Universidade do Minho

Escola de Engenharia

Departamento de Produção e Sistemas

ESTATÍSTICA APLICADA

Caderno de Exercícios

Mestrado Integrado em Engenharia e Gestão Industrial
Mestrado Integrado em Engenharia de Telecomunicações e Informática

Conteúdo

Ano Letivo 2015/2016

Revisto por Lino Costa

Ficha 1 – POPULAÇÃO E AMOSTRA / ESTATÍSTICA DESCRITIVA

1. Das situações a seguir descritas identifique para cada uma delas, a população, a amostra e a variável em estudo, classificando-a.

   A. Pretende-se fazer um estudo sobre o número de pessoas dos agregados familiares de determinada cidade. Para tal efetuou-se um inquérito ao qual responderam 20 agregados familiares.

   B. Registou-se a população, em milhões, de três dos países da União Europeia (UE) referentes a uma publicação de 1999 com vista a obter uma estimativa para a população da UE.

   C. Registaram-se os tempos que cada funcionário de uma fábrica leva a ir para o trabalho de manhã com o objetivo de proceder otimização dos mesmos.

   D. Observou-se durante uma hora os carros que passavam na Ponte Vasco da Gama e classificou-se cada um de acordo com a sua dimensão (pequeno, médio e grande) com o objetivo de caracterizar qual o carro-tipo.

   E. Pretende-se estimar quanto é que as famílias portuguesas gastam mensalmente em telefone, recolhendo-se para tal os dados relativos a 1024 famílias.

   F. Dos 3500 alunos de uma universidade, 1280 responderam à questão se viviam ou não em casa dos pais enquanto estudam.

2. As notas obtidas por 12 alunos num exame de uma disciplina foram as seguintes: 6, 7, 8, 8, 9, 10, 11, 11, 11, 12, 14, 16. Determine
   a) a moda;
   b) a mediana;
   c) a média;
   d) o 1º quartil e o 3º quartil;
   e) a variância e o desvio padrão.
   f) Desenhe um diagrama de extremos e quartis ("Boxplot").

3. Obteve-se uma amostra de 100 tempos de reação a um estímulo em milissegundos:

10	14	11	15	7	7	20	10	14	9	8	6	12	12	10	14	11	13	9	12
13	11	12	10	8	9	14	18	12	10	10	11	7	17	12	9	9	11	7	10
14	12	12	10	9	7	11	9	18	6	12	12	10	8	14	15	12	11	9	9
11	8	11	10	13	8	11	11	13	20	6	13	13	8	9	16	15	11	10	11
20	8	17	12	19	14	17	12	18	16	15	16	10	20	11	19	20	13	11	20

   a) Construa a tabela de frequências dos valores 6,7,8,…,20;
   b) Desenhe um gráfico de frequências;
   c) Calcule a média;
   d) Determine o 1º e o 3º quartil;
   e) Determine o 10º e o 90º percentis da amostra. Explique em termos da forma da distribuição de frequências, por que razão o 90º percentil está mais afastado da mediana do que o 10º;
   f) Determine a amplitude da amostra.
   g) Desenhe um diagrama de extremos e quartis ("Boxplot").

4. O n.º de chamadas telefónicas (por minuto) recebidas em certa empresa foi registado durante um período de 50 minutos, observando-se os seguintes valores:

   1, 0, 1, 1, 0, 0, 2, 2, 0, 1, 1, 0, 1, 4, 0, 3, 0, 1, 0, 2, 0, 0, 1, 1, 0, 0, 1, 2, 1, 0, 0, 1, 3, 1, 0, 1, 1, 4, 0, 1, 2, 1, 1, 0, 2, 0, 0, 1, 2, 1

   a) Classifique a variável quanto ao tipo e escala de medição;
   b) Construa a tabela de frequências da amostra (absolutas, relativas, absolutas acumuladas e relativas acumuladas);
   c) Determine a média, o desvio padrão, a mediana e a moda da amostra;
   d) Desenhe o gráfico de frequências absolutas.

5. 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) Classifique a variável quanto ao tipo e escala de medição;
b) Proceda à organização dos dados, construindo uma tabela onde figurem as frequências absolutas, relativas, absolutas acumuladas e relativas acumuladas;
c) Calcule a média, o desvio padrão, a mediana e a moda;
d) Desenhe o gráfico de frequências absolutas;
e) Será a distribuição de frequências unimodal? Justifique. Que pode concluir sobre a distribuição de opiniões?
f) Calcule a percentagem de pessoas que têm opinião:
   i) Desfavorável;
   ii) Favorável.

6. 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:

1	8	7	2	4	6	4	9	6
8	2	6	5	2	9	7	5	3
5	5	4	4	3	5	4	4	7
5	4	2	4	4	3	8	3	3
5	5	1	10	6	4	2	6	5

a) Identifique o tipo de variável apresentada;
b) Proceda à organização dos dados construindo uma tabela de frequências onde constem as frequências absolutas, relativas, absolutas acumuladas e relativas acumuladas;
c) Construa o gráfico de frequências absolutas;
d) Calcule a média, a variância, a mediana e a moda da amostra.

7. 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) $\bar{x}-s$ e $\bar{x}+s$
   ii) $\bar{x}-2s$ e $\bar{x}+2s$

8. A tabela seguinte dá a distribuição de frequências da duração de 500 lâmpadas:

Duração(horas)	Nº lâmpadas
300 - 500	50
500 - 700	150
700 - 900	275
900 - 1100	25

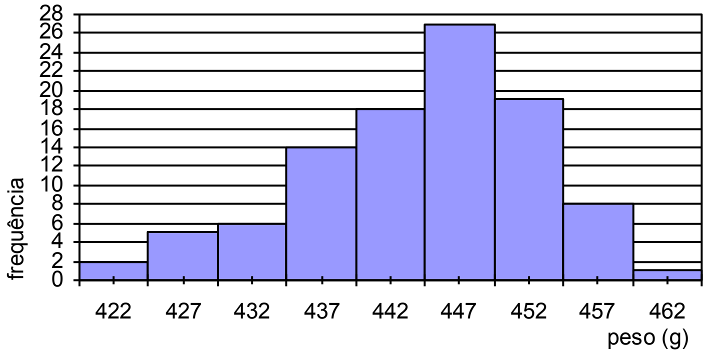
a) Construa o histograma de frequências absolutas;
b) Construa o gráfico de frequências acumuladas;
c) Determine a proporção de lâmpadas que têm uma duração inferior ou igual a 900 horas.

9. Considere a seguinte amostra já classificada

classes	90-93	93-96	96-99	99-102	102-105	105-108	108-111
frequências	1	3	5	7	6	3	3

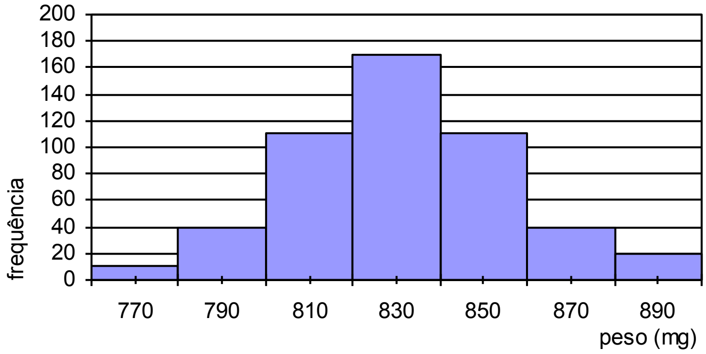
a) Construa o histograma da amostra;
b) Calcule a média e o desvio padrão amostrais.

10. 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?

11. 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?

Ficha 2 - TEORIA DA PROBABILIDADE

1. Tira-se uma carta de um baralho de 52. Considere os acontecimentos relativos à experiência:
A1={extração de um ás do baralho}, A2={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. Existem 3 dados, o primeiro contém 3 senas, o 2º é viciado e o 3º é um dado normal (uma sena). A probabilidade de sair uma sena no 2º dado viciado é de 2/6. Sabendo que se escolheu um dado ao acaso, qual a probabilidade de ao efetuar o lançamento sair uma sena?

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) $P(A \mid B)$, e) $P(B \mid A)$, f) $P(\overline{A} \cap \overline{B})$, g) $P(\overline{A} \mid \overline{B})$, h) $P(\overline{B} \mid \overline{A})$

8. Sejam M1, M2 acontecimentos independentes, tais que $P(M1 \cup M2) = 0.8$ e $P(M1 \mid M2) = 0.2$. Calcule $P(M2)$.

9. O João toma o autocarro com probabilidade 0.3 e o metro com probabilidade 0.7. O João chega atrasado 40% das vezes sempre que toma o autocarro mas só 20% das vezes sempre que usa o metro. Qual é a probabilidade de que o João chegue tarde ao emprego?

10. 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.

11. 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?

12. Um teste de diagnóstico para uma determinada doença é 90% fiável, isto é, se a pessoa tem a doença o teste é positivo com uma probabilidade de 0.9. Se a pessoa não tem a doença, o teste é negativo com uma probabilidade de 0.9. Se só 1% da população tem a doença em questão, qual a probabilidade de que uma pessoa selecionada ao acaso, cujo teste seja positivo, tenha realmente a doença?

13. Três fábricas produzem 20%, 30% e 50% dos circuitos integrados que uma determinada companhia vende. Se as frações de defeituosos são 0.4%, 0.3% e 0.2%, respetivamente. Encontrou um circuito integrado defeituoso, qual a probabilidade de ter vindo da terceira fábrica?

14. Num consultório médico, um cliente pertence a uma das classes A. B ou C de acordo com a propensão para adoecer ao longo do ano, considerando-se que, dentro de classe, a probabilidade de que isso aconteça é respetivamente 0.2, 0.4 e 0.1. Em determinado momento o arquivo do consultório tem a seguinte composição: classe A: 300 clientes; classe B: 520 clientes; classe C: 180 clientes. Admita que se extraviou uma ficha de um cliente doente. Qual a classe em que mais provavelmente o incluiria?

Ficha 3 - DISTRIBUIÇÕES DE PROBABILIDADE / ESPERANÇA MATEMÁTICA

1. Determinar se os valores dados podem ser usados como os valores de uma distribuição 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 distribuições de probabilidade na gama de valores dada.

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

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

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

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

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. Considere $f(x)=\dfrac{x}{6}$, definida para x=1,2,3. Calcule:

   a) a função distribuição acumulada *F(x)* e represente-a graficamente

   b) $P(X \le 2)$

   c) $P(1 < X \le 3)$

6. 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)

7. 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

a) $P(X \le 3)$

b) $P(X=3)$

c) $P(X<3)$

d) $P(-0.4<X<4)$

e) $P(X=5)$

f) $f(x)$

8. Para cada uma das funções, determine a constante c por forma a que a função f(x) satisfaça as propriedades de uma função de distribuição de uma v.a. X,

a) $f(x) = \frac{x}{c}$, x=1, 2, 3, 4 b) $f(x) = cx$, x=1,2,…,10

9. Considere a v.a. X, cuja função de probabilidade é definida por $f(x) = \frac{|x-2|}{7}$ para x=-1,0,1,3.

a) $P(-1 < X \le 3)$ b) $P(X > 0)$ c) Qual o valor esperado de x?

10. Uma cadeia de lojas de desporto gostaria de determinar o número de bicicletas de exercício a ter em cada uma das suas lojas. Assuma que a distribuição de probabilidade da procura diária de bicicletas numa determinada loja é a seguinte.

x	0	1	2	3	4
f(x)	0.15	0.30	0.40	0.10	0.05

a) Qual a probabilidade de uma procura diária superior a 2 bicicletas?
b) E a probabilidade de uma procura diária maior que 1 bicicleta e menor ou igual a 4 bicicletas?
c) Qual o valor esperado da procura diária de bicicletas?

11. O Hotel Bosque Verde tem 160 quartos. O número de unidades de ar condicionado que deve ser substituído durante o Verão tem a seguinte distribuição de probabilidade

x	0	1	2	3	4
f(x)	0.30	0.35	0.20	0.10	0.05

a) Qual o número esperado de unidades de ar condicionado que têm de ser substituídos no Verão?
b) Qual a variância e o desvio padrão do número de unidades substituídas?

12. A procura de um produto varia consideravelmente de mês para mês. Baseado no histórico dos últimos 2 anos, determinou-se a seguinte distribuição de probabilidades da procura mensal desse mesmo produto:

x	300	400	500	600
f(x)	0.20	0.30	0.35	0.15

a) Se a empresa programar ordens de produção mensais baseadas no valor esperado da procura mensal, qual deverá ser a quantidade mensal a planear?
b) Qual a variância e o desvio padrão do número de unidades requeridas?

13. Para cada uma das funções, determine a constante c por forma a que a função f(x) satisfaça as propriedades de uma função de distribuição de uma v.a. X,

a) $f(x) = cx\exp(-x)$, $0 < x < \infty$ b) $f(x) = c(1-x)$, $0 < x < 1$

c) $f(x) = c\sqrt{x}$, $0 \le x \le 4$ d) $f(x) = 4x^c$, $0 \le x \le 1$

14. Seja X uma v.a. com f.d.p. $f(x) = 2x$, definida para $0 < x < 1$. Calcule:

a) $P(1/2 < x < 3/4)$ b) $P(-1/2 < x < 1/2)$

15. Para cada um dos casos seguintes, calcule a função distribuição acumulada, F(x) e represente-a graficamente:

a) $f(x) = 2/x^3$, definida para $1 < x < \infty$

b) $f(x) = 3(1-x)^2$, definida para $0 < x < 1$

c) $f(x) = 1/3$, definida nos intervalos $0 < x < 1$ e $2 < x < 4$

16. Considere a latitude angular θ(x), como função da variável aleatória X, com função de distribuição

$$F_\theta(x) = \begin{cases} 0 & x < 0 \\ 1-\cos(x) & 0 \le x \le \pi/2 \\ 1 & x > \pi/2 \end{cases}$$

Calcule a correspondente função densidade de probabilidade e represente-a graficamente.

17. Dada a função de distribuição acumulada

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

Calcule:

a) $P(-1/2 < X \le 1/2)$

b) $P(X = 0)$

c) $P(2 < X \le 3)$

18. Considere a seguinte função densidade de probabilidade

$$f(x) = \begin{cases} 0.2 & -1 < x \le 0 \\ 0.2 + cx & 0 < x \le 1 \\ 0 & \text{outros valores} \end{cases}$$

a) Determine o valor de c;
b) Determine F(x);
c) Calcule F(-1), F(0) e F(1);
d) Calcule $P\left(0 \le X \le 0.5\right)$;
e) Calcule $P\left(X > 0.5 \mid X > 0.1\right)$

19. Seja X uma v.a. com f.d.p. $f(x) = 2(1-x)$, definida para $0 < x < 1$.Calcule:

a) E[X];
b) $E[X^2]$;
c) V[X].

20. Considere a seguinte função:

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

a) Verifique que se trata de uma função densidade de probabilidade.
b) Determine E[X];
c) Determine V[X].

21. O tempo requerido pelos alunos para completarem um exame de uma hora é uma variável aleatória com função densidade de probabilidade dada por

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

a) Determine o valor de c;
b) Determine F(x);
c) Calcule F(-1), F(0) e F(1);

d) Encontre a probabilidade de um aluno selecionado aleatoriamente termine o exame em menos de meia hora.
e) Qual o tempo médio de realização do exame pelos alunos?

22. O erro de pesagem (g) de uma determinada balança é um fenómeno aleatório com uma função densidade de probabilidade dada por

$$f(x) = \begin{cases} 0 & x \leq 0 \\ x/2 & 0 < x \leq 1 \\ 0 & 1 < x \leq 2 \\ 1/4 & 2 < x \leq 5 \\ 0 & x > 5 \end{cases}$$

Determine:
a) A probabilidade de cometer um erro de pesagem superior a 3 gramas.
b) A probabilidade de cometer um erro de pesagem entre 1 e 3 gramas.
c) O erro esperado de pesagem.
d) A variância do erro de pesagem.

23. Um conhecido fabricante de máquinas fotográficas desenvolveu um novo tipo de flash para um novo modelo de câmara fotográfica. Com base nos testes laboratoriais a equipa de I&D indicou a seguinte função densidade de probabilidade para a duração (horas) do novo flash:

$$f(t) = \begin{cases} \dfrac{t.\exp\left(-\dfrac{t}{10}\right)}{100} & 0 < x < \infty \\ 0 & x \leq 0 \end{cases}$$

Calcule:
a) A probabilidade do flash durar mais de 20 horas.
b) A probabilidade da duração do flash variar entre 10 e 20 horas.
c) a duração esperada do flash.

24. Um gerente de um banco está muito preocupado com a qualidade de atendimento ao balcão em termos de filas de espera. Com base numa observação sistemática determinou que o tempo que um cliente tem de esperar para ser atendido segue uma variável aleatória com a seguinte função densidade:

$$f(t) = \begin{cases} 1/2 & 0 < t < 1 \\ 1/4 & 2 < t < 4 \end{cases}$$

Determine:
a) A probabilidade de um cliente esperar mais de um minuto para ser atendido.
b) A probabilidade de um cliente esperar entre 2 a 3 minutos para ser atendido.
c) A probabilidade de um cliente esperar pelo menos 3 minutos para ser atendido.
d) A probabilidade de esperar no máximo 3 minutos para ser atendido, sabendo que um cliente já esperou pelo menos 1 minuto.
e) O tempo médio de espera de um cliente do banco.
f) O desvio padrão do tempo de espera de um cliente do banco.

Ficha 4 - FAMÍLIAS DE DISTRIBUIÇÕES DE PROBABILIDADES

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 uma ser defeituosa?
   b) Se forem examinadas cinco 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 tal modo que a probabilidade de não haver teares a aguardar reparação seja pelo menos de 90%.

9. As chamadas telefónicas chegam a uma central telefónica a uma média de 4 por minuto. Determine a probabilidade de que num intervalo de 15 segundos ocorram 3 ou mais chamadas?

10. Pequenos defeitos ocorrem na produção de uma fita de seda à média de um por 300 m. Supondo que o número de defeitos num dado comprimento de fita segue a distribuição de Poisson, qual a probabilidade de que:

a) Um rolo de 720 m tenha quanto muito 2 defeitos?

b) Um rolo com 360 m não tenha defeitos?

11. 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?

12. 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?

13. Somente 3% dos estudantes de uma cidade têm coeficiente de inteligência igual ou superior a 130. Com uma amostra aleatória de 50 estudantes calcule:

a) $P(X = 2)$

b) $P(X \geq 3)$

14. Um teste de múltipla escolha consiste em oito questões com três respostas possíveis a cada questão (em que apenas uma é correta). Se um estudante responder a cada questão com o lançamento de um dado perfeito e assinalar a primeira resposta se obtém um 1 ou 2, a segunda resposta se obtém um 3 ou 4, e a terceira se obtém um 5 ou 6, qual a probabilidade de ter exatamente quatro respostas corretas?

15. 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.

16. 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?

17. Uma empresa têxtil inglesa desenvolveu uma máquina energeticamente eficiente que combina a pressão de lavagem com a limpeza a vapor. Desenvolvida para gastar 7 galões de detergente por minuto a uma pressão de lavagem de 1000 psi, na realidade a máquina consome entre 6.5 e 7.5 galões de detergente por minuto. Assuma que y, a quantidade de detergente consumida, é uma variável aleatória uniforme e encontre a probabilidade da máquina gastar:

a) Mais de 7.2 galões/minuto de detergente;

b) Menos de 7 galões/minuto de detergente.

18. 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

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

Calcule a probabilidade de um componente durar

a) menos de 4×10^3 horas.

b) entre 5×10^3 e 10×10^3 horas.

19. 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 θ=40. Encontre as probabilidades de um desses pneus dure.

a) no mínimo 20 000 quilómetros;

b) no máximo 30 000 quilómetros.

20. O tempo que um relógio de pêndulo trabalha sem necessidade de dar corda é uma variável aleatória com uma distribuição exponencial com θ=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.

21. Se Z e uma variável aleatória seguindo a distribuição normal padrão, encontre

a) $P(Z<1.33)$;

b) $P(Z \leq -0.79)$;

c) $P(0.55<Z<1.22)$;

d) $P(-1.90 \leq Z \leq 0.44)$.

22. Encontre z se a área debaixo da curva normal padrão

a) entre 0 e z é 0.4726;

b) à esquerda de z é 0.9868;

c) à direita de z é 0.1314;

d) entre -z e z é 0.8502.

23. Se Z é uma variável aleatória seguindo a distribuição normal padrão, encontre os respetivos valores z_1, z_2, z_3 e z_4 tal que

a) $P(0 < Z < z_1) = 0.4306$;

b) $P(Z \geq z_2) = 0.7704$;

c) $P(Z > z_3) = 0.2912$;

d) $P(-z_4 \leq Z < z_4) = 0.9700$.

24. 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?

25. 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 satisfazem a garantia?

b) O fabricante decidiu substituir apenas 1% de todos os sensores. Qual deverá ser a duração da garantia (em meses)?

26. 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?

27. 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.

28. Num processo fotográfico, o tempo de processamento da imagem pode ser visto como uma variável aleatória que segue uma distribuição normal com μ=15.40 segundos e σ=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.

29. Uma fábrica de sapatos sabe que a medida dos pés dos clientes (senhoras) segue a lei normal, cuja média é 36 cm 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?

30. 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.

31. Sabe-se que 30% dos estudantes de uma universidade frequentaram colégios particulares. Assuma uma amostra aleatória de 50 estudantes dessa universidade.

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?

32. De um questionário realizado 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 5 - DISTRIBUIÇÕES AMOSTRAIS E ESTIMAÇÃO PONTUAL

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) $P(320 \le \bar{X} \le 322)$
   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 μ. 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 $\mu = 25$, qual a probabilidade de a amostra indicar que a máquina não está controlada?
   b) Se a média mudar para $\mu = 25.3$, 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) $Var\left[\bar{X}_1 - \bar{X}_2\right]$
   b) $P\left(\left|\bar{X}_1 - \bar{X}_2\right| > 4\right)$

6. Dada a f.d.p. da variável aleatória X:

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

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

7. Baseando-se numa amostra de tamanho 3 considere três estimativas possíveis para μ (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.

8. 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 μ:

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

   Qual dos dois estimadores deve preferir?

Ficha 6 - INTERVALOS DE CONFIANÇA

1. Se uma amostra aleatória de tamanho n=20 duma população normal com variância σ^2=225 tem média $\bar{x} = 64.3$:
   a) Construa o intervalo de confiança de 95% para a média da população μ.
   b) Construa também o intervalo de 90% de confiança.
   c) Qual seria o tamanho mínimo da amostra para que o erro não ultrapassasse 3.5 com 95% 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. Assuma 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. Para estudar o crescimento das árvores de pinheiro, um trabalhador registou 40 medições das alturas de árvores de 1 ano de idade. Os valores obtidos foram:

2.6	1.9	1.8	1.6	1.4	2.2	1.2	1.6
1.6	1.5	1.4	1.6	2.3	1.5	1.1	1.6
2.0	1.5	1.7	1.5	1.6	2.1	2.8	1.0
1.2	1.2	1.8	1.7	0.8	1.5	2.0	2.2
1.5	1.6	2.2	2.1	3.1	1.7	1.7	1.2

   a) Dê uma estimativa pontual das médias das alturas da população dos pinheiros.
   b) Com 95.4% de certeza, qual o limite de erro conhecido?

4. 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 μ.

5. A intensidade da corrente, em amperes, num dado circuito elétrico é uma variável aleatória normalmente distribuída.
   a) Foram efetuadas 35 medições independentes da intensidade de corrente no circuito, tendo-se obtido, para esta amostra, uma média de 3.1 amperes e um desvio padrão de 0.5 amperes. Construa um intervalo de confiança a 95% para o valor médio da intensidade de corrente no circuito. Interprete o seu resultado.
   b) Para um nível de confiança de 95%, qual seria o tamanho de amostra mínimo para que o erro da estimativa não ultrapassasse os 0.1 amperes, supondo que a intensidade da corrente no circuito elétrico tem um desvio padrão de 0.5 amperes.

6. Uma amostra aleatória de tamanho *n* é retirada duma população com média μ e desvio padrão σ. A média e o desvio padrão da amostra são $\bar{x}$ =45 e s=5.8. Calcule o intervalo de confiança de 95% para μ usando os seguintes tamanhos da amostra:
   a) *n*=30 b) *n*=60 c) *n*=90
   Compare os tamanhos dos três intervalos.

7. 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 $P\left[\bar{x}_1-\bar{x}_2-\frac{\sigma}{5}<\mu_1-\mu_2<\bar{x}_1-\bar{x}_2+\frac{\sigma}{5}\right]=0.90$.

8. 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 μ_x for a capacidade média da população antes do tratamento e μ_y for a capacidade média da população depois do tratamento, construa um intervalo de confiança que tenha 90% de probabilidade de conter $\mu_y - \mu_x$. Defina as condições de aplicabilidade do intervalo.

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

9. 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?

10. Duas amostras independentes de tamanho n_1 e n_2 foram retiradas de duas populações com médias μ_1 e μ_2 e desvios padrão σ_1 e σ_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?

11. 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.

c) Qual a dimensão da amostra a ser recolhida para que o erro não ultrapasse os 3% para uma probabilidade de 95.4%?

12. Suponha que pretende estimar a percentagem de estudantes que são contra a introdução de propinas. Para o efeito recolhe uma amostra de 225 estudantes dos quais 144 manifestam a sua oposição.

a) Estabeleça um intervalo de confiança de 95% para a percentagem de estudantes que se opõem à introdução de propinas.

b) Qual o erro da estimativa calculada na alínea anterior?

c) Suponha que pretende um erro da estimativa inferior a 5%. Qual a dimensão da amostra a ser recolhida, para uma confiança de 95%?

13. 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.

14. 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.

15. 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.

16. 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.

17. 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.

18. 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.

19. Um relojoeiro pretende conhecer as variações do produto que fabrica. Para construir um intervalo de confiança para σ, 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 σ.

20. 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 que 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.

21. 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 σ_1^2 / σ_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 7 - TESTES DE HIPÓTESES

1. 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.

2. 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.

3. Numa fábrica o chefe de produção afirma que 40% das máquinas de escrever vendidas naquela região, são produtos da sua fábrica. Considerando $p = 0.4$ como a hipótese nula, o chefe decide considerar a sua afirmação como aceitável, a não ser que dentre 19 máquinas se verificar que $X \leq 3$ ou $X \geq 12$, sendo X o número de máquinas de escrever vendidas pela sua fábrica.
a) Determine o nível de significância deste teste.
b) Determine a potência do teste para vários valores de p, desde 0.1 a 0.9.

4. 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.

5. 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 θ=2, 4, 6, 8, 12, 16 e 20
c) Apresente o gráfico da função potência para este teste.

6. 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 μ = 37.0, 38.0, 39.0 e 40.0.
b) Encontre as probabilidades de erros do tipo II, quando μ = 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.

7. 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?

8. 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?

9. 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.

10. 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 $\alpha = 0.01$.

11. 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 $\alpha = 0.05$.

12. 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.

13. 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.

14. 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.

15. 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?

16. 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 que 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 $\alpha = 0.05$.

17. 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.

18. 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?

19. 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 (α=0.05).

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

21. 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 α=0.01.

22. 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 α=0.05.

23. 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 controlada 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%.

24. 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.

25. 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 8 - ANÁLISE DE 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 α=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. O aumento de peso de mulheres grávidas parece ter um efeito importante no peso dos bebés. Se o aumento de peso não é adequado, a criança tem maiores probabilidades de ser pequena e tenderá a ser menos saudável. Num estudo conduzido em 3 países, registaram-se os aumentos de peso (em kg) das mulheres durante o 3º trimestre de gravidez:

País	n_j	$\overline{Y}_j$	s_j
Egipto	46	3.7	2.5
Kénia	111	3.1	1.8
México	52	2.9	1.8

Suponha que a variável *Y≡aumento de peso* segue a distribuição normal.

a) Calcule a variância residual $s_r^2 = \dfrac{\sum_{j=1}^{k}(n_j-1)s_j^2}{\sum_{j=1}^{k}(n_j-1)}$ onde k é o número de tratamentos.

b) A variância residual é um estimador não tendencioso da variância σ^2 da distribuição, isto é, da média dos quadrados dos resíduos (MQR). Tendo em conta a variância residual calculada, teste a hipótese nula de que em média o aumento de peso é o mesmo nas mulheres grávidas dos três países observados.

5. 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

a) Teste ao nível de significância de α=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.

b) Estime também os parâmetros μ, α_1, α_2, α_3 para o modelo usado na análise.

6. Um serviço de apoio ao consumidor, pretende testar a precisão de três tipos de ferros elétricos, à temperatura de 480ºF. As leituras obtidas, através de um termóstato, foram as seguintes:

Ferro X	474	496	467	471
Ferro Y	492	498		
Ferro Z	460	495	490	

Considere as leituras normalmente distribuídas. Teste ao nível de significância de α=0.05 se as diferenças entre as três médias podem ser atribuídas ao acaso.

7. Um estudo, em doentes com diabetes dependente de insulina, foi conduzido para investigar os efeitos do fumo em complicações ao nível dos rins e da retina. Antes de examinar os resultados do estudo, pretende-se comparar as pressões sistólicas nos quatro grupos: não fumadores, atuais fumadores, ex-fumadores e consumidores de tabaco de marcar. Uma amostra é selecionada de cada grupo; os dados relevantes são fornecidos na tabela. As médias e os desvios padrões são medidos em miligramas de mercúrio:

	n	$\bar{x}$	s
Grupo 1	269	115	13.4
Grupo 2	53	114	10.1
Grupo 3	28	118	11.6
Grupo 4	9	126	12.2

Calcule a variância residual e teste a hipótese de que as pressões sistólicas médias dão iguais nos quatro grupos, assumindo as pressões normalmente distribuídas.

8. Uma associação de consumidores quis comparar o preço (em unidades monetárias, u.m.) de um brinquedo particular em 3 tipos de lojas: hipermercados, bazares, e lojas de brinquedos. Selecionaram-se aleatoriamente 5 hipermercados, 5 bazares e 5 lojas de brinquedos e os preços foram registados para cada uma delas. Realizada uma Análise de Variância chegou-se à seguinte tabela ANOVA:

Fonte de variação	***SQ***	***Gl***	***MQ***	***F***
Tratamentos	______	______	**9.867**	______
Resíduos	______	______	______	
Total	**47.733**	______		

a) Formule as hipóteses do teste estatístico.
b) Complete a tabela ANOVA.
c) O que pode concluir? ($\alpha = 0.05$)
d) Indique quais os pressupostos do teste realizado.

9. 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

10. Uma experiência realizada numa estação agronómica consiste em testar os efeitos de 5 níveis diferentes de aplicação de potassa nos campos sobre as propriedades do algodão. A medida escolhida para testar essas propriedades foi o índice de Pressley. Tomaram-se amostras de algodão de cada um dos campos e efetuaram-se 4 determinações da resistência por cada amostra. Os resultados apresentados na tabela correspondem aos valores médios dessas determinações. A experiência foi efetuada em três blocos casuais de cinco campos cada um.

Campos	Qdade de potassa	B1	B2	B3
I	36	7.62	8.00	7.93
II	54	8.14	8.15	7.87
III	72	7.76	7.73	7.74
IV	108	7.17	7.57	7.80
V	144	7.46	7.68	7.21

Face a estes resultados, verifique se existem diferenças significativas na resistência do algodão, devido à aplicação de níveis diferentes de potassa. Considere a variável resistência normalmente distribuída.

11. 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?

12. 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 conclusões pode tirar? Justifique
b) Face às conclusões que tirou em a) que modelo populacional lhe parece que seria o mais adequado?

13. A tabela apresenta os teores de colesterol, em miligramas por embalagem, que quatro laboratórios obtiveram para embalagens de três alimentos dietéticos semelhantes:

	Alim A	Alim B	Alim C
Lab 1	3.4	2.6	2.8
Lab 2	3.0	2.7	3.1
Lab 3	3.3	3.0	3.4
Lab 4	3.5	3.1	3.4

Teste, através duma análise de variância de duas entradas, as hipóteses relativas aos laboratórios e aos alimentos dietéticos, ao nível de significância de $\alpha=0.05$

14. Antes de apresentarem uma proposta de orçamento, os engenheiros prepararam uma análise detalhada dos custos estimados de trabalho e materiais, necessários à conclusão da obra. Esta estimativa dependerá do engenheiro que faz a análise. Uma estimativa se é muito elevada, reduzirá as hipóteses de aceitação do orçamento, e se muito baixa, reduzirá o lucro da companhia ou poderá conduzir a um prejuízo. Uma companhia que emprega três engenheiros, quer comparar os seus níveis médios de estimativas. Isto foi feito fazendo cada engenheiro estimar o custo das mesmas quatro obras. Os custos em milhares de euros são apresentados na tabela. Faça uma análise da variância e teste se há evidência suficiente para indicar diferenças entre os orçamentos feitos pelos engenheiros nas quatros obras. Suponha a variável orçamento normalmente distribuída.

	Obra 1	Obra 2	Obra 3	Obra 4
Eng 1	4.6	6.2	5.0	6.6
Eng 2	4.9	6.3	5.4	6.8
Eng 3	4.4	5.9	5.4	6.3

Ficha 9 - REGRESSÃO E CORRELAÇÃO

1. Pensa-se que a frequência do chilrear de um grilo está relacionada com a temperatura. Isto sugere a possibilidade da temperatura poder ser prevista a partir da frequência do chilrear. Registaram-se os seguintes valores de frequência e da temperatura para o grilo listado:

freq. do chilrear / seg, (X)	20	16	20	18	17	16	15
temperatura, °F, (Y)	89	72	93	84	81	75	70

a) Ajuste um modelo linear aos dados.

b) Estime a temperatura para uma frequência de 19.

c) Calcule o coeficiente de determinação e represente graficamente os resíduos. O que conclui sobre o modelo?

2. No âmbito de uma auditoria internacional a uma Agência Governamental, tentou verificar-se a fiabilidade do Índice de Preços no Consumidor publicado pela Agência (IPC oficial). Para o efeito, recalculou-se aquele índice seguindo uma metodologia considerada. Na tabela seguinte apresentam-se os valores dos dois índices, o "correto" e o oficial, para um período de seis anos.

IPC correto (X)	IPC oficial (Y)
112.2	108.8
123.5	119.7
131.0	128.0
138.6	135.8
145.7	142.5
149.4	147.0

a) Ajuste um modelo de regressão linear simples.

b) Teste a hipótese de que não existe relação entre os índices "correto" e oficial. O que pode concluir?

c) Estime o IPC oficial para um IPC correto de 120.

d) Calcule o coeficiente de determinação e represente graficamente os resíduos. O que conclui sobre o modelo?

3. Numa empresa de telecomunicações realizaram-se testes a 12 funcionários com o objetivo de averiguar se existe alguma relação entre a eficiência em pesquisas na Internet e a classificação obtida num teste de aptidão. Na tabela seguinte encontram-se os respetivos resultados obtidos na eficiência, X, e no teste de aptidão, Y.

X	86	70	75	68	91	72	77	92	69	71	88	87
Y	89	74	76	64	96	73	65	90	66	80	81	79

a) Estime o modelo de regressão linear.

b) Verifique se existe relação linear entre o teste de aptidão e a eficiência em pesquisas na Internet.

c) Calcule o coeficiente de determinação e represente graficamente os resíduos. O que conclui sobre o modelo?

d) Indique uma estimativa de Y para X=80 e X=60.

4. Os dados representam o rendimento per capita (milhares de dólares) em função do número de habitantes de alguns países (em milhões).

Pop X	54	42	28	38	25	70	48	41	20	52	65
Rend Y	6	16	33	18	41	3	10	14	45	9	5

a) Ajuste um modelo do tipo: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 \dfrac{1}{X_i}$.

b) Estime o rendimento per capita em milhares de dólares para um país com 50 milhões de habitantes.

c) Teste a hipótese de que não existe relação entre o rendimento e a população. O que pode concluir?

d) Calcule o coeficiente de determinação e represente graficamente os resíduos. O que conclui sobre o modelo?

5. A tabela apresenta a relação entre as temperaturas máximas do dia e a procura máxima de eletricidade, respetivamente. Pretende-se investigar se há uma relação linear entre o pico de procura de eletricidade e a temperatura máxima diária observada.

X (ºC)	33	29	35	39	31	36
Y (MW)	207	139	211	273	156	244

a) Ajuste um modelo do tipo: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$

b) Estime o pico de procura de eletricidade para uma temperatura máxima diária observada de 30 ºC.

c) Teste a hipótese de que não existe relação entre o pico de procura de eletricidade e a temperatura máxima diária.

d) Calcule o coeficiente de determinação e represente graficamente os resíduos. O que conclui sobre o modelo?

6. Considere a seguinte tabela de valores:

X	68	66	72	73	66
Y	64	66	71	70	69

Calcule o coeficiente de correlação da amostra. Será significativa para 5%?

7. Um casal que costuma jogar “bowling” registou os pontos ganhos por cada um, em 10 tentativas, para saber se haveria alguma correlação entre os pontos. Os valores obtidos na experiência foram:

tentativa	1	2	3	4	5	6	7	8	9	10
pontos dele	147	158	131	142	183	151	196	129	155	158
pontos dela	122	128	125	123	115	120	108	143	124	123

Calcule o coeficiente de correlação desta amostra. Será significativa para 5%?

Ficha 10 - TESTE DE BOM AJUSTE DE QUI-QUADRADO

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 (α=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 α=0.01 e α=0.05.

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 λ=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	≥ 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. Fez-se um estudo relativo aos defeitos apresentados por peça de um tecido e obtiveram-se os seguintes valores:

Nº Defeitos	*0*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Frequência	8	10	15	12	10	9	4	1	0	1	0

Faça uma análise estatística destes resultados.

7. 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

8. 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.

9. 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

10. 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.