



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

ESTATÍSTICA APLICADA

2010/2011

Profª Ana Cristina Braga

17-09-2010

ÍNDICE

FICHA Nº 1 - DESCRITIVA	3
FICHA Nº 2 - PROBABILIDADES.....	6
FICHA Nº 3 - DISTRIBUIÇÕES DE PROBABILIDADE	8
FICHA Nº 4 - ESPERANÇA MATEMÁTICA.....	11
FICHA Nº 5 - FAMÍLIAS DE DISTRIBUIÇÕES	12
FICHA Nº 6 - DISTRIBUIÇÕES AMOSTRAIS	16
FICHA Nº 7 - ESTIMADORES PONTUAIS	17
FICHA Nº 8 - INTERVALOS DE CONFIANÇA	19
FICHA Nº 9 - TESTES DE HIPÓTESES	22
FICHA Nº 10 - QUI-QUADRADO	27
FICHA Nº 11 - ANÁLISE DA VARIÂNCIA	30
FICHA Nº 12 - TESTES NÃO PARAMÉTRICOS	33
FICHA Nº 13 - REGRESSÃO E CORRELAÇÃO	38

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

- Proceda à organização dos dados, construindo uma tabela onde figurem as frequências absolutas, relativas, absolutas acumuladas e relativas acumuladas;
 - Calcule a média, o desvio padrão, a mediana e a moda;
 - Desenhe o histograma de frequências absolutas;
 - Será a distribuição de frequências unimodal? Justifique.
 - Que pode concluir sobre a distribuição de opiniões?
 - Calcule a percentagem de pessoas que têm opinião:
 - Desfavorável;
 - 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:

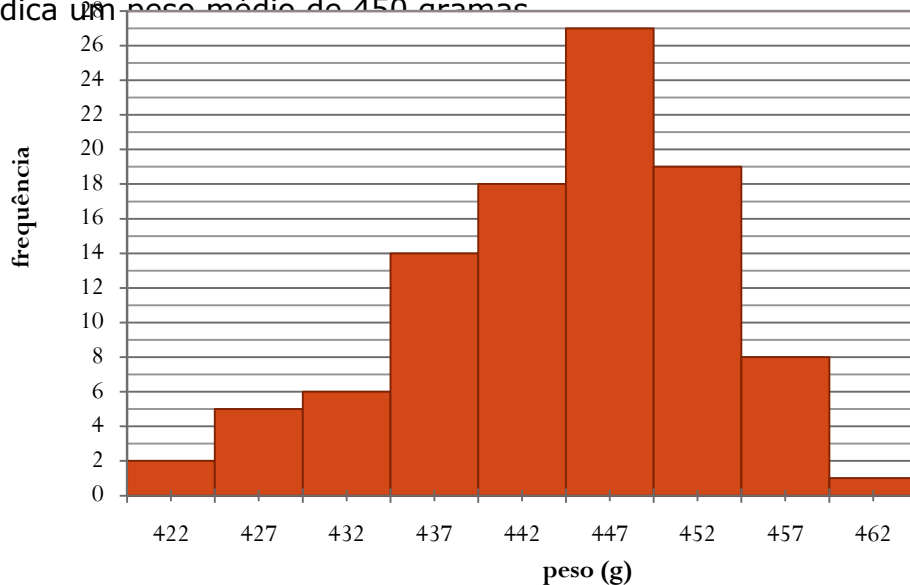
1 8 7 2 4 6 4 9 6
 8 2 6 5 2 9 7 5 3
 5 5 4 4 3 5 4 4 7
 5 4 2 4 4 3 8 3 3
 5 5 1 10 6 4 2 6 5

- Identifique o tipo de variável apresentada;
 - Proceda à organização dos dados construindo uma tabela de frequências onde constem as frequências absolutas, relativas, absolutas acumuladas e relativas acumuladas;
 - Construa o gráfico de frequências absolutas;
 - Calcule a média, a variância, a mediana e a moda da amostra.
3. O vencimento/hora de 100 operários é dado pela tabela

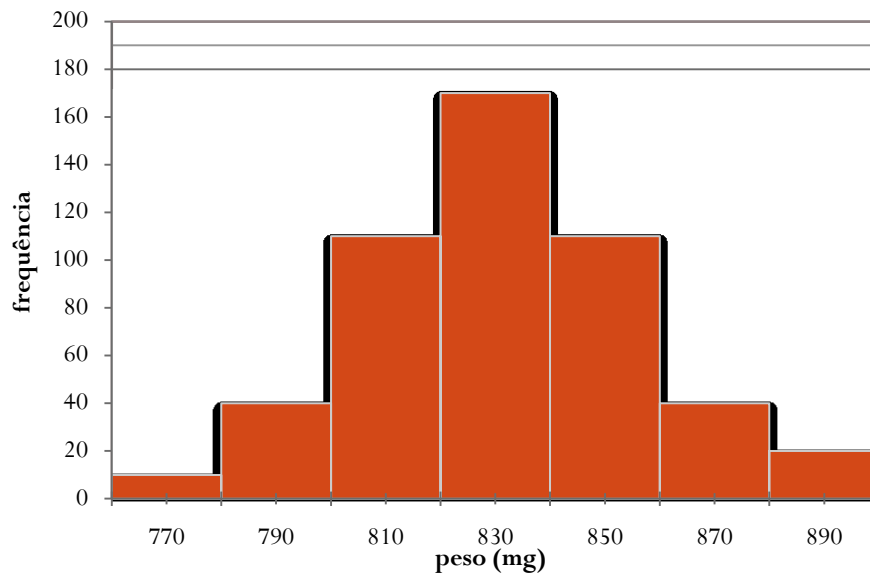
Vencimento	Nº operários
[120 ;125[10
[125 ;130[20
[130 ;135[38
[135 ;140[25
[140 ;145[7

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

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



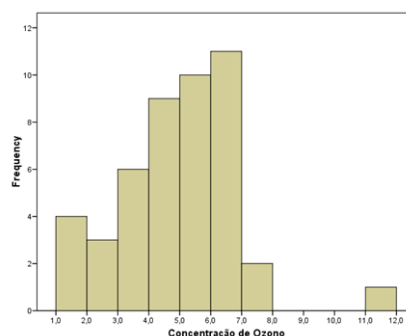
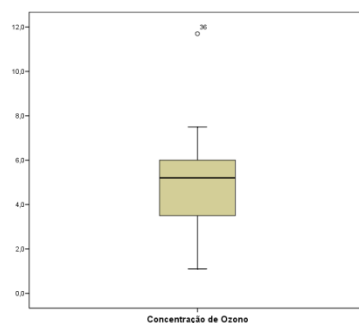
- Apresente a tabela das frequências relativas;
 - Calcule a média da amostra e o desvio padrão das observações;
 - Qual a proporção de latas com peso superior a 450 g?
 - O que pode concluir sobre a forma da distribuição?
5. A distribuição dos pesos (mg) de 500 cigarros de uma determinada marca está representada no seguinte histograma



- Calcule a média, a mediana, a moda e a variância da amostra;
 - Qual a proporção de cigarros com menos de 820 mg?
 - Qual a proporção de cigarros cujo peso está compreendido entre 800 e 880 mg?
6. Considere os dados referentes a doentes com cancro colorectal e respectivo estadio da doença

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			

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

- Classifique a variável (tipo e escala) e comente o quadro das estatísticas resumo.
- Comente as representações gráficas (a) e (b). Que pode afirmar quanto à simetria da distribuição? Justifique.
- 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: $A_1 = \{\text{extração de um ás do baralho}\}$, $A_2 = \{\text{extração de uma carta de espadas do baralho}\}$
 - a) A_1 e A_2 são independentes?
 - b) A_1 e A_2 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
 $A = \text{"soma dos resultados igual a 7"}$
 $B = \text{"ambos os resultados ímpares"}$
 $C = \text{"produto dos resultados igual a 12"}$
Determine: a) $P(A \cup C)$, b) $P(A \cup B)$
4. Dois indivíduos A e B estão afectados 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, respectivamente 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 $\frac{2}{3}$ de probabilidades de ser atingido pelo mau tempo, e, independentemente disso, $\frac{1}{4}$ de probabilidades de ter uma colisão com um petroleiro. Definindo os seguintes acontecimentos: $M = \{\text{ser atingido pelo mau tempo}\}$ e $C = \{\text{colisão com um petroleiro}\}$, calcule
 - a) $P(M \cap C)$, b) $P(M \cap \bar{C})$, c) $P(\bar{M} \cap C)$, d) $P(M \cup C)$
6. Seja A e B acontecimentos independentes e $P(A) = \frac{1}{6}$ e $P(B) = \frac{1}{4}$. Determine
 - a) $P(A \cap B)$, b) $P(A \cup B)$, c) $P(\bar{A} \cap \bar{B})$, d) $P(A \cap \bar{B})$
7. Sejam A e B dois acontecimentos tais que $P(A) = \frac{1}{4}$, $P(B) = \frac{2}{3}$ e $P(A \cap B) = \frac{1}{6}$. Determine
 - a) $P(A \cup B)$, b) $P(\bar{A})$, c) $P(\bar{B})$, d) $P(A|B)$, e) $P(B|A)$, f) $P(\bar{A} \cap \bar{B})$, g) $P(\bar{A}|\bar{B})$,
h) $P(\bar{B}|\bar{A})$
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 correctamente atribuído a falha humana é de 80% e erradamente atribuído a essa causa é de 4%. Calcule a 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 respectivamente 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, respectivamente. 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 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) $f(x) = \frac{x-2}{5}$ $x=1,2,3,4,5$

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

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

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

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

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

Apresente o respectivo gráfico.

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

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

Apresente o respectivo gráfico. Calcule

- a)** $P(2 < X \leq 6)$;
- b)** $P(X=4)$;
- c)** $f(x)$.

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

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

Apresente o respectivo gráfico. Calcule: $P(X \leq 3)$; $P(X=3)$; $P(X < 3)$; $P(-0.4 < X < 4)$; $P(X=5)$; $f(x)$.

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

- a)** $f(x) = cx \quad x = 1, 2, 3, 4, 5$
- b)** $f(x) = cx^2 \quad x = 1, 2, 3, \dots, k$

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

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

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

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

Calcule $P(X \leq 5)$ 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

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

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

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

4. Considere a seguinte função de probabilidade

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

Encontre:

a) $E[X]$, $E[X^2]$ e $V[X]$

b) $E[(3X+2)^2]$

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) $E[X]$, $E[X^2]$, $E[X^3]$ e $V[X]$;

b) $E[X^3 + 2X^2 - 3X + 1]$.

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) Exactamente 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 correcta 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 correctas?
 - b) Qual a probabilidade de obter duas ou mais questões correctas?
 - 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 correcto 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 detecção de mísseis e radares militares permitem avisar contra ataques inimigos. Uma questão importante está relacionada com a capacidade do sistema em identificar e avisar correctamente um ataque. Assuma que um sistema particular de detecção tem 90% de probabilidades de detectar um ataque de míssil.
 - a) Qual a probabilidade de que um único sistema detecte o ataque?
 - b) Se na mesma área forem instalados dois sistemas de detecção com funcionamento independente, qual a probabilidade de pelo menos um deles detectar o ataque?
 - c) Se forem instalados três sistemas, qual a probabilidade de pelo menos um deles detectar 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
- Qual a probabilidade de não chegar nenhum passageiro durante um minuto?
 - Qual a probabilidade de chegarem 3 ou mais passageiros durante um minuto?
 - Qual a probabilidade de não chegar nenhum passageiro durante 15 segundos?
 - 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 probabilidade de que, durante um mês, se avariem 7 ou mais teares?
 - 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 seleccionada para auditoria.
- Qual a probabilidade de que nenhuma das contas seleccionadas tenha erros?
 - Qual a probabilidade de 4 ou mais das contas conterem erros?
 - Qual a probabilidade de exactamente 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:
- Qual o número esperado de calculadoras a falhar nos primeiros 30 dias de operação?
 - Qual a probabilidade de pelo menos 2 falhem?
 - Qual a probabilidade de que falhem exactamente 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:
- estar entre -0.002 e 0.003 ;
 - 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 fracção 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 4×10^3 horas?
b) Entre 5×10^3 e 10×10^3 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 satisfazem 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 exactamente 10 dos estudantes seleccionados terem frequentado um colégio particular?
 - b) Qual a probabilidade de 20 ou mais dos estudantes seleccionados 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 - 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 seleccionar 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 respectivas 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 respectivamente 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 7 - ESTIMADORES PONTUAIS

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

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

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

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

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

- 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 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? Justifique.

4. Se $X_1, X_2 \dots X_n$ representam os elementos de uma amostra aleatória de uma distribuição com f.d.p.

$$f(x; \theta) = \begin{cases} \theta x^{\theta-1} & \text{para } 0 < x < 1 \text{ e } 0 < \theta < \infty \\ 0 & \text{nos outros casos} \end{cases}$$

determine a estatística de máxima verosimilhança para o parâmetro θ .

5. Se $X_1, X_2 \dots X_n$ são os elementos de uma amostra aleatória tirada da distribuição $N(\theta_1, \theta_2)$ em que $-\infty < \theta_1 < \infty$ e $0 < \theta_2 < \infty$, determine as estatísticas de máxima verosimilhança para os parâmetros θ_1 e θ_2 .
6. Suponha uma distribuição de Weibull, com f.d.p. $f(x; \theta) = \theta \alpha x^{\alpha-1} e^{-\theta x^\alpha}$ para $x > 0$ (sendo α conhecido). Determine uma estatística de máxima verosimilhança para estimar o parâmetro da distribuição, baseada numa amostra de tamanho n .
7. Assuma que uma variável aleatória X tem a distribuição exponencial $f(x; \theta) = \theta e^{-\theta x}$, $x > 0$ e $\theta > 0$, onde θ é o parâmetro da distribuição. Use o método da máxima verosimilhança para estimar θ se 5 observações x fossem $x_1=0.9$, $x_2=1.7$, $x_3=0.4$, $x_4=0.3$ e $x_5=2.4$.
8. A função densidade de probabilidade da v.a. X é

$$f(x; \theta_1, \theta_2) = \frac{1}{\theta_2} \exp\left[-\frac{(x - \theta_1)}{\theta_2}\right] \text{ para } \theta_1 \leq x < \infty, \quad -\infty < \theta_1 < \infty, \quad 0 < \theta_2 < \infty$$

- a) Calcule estimadores de máxima verosimilhança para θ_1 e θ_2 .
- b) Calcule a média da distribuição.
- c) A partir dos estimadores já obtidos em a), determine um estimador para a média da distribuição.

9. Seja X uma variável aleatória cuja f.d.p. é a seguinte

$$f(x, \theta_1, \theta_2) = \begin{cases} \frac{1}{\theta_2 - \theta_1} & \text{se } \theta_1 < x < \theta_2 \\ 0 & \text{nos outros casos} \end{cases}$$

Determine estimadores de máxima verosimilhança para θ_1 e θ_2 .

FICHA Nº 8 - INTERVALOS DE CONFIANÇA

- Se uma amostra aleatória de tamanho $n=20$ duma população normal com variância $\sigma^2=225$ tem média $\bar{x} = 64.3$:
 - Construa o intervalo de confiança de 95% para a média da população μ .
 - Construa também o intervalo de 90% de confiança.

- Um inspector 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:
 - Construa o intervalo de 99% de confiança para o teor médio de impurezas nesta marca de compotas.
 - Construa também o intervalo de 90% e de 95% de confiança.

- 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

- Dê uma estimativa pontual das médias das alturas da população dos pinheiros
 - Com 95.4% de certeza, qual o limite de erro conhecido?
- Uma mostra de $n=100$ empregados de uma companhia foi seleccionada, e, o salário mensal foi registado. A média e o desvio padrão dos seus salários foram respectivamente $\bar{x} = 177500$ e $s=9000$. Construa o intervalo de confiança de 95% para o salário médio da população μ .
 - 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:
 - $n=30$
 - $n=60$
 - $n=90$
 Compare os tamanhos dos três intervalos.

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

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

- Cinco pessoas seleccionadas 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.

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

8. 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:
- Construa o intervalo de 99% para a diferença entre as verdadeiras médias das capacidades caloríficas do carvão das duas minas.
 - Construa também o intervalo de 90%. O que pode concluir?
9. 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_1^2 e σ_2^2 respectivamente. A seguinte informação amostral é conhecida:
Amostra 1: $n_1 = 50$, $\bar{x}_1 = 91.1$, $s_1 = 5.4$
Amostra 2: $n_2 = 50$, $\bar{x}_2 = 92.3$, $s_2 = 7.6$
- Estime a diferença e construa o intervalo de 95% de confiança.
 - Qual o limite de 95% de confiança para o erro na estimação?
10. 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.
- Dê uma estimativa pontual da proporção de sócios que se espera comprem o produto.
 - Calcule, com 95.4% de certeza, um limite do erro cometido.
11. 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.
12. 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.
13. 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.
- Estime p .
 - Construa um intervalo de 95% de confiança indicando o erro da estimativa.

14. 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 seleccionados aleatoriamente, só 140 apoiam o projecto.
- Construa um intervalo de 90% de confiança para a proporção de todos os votantes nesta comunidade que apoiam o projecto.
 - Construa também os intervalos de 95% e de 98% de confiança.
15. 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 actual de votantes das duas freguesias que apoiam o candidato.
16. 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 respectivamente. Estime a diferença na taxa de mortalidade dos dois extintores, quando usados no ambiente de teste.
17. 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 .
18. Um controlador de qualidade numa fábrica de refrigerantes sabe que a quantidade exacta de cada lata variará, uma vez que existem factores incontroláveis que afectam 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 selecciona 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.
19. 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 Nº 9 - TESTES DE HIPÓTESES

- 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 detectar 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 $\sigma = 5$.
- 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 correcta

H_1 : p_1 é a distribuição correcta

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

- Determine a probabilidade de cometer um erro do tipo I;
 - Determine a probabilidade de cometer um erro do tipo II.
- 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 $\alpha = 0.02$
 - Qual é a hipótese nula assumida?
 - Qual o nível que indica o maior grau de significância, 0.05 ou 0.02? Justifique.
 - Para $\alpha = 0.05$, rejeitar-se-á H_1 ? Justifique. Será H_1 rejeitada para $\alpha = 0.02$? Justifique.
 - Ao escolher $\alpha = 0.02$ em vez de $\alpha = 0.05$, aumenta-se o risco de um dos dois tipos de erro? Qual? Justifique.
 - 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

- Calcule as funções α e β , probabilidades associadas com os erros do tipo I e II, respectivamente, se for definida como região de rejeição o conjunto
 - $C_1 = \{b, c\}$
 - $C_2 = \{d\}$

- 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, seleccionados 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 ruptura de um certo tipo de fita seja de 83.9 kg e que 5 peças seleccionadas 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 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 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 $\alpha = 0.05$.

- 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 factor 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 concluir?

- 22.** Um novo tratamento para a esquizofrenia foi testado durante seis meses com 54 doentes seleccionados 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 detectaram 72 das 100 armas falsas levadas por inspectores. Esta taxa de detecção está abaixo da taxa nacional de detecção de 80%. Há evidência suficiente para concluir que a detecção no referido aeroporto está abaixo da taxa nacional? Use $\alpha=0.01$.
- 25.** Em 1990, 371 empresas foram seleccionadas para determinar em que medida disponham de sistemas de informação em logística. Cinco anos mais tarde, em 1995, 459 empresas foram seleccionadas 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 semiconductor 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%.
- 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 – 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 ($\alpha=0.05$).

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 radioactiva 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	≥ 7
Frequência	19	48	66	74	44	35	10	4

5. Caixas de um certo tipo foram expostas ao risco de acidentes sob a acção 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 efectuada, 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, efectuou-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, respectivamente, 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 radioactiva 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 probabilidades 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.

10. Foi escolhido ao acaso um cavalo para correr 80 corridas. Em cada corrida o cavalo foi classificado de acordo com a posição no início da corrida e a posição em que ficou no final. A tabela das frequências observadas é a seguinte:

Posição inicial		Posição final			
		1	2	3	outras
	1 a 4	8	6	8	16
	5 a 9	3	6	5	28

Verifique se os dados são consistentes com a afirmação de que a posição do cavalo no final da corrida não depende da posição inicial dada no início da corrida.

11. Foram encontrados 77 parafusos partidos na estrutura metálica da estação dos comboios da Vila Grande. O investigador do acontecimento considerou como hipótese nula o facto de que a localização e a causa da fractura não estão relacionadas. Os dados registados das fracturas encontradas são os seguintes:

causa da fractura	localização da fractura		
	base	meio	extremidade
fogo	21	8	18
outras	15	11	4

Que conclusões poderá tirar o investigador do estudo, supondo que considera

- a) $\alpha=0.05$
b) $\alpha=0.025$

12. Num estudo sobre a opinião dos pais acerca dum curso em educação sexual, 360 pais, uma amostra aleatória, foram classificados de acordo com o facto de terem uma, duas ou mais crianças no sistema escolar, e, também se achavam o curso fraco, adequado ou bom. Baseado nestes resultados, apresentados na tabela, teste, ao nível de significância de 0.05, se existe uma relação entre a reacção dos pais ao curso e o número de crianças que têm no sistema escolar.

Opinião Curso	Nº de Crianças		
	1	2	3 ou mais
Fraco	48	40	12
Adequado	55	53	29
Bom	57	46	20

13. Os dados seguintes dizem respeito a encomendas recebidas por uma grande firma de três vendedores diferentes:

	Nº Encomendas		
	Rejeitadas	Imperfeitas mas aceitáveis	Perfeitas
Vendedor A	12	23	89
Vendedor B	8	12	62
Vendedor C	21	30	119

Teste ao nível de significância de 0.01, se os três vendedores enviam produtos de igual qualidade.

FICHA Nº 11 - 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 objectivo 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, injectados, respectivamente 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, seleccionou 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:

trabalhador	unidades produzidas	
	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 conclusões pode tirar? Justifique
- b) Face às conclusões que tirou em a) que modelo populacional lhe parece que seria o mais adequado?

8. Uma experiência foi realizada para avaliar o efeito de 4 tipos de fuel e três lançadores de foguetões no alcance, em quilómetros, dum determinado tipo de foguetão. Teste, baseado nos resultados, em quilómetros, se existe um efeito significativo devido a diferenças nos tipos de fuel e se existe um efeito significativo devido a diferenças nos lançadores.

	Fuel 1	Fuel 2	Fuel 3	Fuel 4
Lançador 1	45.9	57.6	52.2	41.7
Lançador 2	46.0	51.0	50.1	38.8
Lançador 3	45.7	56.9	55.3	48.1

9. A seguinte experiência foi executada para determinar o efeito de duas campanhas de publicidade em três qualidades de bolos. As vendas de cada qualidade foram registadas depois da primeira campanha e depois da segunda campanha publicitária. A experiência foi repetida 3 vezes para cada campanha com os seguintes resultados:

	Campanha 1	Campanha 2
Q1	574, 564, 550	1092, 1086, 1065
Q2	524, 573, 551	1028, 1073, 998
Q3	576, 540, 592	1066, 1045, 1055

Assumindo que os dados são normalmente distribuídos, que conclusões pode retirar da experiência. Justifique.

10. Uma firma pretende adquirir um novo processador de texto para um novo escritório da companhia. Para o efeito seleccionou três softwares e planeou um teste comparativo de medição do tempo (em minutos) necessário para digitar o texto padrão. Três operadores não familiarizados com o software são aleatoriamente seleccionados. A cada operador é dado o mesmo tempo de instrução em cada tipo de software. O tempo é medido para cada operador em cada um dos três tipos de software escolhidos. Que conclusões pode tirar? Admita que os dados seguem uma distribuição normal ($\alpha=0.05$).

Operador	Software		
	1	2	3
1	42	45	45
	43	43	44
	39	45	47
2	37	36	40
	36	37	41
	36	34	42
3	36	39	40
	34	39	42
	37	37	38

{cálculos auxiliares: $\bar{Y} = 39.78$; $STQ = 348.6668$, $SQR = 38.66667$, $T^2 = 1153476$ }

FICHA Nº 12 - TESTES NÃO PARAMÉTRICOS

1. Numa análise da regularidade de fibras de lã consideram-se três bobinas e pesaram-se cinco comprimentos de 1 metro, espaçados ao longo das bobinas. Os valores obtidos foram:

bobina 1	bobina2	bobina 3
15.0	17.8	17.3
14.3	18.3	16.7
14.3	18.4	15.6
14.2	17.7	16.0
14.6	17.2	15.4

Face a estes valores, parece-lhe que existam diferenças significativas entre as bobinas?

2. A absorção de água de um material têxtil em 24 horas é dada na tabela seguinte para 5 amostras. Pretende-se saber se as amostras podem ser consideradas do mesmo material.

A	6,7	5,8	5,8	5,5
B	5,1	4,7	5,1	5,2
C	4,4	4,9	4,6	4,5
D	6,7	7,2	6,8	6,3
E	6,5	5,8	4,7	5,9

- Formule a hipótese nula do problema;
- Análise os dados pelo método que achar mais adequado;
- Faça uma interpretação dos resultados.

3. Ao longo do ano lectivo um aluno do 11º ano obteve as seguintes notas em várias matérias como mostra a tabela que se segue.

Matemática	72	80	83	75
Química	81	74	77	
Inglês	88	82	90	87
Biologia	74	71	77	70

Determine se existem diferenças entre as notas para um nível de significância $\alpha = 0.05$.

4. Num estudo para determinar a eficiência de um medicamento A na redução de BVP (batidas ventriculares prematuras), considerou-se uma amostra de 9 doentes do coração e administrou-se-lhes o medicamento (na proporção 2 mg/kg). Os resultados de BVP/min obtidos 20 minutos depois do tratamento, juntamente com os valores antes do tratamento, foram registados na seguinte tabela:

Doente Nº	1	2	3	4	5	6	7	8	9
antes	6	9	17	22	7	5	5	14	9
depois	5	2	0	0	2	1	0	0	0

Tomando o nível de significância do teste igual a 0.1, o que pode dizer sobre a eficiência do medicamento?

5. A freguesia de Sto. António da cidade de Vila Velha tem dois bairros habitacionais. De cada bairro, seleccionaram-se aleatoriamente 20 casas e foram-lhe atribuídas classificações de 0 a 10 dependendo do estado de degradação das casas (0 = estado novo, ..., 10 = casa degradada). Os resultados da experiência foram os seguintes:

Bairro da Sé	0	4	4	6	0	8	7	4	1	5	3	2	4	4	0	3	6	0	8	9
Bairro Canoa	7	4	7	0	3	6	1	4	1	2	9	7	9	3	3	8	2	7	8	3

Teste a hipótese de que não existem diferenças significativas entre o estado de degradação das casas dos dois bairros.

6. Fez-se uma experiência num laboratório para o controlo da temperatura em salas, que envolveu 6 homens e 6 senhoras. Pediu-se-lhes que indicassem a temperatura na sala que os fizesse sentir mais confortáveis. Os resultados obtidos foram os seguintes:

Homens	Senhoras
74	75
74	77
75	78
80	79
81	77
78	75
$\bar{X}_1 = 77$	$\bar{X}_2 = 76.8$
$s_1 = 3.1$	$s_2 = 1.6$

Acha que as temperaturas consideradas mais confortáveis pelos homens são as mesmas do que as das senhoras? Justifique. Use $\alpha=0.05$.

7. Pretende-se estudar o efeito da alimentação intermitente na pressão arterial dos ratos. Oito ratos foram alimentados intermitentemente por um período de várias semanas, no fim do qual as pressões arteriais dos ratos foram medidas. Os resultados em mm hg

170 168 115 181 162 199 207 162

Estes números pouco significariam na ausência de dados para comparação. A pressão arterial de sete ratos alimentados normalmente foi também medida, produzindo os seguintes resultados:

169 155 134 152 133 108 145

A questão que se coloca é decidir qual a explicação mais viável para a discrepância dos dados - que a alimentação tem efeito na pressão arterial, ou que não tem efeito e que a variabilidade amostral é responsável pela discrepância entre as duas amostras.

- Compare os dados aos nível das médias supondo que os dados seguem uma distribuição normal.
- Compare as distribuições supondo desconhecida a distribuição amostral dos dados.

8. Foram feitas medições de viscosidade de uma certa substância em dois dias diferentes. Os resultados obtidos foram:

1º Dia	37.0	31.4	34.4	33.4	34.9	36.2	31.0	33.5	33.7	33.4	34.8	30.8
2º Dia	28.4	31.3	28.7	32.1	31.9	32.8	30.2	30.2	32.4	30.7		

Poder-se-á dizer que a população (descrita pela variável $X \equiv$ viscosidade da substância) mudou de um dia para o outro? Justifique a utilização do teste estatístico.

9. Três grupos de pessoas pertencentes às chamadas classes baixa, média e alta foram entrevistadas e foi-lhes solicitado que indicassem numa escala a sua opinião relativa à concessão de mais direitos à mulher na sociedade. Os resultados obtidos foram os seguintes:

									Média	Desvio	Variância
classe Baixa	5	5	6	4	5	6	7	4	5,25	1,035098	1,071429
classe Média	8	10	6	6	5	4	3	9	6,375	2,445842	5,982143
classe Alta	10	12	9	9	10	10	6	12	9,75	1,908627	3,642857

Tire as conclusões que os dados permitem, sabendo-se que quanto maior é o valor da escala mais favorável é a atitude do entrevistado em relação ao tema. Justifique a utilização do teste estatístico.

10. Verifique se os dados apresentados são indicação da existência de diferenças nos comprimentos das palavras em latim. As observações representam o número de letras das palavras em latim, seleccionadas aleatoriamente dos três géneros: masculino, feminino e neutro.

Masculino	Feminino	Neutro
5	4	7
7	8	10
6	5	7
7	6	8
5	3	7
9	6	12

11. Num concurso musical, o júri era composto por três pessoas. As pontuações atribuídas aos 8 concorrentes foram as seguintes (mínimo:1, máximo: 6)

	Concorrentes							
Júri	1	2	3	4	5	6	7	8
Pessoa 1	5	4	4	6	5	4	6	4
Pessoa 2	3	4	3	5	5	3	6	4
Pessoa 3	2	2	2	4	6	4	5	4

Que conclusões pode tirar sobre as distribuições das pontuações atribuídas pelos elementos do júri. Considere $\alpha=0.05$ e $\alpha=0.10$.

12. Para verificar se um período maior entre o último dia de aulas e o dia do exame final, afecta significativamente o desempenho dos alunos no exame, fez-se uma experiência com os 48 alunos de uma turma. Estes foram divididos aleatoriamente em quatro grupos de 12 estudantes cada. Para o 1º grupo deixou-se passar dois dias de intervalo. O 2º grupo fez exame 4 dias após o fim das aulas. Ao 3º grupo foi dado um intervalo de 6 dias e, ao 4º grupo um intervalo de 8 dias. As classificações obtidas no exame foram as seguintes (de 0 a 100)

1º grupo	48	61	67	68	71	74	75	79	80	82	87	89
2º grupo	48	42	67	62	71	70	75	73	81	77	92	89
3º grupo	38	58	71	70	73	74	79	75	83	87	94	90

Que conclusões pode tirar sobre as distribuições das notas obtidas por cada grupo?

13. De uma amostra aleatória de tamanho 10, obtiveram-se os seguintes resultados, $x_1 = 0.621$, $x_2 = 0.503$, $x_3 = 0.203$, $x_4 = 0.477$, $x_5 = 0.710$, $x_6 = 0.581$, $x_7 = 0.329$, $x_8 = 0.480$, $x_9 = 0.554$ e $x_{10} = 0.382$. Acha que estes dados se ajustam à distribuição Uniforme

representada por:

$$F^*(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

- 14.** A professora de Estatística Aplicada queria testar a hipótese de que a hora de chegada dos seus alunos à aula teórica das 14:00 às 16:00 de 3ª feira, segue uma distribuição normal com desvio padrão igual a 4 minutos. Num dos dias de aula registou as horas de chegada dos alunos que assistiram àquela aula:

13:53	13:56	13:57	13:57	13:59	13:59
14:00	14:01	14:02	14:02	14:02	14:03
14:04	14:05	14:05	14:07	14:07	14:11

O que poderá ele concluir em relação à distribuição das horas de chegada dos seus alunos?

- 15.** Após longos anos de experiência, verificou-se que o número de plaquetas no sangue de pessoas saudáveis do sexo masculino segue uma distribuição normal com média 235 000 por mm³ e desvio padrão igual a 44 600 por mm³. Os números de plaquetas registadas a partir de doentes com cancro nos pulmões foram os seguintes (em unidades de 1 000 por mm³):

173	189	196	207	215	237	275	282	293	300
305	316	346	382	395	399	401	437	480	504
524	634	682	882	999					

Usando $\alpha=0.01$ decida se estas observações podem ser consideradas provenientes da população definida pelo número de plaquetas no sangue de pessoas saudáveis do sexo masculino.

- 16.** Uma empresa acredita que os seus televisores a cores têm um tempo de vida médio de 18 meses (a funcionar de uma forma contínua). Esta empresa, através do departamento de controlo de qualidade, efectuou testes a 12 televisores, tendo registado os seguintes tempos de vida (em meses):

5,6	12,2	14,8	16,4	16,8	19,0	19,4	21,0	22,4	23,5	23,7	24,9
-----	------	------	------	------	------	------	------	------	------	------	------

Verifique se os dados se ajustam a uma distribuição exponencial, com tempo de vida médio previsto pela empresa (utilize $\alpha=0.1$)

- 17.** Os resultados de um investimento feito em 20 tipos distintos de "stocks", seleccionados aleatoriamente de um armazém, passados 12 meses, foram os seguintes:

9.1	5.0	7.3	7.4	5.5	8.6	7.0	4.3	4.7	8.0
4.0	8.5	6.4	6.1	5.8	9.5	5.2	6.7	8.3	9.2

Teste a hipótese nula de que o resultado do investimento segue uma distribuição normal.

- 18.** Um estudo dos efeitos das descargas de lixo no mar produziu os seguintes resultados para concentrações de zinco (mg/kg) em 13 pontos num raio de 1km do ponto de descarga

13.5	23.8	23.3	20.9	23.8	29.0	
20.9	24.4	16.4	18.3	17.6	25.4	23.3

Teste se os dados seguem uma distribuição normal com $\mu=21.6$.

- 19.** Foram testadas 20 válvulas electrónicas relativamente à sua duração de vida em horas. Foram registados os seguintes valores da duração de vida:

7,2	37,8	49,6	21,4	67,2	41,1	3,8	8,1	23,2	72,1
11,4	17,5	29,8	57,8	84,6	12,8	2,9	42,7	7,4	33,4

Verifique se estes dados são consistentes com a hipótese de que a variável aleatória X , duração de vida em horas, segue a distribuição exponencial.

- 20.** As chamadas de longa distância que passam por uma central telefónica, formam um processo aleatório, em que o intervalo de tempo entre as chamadas parece seguir uma distribuição exponencial. As primeiras dez chamadas de uma 2ª feira, depois das 13:00, ocorreram às 13:06, 13:08, 13:16, 13:22, 13:23, 13:34, 13:44, 13:47, 13:51 e 13:57. Face a estes valores, teste a hipótese nula formulada em termos da distribuição exponencial.

- 21.** Um gerente de armazém quer testar a hipótese de que os clientes chegam aleatoriamente ao armazém. Para isso registou os tempos entre chegadas sucessivas ao longo de uma manhã. Esses tempos foram os seguintes (em minutos):

3.6	22.1	38.0	3.3	10.1
14.2	1.4	6.1	12.7	4.6
10.8	6.2	4.2	3.8	8.2

Teste a hipótese de que os tempos entre chegadas seguem uma distribuição exponencial.

FICHA Nº 13 – REGRESSÃO E CORRELAÇÃO

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 β_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 a diferença de potencial for 2.3?.
2. Determine a relação existente entre o calor envolvido no endurecimento, representado pela variável Y e os pesos de duas substâncias X_1 e X_2 , 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 , $i=1,2,\dots,6$, conhecidos. As alturas máximas 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 σ^2 .

Estime os valores dos parâmetros k_1 e k_2 que definem o modelo, considerando os resultados obtidos numa experiência:

d (μ)	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 ($^{\circ}\text{C}$)	-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

- Calcule os coeficientes da recta de regressão da taxa de consumo do oxigénio em função da temperatura.
 - Teste a hipótese, $H_0 : \beta_1 = 0$.
 - Calcule o coeficiente de determinação.
 - Calcule o intervalo de confiança de 95% para o declive.
6. Uma experiência foi realizada para estudar o efeito do 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

- Calcule os coeficientes da recta de regressão do decrescimento da tensão em função da percentagem de enxofre. O que pode concluir acerca do ajuste?
 - Calcule os coeficientes da recta 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

- Assumindo uma relação linear, estime os valores de β_0 , β_1 e β_2 .
 - 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
Intellectual	77	85	105	86	108	116	96	100	100

9. A tabela apresenta os resultados de um estudo, conduzido por Howell e Huesy (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 seleccionados, 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

- Produza as estatísticas descritivas para os dados.
- Apresente o diagrama de dispersão.
- 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?

11. A tabela apresenta os teores de colesterol (mg/100ml) e deposição de cálcio nas artérias (mg/100g de tecido seco) em doze animais.

Cálcio	59	52	42	59	24	24	40	32	63	57	36	24
Colesterol	298	303	233	287	236	245	265	233	286	290	264	239

Calcule o coeficiente de correlação de Pearson e de Spearman.

12. 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

a) Calcule o coeficiente de correlação desta amostra.

b) Teste a hipótese de independência, usando um teste bilateral, baseado na estatística calculada em a).

- 13.** Os estudantes de um curso de Física tiveram de realizar uma experiência que consistia em atirar uma bola de uma certa altura e medir a velocidade, em metros por segundo, em vários pontos do trajecto. A velocidade, teoricamente, deve aumentar linearmente com o tempo, pois a bola está sujeita à gravidade. Não sendo fácil medir a velocidade da bola, depois de um período de tempo fixo, as observações dos estudantes não caem necessariamente sobre a recta. Teste o tipo de relação sugerida teoricamente, a um nível de significância de 0.05. Os resultados obtidos foram os seguintes:

tempo (segundos)	0	0.2	0.4	0.6	0.8
velocidade (m/seg)	0	1.8	3.58	6.01	7.88

- 14.** Num curso de especialização para licenciados, a disciplina de Estatística foi leccionada por dois docentes. O docente A deu o 1º módulo e o docente B deu o 2º módulo. No fim do ano lectivo foi feito um teste de avaliação composto por duas partes. A primeira, com questões do 1º módulo foi corrigida pelo docente A. A segunda parte continha questões do 2º módulo e foi corrigida pelo docente B. No curso estavam inscritos 9 alunos. As pontuações atribuídas foram: (1-mínimo, 5-máximo)

docente A	5	4	3	3	4	5	3	5	3
docente B	4	5	4	2	4	4	4	4	3

Que conclusões pode tirar da relação entre as pontuações atribuídas pelos dois docentes da disciplina? Use $\alpha=0.05$.

- 15.** O Sr. José foi seleccionado para júri num concurso de beleza. Concorreram oito jovens, e os resultados foram os seguintes:

		lugar
Amélia	17 anos	1
Bela	16 anos	2
Carolina	18 anos	2
Deolinda	20 anos	4
Eva	18 anos	5
Francisca	18 anos	6
Georgina	20 anos	7
Helena	23 anos	7

As idades são também apresentadas porque se suspeita que o Sr. José tem um favoritismo em relação às mais jovens. Verifique se existe ou não favoritismo.

