.0, contra a alternativa de que é maior que 40.0. A hipótese é rejeitada, se e só se, a média da amostra aleatória exceder 43.5.

a) Encontre as probabilidades de erros do tipo I, quando $\mu$ = 37.0, 38.0, 39.0 e 40.0.
b) Encontre as probabilidades de erros do tipo II, quando $\mu$ = 41.0, 42.0, 43.0, 44.0, 45.0, 46.0, 47.0 e 48.0.
c) Apresente o gráfico da função potência deste teste.

**10.** Baseado em determinado conjunto de dados, a hipótese nula é rejeitada ao nível de significância de 0.05. Seria também rejeitada ao nível de significância de:
a) 0.01?
b) 0.10?

**11.** Num determinado teste de hipótese, o valor p correspondente à estatística é de 0.0316. Pode a hipótese nula ser rejeitada ao nível de significância de:
a) 0.01?
b) 0.05?
c) 0.10?

**12.** De acordo com as normas estabelecidas para um teste de compreensão de leitura, os alunos do oitavo ano devem ter uma média de 84.3 com um desvio padrão de 8.6. Se 45 alunos, selecionados aleatoriamente num certo distrito, têm uma média de 87.7, teste a hipótese nula $\mu = 84.3$ contra a hipótese alternativa $\mu > 84.3$ a um nível de significância de 0.01.

**13.** Suponha que é conhecido pela experiência que o desvio padrão do peso de 8 g de bolos fabricados por uma certa padaria é 0.18 g. Para verificar se a produção está sobre controlo, isto é, para verificar se o verdadeiro peso médio dos pacotes é de 8 g, foi extraída uma amostra aleatória de 25 pacotes sendo a sua média $\bar{x} = 8.172g$. Uma vez que a padaria perde dinheiro quando $\mu > 8$ e os clientes o perdem quando $\mu < 8$, teste a hipótese nula $\mu = 8$ contra a hipótese alternativa $\mu \neq 8$ usando $\alpha = 0.01$.

**14.** Suponha que é necessário que a resistência à rutura de um certo tipo de fita seja de 83.9 kg e que 5 peças selecionadas aleatoriamente de diferentes rolos têm uma resistência média de 83.05 kg com um desvio padrão de 3.72 kg. Assumindo que os dados provêm de uma amostra aleatória de uma população normal, teste a hipótese nula $\mu = 83.9$ contra a hipótese alternativa $\mu < 83.9$ a um nível de significância de $\alpha = 0.05$.

**15.** Em doze corridas numa pista, um novo barco gastou um tempo médio de 33.6 segundos com um desvio padrão de 2.3 segundos. Assumindo ser razoável tratar os dados como uma amostra aleatória duma população normal, teste a hipótese nula $\mu = 35$ contra a alternativa $\mu < 35$ ao nível de significância de 0.05.

**16.** Na comparação de dois tipos de tinta constatou-se que com 4 latas de tinta de uma marca se pintou uma superfície de 512 $cm^2$ com um desvio padrão de 31 $cm^2$, enquanto que com a mesma quantidade de outra tinta se conseguiu pintar uma superfície de 492 $cm^2$ com um desvio padrão de 26 $cm^2$. Teste a hipótese nula

$\mu_1 - \mu_2 = 0$ contra a hipótese alternativa $\mu_1 - \mu_2 \neq 0$, a um nível de significância $\alpha = 0.05$. Considere que as duas populações são normais e têm variâncias iguais.

**17.** Os dados registam o número médio de horas-homem perdidas devidas a acidentes em 10 fábricas, antes e depois de um programa de higiene e segurança ter sido implementado:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Antes | 45 | 73 | 46 | 124 | 33 | 57 | 83 | 34 | 26 | 17 |
| Depois | 36 | 60 | 44 | 119 | 35 | 51 | 77 | 29 | 24 | 11 |

Use o nível de significância de 0.05 para testar se o programa de higiene e segurança é eficaz.

**18.** Experimentou-se uma nova máquina de enchimento estéril de frascos de antibióticos, obtendo-se para os 33 frascos, o peso médio de 1093 mg e um desvio padrão de 36 mg. Pelo processo de enchimento manual, uma amostra de 30 frascos deu o peso médio de 1122 mg e um desvio padrão de 23 mg. Acha que existe uma diferença significativa entre as médias dos pesos obtidos pelos dois processos?

**19.** Os teores de nicotina de duas marcas de cigarros estão a ser medidos. Se numa experiência 50 cigarros da marca A têm um teor médio de nicotina de $\bar{y}_1 = 2.61$ mg com um desvio padrão de $s_1 = 0.12$ mg, enquanto os 40 cigarros da marca B têm um teor médio de nicotina de $\bar{y}_2 = 2.38$ mg com um desvio padrão de $s_2 = 0.14$ mg, teste a hipótese nula $\mu_1 - \mu_2 = 0.2$ contra a hipótese alternativa $\mu_1 - \mu_2 \neq 0.2$, usando $\alpha = 0.05$.

**20.** Um estudo pretende comparar a atitude das pessoas sobre o feminismo com o seu grau de autoritarismo. Foram usadas duas amostras: uma de 30 pessoas que foram classificadas de muito autoritárias, e outra de 31, classificadas de pouco autoritárias. A cada pessoa foi dado um questionário com 18 questões, e as pontuações finais registadas variavam desde o 18 ao 90 (pontuações altas indicavam uma atitude pro-feminismo). Do estudo obtivemos as seguintes estatísticas:

| Autoritarismo | n | $\bar{x}$ | s |
|---|---|---|---|
| elevado | 30 | 67.7 | 11.8 |
| baixo | 31 | 52.4 | 13.0 |

Teste a hipótese nula de que o autoritarismo não é um fator que influencia a atitude da pessoa em relação ao feminismo.

**21.** Um estudo, sobre o número de almoços que executivos nos seguros e na banca apresentam como despesas dedutíveis num mês, foi baseado em amostras aleatórias que produziram os seguintes resultados:

$n_1 = 40$, $\bar{x}_1 = 9.1$, $s_1 = 1.9$

$n_2 = 50$, $\bar{x}_2 = 8.0$, $s_2 = 2.1$

O que pode conclui?

**22.** Um novo tratamento para a esquizofrenia foi testado durante seis meses com 54 doentes selecionados aleatoriamente. Ao fim desse período foi dada alta a 25 doentes. A proporção usual em seis meses é de 1/3. Usando uma aproximação normal à distribuição binomial, determine se o novo tratamento resultou em maior número de altas que o tratamento anterior ($\alpha$=0.05).

**23.** Dentre as 60 lâminas testadas somente 7 lâminas do rotor de uma turbina a gaz falharam. Até agora e em testes idênticos costumavam falhar 20% das lâminas. Serão agora as lâminas testadas significativamente melhores que as usadas anteriormente?

**24.** Num determinado país, uma série de testes conduzidos num aeroporto mostraram que os seguranças só detetaram 72 das 100 armas falsas levadas por inspetores. Esta taxa de deteção está abaixo da taxa nacional de deteção de 80%. Há evidência suficiente para concluir que a deteção no referido aeroporto está abaixo da taxa nacional? Use $\alpha$=0.01.

**25.** Em 1990, 371 empresas foram selecionadas para determinar em que medida disponham de sistemas de informação em logística. Cinco anos mais tarde, em 1995, 459 empresas foram selecionadas para determinar a evolução do uso destes sistemas de informação. Assim, a percentagem varia de 1990 para 1995, de 25% para 33%. Permitem os dados concluir que houve um aumento significativo de empresas que dispõem de sistemas de informação em logística? Use $\alpha$=0.05.

**26.** Suponha que a espessura de um componente usado num semicondutor é a sua dimensão crítica e que as medidas da espessura, de uma amostra aleatória de 18 desses componentes, têm variância igual a 0.68 cm. Considera-se que o processo está controlado se a variância da espessura não é superior a 0.36 cm. Assumindo que as medições constituem uma amostra aleatória de uma população normal, teste a hipótese nula contra a hipótese alternativa $\sigma^2 > 0.36$ a um nível de significância de 5%.

**27.** Nove determinações do calor específico do ferro apresentaram um desvio padrão de 0.0086. Assumindo que estas determinações constituem uma amostra aleatória duma população normal, teste a hipótese nula $\sigma = 0.0100$ contra a hipótese alternativa $\sigma < 0.0100$ ao nível de significância de 0.05.

**28.** Ao comparar a variabilidade da tensão em dois tipos de aço, uma experiência conduziu aos seguintes resultados:

$n_1 = 13$, $s_1^2 = 19.2$

$n_2 = 16$ e $s_2^2 = 3.5$

onde as unidades são em 100 psi (libras por polegada quadrada). Assumindo que as medidas constituem amostras aleatórias independentes provenientes de duas populações normais, teste a hipótese nula $\sigma_1^2 = \sigma_2^2$ contra a hipótese alternativa $\sigma_1^2 \neq \sigma_2^2$ a um nível de significância de 5%.

# FICHA Nº 10 - ANÁLISE DA VARIÂNCIA

**1.** Pretende-se fazer um teste de tensão a uma peça de alumínio. Para o teste foram usadas três máquinas A, B e C. Para testar a existência de efeitos devidos às máquinas foram usadas cinco peças de cada uma das máquinas. Os resultados obtidos na experiência foram:

| máquina A | máquina B | máquina C |
|---|---|---|
| 3,2 | 4,9 | 3,0 |
| 4,1 | 4,5 | 2,9 |
| 3,5 | 4,5 | 3,7 |
| 3,0 | 4,0 | 3,5 |
| 3,1 | 4,2 | 4,2 |

**a)** Teste a hipótese nula de não existirem diferenças significativas nos efeitos das máquinas. Considere a variável resposta normalmente distribuída.

**b)** Determine um intervalo de confiança com 90% de probabilidade de conter a diferença entre as médias das máquinas B e C.

**2.** Foram testadas três marcas diferentes de lâmpadas A, B e C com o objetivo de determinar o tempo de duração. Os resultados da experiência foram os seguintes:

| | | | | | |
|---|---|---|---|---|---|
| **A** | 73 | 64 | 67 | 62 | 70 |
| **B** | 84 | 80 | 81 | 77 | |
| **C** | 82 | 79 | 71 | 75 | |

Baseando-se nesta amostra, acha que os resultados indicam alguma diferença entre o tempo de duração das marcas (para $\alpha=0.05$)? Use a análise da variância.

**3.** Retiraram-se seis amostras de algodão de sete fardos para ser analisado o índice *micronaire*, que se supõe segue a distribuição normal. Pretende-se saber se existem diferenças significativas entre os fardos de algodão

| **Fardos** | | | | | | |
|---|---|---|---|---|---|---|
| **1** | 3.72 | 3.75 | 3.67 | 3.67 | 3.70 | 3.70 |
| **2** | 3.70 | 3.77 | 3.77 | 3.87 | 3.85 | 3.70 |
| **3** | 3.87 | 3.95 | 3.90 | 3.82 | 3.77 | 3.92 |
| **4** | 4.02 | 4.02 | 4.02 | 3.85 | 3.92 | 3.87 |
| **5** | 4.37 | 4.35 | 4.00 | 4.10 | 3.92 | 3.95 |
| **6** | 3.90 | 3.77 | 3.75 | 3.72 | 3.57 | 3.55 |
| **7** | 3.90 | 3.97 | 3.90 | 4.00 | 4.15 | 4.10 |

**4.** Três grupos de cobaias (ratos) foram, cada um, injetados, respetivamente com 0.05mg, 1.0mg e 1.5mg de um novo tranquilizantes, e os tempos que demoraram a adormecer foram os seguintes, em minutos:

| | | | | | | |
|---|---|---|---|---|---|---|
| 0.5 mg | 21 | 23 | 19 | 24 | 25 | 23 |
| 1.0 mg | 19 | 21 | 20 | 18 | 22 | 20 |
| 1.5 mg | 15 | 10 | 13 | 14 | 11 | 15 |

Teste ao nível de significância de $\alpha=0.05$, se se pode rejeitar a hipótese nula de que a diferença na dose não têm efeito. Assuma que a variável resposta é normalmente distribuída.

**5.** Para comparar as velocidades de corte de 4 máquinas, arranjaram-se peças com cinco graus de dureza diferentes: Formaram-se cinco blocos experimentais, cada um com peças do mesmo grau de dureza. Os resultados da experiência foram (em segundos):

| bloco | M1 | M2 | M3 | M4 |
|---|---|---|---|---|
| 1 | 12 | 20 | 13 | 11 |
| 2 | 2 | 14 | 7 | 5 |
| 3 | 8 | 17 | 13 | 10 |
| 4 | 1 | 12 | 8 | 3 |
| 5 | 7 | 17 | 14 | 6 |

**a)** Teste a hipótese nula de que não existem diferenças significativas entre as velocidades de corte das diferentes máquinas, supondo que o tempo segue a distribuição normal

**b)** Teste os efeitos resultantes dos blocos

**6.** O Eng. José Costa da empresa *Jotex*, Lda. está preocupado com os baixos níveis de produção dos seus trabalhadores. Para ver se aumentava a produção resolveu estudar um esquema de incentivos. Na realização da experiência, selecionou aleatoriamente seis trabalhadores a quem propôs o esquema de incentivos. A produção mensal obtida antes e a conseguida depois da introdução do esquema foram as seguintes:

| | unidades produzidas | |
|---|---|---|
| trabalhador | antes | depois |
| Luís Mota | 80 | 85 |
| Ana Lopes | 75 | 75 |
| Cristina Pinto | 65 | 71 |
| Joana Silva | 82 | 79 |
| José Gonçalves | 70 | 86 |
| Maria Cruz | 56 | 68 |

**a)** Formule as hipóteses, supondo que esta produção mensal segue a distribuição normal.

**b)** O que se pode concluir desta experiência?

**7.** Os seguintes dados representam o tempo (em minutos) que uma pessoa demorou a conduzir até ao seu local de emprego, de segunda a sexta, por quatro caminhos diferentes:

| | Seg | Ter | Qua | Qui | Sex |
|---|---|---|---|---|---|
| caminho 1 | 22 | 26 | 25 | 25 | 31 |
| caminho 2 | 25 | 27 | 28 | 26 | 29 |
| caminho 3 | 26 | 29 | 33 | 30 | 33 |
| caminho 4 | 26 | 28 | 27 | 30 | 30 |

Considere os dados normalmente distribuídos.

**a)** Existem diferenças significativas nos tempos médios de condução para os quatro caminhos diferentes? Que conclusão pode tirar? Justifique

**b)** Face às conclusões que tirou em a) que modelo populacional lhe parece que seria o mais adequado?

# FICHA Nº 11 – QUI-QUADRADO – "BOM AJUSTE"

**1.** Uma fábrica tem quatro máquinas de produção de moldes. Uma amostra de 500 moldes é retirada de cada máquina e o número de moldes defeituosos encontrados foi:

| máquina | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Nº defeitos/500 | 10 | 25 | 0 | 5 |

Use um teste estatístico apropriado para comparar as 4 máquinas em termos do número de defeitos produzidos ($\alpha$=0.05).

**2.** 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 dados obtidos na experiência contradizem a afirmação.

| Trimestre | Jan-Mar | Abr-Jun | Jul-Set | Out-Dez |
|---|---|---|---|---|
| Nº nascimentos | 110 | 57 | 53 | 80 |

**3.** Ao examinar os registos de uma agência de venda de automóveis (camiões) verificou-se que em 70 dias houve vendas diárias de um só camião, em 60 dias venderam-se 2 camiões, em 40 dias venderam-se diariamente 3 camiões e em 30 dias 4 camiões.

| Camiões vendidos por dia | Nº de dias |
|---|---|
| 1 | 70 |
| 2 | 60 |
| 3 | 40 |
| 4 | 30 |

Teste a hipótese de que a procura de camiões é uniformemente distribuída considerando $\alpha$=0.01.

**4.** Deseja-se testar se o número de raios gama emitidos por segundo por uma determinada substância radioativa é uma variável aleatória seguindo a distribuição de Poisson com $\lambda$=2.4. Use os dados obtidos por 300 intervalos de um segundo para testar esta hipótese nula ao nível de significância de 0.05.

| Nº Raios Gama | 0 | 1 | 2 | 3 | 4 | 5 | 6 | $\geq 7$ |
|---|---|---|---|---|---|---|---|---|
| Frequência | 19 | 48 | 66 | 74 | 44 | 35 | 10 | 4 |

**5.** Caixas de um certo tipo foram expostas ao risco de acidentes sob a ação de tempestades, gelo, fogo, queda, etc., por um período de 400 dias. O número de acidentes com cada caixa é uma variável aleatória X que se afirma seguir a distribuição de Poisson. Verifique se os dados da experiência efetuada, registados na tabela, fundamentam a afirmação.

| Nº Acidentes, X | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| Nº Caixas com X acidentes | 1448 | 805 | 206 | 34 | 4 | 2 | 1 |

**6.** No estudo da velocidade dos fios, efetuou-se a contagem do número de fibras soltas por mm de comprimento. Verifique se a distribuição dos comprimentos das fibras soltas de um fio de lã usado na experiência segue uma distribuição exponencial. A tabela dos valores observados é a seguinte:

| Comprimento valor médio da classe | frequências observadas |
|---|---|
| 2.5 | 55 |
| 7.5 | 19 |
| 12.5 | 6 |
| 17.5 | 20 |

**7.** Quatro moedas são lançadas 160 vezes e o número de 0, 1, 2, 3 ou 4 caras observadas foi, respetivamente, 19, 54, 58, 23 e 6. Use o nível de significância de 0.05 para testar se é razoável supor que as moedas são equilibradas e lançadas aleatoriamente.

**8.** Em cada dia, de segunda a sexta, um padeiro produz três grandes bolos de chocolate, e os que não são vendidos no mesmo dia são dados a um banco alimentar. Use os dados apresentados no quadro seguinte para testar, ao nível de significância de 0.05, se podem ser considerados como valores duma variável aleatória binomial.

| | | | | |
|---|---|---|---|---|
| Nº de bolos | 0 | 1 | 2 | 3 |
| Nº de dias | 1 | 16 | 55 | 228 |

**9.** Os dados seguintes apresentam a distribuição das leituras com um contador Geiger, do número de partículas emitidas por uma substância radioativa em 100 intervalos sucessivos de 40 segundos:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Nº de partículas | 5-9 | 10-14 | 15-19 | 20-24 | 25-29 | 30-34 | 35-39 |
| Frequência | 1 | 10 | 37 | 36 | 13 | 2 | 1 |

**a)** Verifique que a média e o desvio padrão desta distribuição são $\bar{x} = 20$ e $s = 5$.

**b)** Encontre as probabilidade de que uma variável aleatória, seguindo uma distribuição normal com $\mu = 20$ e $\sigma = 5$, tome um valor:

i) menor que 9.5;
ii) entre 9.5 e 14.5;
iii) entre 14.5 e 19.5;
iv) entre 19.5 e 24.5;
v) entre 24.5 e 29.5;
vi) entre 29.5 e 34.5;
vii) maior que 34.5.

**c)** Encontre as frequências esperadas para as várias classes, multiplicando as probabilidades obtidas na alínea b) pela frequência total e, então teste ao nível de significância de 0.05 se os dados podem ser vistos como uma amostra aleatória duma população normal.

# FICHA Nº 12 – REGRESSÃO E CORRELAÇÃO (SPSS)

**1.** A Lei de Ohm diz que a intensidade de corrente I num fio de metal é proporcional à diferença de potencial V aplicada nos seus extremos e, inversamente proporcional à resistência R no fio.

Usando uma equação, a lei de Ohm é descrita por $I = \frac{V}{R}$.

Num laboratório, os estudantes realizaram várias experiências para estudar esta lei. Variaram a diferença de potencial V e para cada valor leram o valor da intensidade I. Pretendiam assim determinar o valor de R para aquele fio.

Podemos escrever a Lei de Ohm na forma $I = \beta_0 + \beta_1 V$, com $\beta_0 = 0$ e $\beta_1 = \frac{1}{R}$.

Os dados obtidos a partir das experiências foram:

| **V** | **0,5** | **1,0** | **1,5** | **1,8** | **2,0** |
|---|---|---|---|---|---|
| **I** | 0,52 | 1,19 | 1,62 | 2,00 | 2,40 |

**a)** Qual a estimativa de $\frac{1}{R}$ para aquele cabo?

**b)** Construa um intervalo de confiança de 95% para $\frac{1}{R}$.

**c)** Como a Lei de Ohm define no modelo o valor de $\beta_0$ é igual a zero. Faça um teste estatístico em relação a esta hipótese.

**d)** Calcule a estimativa para a resistência R e determine um intervalo de confiança de 95% para R.

**e)** Estime o valor esperado da intensidade I, para uma diferença de potencial de 1.2. E se essa estimativa for de 2.2?.

**2.** Determine a relação existente entre o calor envolvido no endurecimento, representado pela variável Y e os pesos de duas substâncias X1 e X2, tendo em consideração os seguintes valores obtidos numa experiência:

| **Y** | 79 | 74 | 104 | 88 | 96 | 109 | 103 | 73 | 93 | 116 |
|---|---|---|---|---|---|---|---|---|---|---|
| **X1** | 7 | 1 | 11 | 11 | 7 | 11 | 3 | 1 | 2 | 21 |
| **X2** | 26 | 29 | 59 | 31 | 52 | 55 | 71 | 31 | 54 | 47 |

**3.** Amostras de solo seco a diferentes temperaturas, X, perdem proporções diferentes de mistura, Y. Ajuste um modelo do tipo $Y = b_0 + b_1 x + b_2 x^2 + b_3 x^3$, considerando os seguintes valores obtidos numa experiência:

| **percentagem da perda de peso, Y** | 3,71 | 3,81 | 3,86 | 3,93 | 3,96 | 4,20 | 4,34 | 4,51 | 4,73 | 0,35 |
|---|---|---|---|---|---|---|---|---|---|---|
| **Temperatura, X** | 100 | 105 | 110 | 115 | 121 | 132 | 144 | 153 | 163 | 179 |

**4.** Para calcular a capacidade de um aparelho "*air flow*" foram recolhidas seis amostras de lã de diâmetros $d_i$, I=1,2,…,6, conhecidos. As alturas mano métricas do aparelho, h, estão relacionadas com os diâmetros das fibras de lã utilizadas, segundo a expressão $h_i = k_1 d_i^{k_2} u_i$, em que $u_i$ são os erros casuais de observação. O logaritmo decimal da variável u segue uma distribuição normal com média zero e variância $\sigma^2$.

Estime os valores dos parâmetros k1 e k2 que definem o modelo, considerando os resultados obtidos numa experiência:

| **d ($\mu$ )** | 19,84 | 20,95 | 22,25 | 24,46 | 26,30 | 30,18 |
|---|---|---|---|---|---|---|
| **h (mm)** | 335,0 | 330,3 | 293,5 | 239,3 | 205,9 | 160,2 |

**5.** Os dados da tabela apresentam as taxas de consumo de oxigénio de pássaros a diferentes temperaturas ambientais.

| **Temperatura (ºC)** | -18 | -15 | -10 | -5 | 0 | 5 | 10 | 19 |
|---|---|---|---|---|---|---|---|---|
| **Oxigénio (ml/g/hr)** | 5.2 | 4.7 | 4.5 | 3.6 | 3.4 | 3.1 | 2.7 | 1.8 |

**a)** Calcule os coeficientes da reta de regressão da taxa de consumo do oxigénio em função da temperatura.

**b)** Teste a hipótese, $H_0 : \beta = 0$.

**c)** Calcule o coeficiente de determinação.

**d)** Calcule o intervalo de confiança de 95% para o declive.

**6.** Uma experiência foi realizada para estudar o efeito do efeito de enxofre dissolvido na tensão superficial de cobre líquido. O decrescimento na tensão foi medido para diferentes percentagens de enxofre, de acordo com os dados da tabela.

| **Percentagem $S_2$ (X)** | 0.034 | 0.093 | 0.301 | 0.399 | 0.613 | 0.827 |
|---|---|---|---|---|---|---|
| **Decrescimento (deg/cm) (Y)** | 308 | 426 | 590 | 624 | 649 | 727 |

**a)** Calcule os coeficientes da reta de regressão do decrescimento da tensão em função da percentagem de enxofre. O que pode concluir acerca do ajuste?

**b)** Calcule os coeficientes da reta de regressão do decrescimento da tensão em função do logaritmo neperiano da percentagem de enxofre. O que pode concluir acerca do ajuste?

**7.** Os dados representam os resultados obtidos por 10 alunos num exame, os seus Coeficientes de Inteligência e o número de horas de estudo para o exame:

| Q.I. | 112 | 126 | 100 | 114 | 112 | 121 | 110 | 103 | 111 | 124 |
|---|---|---|---|---|---|---|---|---|---|---|
| Horas | 5 | 13 | 3 | 7 | 11 | 9 | 8 | 4 | 6 | 2 |
| Resultado | 79 | 97 | 51 | 65 | 82 | 93 | 81 | 38 | 60 | 86 |

**a)** Assumindo uma relação linear, estime os valores de $\beta_0$, $\beta_1$ e $\beta_2$.

**b)** Preveja o resultado de um estudante com um Coeficiente de Inteligência de 108 que estudou 6 horas para o exame.

**8.** O Índice de Desenvolvimento de Griffiths é uma medida agregada destinada a avaliar o desenvolvimento psico-motor de crianças. Este índice é calculado através da avaliação de determinadas tarefas motoras e intelectuais. Os dados representam as avaliações, motora e intelectual para 9 crianças com a idade de 4 anos. Calcule o coeficiente de correlação.

| Motor | 84 | 73 | 101 | 74 | 88 | 100 | 86 | 95 | 82 |
|---|---|---|---|---|---|---|---|---|---|
| **Intelectual** | 77 | 85 | 105 | 86 | 108 | 116 | 96 | 100 | 100 |

**9.** A tabela apresenta os resultados de um estudo, conduzido por Howell e Huessy (1981), sobre a relação entre o Coeficiente de Inteligência (QI) (Teste de Capacidade Mental de Otis-Lennon) e a classificação média (CM) no nono ano. O estudo foi conduzido em 30 sujeitos aleatoriamente selecionados, tendo o teste sido administrado enquanto os sujeitos se encontravam entre o quinto e o oitavo ano.

| QI | CM | QI | CM | QI | CM |
|---|---|---|---|---|---|
| 102 | 2.75 | 115 | 4.00 | 111 | 3.00 |
| 108 | 4.00 | 92 | 2.23 | 95 | 1.50 |
| 109 | 2.25 | 95 | 2.50 | 106 | 3.75 |
| 118 | 3.00 | 90 | 2.50 | 83 | 0.67 |
| 79 | 1.67 | 106 | 2.75 | 81 | 1.50 |
| 88 | 2.25 | 85 | 2.75 | 112 | 3.00 |
| 100 | 2.50 | 95 | 2.75 | 85 | 1.75 |
| 92 | 3.50 | 97 | 2.67 | 115 | 3.75 |
| 131 | 3.75 | 93 | 2.00 | 86 | 1.00 |
| 83 | 2.75 | 81 | 2.00 | 85 | 2.50 |

**a)** Produza as estatísticas descritivas para os dados.
**b)** Apresente o diagrama de dispersão.
**c)** Calcule o coeficiente de correlação.

**10.** Calcule o coeficiente de correlação de Pearson para o conjunto de valores X Y apresentados na tabela.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **X** | 1.40 | 4.31 | 6.12 | 2.91 | 1.56 | 7.00 | 3.46 | 4.46 | 0.52 | 1.03 |
| **Y** | 11.91 | 0.16 | 0.11 | 3.16 | 10.62 | 4.16 | 6.16 | 0.01 | 16.12 | 25.32 |
| **X** | 1.41 | 0.43 | 6.22 | 1.54 | 7.15 | 9.28 | 5.78 | 2.62 | 7.25 | 0.37 |
| **Y** | 9.35 | 15.67 | 3.22 | 13.56 | 4.06 | 20.27 | 0.14 | 6.90 | 4.97 | 16.45 |
| **X** | 1.01 | 7.33 | 2.25 | 2.21 | 6.10 | 9.54 | 2.37 | 3.39 | 0.53 | 7.38 |
| **Y** | 12.11 | 4.01 | 7.74 | 4.05 | 0.91 | 18.65 | 7.46 | 3.98 | 20.69 | 8.13 |

Que pode concluir sobre a associação entre as duas variáveis? Quais as condições subjacentes a uma análise deste tipo?