Escola Superior de Tecnologia e Gestão de Mirandela
2007 / 2008

ESTATÍSTICA

ESTGM

Setembro de 2007

Escola Superior de Tecnologia e Gestão de Mirandela
2007 / 2008

Apontamentos da disciplina de Estatística

do 2º ano dos cursos de:

- Marketing
- Tecnologias da Comunicação

ESTGM
Setembro de 2007

Índice

1 Estatística descritiva **1**
1.1 Distribuições de frequências. Histogramas . . . 2
1.2 Representações gráficas . . . 4

2 Características numéricas. **5**
2.1 Estatísticas de localização. . . . 5
2.2 Estatísticas de dispersão. . . . 6
2.3 Outras estatísticas. . . . 7
2.4 Caracterização de amostras bivariadas. . . . 8

3 Teoria elementar da probabilidade. **10**
3.1 Definições. . . . 10
3.2 Medida de probabilidade. . . . 12
3.3 Interpretações do conceito de probabilidade. . . . 12
3.4 Análise combinatória. . . . 13
3.5 Probabilidade condicionada. . . . 14
3.6 Teorema da probabilidade total. . . . 15
3.7 Fórmula de Bayes. . . . 15
3.8 Acontecimentos independentes. . . . 16

4 Variáveis aleatórias e distribuição de probabilidade. **17**
4.1 Distribuições de variáveis discretas. . . . 17
4.2 Distribuições de variáveis contínuas. . . . 18
4.3 Parâmetros das distribuições. . . . 19
4.4 Transformação linear de uma variável. . . . 20

5 Distribuições de variáveis discretas. **21**
5.1 Prova de Bernoulli. Distribuição binomial. . . . 21
5.2 Distribuição hipergeométrica. . . . 22
5.3 Relação entre as distribuições hipergeométrica e binomial. . . . 23
5.4 Distribuição de Poisson. . . . 23
5.5 Relação entre a distribuição de Poisson e a distribuição binomial. . . . 25

6 Distribuições de variáveis contínuas. **26**
6.1 Distribuição uniforme, $U(a, b)$. . . . 26
6.2 Distribuição exponencial, $\mathrm{Ex}(\lambda)$. . . . 27
6.3 Distribuição normal, $N(\mu, \sigma^2)$. . . . 28

7 Distribuições discretas e contínuas multivariadas. **34**
7.1 Estudo de pares aleatórios discretos (X, Y). . . . 34
7.2 Distribuições contínuas. . . . 35
7.3 Covariância e correlação. . . . 35

8 Amostragem. Distribuições amostrais. **37**
8.1 População e amostra. Inferência estatística. 37
8.2 Amostragem aleatória. 37
8.3 Distribuições amostrais. 38
8.4 Teorema do limite central. 38

9 Amostragem. Estimação pontual. **40**
9.1 Propriedades dos estimadores por pontos. 40

10 Estimação de intervalos. **43**
10.1 Conceito de intervalo de confiança. 43
10.2 Simetria *versus* assimetria do intervalo de confiança relativamente ao valor da estimativa pontual. 44
10.3 Especificação do intervalo. 44
10.4 Dimensionamento de amostras. 46
10.5 Intervalos de confiança a $(1-\alpha)100\%$ para a diferença dos valores esperados, $\mu_1 - \mu_2$, quando as amostras não podem ser consideradas independentes (observações emparelhadas). 47

11 Testes de hipóteses. **48**
11.1 Valor de prova. 49
11.2 Erro do Tipo II- potência de teste. 50
11.3 α *versus* β. 50
11.4 Relação entre testes de hipóteses e intervalos de confiança. 50

1 Estatística descritiva

Nas áreas económicas e de gestão, a Estatística pode ser utilizada com três objectivos:

- Descrever e compreender relações entre diferentes características de uma população;
- Tomar decisões mais correctas;
- Fazer face à mudança.

Método estatístico de resolução de um problema:

- Identificar correctamente o problema em análise;
- Recolher a informação necessária, relevante para o problema em estudo, em tempo útil e tão completa quanto possível;
- Classificar e ordenar os dados;
- Análise dos dados e apresentação dos resultados;
- Tomar a decisão mais adequada.

População: conjunto de elementos cujos atributos são objecto de determinado estudo.
Constituem exemplos de populações:

- O conjunto de empresas de turismo em Setembro de 2005;
- O conjunto dos cidadãos eleitores nacionais;
- O conjunto das temperaturas instantâneas que se registaram no Verão ao longo do dia 16 de Agosto.

Os elementos da população são em geral designados por **unidades estatísticas**.
Para conhecer completamente a população realiza-se um **censo** ou **indagação completa**.
Como se torna difícil analisar todos os elementos de uma população, o estudo dos seus atributos tem de ser feito sobre um seu conjunto finito - **amostra**.
Inferência Estatística: A partir de dados ou observações procura-se saber ou inferir alguma coisa sobre a população.
Estatística Descritiva:
Ramo da estatística que se preocupa com a recolha, apresentação e interpretação de dados através de **quadros, gráficos, indicadores numéricos**.

Não há necessidade de tirar quaisquer conclusões.

Variáveis estatísticas

- Qualitativas ou categorias
  Ex.: cor dos olhos.

- Quantitativas discretas - a variável só pode tomar um número finito de valores ou uma quantidade numerável.
  Ex.: número de eleitores em qualquer eleição.
- Quantitativas contínuas - a variável pode tomar valores dentro de um conjunto infinito não numerável.
  Ex.: temperatura em ^{o}C.
-

1.1 Distribuições de frequências. Histogramas

Dados não agrupados

- Frequência absoluta (F_i): número de vezes que o valor é observado.
- Frequência relativa (f_i): quociente entre a frequência absoluta e a dimensão da população.
- Frequência absoluta acumulada: $A_i = \sum_{j=1}^{i} F_j$
- Frequência relativa acumulada: $a_i = \sum_{j=1}^{i} f_j$

Dados agrupados
Classes C_k, $k \in \{1, \cdots, m\}$

$$C_i \cap C_j = \phi, \;\; i \neq j$$

$$[\min(x_i), \max(x_i)] \subseteq \bigcup_{k=1}^{m} C_i$$

Definição de classe

$$C_k = [l_{k-1}, l_k[\,, \;\; k \in \{1, \cdots, m-1\}$$

$$C_m = [l_{m-1}, l_m]$$

Amplitude de cada classe C_k: $\Delta C_k = l_k - l_{k-1}$
Marca de cada classe C_k: $x' = \frac{l_{k-1}+l_k}{2}$

Nota: É vantajoso considerar classes com amplitude constante, pois torna mais sugestiva a comparação das frequências das diversas classes.

Exemplo 1.1.1 *Foi feito um inquérito a 40 pessoas que compraram um carro novo com o objectivo de se saber quantas reparações, ou substituições de peças, foram feitas durante o primeiro ano de utilização dos carros. Obtiveram-se os seguintes resultados:*

1	*4*	*1*	*2*	*2*	*3*	*3*	*2*	*1*	*2*
3	*2*	*3*	*1*	*0*	*1*	*2*	*7*	*4*	*3*
5	*1*	*2*	*4*	*2*	*1*	*3*	*1*	*0*	*1*
2	*1*	*1*	*3*	*1*	*0*	*4*	*2*	*3*	*5*

x_i	F_i	f_i	A_i	a_i
0	*3*	*3/40=0,075*	*3*	*0,075*
1	*12*	*12/40=0,3*	*15*	*0,375*
2	*10*	*10/40=0,25*	*25*	*0,625*
3	*8*	*8/40=0,2*	*33*	*0,825*
4	*4*	*4/40=0,1*	*37*	*0,925*
5	*2*	*2/40=0,05*	*39*	*0,975*
6	*0*	*0*	*39*	*0,975*
7	*1*	*1/40=0,025*	*40*	*1*
$\Sigma =$	*40*	*1*		

Exemplo 1.1.2 *Os dados listados abaixo são relativos ao peso, expresso em gramas, do conteúdo de uma série de 100 garrafas que acabaram de sair de uma linha de enchimento automático.*

302,50	*303,76*	*300,86*	*297,99*	*297,83*	*299,83*	*298,34*	*299,48*	*301,19*	*298,86*
299,16	*298,29*	*299,08*	*299,70*	*303,91*	*298,56*	*301,35*	*302,83*	*298,72*	*301,99*
299,20	*298,34*	*301,33*	*305,13*	*300,02*	*299,07*	*298,36*	*299,26*	*301,05*	*302,45*
298,65	*297,84*	*298,23*	*299,13*	*297,87*	*298,35*	*298,16*	*299,75*	*298,85*	*299,95*
299,83	*298,00*	*300,24*	*302,10*	*302,00*	*299,60*	*299,57*	*301,11*	*301,59*	*303,49*
299,23	*299,76*	*299,38*	*298,49*	*297,72*	*299,68*	*301,63*	*298,94*	*297,59*	*302,39*
299,75	*303,66*	*302,52*	*304,22*	*300,36*	*297,47*	*300,86*	*298,11*	*300,84*	*300,20*
300,52	*302,85*	*298,73*	*302,17*	*300,12*	*300,83*	*298,35*	*300,14*	*300,48*	*298,07*
299,07	*298,61*	*299,55*	*302,04*	*303,07*	*299,12*	*299,19*	*301,85*	*300,80*	*300,96*
302,45	*299,70*	*298,98*	*299,62*	*298,76*	*300,73*	*301,81*	*299,07*	*299,99*	*299,18*

Classe	x_i'	F_i	f_i	A_i	a_i
$[297, 298[$	*297,5*	*7*	*0,07*	*7*	*0,07*
$[298, 299[$	*298,5*	*22*	*0,22*	*29*	*0,29*
$[299, 300[$	*299,5*	*28*	*0,28*	*57*	*0,57*
$[300, 301[$	*300,5*	*15*	*0,15*	*72*	*0,72*
$[301, 302[$	*301,5*	*10*	*0,1*	*82*	*0,82*
$[302, 303[$	*302,5*	*11*	*0,11*	*93*	*0,93*
$[303, 304[$	*303,5*	*5*	*0,05*	*98*	*0,98*
$[304, 305[$	*304,5*	*1*	*0,01*	*99*	*0,99*
$[305, 306]$	*305,5*	*1*	*0,01*	*100*	*1*
	$\Sigma =$	*100*	*1*		

1.2 Representações gráficas

- **Diagramas de Barras**: utiliza-se para a representação gráfica de distribuições de frequências no caso da variável ser **discreta**, representando-se no eixo das abcissas os valores de x e no eixo das ordenadas as frequências (relativas ou absolutas), com que os valores foram observados.
- **Polígono de frequências**
- **Histograma**: Utiliza-se para a representação gráfica de distribuições de frequências de variáveis contínuas e é formado por uma sucessão de rectângulos.
- Representação gráfica das funções cumulativas das frequências relativas ($s(x)$).
- **Gráficos circulares**.

Ainda relativamente ao Exemplo 1.1.1:

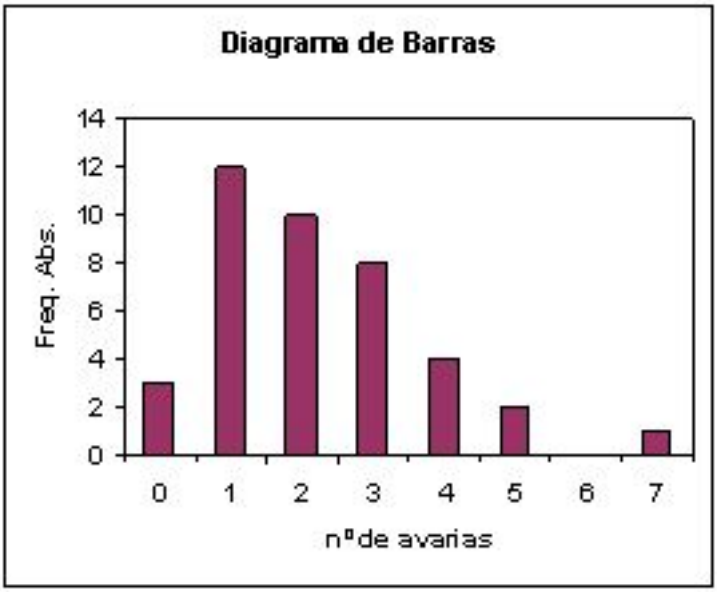

Figura 1: Diagrama de barras

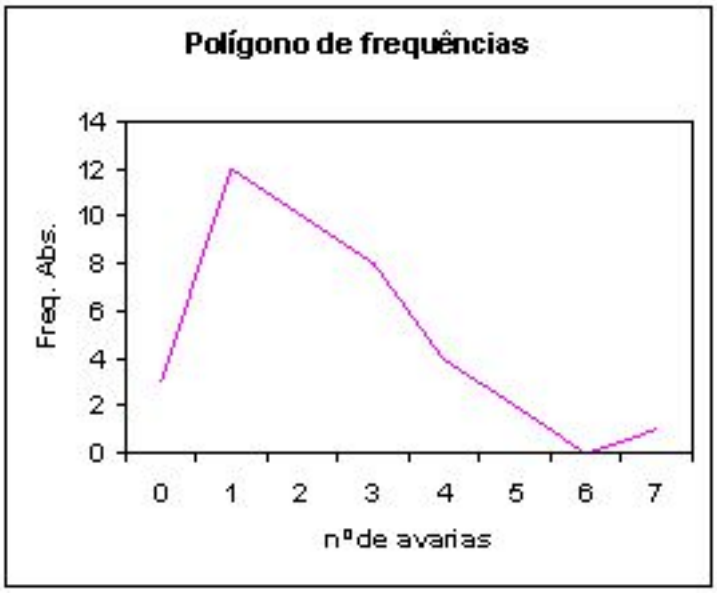

Figura 2: Polígono de frequências

2 Características numéricas.

2.1 Estatísticas de localização.

Média aritmética: valor a partir do qual se distribuem os valores da colecção.

$$\overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Propriedades da média aritmética:

Sendo a uma constante e x e y variáveis estatísticas, tem-se:

1. $\overline{a+x} = a + \overline{x}$;
2. $\overline{ax} = a\overline{x}$;
3. $\overline{x+y} = \overline{x} + \overline{y}$.

Mediana (m_e): valor da colecção que tem **50%** de observações anteriores e **50%** de observações posteriores.

Quantis: q_α, quantil de ordem α ($0<\alpha<1$) é o valor da colecção que tem α**100%** de observações anteriores e **(1-**α**)100%** de observações posteriores.

Nota: $m_e = q_{0.5}$

Exemplos de quantis:

- Percentis: quando $\alpha = 1\%, 2\%, \cdots, 99\%$.
- Decis: quando $\alpha = 10\%, 20\%, \cdots, 90\%$.
- Quartis: quando $\alpha = 25\%, 50\%, \cdots, 75\%$.

Moda (m_0)

Dados discretos: valor que ocorre com mais frequência.

Dados contínuos: classe modal: classe com maior frequência.

2.2 Estatísticas de dispersão.

Amplitude

Dados não agrupados: $\alpha = \max(x_i) - \min(x_i)$

Dados agrupados: $\alpha = l_{m+1} - l_1$

Amplitude inter-quartis: $q = q_{0.75} - q_{0.25}$

Variância amostral (S^2):

$$S^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Variância amostral corrigida ($\hat{S}^2$):

$$\hat{S}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

Desvio padrão (S)

$$S = \sqrt{\text{variância}}$$

Erro Padrão ($\hat{S}$):

$$\hat{S} = \sqrt{\text{variância amostral corrigida}}$$

2.3 Outras estatísticas.

Coeficiente de assimetria: afere como os valores da colecção se comportam em torno da respectiva média.

$$C_A = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{S^3}$$

Notas:

Se os valores da colecção se distribuem de forma simétrica em relação à média, há sempre compensação entre os desvios positivos e os desvios negativos, pelo que $C_A = 0$.

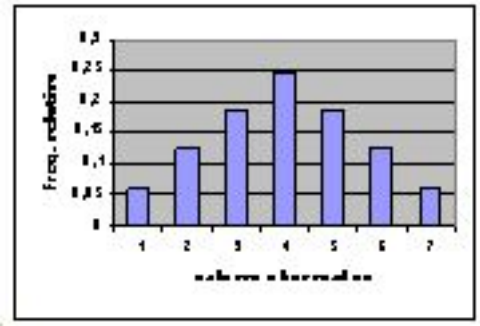

Figura 3:

Exemplo:

- Se $m_0 < m_e < \overline{x}$, os desvios positivos dominam e tem-se $C_A > 0$.
  Ex.: Série 2$\rightarrow$ $m_0 = 1,\ m_e = 1.5,\ \overline{x} = 1.9844,\ C_A = 1.6635$

- Se $\overline{x} < m_e < m_0$, os desvios negativos dominam e tem-se $C_A < 0$.
  Ex.: Série 1$\rightarrow$ $m_0 = 7,\ m_e = 6.5,\ \overline{x} = 6.0156,\ C_A = -1.6635$

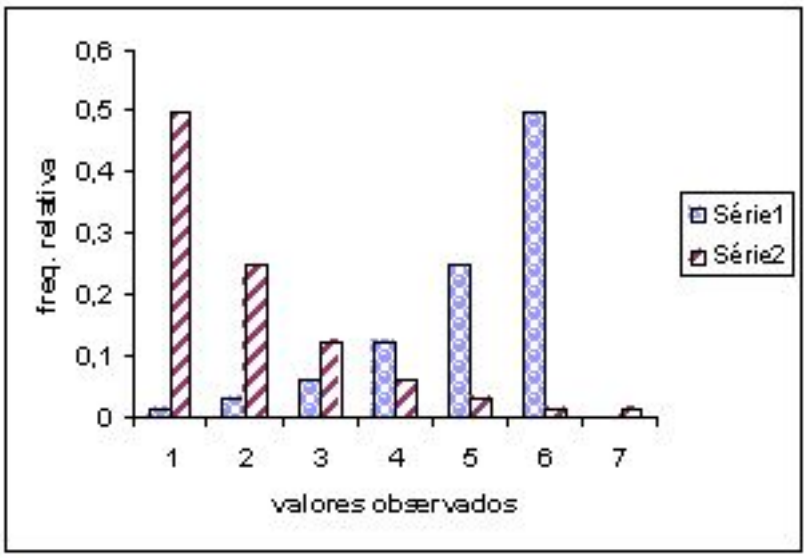

Figura 4: Amostra bivariada.

2.4 Caracterização de amostras bivariadas.

Associação entre fenómenos de natureza quantitativa.
Diagrama de dispersão (nuvem de pontos)
Inexistência de correlação linear Correlação linear positiva Correlação linear negativa
Covariância

$$S_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$$

Coeficiente de correlação amostral, r mede o grau de relacionamento linear entre as variáveis.

$$\boxed{r = \frac{S_{xy}}{S_x \cdot S_y}}$$

onde

$$S_x = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

e

$$S_y = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

Recta a estimar: $y = ax + b$
Aplicando o método dos mínimos quadrados tem-se:
Recta de regressão de y em x

$$y - \overline{y} = \frac{S_{xy}}{S_x^2}(x - \overline{x})$$

Recta de regressão de x em y

$$x - \overline{x} = \frac{S_{xy}}{S_y^2}(y - \overline{y})$$

Exemplo 2.4.1 *Diagrama de Dispersão*

Aluno n.º	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Horas de estudo(x)	*4*	*2*	*8*	*0*	*7*	*6*	*11*	*14*	*12*	*10*	*13*	*9*	*2*	*8*	*14*	*4*
Notas no teste de Biologia (y)	*11*	*10*	*13*	*6*	*10*	*11*	*16*	*18*	*17*	*15*	*14*	*12*	*8*	*12*	*16*	*15*

$S_{xy} = 11.9375 r = 0.85$
Recta a estimar:

$$y - 12.75 = \frac{11.9375}{4.323^2}(x - 7.75) \Leftrightarrow y = 0.6388x + 7.8$$

Exemplo 2.4.2 *No quadro que se segue apresentam-se os valores da massa monetária M_1^- (circulação monetária mais depósitos à ordem) em milhões de contos e as taxas de juro (em percentagem) dos Boletins do Tesouro de 181 dias a 1 ano cedidos sem recurso pelos bancos. Os dados constam no Boletim Estatístico do Banco de Portugal.*

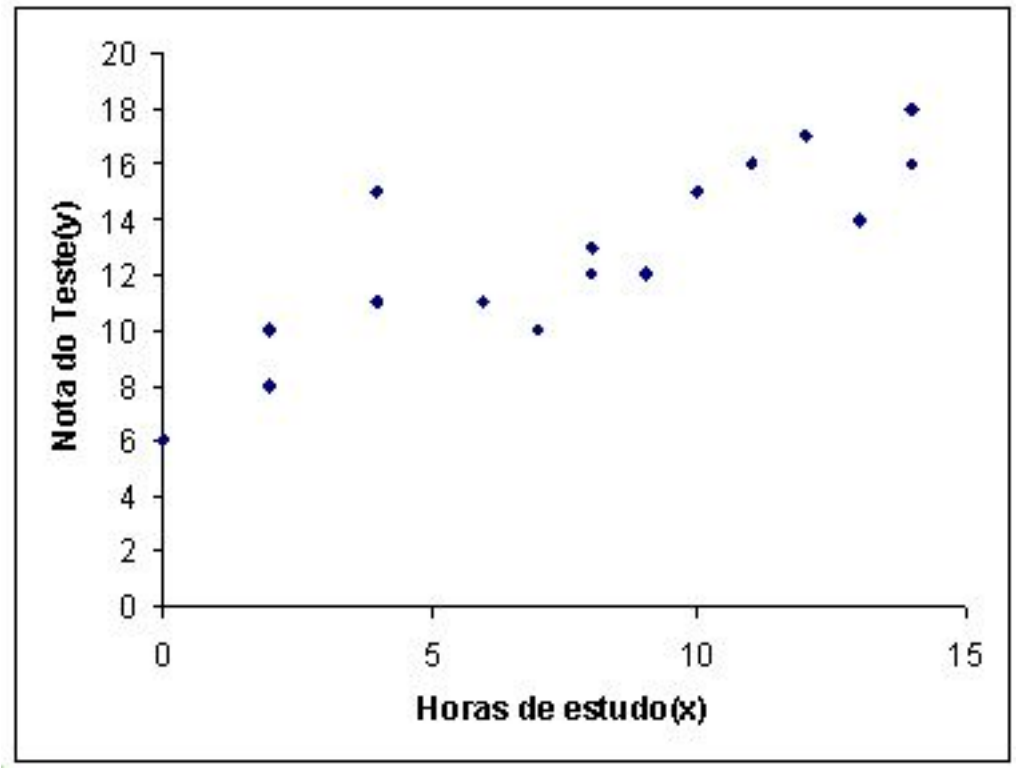

Figura 5:

Mês	M_1^-	*Taxa Juro*	*Mês*	M_1^-	*Taxa Juro*	*Mês*	M_1^-	*Taxa Juro*
Abr-95	*3422*	*10,9*	*Jun-96*	*3843,8*	*7*	*Ago-97*	*4561,3*	*5,2*
Mai-95	*3421,4*	*10,7*	*Jul-96*	*4070,5*	*7,1*	*Set-97*	*4591*	*5*
Jun-95	*3546,9*	*10,2*	*Ago-96*	*4013,7*	*7,1*	*Out-97*	*4558,1*	*5,1*
Jul-95	*3899,6*	*9,8*	*Set-96*	*3917,9*	*7*	*Nov-97*	*4880,5*	*4,9*
Ago-95	*3789,3*	*9,7*	*Out-96*	*3862,4*	*6,8*	*Dez-97*	*4901,9*	*4,6*
Set-95	*3731*	*9*	*Nov-96*	*4042,6*	*6,7*	*Jan-98*	*4735,1*	*4,5*
Out-95	*3672,8*	*9,1*	*Dez-96*	*4301,9*	*6,3*	*Fev-98*	*4698,1*	*4,4*
Nov-95	*3772,9*	*8,8*	*Jan-97*	*4184*	*6*	*Mar-98*	*4840*	*4,1*
Dez-95	*3885*	*8,7*	*Fev-97*	*4088,3*	*5,9*	*Abr-98*	*4841,1*	*4,3*
Jan-96	*3760,3*	*8,3*	*Mar-97*	*4153*	*6*	*Mai-98*	*4901,2*	*4,2*
Fev-96	*3702,8*	*7,9*	*Abr-97*	*4148,1*	*5,8*	*Jun-98*	*5131,6*	*4*
Mar-96	*3624,1*	*7,7*	*Mai-97*	*4238,6*	*5,6*	*Jul-98*	*5494,2*	*3,9*
Abr-96	*3698,4*	*7,4*	*Jun-97*	*4482,6*	*5,5*	*Ago-98*	*5284,3*	*3,8*
Mai-96	*3725,3*	*7*	*Jul-97*	*4655,6*	*5,4*	*Set-98*	*5322,3*	*3,8*

Nesta situação pode afirmar-se que os aumentos da taxa de juro implicam, ***em média****, diminuições da massa monetária. Trata-se de uma relação estatística.*

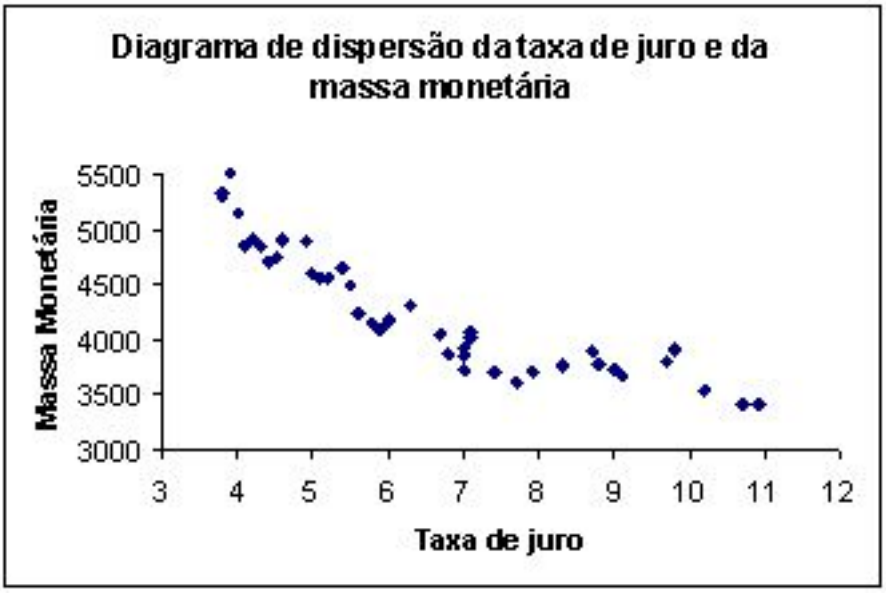

Figura 6: Diagrama de dispersão da taxa de juro.

3 Teoria elementar da probabilidade.

3.1 Definições.

A Teoria das Probabilidades surge com a necessidade de analisar fenómenos ou experiências não deterministas.

Experiência aleatória:

1. Pode ser repetida em condições análogas;
2. É conhecido o conjunto de todos os resultados possíveis da experiência;
3. Em cada realização da experiência não se sabe qual dos resultados possíveis irá ocorrer.

Exemplos:

1. Lançar uma moeda (com faces cara e escudo);
2. Lançar um dado;
3. Fazer dois lançamentos de uma moeda;
4. Momentos de entrada e saída de um cliente numa loja que abre às 9 horas e encerra às 19 horas.

Espaço Amostral (Ω)- Conjunto de todos os resultados possíveis da experiência (esses resultados não são decomponíveis).

Acontecimento- subconjunto do espaço amostral, ou seja, um conjunto formado por resultados elementares.

Exemplo 3.1.1 *No lançamento de duas moedas,onde* $\Omega = \{FF, FC, CF, CC\}$*, há quatro resultados elementares:* $\{FF\}, \{FC\}, \{CF\}, \{FC\}$.
Dê exemplos de outros acontecimentos.

Note-se que, como os *acontecimentos* são conjuntos, podem efectuar-se as operações de:

- reunião ($A \cup B$);
- intersecção ($A \cap B$);
- diferença ($A - B$ e $B - A$);
- complementaridade ($\overline{A}$ e $\overline{B}$).

Seja A um acontecimento. Diz-se que A é um:

- *acontecimento impossível* se não contém nenhum elemento do espaço amostral ($A = \phi$).
- *acontecimento certo* se contém todos os elementos do espaço amostral ($A = \omega$).

Sejam A e B dois acontecimentos. Diz-se que A e B são:

- *acontecimentos incompatíveis* se não têm nenhum acontecimento elementar comum, isto é, se:
$$A \cap B = \phi$$
- *acontecimentos mutuamente exclusivos* se a realização de um **implica** a não realização do outro, isto é, se:
$$A \cap B = \phi$$
e se
$$A \cup B = \Omega$$

3.2 Medida de probabilidade.

A medida de probabilidade é uma função P que a cada acontecimento A faz corresponder um número real, $P(A)$, probabilidade do acontecimento A, que verifica os três axiomas seguintes:

1. $P(A) \geqslant 0$
2. $P(\Omega) = 1$
3. Se A e B forem acontecimentos incompatíveis, $A \cap B = \phi$, então

$$P(A \cup B) = P(A) + P(B)$$

Propriedades da medida de probabilidade

1. $P(A) = 1 - P(\overline{A})$
2. $P(\phi) = 0$
3. $P(B - A) = P(B) - P(A \cap B)$
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
5. $P(A) \leqslant 1$

Exemplo 3.2.1 *No País* **Lê-se Tudo**, *9,8% das pessoas adquirem a revista* **Lê-se Bem**, *22,9% a revista* **Lê-se Melhor** *e 5,1% ambas as revistas. Admite-se que a medida de probabilidade é a proporção dos indivíduos da população que adquirem as revistas.*
Calcule a probabilidade de uma pessoa escolhida ao acaso:

1. *adquirir somente a revista* **Lê-se Bem**.
2. *adquirir pelo menos uma das revistas.*
3. *não adquirir nenhuma das revistas.*

3.3 Interpretações do conceito de probabilidade.

1. Interpretação Clássica (Laplace)

$$\boxed{P(A) = \frac{\text{n.º de casos favoráveis a } A}{\text{n.º de casos possíveis}}}$$

2. Interpretação frequencista

$$\boxed{P(A) = \lim_{N \to +\infty} \frac{N_A}{N},}$$

sendo:

- N_A n.º de ocorrências do acontecimento A
- N n.º de experiências.

3. Interpretação Subjectiva

Exemplo 3.3.1 *A ideia básica da interpretação frequencista é bem perceptível na seguinte frase:*

"Em Portugal, durante a década de 70, a ■probabilidade■ de um nado-vivo ser do sexo masculino cifrava-se em cerca de 0,5165". Para proferir tal afirmação registou-se no fim de vários anos a frequência relativa de nados-vivos do sexo masculino. De notar que de ano para ano as frequências relativas pouco diferem entre si, aproximando-se do valor calculado para o conjunto dos seis anos (ver quadro 1).

Anos	Número de nados-vivos	Do sexo masculino	Frequência relativa
1972	174 695	90 332	0.5171
1973	172 324	88 697	0.5147
1974	171 979	88 062	0.5121
1975	179 628	93 089	0.5182
1976	186 712	96 582	0.5172
1977	181 064	94 040	0.5172
Total	1 066 402	550 082	0.5165

Tabela 1: Número de nados-vivos total e do sexo masculino, em Portugal.

3.4 Análise combinatória.

Análise combinatória (ou cálculo combinatório) é um conjunto de técnicas de contagem do número de elementos de certo tipo de conjuntos. Tem um papel fundamental na definição clássica de probabilidade de Laplace.

0.5cm

Considere-se um conjunto de n elementos todos distintos. Vamos seleccionar alguns (ou todos) os elementos de entre os n disponíveis e formar uma sequência **linear e ordenada** desses elementos. Nestas condições, diz-se que:

⋈ **Permutação** é qualquer sequência formada pelos n elementos. O número total de permutações possíveis com n elementos é:

$$\boxed{P_n = n! = n \times (n-1) \times \cdots \times 2 \times 1}.$$

⋈ **Arranjo** é qualquer sequência formada com apenas p elementos ($p \leqslant< n$). O número total de arranjos possíveis com p elementos é:

$$\boxed{A_p^n = \frac{n!}{(n-p)!}}$$

Note-se que se $p = n$ então o arranjo é uma permutação, pois

$$A_n^n = n! = P_n.$$

Exemplo 3.4.1 *Quantas palavras (com ou sem sentido) de 4 letras podem formar-se com as letras da palavra* **LISBOA***?*

Exemplo 3.4.2 *De entre os 20 alunos de uma turma vão ser escolhidos três para representar a turma em três países diferentes: o primeiro irá para Inglaterra, o segundo para França e o terceiro a Itália. De quantos modos distintos pode ocorrer a escolha?*

Exemplo 3.4.3 *Quantas classificações gerais diferentes podem ocorrer com os* 13 *cavalos que participam numa corrida?*

$\bowtie$ **Combinação** de elementos é qualquer subconjunto formado por p elementos escolhidos de entre os de um conjunto dado. O número total de combinações possíveis com p elementos é:

$$\boxed{C_p^n = \frac{n!}{p!(n-p)!}}$$

Exemplo 3.4.4 *Das dez raparigas de uma turma vão ser escolhidas quatro para dançar numa quadrilha. Quantas são as diferentes escolhas possíveis?*

3.5 Probabilidade condicionada.

Dados dois acontecimentos, A e B, a probabilidade de A se realizar sabendo-se que B se realizou - ou probabilidade de A condicionada a B, designada por $P(A|B)$ - é definida por

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)} \text{ se } P(B) > 0}$$

Exemplo 3.5.1 *Considere duas caixas: caixa A e caixa B.*
A caixa A contém duas bolas verdes e cinco bolas amarelas.
A caixa B contém seis bolas veres e uma bola amarela.

Lança-se um dado equilibrado, com as faces numeradas de 1 a 6.
Se sair face 1, tira-se, ao acaso uma bola da caixa A.
Caso contrário, tira-se, ao acaso uma bola da caixa B.

Considere os acontecimentos:
X*: Sair face par no lançamento do dado*
Y*: Sair Bola verde*

Sem aplicar a fórmula da probabilidade condicionada, *indique, justificando a sua resposta, o valor de* $P(Y|X)$.
Nota: *Comece por indicar o significado de* $P(Y|X)$ *no contexto da situação descrita.*

Exemplo 3.5.2 *O sangue humano está classificado em quatro grupos distintos: A, B, AB e O. Independentemente do grupo, o sangue pode possuir, ou não, o factor Rhésus. Se o sangue de uma pessoa possui este factor, diz-se Rhésus positivo (Rh^+); se não possui este factor, diz-se Rhésus negativo (Rh^-).*
Na população portuguesa, os grupos sanguíneos e os respectivos Rhésus estão repartidos da seguinte forma:

	A	B	AB	O
Rh^+	40,0%	6,9%	2,9%	35,4%
Rh^-	6,5%	1,2%	0,4%	6,7%

1. *Escolhido um português ao acaso, qual é a probabilidade de o seu grupo sanguíneo não ser o O?*

2. *Escolhido um português ao acaso, e sabendo que é Rhésus negativo, qual a probabilidade de o seu grupo sanguíneo ser o A?*

3.6 Teorema da probabilidade total.

Se os acontecimentos $A_1, A_2, \cdots, A_n$ são mutuamente exclusivos, dois a dois, e a sua reunião é Ω, então:

$$\begin{aligned} P(B) &= \sum_{i=1}^{n} P(B|A_i)P(A_i) \\ &= P(B|A_1) \times P(A_1) + P(B|A_2) \times P(A_2) + \cdots \\ &\quad + \cdots + P(B|A_n) \times P(A_n) \end{aligned}$$

3.7 Fórmula de Bayes.

Se os acontecimentos $A_1, A_2, \cdots, A_n$ são mutuamente exclusivos, dois a dois, e a sua reunião é Ω, então, para B definido em Ω:

$$P(A_j|B) = \frac{P(B|A_j) \times P(A_j)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}, \text{ para } j = 1, \cdots, n$$

Exemplo 3.7.1 *As famílias da cidade* **Passatempo** *escolhem uma das três alternativas para fazer férias: praia, campo ou ficar em casa.*

Durante a última década do século XX, verificou-se que escolhiam aquelas alternativas respectivamente 50%, 30% e 20% das famílias.

A probabilidade de descansar durante as férias está relacionada com a alternativa escolhida: 0,6; 0,75 e 0,35 conforme se tenha ido para a praia, para o campo ou ficado em casa.

1. *Qual a probabilidade de uma família da cidade* **Passatempo** *descasar durante as férias.*

2. *A família* **Boaviagem** *descansou durante as férias. Qual a alternativa mais provável de ter sido escolhida por esta família?*

3.8 Acontecimentos independentes.

Diz-se que os acontecimentos A e B são **independentes** se o facto de um deles se ter realizado não influenciou a realização ou não do outro. Nesse caso

$$P(A \cap B) = P(A)P(B)$$

se A e B forem acontecimentos independentes, então

$$P(A|B) = P(A), \text{ se } P(B) > 0,$$
$$P(B|A) = P(B), \text{ se } P(A) > 0.$$

$A_1, A_2, \cdots, A_n$ dizem-se independentes se

- $P(A_i \cap A_j) = P(A_i)P(A_j)$, $i \neq j$
- $P(A_i \cap A_j \cap A_k) = P(A_i)P(A_j)P(A_k)$, $i \neq j, i \neq k, j \neq k$
- $\cdots$
- $P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i)$

Exemplo 3.8.1 *Suponha que no fabrico de certo artigo podem ocorrer, independentemente um do outro, dois (e só dois) tipos de defeito: de tipo I, com probabilidade 0,1, e de tipo II, com probabilidade 0,2. Qual a probabilidade de um artigo escolhido ao acaso ser defeituoso?*

4 Variáveis aleatórias e distribuição de probabilidade.

Uma **variável aleatória** (v.a.) utiliza-se para expressar os resultados de uma experiência aleatória, sendo uma função $X : \Omega \to \mathbb{R}$ que a cada $\omega \in \Omega$ faz corresponder um número real $X(\omega)$. Denota-se em geral por uma letra maiúscula X, Y, etc.

Exemplo 4.0.2 *Sejam N e M as v.a.'s que representam respectivamente o número de "caras" e o número de mudanças de face em três lançamentos de uma moeda equilibrada. Complete os espaços em branco.*

ω	$N(\omega)$	$M(\omega)$
$C_1C_2C_3$	...	...
$C_1C_2E_3$	...	...
$C_1E_2C_3$	...	...
$C_1E_2E_3$	...	...
$E_1C_2C_3$	...	...
$E_1C_2E_3$	...	...
$E_1E_2C_3$	...	...
$E_1E_2E_3$	...	...

$P(N = 0) = \ldots\ldots$ $\quad P(N = 1) = \ldots\ldots$ $\quad P(1 < N \leqslant 3) = \ldots\ldots$

$P(M = 0) = \ldots\ldots$ $\quad P(M = 1) = \ldots\ldots$ $\quad P(M > 1) = \ldots\ldots$

4.1 Distribuições de variáveis discretas.

Caracterizam-se por terem contradomínio finito ou numerável, ou seja, que se pode representar na forma $\{x_1, x_2, x_3, \cdots\}$. Neste texto designa-se por p_i a probabilidade $P(X = x_i)$.

Chama-se **função de probabilidade** à função que a cada x_i faz corresponder p_i.

$$X : \begin{cases} x_1 & x_2 & x_3 & \cdots \\ \\ p_1 & p_2 & p_3 & \cdots \end{cases}$$

A função de probabilidade satisfaz as seguintes condições:

1. $0 < p_i < 1$
2. $\sum_i p_i = 1$

e pode representar-se graficamente.

Exemplo 4.1.1 *Num lote de cinco componentes electrónicos sabe-se que dois deles apresentam defeitos. Selecciona-se uma amostra de dois componentes, sem reposição.*

Sendo X a v.a. que representa o número de componentes com defeito na amostra, indique a função de probabilidade de X e a correspondente representação gráfica.

Chama-se **função de distribuição**(f.d.) de uma v.a. X a uma função real de variável real, onde

$$F(x) = P(X \leqslant x) = \sum_{x_i \leqslant x} P(X = x_i)$$

A função de distribuição é uma função em escada, não decrescente, descontínua à esquerda e goza das seguintes propriedades:

1. $0 \leqslant F(x) \leqslant 1$
2. $P(a < X \leqslant b) = F(b) - F(a)$

Continuação do exemplo 4.1.1: Determine a função de distribuição de X e represente-a graficamente. Calcule a probabilidade de X ser menor ou igual a 1.

Exemplo 4.1.2 *Um saco tem cinco bolas, numeradas de 1 a 5. Extraem-se duas bolas sem reposição. Seja X a variável aleatória que representa o maior valor observado.*

1. *Determine a função de probabilidade e a função distribuição de X.*
2. *Qual a probabilidade do maior valor observado ser superior a 3?*

4.2 Distribuições de variáveis contínuas.

Seja X uma v.a. com um contradomínio infinito não numerável (por exemplo um intervalo, $\mathbb{R}^+$, $\mathbb{R}$, etc).

Chama-se **função densidade de probabilidade** da v.a. X, designando-se por $f(x)$, a uma função tal que

$$P(X \in [x, x + dx]) = f(x) \times dx.$$

Dos axiomas de probabilidade e da definição de $f(x)$ resulta que:

1. $f(x) \geqslant 0$;
2. $P(a \leqslant X \leqslant b) = \int_a^b f(x)dx$;
3. $\int_{-\infty}^{+\infty} f(x) = 1$.

Função distribuição de probabilidade

Chama-se **função distribuição de probabilidade** da v.a. X e designa-se por $F(x)$ a função que associa a cada valor de x a probabilidade de X ser inferior ou igual a x.

$$F(x) = P(X \leqslant x) = \int_{-\infty}^{x} f(t)dt$$

Graficamente $F(x)$ é a área da figura compreendida entre o eixo das abcissas e a curva $f(t)$, para $t \leqslant x$ (ver figura).

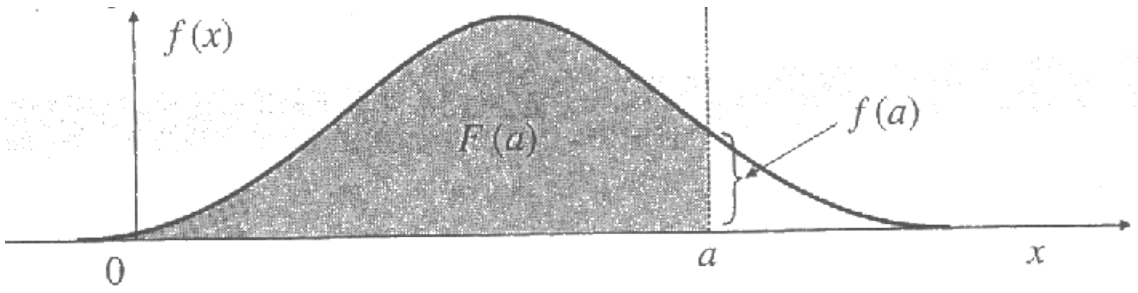

Figura 7: Função de distribuição.

A função de distribuição satisfaz as seguintes propriedades:

1. $P(X \leqslant a) = F(a)$;
2. $P(a \leqslant X \leqslant b) = F(b) - F(a)$.

Exemplo 4.2.1 *1. Determine a constante k de modo que a função definida por*

$$f(x) = \begin{cases} kx, & 0 < x < 4 \\ 0, & \textit{caso contrário} \end{cases}$$

represente uma função densidade de probabilidade.

2. Calcule $P(1 \leqslant X \leqslant 3)$.

3. Determine a função distribuição de probabilidade.

4. Utilizando o resultado anterior calcule $P(1 \leqslant X \leqslant 3)$.

5. Represente graficamente as funções densidade e distribuição de probabilidade.

4.3 Parâmetros das distribuições.

	Variáveis discretas	Variáveis contínuas
valor esperado	$\mu_Y = E(Y) = \sum_i y_i p_i$	$E(X) = \int_{-\infty}^{+\infty} x \times f(x)dx$
Mediana	$m_e = \chi_{\frac{1}{2}} = \inf\{x : F(x) \geqslant 0.5\}$	$m_e = x$ tal que $F(x) = 0.5$
Moda	$Mod(Y)$: valo(res) de y_i para os quais p_i é máximo	$Mod(X)$: valor(es) de x para os quais $f(x)$ é máximo
Quartis ($p \in \{0.25, 0.5, 0.75\}$)	$\chi_p = \inf\{x : F(x) \geqslant p\}$	$\chi_p = x$ tal que $F(x) = p$

Parâmetros de dispersão

	variáveis discretas	variáveis contínuas
variância	$\sigma_Y^2 = Var(Y) = \sum_i (y_i - \mu_Y)^2 p_i$	$\sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \cdot f(x)dx$
desvio padrão	$\sigma_Y = \sqrt{\sigma_Y^2}$	$\sigma_X = \sqrt{\sigma_X^2}$

Exemplo 4.3.1 *Calcule os parâmetros de localização e dispersão das v.a.'s presentes nos exercícios 4.1.2 e 4.2.1.*

coeficiente de assimetria

variáveis discretas	variáveis contínuas
$C_A = \frac{\sum_i (y_i - \mu_Y)^3 \cdot p_i}{\sigma_Y^3}$	$C_A = \frac{\int_{-\infty}^{+\infty} (x - \mu_X)^3 \cdot f(x)dx}{\sigma_X^3}$

Parâmetros populacionais $\neq$ Medidas amostrais

Os parâmetros são calculados com base em probabilidades, enquanto as medidas amostrais são calculadas com base em frequências observadas numa amostra.

De referir que os parâmetros são fixos e as medidas amostrais variam de amostra para amostra.

4.4 Transformação linear de uma variável.

Seja Z uma variável qualquer (discreta ou contínua) à qual se aplica uma transformação linear.

$$W = a + b \cdot Z \Rightarrow \begin{cases} E(W) = a + b \cdot E(Z) \\ \\ VAR(W) = b^2 \cdot VAR(Z) \end{cases}$$

Exemplo 4.4.1 *Considere a v.a. X com $E(X) = 4$ e $VAR(X) = 0.5$. Calcule o valor esperado e a variância da v.a. $Y = 2X + 4$.*

5 Distribuições de variáveis discretas.

5.1 Prova de Bernoulli. Distribuição binomial.

Prova de Bernoulli - Experiência aleatória em que se observa a realização ou não realização de determinado acontecimento A com probabilidade p.

A realização de A diz-se um **sucesso** e a sua não realização, que ocorre com probabilidade $q = 1 - p$ diz-se um **insucesso**.

Exemplo 5.1.1 *Qual a probabilidade de, em cinco lançamentos de um dado ■perfeito■, se obter duas vezes a sena?*

Consideremos n provas de Bernoulli independentes, isto é, n experiências aleatórias em cada uma das quais se observa a realização ou não realização de determinado acontecimento A, com probabilidade p, constante de prova em prova.

A v.a. X representa o número de vezes que ocorre "sucesso". Dizemos que X tem distribuição binomial com parâmetros n e p, abreviadamente

$$X \sim B(n, p)$$

Tem-se

$$\boxed{P(X = k) = C_k^n p^k (1-p)^{n-k}, \text{ para } k = 0, 1, \cdots, n}$$

$E(X)$	$\mu = np$
$Var(x)$	$\sigma^2 = npq$
	$\sigma = \sqrt{npq}$

Exemplo 5.1.2 *Numa população, 20% dos trabalhadores são agrícolas. Numa amostra aleatória de 6 pessoas dessa população, determine a probabilidade de:*

1. *não haver trabalhadores agrícolas;*
2. *haver um único agricultor;*
3. *existirem pelo menos dois trabalhadores agrícolas.*

Exemplo 5.1.3 *Numa lotaria com um milhão de bilhetes, numerados de* 000000 *a* 999999*, qual a probabilidade de o primeiro prémio recair num número formado totalmente por "números ímpares".*

Exemplo 5.1.4 *Num armazém estão preparadas para distribuição 500 embalagens de um produto dos quais exactamente 50 estão deterioradas. É efectuada uma inspecção sobre uma amostra de 10 embalagens escolhidas ao acaso com reposição. A inspecção rejeita a comercialização quando encontra pelo menos 3 três embalagens deterioradas.*

1. *Qual a probabilidade da inspecção rejeitar o lote?*
2. *Qual a dimensão mínima que a amostra deve ter, de modo que a probabilidade de rejeição seja superior a 5%?*

5.2 Distribuição hipergeométrica.

Consideremos uma população com N elementos, dos quais uma certa proporção p possui uma determinada característica, ou seja, $N \cdot p$ elementos da população têm essa característica e $N \cdot (1-p)$ não a têm. Admitamos que queremos retirar ao acaso, ***sem reposição***, uma amostra de n elementos dessa população. A v.a. X representa o n.º de sucessos ocorridos em n extracções sem reposição. Diz-se que X tem distribuição hipergeométrica com parâmetros $M \times p$, $M \times q$ e n, abreviadamente

$$X \sim HG(M \times p, M \times q, n),$$

onde

M-n.º total de objectos/ indivíduos que constituem a população.
n-n.º de elementos da amostra
$q = 1 - p$.
Tem-se

$$P(X = k) = \frac{(C_k^{Mp})(C_{n-k}^{Mq})}{(C_n^M)}.$$

$E(X)$	$\mu = np$
$Var(x)$	$\sigma^2 = npq\frac{M-n}{M-1}$
	$\sigma = \sqrt{npq\frac{M-n}{M-1}}$

Exemplo 5.2.1 *Admita que, de um lote de 20 peças, das quais 2 são defeituosas, se extrai uma amostra de 5 peças sem reposição.*

Qual a probabilidade da amostra apenas conter uma peça defeituosa?

Exemplo 5.2.2 *Num concurso de totoloto qual a probabilidade de serem extraídas 4 bolas com número ímpar e três com número par?*

5.3 Relação entre as distribuições hipergeométrica e binomial.

Quando M é grande comparado com n

$$B(n,p) \approx HG(Mp, Mq, n)$$

Regra prática: $M \geqslant 10n$

Exemplo 5.3.1 *De um grupo de 1000 explorações agrícolas de uma certa região há 20 que produzem espargos. Recolhem-se ao acaso uma amostra de 100 explorações, com e sem reposição.*

Quadro 1: Comparação entre as distribuições binomial e hipergeométrica($M = 100; Mp = 20($ logo $p = 0.02); n = 100$)

x	Binomial	Hipergeométrica
0	0.1326	0.1190
1	0.2707	0.2702
2	0.2734	0.2881
3	0.1823	0.1918
4	0.0902	0.0895
5	0.0353	0.0311
6	0.0114	0.0083
7	0.0031	0.0018
8	0.0007	0.0003
9	0.0002	0.0000
...	...	...

5.4 Distribuição de Poisson.

A distribuição de Poisson permite descrever certos tipos de fenómenos ou acontecimentos cuja ocorrência se repete no tempo ou no espaço.

Exemplo 5.4.1 *Constituem exemplos de situações que se adequam a uma distribuição de Poisson:*

- *o número de doentes que afluem em cada domingo nas urgências do Hospital de Mirandela;*
- *o número de avarias de uma máquina durante o dia;*
- *o número de automóveis que chegam a um parque de estacionamento entre as 8h30m e as 10h30m.*

A **distribuição de Poisson** aplica-se quando se verificam as seguintes condições:

- os acontecimentos ocorrem um a um e não em grupos;

- a probabilidade de um acontecimento ocorrer num intervalo (de tempo ou de espaço) de amplitude arbitrariamente pequena é proporcional à dimensão do intervalo e independente de ocorrências verificadas em intervalos anteriores.

Função de probabilidade

X: contagem do número de ocorrências de um acontecimento num intervalo unitário (uma hora, um mês, um ano)

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

onde λ representa o parâmetro da distribuição. Simbolicamente escreve-se

$$X \sim Poisson(\lambda)$$

$E(X)$	$\mu = \lambda$
$Var(x)$	$\sigma^2 = \lambda$
	$\sigma = \sqrt{\lambda}$
coef. assimetria	$\gamma_1 = \frac{1}{\sqrt{\lambda}}$

Nota: Sejam X_1 e X_2 duas variáveis aleatórias independentes tais que $X_1 \sim Poisson(\lambda_1)$ e $X_2 \sim Poisson(\lambda_2)$. Então

$$X_1 + X_2 \sim Poisson(\lambda_1 + \lambda_2).$$

Exemplo 5.4.2 *Numa fábrica de têxteis existem numerosos teares do mesmo tipo. Depois de muitas observações chegou-se à conclusão de que o número de teares que se avariam mensalmente é uma v.a. com distribuição de* $Poisson(\lambda = 3)$.

1. *Qual a probabilidade de no mês de Dezembro se avariem dois teares?*
2. *Qual a probabilidade de no mesmo período se avariem pelo menos sete teares?*
3. *Determine a capacidade mínima mensal disponível, C, da oficina de reparação de modo a ser pelo menos 0.9 a probabilidade de não haver teares a aguardar reparação.*

Exemplo 5.4.3 *Admita que o número de camiões TIR que, por hora, atravessam a ponte 25 de Abril é uma v.a. com distribuição de Poisson e desvio padrão igual a 3.*

1. *Qual a probabilidade de que, numa hora, exactamente 8 camiões TIR atravessem a ponte?*
2. *Calcule a probabilidade de que, em duas horas, pelo menos seis camiões TIR atravessem a ponte?*

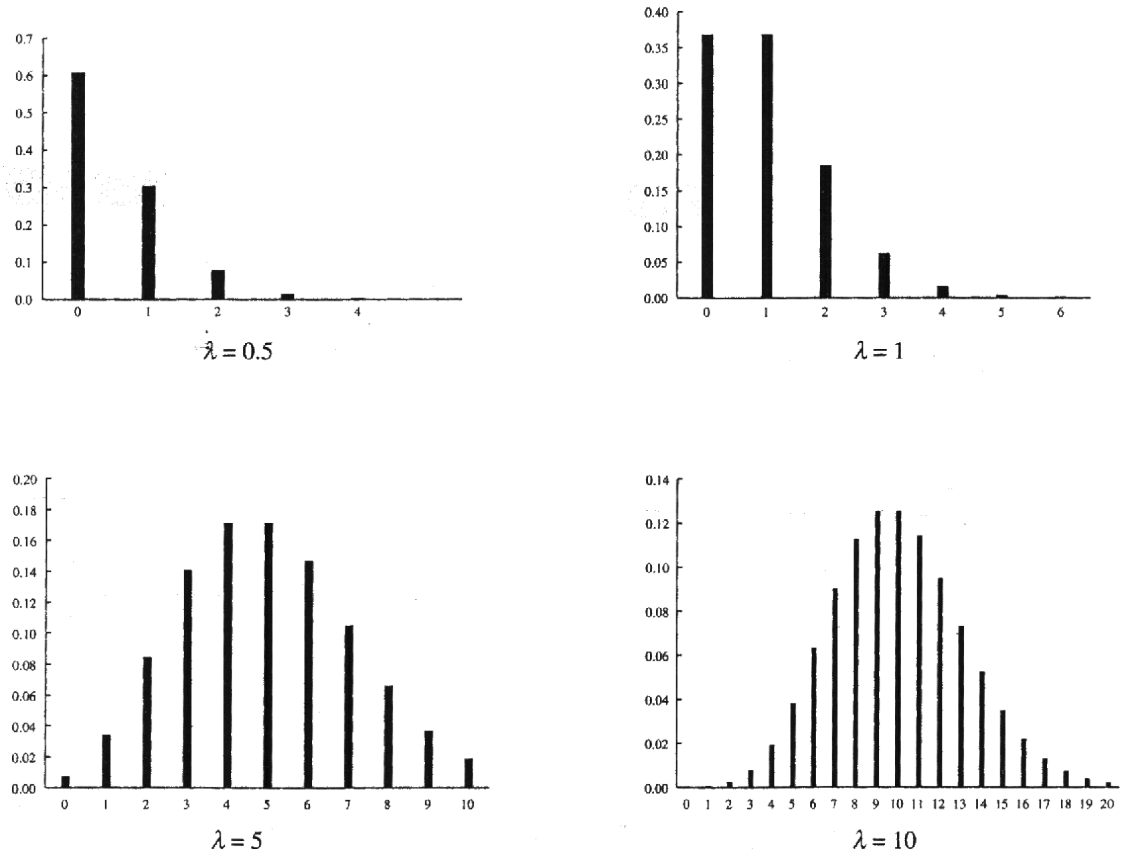

Figura 8: Função de probabilidade da distribuição de Poisson para $\lambda = 0.5$, $\lambda = 1$, $\lambda = 5$ e $\lambda = 10$.

5.5 Relação entre a distribuição de Poisson e a distribuição binomial.

$$\lim_{\substack{n\to+\infty \\ p\to 0}} B(n,p) \equiv Poisson(\lambda = np)$$

Assim, se X segue uma distribuição $B(n,p)$ em que n é grande e p é pequeno, então X tem *aproximadamente* distribuição $Poisson(np)$.

Nesta situação o "sucesso" é um acontecimento raro, dado que o valor de p é pequeno, o que justifica que a distribuição de Poisson seja também conhecida por *"lei dos acontecimentos raros"*.

Tal aproximação é aconselhável quando:

- $n > 20$;
- $p \leqslant 0.1$ (o "sucesso" em causa é raro") ou $p \geqslant 0.9$ (o acontecimento complementar é raro)

Exemplo 5.5.1 *A probabilidade de determinada peça, produzida por certa máquina, ser defeituosa é de 0.001. Qual a probabilidade de, num lote de 1000 peças, haver mais que uma defeituosa?*

Exemplo 5.5.2 *Admita que 2% dos indivíduos de uma certa população são canhotos. Numa amostra de 100 indivíduos, escolhidos aleatoriamente, calcule a probabilidade de haver exactamente 5 canhotos.*

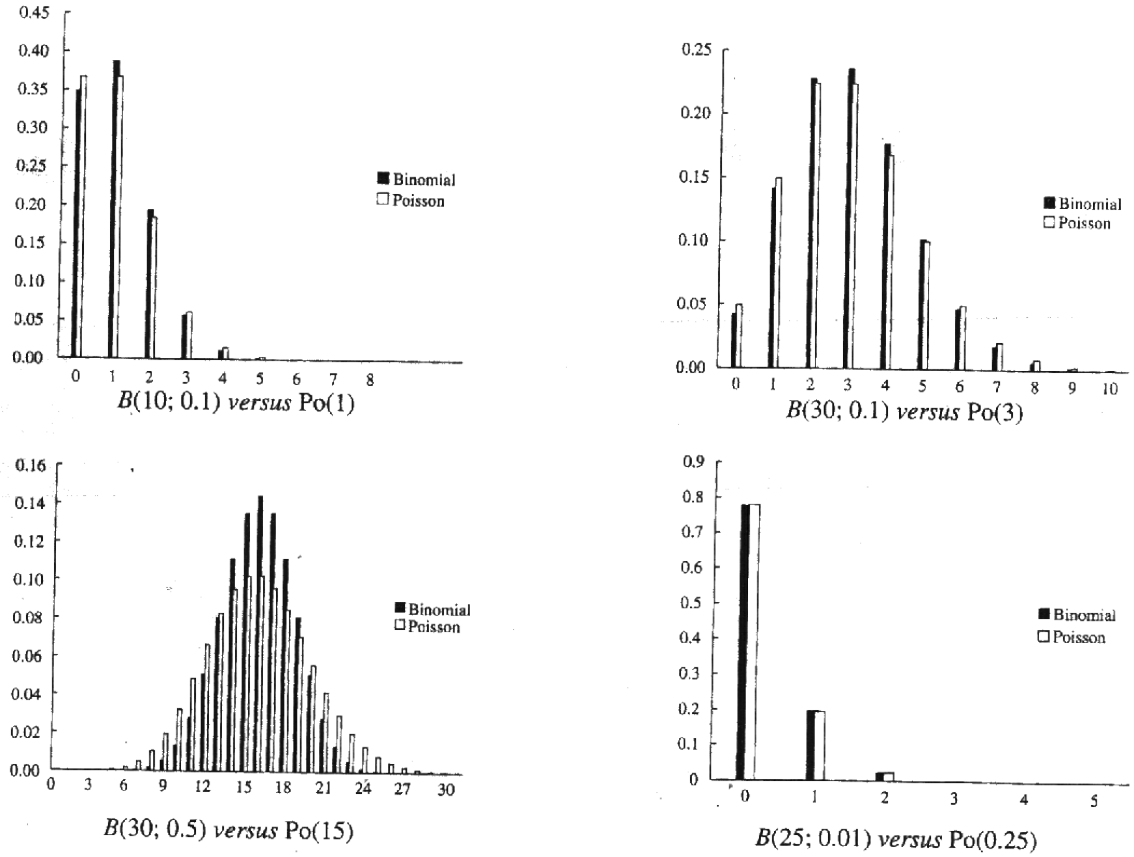

Figura 9: Aproximação da dist. de Poisson à dist. binomial.

6 Distribuições de variáveis contínuas.

6.1 Distribuição uniforme, $U(a, b)$.

Diz-se que uma variável aleatória, X, tem distribuição uniforme no intervalo (a, b), com $a < b$, quando a função densidade é constante nesse intervalo e nula fora dele, ou seja,

$$\boxed{f(x) = \begin{cases} c & , \text{ se } a \leqslant< x <\leqslant b, \\ 0 & , \text{ caso contrário.} \end{cases}}$$

Simbolicamente $X \sim U(a, b)$.

Qual o valor de c? A função distribuição de $f(x)$ é

$$F(x) = \begin{cases} 0 & , \text{ se } x < a \\ \frac{x-a}{b-a} & , \text{ se } a \leqslant x < b \\ 1 & , \text{ se } x \geqslant b \end{cases}$$

$E(X)$	$\mu = \frac{a+b}{2}$
$Var(x)$	$\sigma^2 = \frac{(b-a)^2}{12}$

Exemplo 6.1.1 *Represente graficamente a função de distribuição e a função densidade de probabilidade (f.d.p.) de uma variável* $U(0,6)$.

Exemplo 6.1.2 *O verdadeiro conteúdo de pacotes de leite de certa marca é uma v.a. com distribuição uniforme entre 0.85 l e 1.05 l.*

1. *Indique a f.d.p. correspondente e a função de distribuição.*
2. *Qual a probabilidade de um pacote de leite apresentar um volume inferior a 1 litro?*
3. *Calcule a probabilidade de um pacote escolhido arbitrariamente apresentar um volume entre 0,75 l e 0,95 l.*

6.2 Distribuição exponencial, Ex(λ).

A distribuição de exponencial com parâmetro $\lambda > 0$ é definida pela f.d.p.

$$\boxed{f(x) = \begin{cases} \lambda e^{-\lambda \cdot x} & , \quad \text{se } x > 0 \\ 0 & , \quad \text{se } x \leqslant 0 \end{cases}}$$

e com correspondente f.d.

$$F(x) = \begin{cases} 0 & , \quad \text{se } x < 0 \\ 1 - e^{-\lambda \cdot x} & , \quad \text{se } x \geqslant 0 \end{cases}$$

Tem-se ainda

$$P(a < X < b) = F(b) - F(a) = e^{-\lambda a} - e^{-\lambda b}.$$

$E(X)$	$\mu = \frac{1}{\lambda}$
$Var(x)$	$\sigma^2 = \frac{1}{\lambda^2}$

Existe uma relação estreita entre a *distribuição exponencial* e a distribuição de Poisson.

Admitindo que λ é o parâmetro da distribuição de Poisson que descreve o número de ocorrências por unidade de tempo (ou de espaço), a variável tempo (distância) entre ocorrências sucessivas segue uma distribuição exponencial Ex(λ).

A duração de vida (ou tempo decorrido até à falha) de certos aparelhos ou componentes, os tempos decorridos entre chamadas telefónicas numa central e entre a emissão de partículas radioactivas constituem exemplos de fenómenos reais bem descritos por esta distribuição.

Exemplo 6.2.1 *A chegada de clientes a uma lavandaria segue uma distribuição de Poisson em que o ritmo médio de afluência é de 20 clientes por hora. Após abrir a loja, qual a probabilidade de ter de esperar mais de 3 minutos pelo primeiro cliente?*

Exemplo 6.2.2 *Admita que o tempo de vida útil de um dado componente electrónico é uma v.a. com distribuição exponencial de média igual a 600 horas.*

1. *Calcule a probabilidade de o componente durar mais de 700 horas.*
2. *Sabendo que o componente já durou 400 horas, qual a probabilidade de durar ainda mais 700 horas?*

6.3 Distribuição normal, $N(\mu, \sigma^2)$.

A *distribuição normal* é uma das mais importantes das distribuições contínuas.

Por um lado é aquela que frequentemente se utiliza para descrever fenómenos físicos, processos físicos ou características humanas (peso, altura, etc).

Por outro existem variáveis aleatórias que não seguem a distribuição mas aproximam-se muito desta distribuição.

A correspondente f.d.p. apresenta o seguinte gráfico:

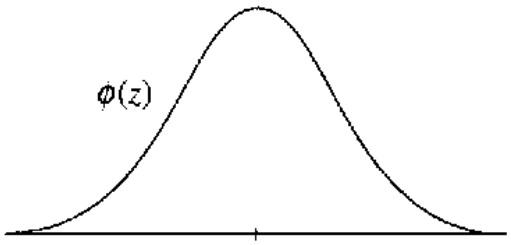

Figura 10: Função densidade de probabilidade normal.

Qualquer distribuição normal é definida por duas medidas:

- a média μ (localiza o centro da distribuição);
- o desvio padrão (mede a variabilidade de X em torno da sua média);

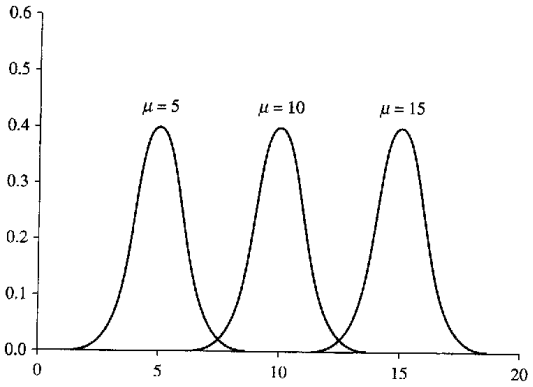

Figura 11: Funções densidade da distribuição normal com a mesma variância ($\sigma^2 = 1$) e diferentes médias.

Cálculo de probabilidades na distribuição normal

Utiliza-se a distribuição normal reduzida ou normal standard $Z \sim N(0, 1)$. Neste caso

- $E(Z) = 0$
- $Var(Z) = 1$

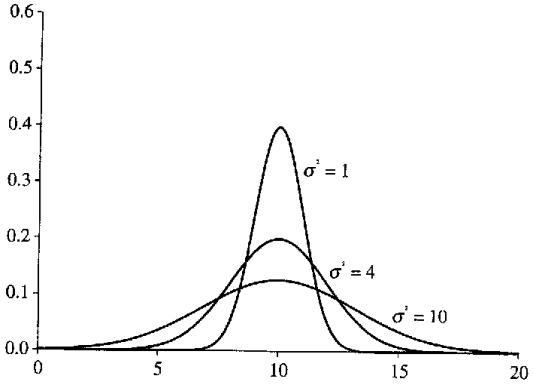

Figura 12: Funções densidade da distribuição normal com a mesma média ($\mu = 10$) e diferentes variâncias.

- Função distribuição

$$\boxed{P(Z \leqslant x) = \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt}$$

Esta função encontra-se tabelada. Complete os espaços em branco:

- $P(Z \leqslant 0) = \Phi(0) =$
- $P(Z \leqslant 1.15) = \Phi(1.15) =$

De notar que

$$\Phi(-x) = 1 - \Phi(x),$$

pois $P(Z \leqslant -x) = P(Z \geqslant x)$.

Diz-se que uma v.a. tem distribuição normal com parâmetros μ e σ^2 ou, abreviadamente, $N(\mu, \sigma^2)$ se

$$\frac{X - \mu}{\sigma} \sim N(0, 1)$$

Exemplo 6.3.1 *Utilizando a tabela da distribuição normal padrão, calcule as seguintes probabilidades:*

1. $P(0 < Z < 1.32)$*;*
2. $P(Z > -0.87)$*;*
3. $P(-1.96 < Z < 1.96)$*;*
4. $P(24 < X < 30)$*, sendo* X *uma v.a. com distribuição* $N(28, \sigma^2 = 4)$*.*

No que se segue, apresentam-se resultados importantes sobre a distribuição normal.

Teorema 6.3.2 *Sejam* n *variáveis independentes* X_i*,* $i = 1, \cdots, n$*. Se*

$$X_i \sim N(\mu_i, \sigma_i^2),$$

então

$$T = \sum_{i=1}^{n} X_i \sim N(\mu, \sigma^2)$$

em que

$$\mu = \mu_1 + \mu_2 + \cdots + \mu_n \text{ e } \sigma^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2.$$

Corolário 6.3.3 *Se $X_i (i = 1, 2, \cdots, n)$ são variáveis aleatórias independentes em que $X_i \sim N(\mu, \sigma^2)$, então*

$$S_n = \sum_{i=1}^{n} X_i \sim N(n \times \mu, n \times \sigma^2).$$

Corolário 6.3.4 *Se $X_i (i = 1, 2, \cdots, n)$ são variáveis aleatórias independentes em que $X_i \sim N(\mu, \sigma^2)$, então*

$$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$$

onde $\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$.

Corolário 6.3.5 *Se $X_1 \sim N(\mu_1, \sigma_1^2)$ e $X_2 \sim N(\mu_2, \sigma_2^2)$ são variáveis aleatórias independentes, então*

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$
$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2).$$

Exemplo 6.3.6 *Considere que o peso de um homem é uma variável aleatória com distribuição $N(75, 5)$. Calcule a probabilidade do peso de 4 homens (com pesos independentes) exceder 320 Kg.*

Exemplo 6.3.7 *A resistência à compressão de amostras de cimento de um certo tipo pode ser modelada por uma distribuição normal com valor médio 6000 Kg/cm^2 e desvio padrão 100 Kg/cm^2.*

1. *Calcule a probabilidade de que uma amostra de cimento tenha resistência superior a 6150 Kg/cm^2.*
2. *Qual a probabilidade de que uma amostra tenha resistência entre 5900 Kg/cm^2 e 5950 Kg/cm^2.*
3. *Calcule a resistência R de modo que a probabilidade de exceder R seja 90%.*

Exemplo 6.3.8 *O serviço de expedição e entrega de uma certa unidade fabril verificou que os volumes de vendas (em m^3) entregues semanalmente aos clientes eram essencialmente de dois tipos:*

- *Tipo A: com distribuição* $N(\mu = 50, \sigma^2 = 100)$.
- *Tipo B: com distribuição* $N(\mu = 15, \sigma^2 = 25)$.

O volume de entregas semanais é de 20 encomendas do tipo A e 100 encomendas do tipo B. O responsável do serviço de expedição e entrega negociou com uma empresa transportadora o transporte máximo de 3000 m^3 *semanais.*

Comente tal decisão.

Relação entre a distribuição Normal e a distribuição binomial

Se $X \sim B(n, p)$, tal que

- n é grande (na prática $n \geqslant 20$)
- p tome um valor suficientemente grande para que a distribuição seja simétrica (na prática $np \geqslant 7$), então

$$X \overset{a}{\sim} N(\mu = np, \sigma^2 = np(1-p)),$$

ou seja

$$\frac{X - np}{\sqrt{np(1-p)}} \overset{a}{\sim} N(0, 1).$$

As figuras seguintes mostram a aproximação da **normal à binomial** para $n = 10, 25, 100$, considerando-se sempre $p = 0.25$.

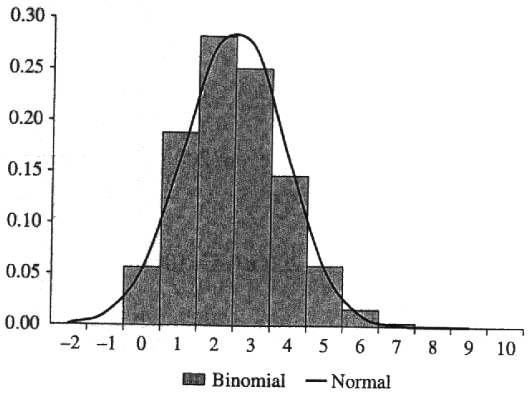

Figura 13: Aproximação da normal à binomial: $n = 10$, $p = 0.25$.

A distribuição normal como aproximação da distribuição de Poisson

Se $X \sim Poisson(\lambda)$, com λ grande (em termos práticos quando $\lambda > 20$), então

$$X \overset{a}{\sim} N(\mu = \lambda, \sigma^2 = \lambda),$$

ou seja

$$\frac{X - \lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0, 1).$$

De referir que quando se utiliza a distribuição normal (distribuição contínua) como aproximação de uma v.a. discreta é necessário fazer a chamada ***correcção de continuidade***.

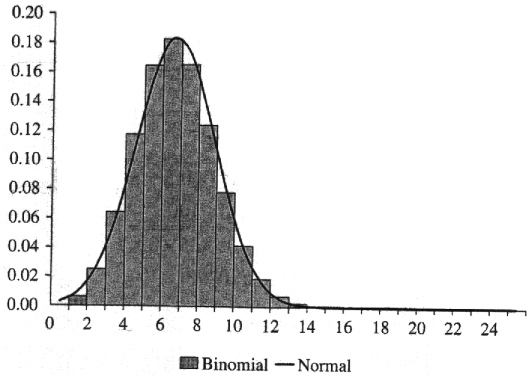

Figura 14: Aproximação da normal à binomial: $n = 25$, $p = 0.25$.

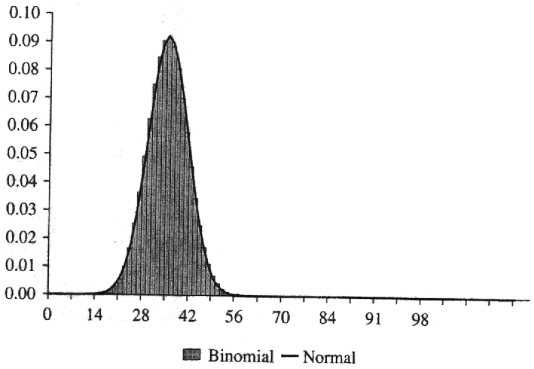

Figura 15: Aproximação da normal à binomial: $n = 100$, $p = 0.25$.

Exemplo 6.3.9 *Sendo* $X \sim B(200; 0.5)$*, calcule*

$$P(95 \leqslant X \leqslant 105).$$

O valor exacto é 0.56325. Tem-se

$$X \overset{a}{\sim} N(\mu = 100, \sigma^2 = 50)$$

Sem correcção de continuidade...

$$\begin{aligned} P(95 \leqslant X \leqslant 105) &\approx P(\tfrac{95-100}{\sqrt{50}} \leqslant \tfrac{X-100}{\sqrt{50}} \leqslant \tfrac{105-100}{\sqrt{50}}) \\ &= \Phi(0.7071) - \Phi(-0.7071) \\ &= 0.520498 \end{aligned}$$

Com correcção de continuidade...

$$\begin{aligned} P(95 \leqslant X \leqslant 105) &\approx P(95 - \tfrac{1}{2} \leqslant X \leqslant 105 + \tfrac{1}{2}) \\ &= P(94.5 \leqslant X \leqslant 105.5) \\ &= \Phi(0.7778) - \Phi(-0.7778) \\ &= 0.563316 \end{aligned}$$

Exemplo 6.3.10 *O número de avarias diárias de uma máquina de tecelagem é uma v.a. com distribuição* $Poisson(\lambda = 0.2)$.

Calcule a probabilidade de a referida máquina ter durante um ano (365 dias) exactamente 75 avarias.

7 Distribuições discretas e contínuas multivariadas.

7.1 Estudo de pares aleatórios discretos (X, Y).

$$p_{ij} = P(X = x_i, Y = y_j)$$

- $p_{ij} \geqslant 0$
- $\sum_i \sum_j p_{ij} = 1$

No caso de ser finito o contradomínio do par (X, Y) é usual representar a função de probabilidade numa tabela da seguinte forma:

X^Y	y_1	y_2	$\cdots$	y_m	
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}	$p_{2\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n2}	$\cdots$	p_{nm}	$p_{n\cdot}$
	$p_{\cdot 1}$	$p_{\cdot 2}$	$\cdots$	$p_{\cdot m}$	1

Tem-se

$$p_{i\cdot} = \sum_{j=1}^{m} p_{ij} = \sum_{j=1}^{m} P(X = x_i, Y = y_j) = P(X = x_i)$$

$$p_{\cdot j} = \sum_{i=1}^{n} p_{ij} = \sum_{i=1}^{n} P(X = x_i, Y = y_j) = P(Y = y_j)$$

Estas são as **f.d.p. marginais de X e de Y**, respectivamente.
$\rightsquigarrow$ **f.d.p. condicional** de uma das v.a.'s em relação à outra.

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{\cdot j}}$$

$$P(Y = y_j | X = x_i) = \frac{P(X = x_i \cap Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i\cdot}}$$

O conceito de independência pode ser alargado para as v.a.'s. Diz-se que as v.a.'s X e Y são **independentes** se para quaisquer subconjuntos A e B de $\mathbb{R}$ tivermos

$$P(x \in A, y \in B) = P(x \in A)P(y \in B)$$

o que é equivalente a exigir, para todo i e j

$$p_{ij} = p_{i\cdot}p_{\cdot j}$$

Exemplo 7.1.1 *Em três lançamentos de uma moeda equilibrada seja N a v.a. que representa o número de caras e M a que representa o número de mudanças de face.*

1. *Calcule a distribuição conjunta de probabilidade.*
2. *Defina a função de probabilidade marginal de M.*
3. *Determine a função de probabilidade condicional de N dado $M = 1$.*

7.2 Distribuições contínuas.

$f : \mathbb{R}^2 \to \mathbb{R}$ é uma f.d.p. se

1. $f(x, y) \geqslant 0$
2. $\displaystyle\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(u, v)dudv = 1$

Para um subconjunto $A \subseteq \mathbb{R}^2$

$$P(x \in A) = \int\int_A f(u, v)dudv$$

$\rightsquigarrow$ f.d.p. marginal de duas variáveis contínuas X e Y

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)dy$$

$$f_X(y) = \int_{-\infty}^{+\infty} f(x, y)dx$$

$\rightsquigarrow$ Variáveis aleatórias independentes
As v.a.'s contínuas X e Y são independentes se e só se, para todo $x, y \in \mathbb{R}$,

$$f(x, y) = f_X(x)f_Y(y).$$

7.3 Covariância e correlação.

Duas variáveis são **independentes** se não existe qualquer tipo de relação entre elas. Num exemplo apresentado no cap. 2 verificou-se que as v.a.'s tempo de estudo e notas no teste de Biologia não eram independentes. À medida que aumenta o estudo, a nota no teste tenderá também a aumentar. Existem várias formas de relacionamento: relação linear, quadrática, exponencial, etc.

A **Covariância** permite medir a variação conjunta das duas variáveis.
Caso discreto:

$$\begin{aligned} cov(X, Y) &= E((X - \mu_X)(Y - \mu_Y)) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m}(x_i - \mu_X)(y_j - \mu_Y)p_{ij}. \end{aligned}$$

Caso contínuo:

$$cov(X, Y) = \int\int (x - \mu_X)(y - \mu_Y)f(x, y)dxdy.$$

Nota: Se X e Y são v.a.'s independentes, então $cov(X, Y) = 0$. No entanto a recíproca não é verdadeira.

Exemplo 7.3.1 *Seja (X, Y) um par aleatório com f.d.p. conjunta*

X^Y	*0*	*1*	
0	$\frac{1}{3}$	*0*	$\frac{1}{3}$
1	*0*	$\frac{1}{3}$	$\frac{1}{3}$
2	$\frac{1}{3}$	*0*	$\frac{1}{3}$
	$\frac{2}{3}$	$\frac{1}{3}$	*1*

As variáveis X e Y não são independentes mas $cov(X, Y) = 0$. Assim, a independência implica não correlação, mas a inversa não é verdadeira.

$\rightsquigarrow$ **Correlação**

O coeficiente de correlação entre as variáveis aleatórias X e Y é dado por:

$$\boxed{\rho = \frac{cov_(x, y)}{\sqrt{Var(X) \cdot Var(Y)}} = \frac{cov_(x, y)}{\sigma_X \sigma_Y}}$$

Verifica-se que:

- $|\rho| \leqslant 1$
- $\rho = 0$ Não existe uma relação linear entre X e Y; podem existir outras relações.
- $\rho = 1$ Relacionamento linear perfeito ($Y = \alpha + \beta X$) e $\beta > 0$.
- $\rho = -1$ Relacionamento linear perfeito ($Y = \alpha + \beta X$) e $\beta < 0$

Covariância Positiva - Se uma das variáveis se desvia significativamente do seu valor esperado a outra tenderá a desviar-se no mesmo sentido o que levará a um aumento da dispersão da soma das duas variàveis.

Covariância negativa - Se uma das variáveis se desvia significativamente do seu valor esperado a outra tenderá a desviar-se no sentido contrário conduzindo a um diminuição da dispersão da sua soma.

Sendo X e Y variáveis aleatórias, tem-se

$$\boxed{Var(X \pm Y) = Var(X) \pm 2Cov(X, Y) + Var(Y)}.$$

Se as variáveis X e Y forem independentes tem-se $Cov(X, Y) = 0$ e, então, a fórmula anterior reduz-se a

$$\boxed{Var(X \pm Y) = Var(X) + Var(Y)}.$$

8 Amostragem. Distribuições amostrais.

8.1 População e amostra. Inferência estatística.

Inferência estatística: processo de generalização de conclusões obtidas numa amostra para a população.

Exemplo 8.1.1 *Considere-se o número de pessoas inscritas no ano lectivo 2004/2005, entre os quais há uma proporção θ (desconhecido) que pratica desporto. O observador escolhendo ao acaso uma amostra, utiliza o resultado nela obtido para tirar conclusões sobre a proporção de praticantes no IPB. Trata-se de um problema de inferência estatística.*

Neste exemplo, o observador pode ter interesse em

1. *estimar ou "adivinhar", com base na amostra, a proporção de praticantes de desporto no IPB (***estimação pontual***);*
2. *construir, a partir da observação da amostra, um intervalo que, com razoável confiança, contenha o parâmetro θ (***estimação intervalar***);*
3. *proposta a hipótese de que a proporção de praticantes de desporto é inferior a 20%, averiguar se os dados obtidos na amostra suportam essa proporção (***testes de hipóteses***).*

8.2 Amostragem aleatória.

Para que o processo de inferência estatística seja válido, é necessário que as amostras estudadas sejam representativas da população a partir da qual foram obtidas.

Existem, essencialmente, dois tipos de amostragem:

1. Amostragem probabilística ou aleatória;
2. Amostragem não probabilística ou não aleatória.

Na amostragem aleatória as amostras são obtidas de forma aleatória, ou seja, a probabilidade de cada elemento da população participar no processo de amostragem é igual para todos os elementos.

A amostragem aleatória é preferida pois é aquela que evita o enviesamento na selecção. Tal enviesamento existirá sempre que ocorrer uma tendência sistemática para sobrerepresentar ou subrepresentar uma parte da população.

De referir que se designa por *amostra aleatória de dimensão n* ao vector aleatório $(X_1, X_2, \cdots, X_n)$ e que se designa por *amostra aleatória observada* ao correspondente valor observado $(x_1, x_2, \cdots, x_n)$ daquele vector aleatório.

Considere-se uma população formada por N elementos (ex.: alunos da ESTGM) e uma variável aleatória X representando uma característica da população a estudar (ex.: altura).

- x_1- altura do 1º indivíduo seleccionado- valor observado da variável aleatória X_1;
- x_2- altura do 2ºindivíduo seleccionado- valor observado da variável aleatória X_2; $\vdots$
- x_n- altura do n-ésimo indivíduo seleccionado- valor observado da variável aleatória X_n.

Como obter neste caso uma amostra aleatória?

1. Numerar de 1 até N todos os elementos da população;
2. Colocar bolas numeradas de 1 até N numa máquina do tipo "totoloto";
3. Retirar uma bola, seleccionar o elemento da população com o número correspondente e registar a sua altura;
4. Repor a bola na máquina;
5. Repetir os passos 3. e 4. até terem sido registados, no total, n elementos.

Numa amostra aleatória as variáveis $X_1, X_2, \cdots, X_n$ são identicamente distribuídas (com a mesma função de probabilidade) e independentes(dado que a amostragem é feita com reposição).

8.3 Distribuições amostrais.

Parâmetros populacionais $\neq$ Estatísticas amostrais

Dada uma v.a. (variável aleatória) definida sobre uma determinada população, os parâmetros da distribuição correspondente são fixos e as estatísticas amostrais variam de amostra para amostra.

Uma estatística amostral calculada a partir de $X_1, X_2, \cdots, X_n$ é uma função de várias variáveis aleatórias e é, portanto, ela própria uma variável aleatória.

Exemplo 8.3.1 *Considere a amostra aleatória da pressão sanguínea de 10 homens com um determinado problema circulatório (a amostra foi obtida escolhendo ao acaso 10 homens entre todos os doentes com tal problema de saúde)*

$151,7; 123,2; 137,2; 121,9; 125,5; 145,1; 161,8; 110,9; 138,6; 110,4.$

Um exemplo de uma estatística é

$$S = \frac{1}{2}(X_1 + X_{10})$$

com correspondente valor observado

$$s = \frac{1}{2}(151,7 + 110,4) = 131.25.$$

Outros exemplos de estatísticas são $T = \frac{1}{10}(X_1 + \cdots + X_{10})$ *(média amostral) com correspondente valor observado* $t = 132,01$*,* $U = \max(X_1, \cdots, X_{10})$ *com valor observado* $u = 161,8$*, etc.*

8.4 Teorema do limite central.

A importância da distribuição normal advém do seguinte resultado fundamental sobre a soma de variáveis aleatórias independentes e identicamente distribuídas (v.a.'s i.i.d.'s).

Teorema do Limite Central: *Sejam X_1, X_2, $\cdots$, X_n v.a.'s i.i.d.'s com valor médio μ e desvio padrão σ e seja $S_n = X_1 + X_2 + \cdots + X_n$. Então, quando $n \to +\infty$,*

$$\boxed{S_n \overset{a}{\sim} N(n\mu, n\sigma^2)},$$

ou seja,

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \overset{a}{\sim} N(0,1).$$

A distribuição normal aproxima qualquer soma de *um grande número de v.a.'s i.i.d.'s* com variância finita. Assim, se n for grande, a partir deste teorema pode-se calcular valores aproximados para probabilidades do tipo $P(a \leqslant S_n \leqslant b)$ da seguinte forma:

$$\begin{aligned} P(a \leqslant S_n \leqslant b) &= P(\tfrac{a-n\mu}{\sigma\sqrt{n}} \leqslant \tfrac{S_n - n\mu}{\sigma\sqrt{n}} \leqslant \tfrac{b-n\mu}{\sigma\sqrt{n}}) \\ &\simeq \Phi(\tfrac{b-n\mu}{\sigma\sqrt{n}}) - \Phi(\tfrac{a-n\mu}{\sigma\sqrt{n}}) \end{aligned}$$

O valor de n a partir do qual se obtém uma boa aproximação depende da função distribuição das v.a.'s X_i. Em geral, se n for superior a 30 e tais v.a.'s não forem muito assimétricas, a aproximação é boa.

Exemplo 8.4.1 *Considerem-se as v.a.'s i.i.d.'s $X_1, X_2, \cdots, X_{30}$, com distribuição uniforme no intervalo $(0, 10)$. Determine*

$$P(S_{30} < 170).$$

Exemplo 8.4.2 *Num teste de escolha múltipla há 75 perguntas, todas com 4 respostas à escolha, em que apenas uma está correcta. Cada resposta certa vale um ponto e cada resposta errada desconta $\frac{1}{3}$ de ponto. Se um aluno responder ao acaso a todas as perguntas, qual a probabilidade de obter mais de 15 pontos no teste?*

Corolário: *Dada uma sucessão de v.a.'s i.i.d.'s $X_1, X_2, \cdots, X_n$ com média μ e desvio padrão σ, então*

$$\overline{X} \overset{a}{\sim} N(n\mu, \frac{\sigma^2}{n}),$$

ou seja,

$$\frac{\overline{X} - \mu}{\sqrt{\frac{\sigma^2}{n}}} \overset{a}{\sim} N(0,1),$$

onde $\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$.

De notar que $E(\overline{X}) = \mu$ e $var(\overline{X}) = \frac{\sigma^2}{n}$.

Exemplo 8.4.3 *Considere as v.a.'s i.i.d.'s $X_1, X_2, \cdots, X_{45}$ com distribuição exponencial de média 5. Determine $P(\overline{X} < 5.5)$.*

Exemplo 8.4.4 *Admita que a função distribuição dos comprimentos das peças produzidas por determinada empresa, tem valor esperado e variância respectivamente iguais a 1,5 cm e 0,75. Nestas situações, imagine que era contratado para verificar se 300 dessas peças eram suficientes para forrar uma tábua com largura igual à da peça e comprimento igual a 455. Qual seria a sua decisão? Justifique a sua resposta.*

9 Amostragem. Estimação pontual.

As estatísticas que vão ser utilizadas com o propósito de estimar o parâmetro θ designam-se por T, T', T''.

Exemplo 9.0.5 *Admita que o número de "gralhas" por cada página escrita por certo dactilógrafo é uma v.a. (variável aleatória) que segue uma distribuição de Poisson. Normalmente desconhece-se o valor do parâmetro $\lambda = E(X) = var(X)$. Recolhida uma amostra aleatória, intuitivamente, utilizamos a média da amostra observada ou a variância amostral como estimativa de λ.*

No entanto podemos estar apenas interessados em estimar a probabilidade de uma página escolhida ao acaso não ter gralhas $P(X = 0) = e^{-\lambda}$ ou a probabilidade de nenhuma das páginas que compõem a amostra ter gralhas

$$P(\sum_{i=1}^{n} X_i = 0) = e^{-n\lambda}.$$

É importante distinguir dois conceitos:

Estimador: é uma variável aleatória.

Estimativa: é um valor concreto assumido pelo estimador para uma amostra particular.

9.1 Propriedades dos estimadores por pontos.

Estimador centrado: Um estimador T para o parâmetro θ diz-se *centrado* ou *não enviesado* quando

$$E(T) = \theta.$$

Se $E(T) \neq \theta$, então o estimador diz-se enviesado e a diferença

$$Env(T) = E(T) - \theta$$

mede o **enviesamento** de T.

Exemplo 9.1.1 *Para uma amostra aleatória $X_1, X_2, \cdots, X_n$ proveniente da distribuição $N(\mu, 1)$, com μ desconhecido, um exemplo de um estimador centrado é $\overline{X}$.*

Recorde que $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$. Como $E(S^2) = \frac{n-1}{n}\sigma^2$, para valores pequenos de n é preferível usar como estimador de σ^2, a estatística $S'^2 = \frac{n}{n-1}S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

Eficiência: Sejam T e T' dois estimadores centrados para θ. O estimador T é mais eficiente do que T' quando

$$Var(T) \leqslant Var(T').$$

O estimador T é o mais eficiente quando a relação anterior se verifica qualquer que seja outro estimador T' centrado para θ.

Exemplo 9.1.2 *Pretende-se encontrar um estimador para a média através de uma amostra aleatória com dimensão 2. Considere os estimadores seguintes:*

$$T = \frac{1}{2}X_1 + \frac{1}{2}X_2$$

$$T' = \frac{1}{3}X_1 + \frac{2}{3}X_2$$

1. *Os estimadores T e T' são centrados?*
2. *Qual o estimador mais eficiente?*

consistência: Um estimador T_n diz-se consistente quando, para qualquer número real $\epsilon > 0$,

$$\lim_{n\to+\infty} P(\theta - \epsilon < T_n < \theta + \epsilon) = 1$$

Quando se procura um estimador consistente pretende-se que a precisão do estimador aumente quando cresça a dimensão da amostra (no limite a precisão é total). Na verdade, um estimador que não seja consistente é um estimador que não deve ser utilizado.

Método de Máxima Verosimilhança

Dada uma amostra aleatória proveniente de uma função distribuição $F(x;\theta)$, como encontrar um estimador de θ? Vamos estudar um método de estimação pontual: **Método da Máxima Verosimilhança**

Dada uma amostra aleatória $x_1, x_2, \cdots, x_n$ proveniente de uma f.d. $F(x;\theta)$, entende-se por **verosimilhança da amostra** a função de probabilidade

$$L(x_1, x_2, \cdots, x_n) = f(x_1|\theta)f(x_2)|\theta) \cdots f(x_n|\theta),$$

no caso da v.a. X ser contínua com f.d.p. $f(x|\theta)$, ou alternativamente no caso discreto

$$L(x_1, x_2, \cdots, x_n) = P(X = x_1)P(X = x_2) \cdots P(X = x_n).$$

A estimativa de **máxima verosimilhança** de θ é o valor de θ que maximiza L como função de θ. De acordo com este método, depois de observada a amostra deve escolher-se para estimativa do parâmetro o valor que torna a amostra o mais verosímel possível, aquele que maximiza a probabilidade da amostra recolhida.

Exemplo 9.1.3 *Um saco contém bolas pretas e bolas brancas, na proporção de 3 para 1, desconhecendo-se qual a cor predominante. A probabilidade de tirar ao acaso uma bola preta é igual a 3/4 ou 1/4 consoante for preta ou branca a cor predominante. Retirando do saco, com reposição, três bolas a v.a. X, que designa o número de bolas pretas presentes na amostra, tem distribuição binomial $B(3, 1/4)$ ou $B(3, 3/4)$. Trata-se de escolher uma das duas possibilidades. Sabendo que se observou X=1, encontre a estimativa de máxima verosimilhança para o parâmetro.*

O estimador é encontrado derivando L em ordem a θ e igualando a zero. Na prática trabalha-se com $\ln(L)$ em vez de L.

É possível provar que os estimadores de m.v. são consistentes, tendem a ser centrados e normais à medida que a dimensão da amostra cresce e não existem outros estimadores nestas condições que tenham menor variância que estes.

Exemplo 9.1.4 *Admita que o número de pessoas que visita diariamente um museu segue uma distribuição de Poisson com média λ desconhecida. Nos dois últimos dias visitaram o museu* 5 *e* 9 *pessoas. Encontre o estimador de m.v. de λ.*

10 Estimação de intervalos.

10.1 Conceito de intervalo de confiança.

Os métodos de estimação por pontos propõem apenas uma estimativa isolada $\widehat{\theta}$ para θ. Assim existem situações em que é preferível a estimação por intervalos. Na última a estimação faz-se acompanhar de um certo intervalo (t_1, t_2), significando que o verdadeiro valor do parâmetro está muito provavelmente entre t_1 e t_2, atribuindo-se a esse intervalo um certo grau ou coeficiente de confiança. Em muitos casos, o intervalo é da forma $(\widehat{\theta} - \epsilon, \widehat{\theta} + \epsilon)$, em que ϵ pode ser considerado uma medida de erro inerente à estimativa $\widehat{\theta}$.

Exemplo 10.1.1 *O quadro seguinte apresenta uma amostra aleatória de pesos em mg de cigarros SG Filtro com* $n = 20$.

807.71	*814.31*	*861.03*	*838.27*
790.68	*823.45*	*814.30*	*842.45*
818.82	*828.94*	*82.47*	*816.56*
853.42	*8.8.86*	*785.84*	*850.44*
858.59	*797.88*	*833.19*	*838.64*

Admita-se que a variável aleatória "peso dos cigarros SG Filtro" tem distribuição normal com parâmetros μ *(desconhecido) e* $\sigma = 22.5mg$.

Os métodos apresentados no capítulo anterior permitem-nos propor um valor para o peso médio dos cigarros: a média da amostra.

$$\widehat{\mu} = \overline{x} = 825.27mg.$$

Mas não temos qualquer informação sobre o rigor e a confiança da estimativa anterior. O próximo objectivo é determinar um intervalo de confiança a 95% para μ.

A média amostral, $\overline{X}$, *segue uma distribuição normal*

$$\overline{X} \sim N(\mu, \frac{22.5^2}{20}),$$

ou seja,

$$\frac{\overline{X} - \mu}{\frac{22.5}{\sqrt{20}}} \sim N(0, 1).$$

Como a variável aleatória Z *tem distribuição conhecida, independente do parâmetro a estimar, podemos calcular a seguinte probabilidade:*

$$P(-1.96 < Z < 1.96) = \Phi(1.96) - \Phi(-1.96) = 0.95.$$

Assim, tem-se

$$\begin{aligned} & P\left(-1.96 < \frac{\overline{X}-\mu}{\frac{22.5}{\sqrt{20}}} < 1.96\right) && = && 0.95 \\ \Leftrightarrow \quad & P\left(-1.96 \cdot \frac{22.5}{\sqrt{20}} < \overline{X} - \mu < 1.96 \cdot \frac{22.5}{\sqrt{20}}\right) && = && 0.95 \\ \Leftrightarrow \quad & P(\overline{X} - 9.86 < \mu < \overline{X} + 9.86) && = && 0.95 \end{aligned}$$

Deste modo podemos afirmar que o intervalo aleatório

$$\left]\overline{X} - 9.86, \overline{X} + 9.86\right[$$

contém o valor de μ(desconhecido) com probabilidade 95%..

Diz-se que o particular intervalo $]\overline{x} - 9.86, \overline{x} + 9.86[$ é um intervalo de confiança a 95% *para o valor do peso médio dos cigarros.*

Com a amostra considerada neste exemplo, o intervalo de confiança a 95% μ é $]815.41, 835.13[$.

De notar que nesta situação se admitiu que σ^2(variância da população) era conhecida. Normalmente é apenas conhecida uma estimativa.

O valor de μ é fixo (trata-se de um parâmetro da população a estimar), no entanto o intervalo especificado varia de amostra para amostra (por depender de $\overline{X}$).

10.2 Simetria *versus* assimetria do intervalo de confiança relativamente ao valor da estimativa pontual.

O intervalo de confiança a $100(1-\alpha)\%$ de μ (valor esperado)

$$\overline{x} \pm z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$$

é um intervalo simétrico em relação a $\overline{x}$.

Para quaisquer valores α_1, α_2 que satisfaçam condição $\alpha = \alpha_1 + \alpha_2$, os intervalos de confiança

$$\left]\overline{x} - z_{\alpha_1} \cdot \frac{\sigma}{\sqrt{n}}, \overline{x} + z_{\alpha_2} \cdot \frac{\sigma}{\sqrt{n}}\right[$$

são intervalos de confiança a $100(1-\alpha)\%$. De notar que o intervalo simétrico em relação à estatística é normalmente o de menor amplitude e consequentemente o mais interessante.

Exemplo 10.2.1 *Considerando novamente a amostra de dimensão $n = 20$ com $\overline{x} = 825.27$, defina o intervalo de confiança a 95% para as três situações seguintes:*

1. $\alpha_1 = \alpha_2 = 0.025$*;*
2. $\alpha_1 = 0.01$ *e* $\alpha_2 = 0.04$*;*
3. $\alpha_1 = 0.005$ *e* $\alpha_2 = 0.045$.

10.3 Especificação do intervalo.

A especificação de um intervalo de confiança implica o conhecimento de:

- um estimador do parâmetro em causa;
- a sua distribuição amostral;

- um valor particular do estimador, ou seja, uma estimativa pontual.

Exemplo 10.3.1 *Admita que o tempo de vida (em horas) de componentes electrónicas produzidas segundo determinado processo de fabrico tem distribuição normal. Retirou-se uma amostra aleatória de $n = 20$ componentes da produção de determinado dia, verificando-se que $\overline{x} = 1832$ e $s' = 497$. Calcule um intervalo de confiança a*

$$a)95\% \quad b)99\% \quad c)99.9\%$$

para μ (tempo médio de vida das componentes electrónicas).

Exemplo 10.3.2 *Relativamente ao exemplo anterior, determine um intervalo de confiança a 90% para a variância σ^2 do tempo de vida das componentes electrónicas.*

Exemplo 10.3.3 *Para testar, em condições naturais de cultura, dois novos tipos de determinado cereal, uma exploração seleccionou aleatoriamente 8 zonas de terreno mais ou menos homogéneo, e em cada uma delas semeou ambos os tipos em talhões separados. Os resultados da produção (alqueire por hectare) foram os seguintes:*

Amostra	$\overline{x}$	s'
Cereal A	*81,63*	*232,41*
Cereal B	*75,88*	*102,25*

Admitindo que as produções (alqueires por hectare) têm distribuição normal e que não há razões para admitir que $\sigma_A \neq \sigma_B$, construa um intervalo de confiança a 95% para a diferença $\mu_A - \mu_B$.

Será que se pode concluir que as produções médias das populações são diferentes?

Exemplo 10.3.4 *Uma empresa produz peças para motores a jacto. Uma das operações envolve o polimento da superfície de uma componente feita com uma liga de titânio. Dois métodos de polimento podem ser usados e a empresa pretende escolher o que conduz à rugosidade da superfície com menor variância. De modo a tomar uma decisão, obteve uma amostra aleatória de dimensão $n_1 = 10$ dos resultados conseguidos com o primeiro método e uma amostra aleatória de $n_2 = 16$ dos resultados conseguidos com o segundo método. As variâncias das amostras foram, respectivamente, $s_1'^2 = 26.01$, $s_2'^2 = 22.09$. Construa um intervalo de confiança para a razão das variâncias $\frac{\sigma_1^2}{\sigma_2^2}$, partindo do princípio de que a rugosidade obtida por qualquer dos métodos têm distribuição normal.*

Exemplo 10.3.5 *Num inquérito às Receitas e Despesas familiares, 1980-1981, foram inquiridos 8041 agregados familiares. Considere o seguinte quadro:*

Despesa anual	*n^o de agregados*	*n^o de agregados com automóvel*
4000-5000	*526*	*316*
5000-6000	*313*	*210*

Será que os dados referidos no quadro anterior permitem afirmar, com um grau de confiança de 90% que a proporção de agregados com automóvel é a mesma nas duas subpopulações?

10.4 Dimensionamento de amostras.

Qual o interesse em dimensionar convenientemente uma amostra?

- Recolher e tratar uma amostra grande demais para os resultados que se pretendem obter constitui um desperdício de recursos.
- É destituída de interesse uma amostra cuja dimensão não seja suficiente para se poderem tirar conclusões relevantes.

O dimensionamento correcto de uma amostra com base na qual se pretende estimar um ou mais parâmetros passa pela identificação da qual a fiabilidade desejada para as estimativa(s).

Exemplo 10.4.1 *População normal com desvio padrão (σ) conhecido*
O*bjectivo: Qual deverá ser a dimensão da amostra a recolher para que a amplitude do intervalo de confiança não exceda ω?*

Intervalo de confiança $\quad \left]\overline{x} - z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \overline{x} + z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}\right[$

Amplitude $\quad \omega = 2 \cdot z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}$

Dimensão da amostra $\quad n = 4\sigma^2 \left[\frac{z_{\frac{\alpha}{2}}}{\omega}\right]^2$

Exemplo 10.4.2 *Sendo X o tempo, em minutos, que certa tarefa leva a executar, admita que X tem distribuição normal. A partir da experiência anterior determinou-se, com boa aproximação, que $\sigma = 2$ minutos.*

Qual deverá ser a dimensão da amostra de tempos, de modo que a amplitude do intervalo de confiança a 95% de μ (tempo médio) seja inferior a 0.5 minutos?

Exemplo 10.4.3 *Dimensionamento de amostras para a proporção binomial*

Intervalo de confiança $\quad \left]\widehat{p} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\widehat{q}}{n}}, \widehat{p} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\widehat{q}}{n}}\right[$

Amplitude $\quad \omega = 2 \cdot z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{p}\widehat{q}}{n}}$

Dimensão da amostra $\quad n = 4\widehat{p}\widehat{q}\left[\frac{z_{\frac{\alpha}{2}}}{\omega}\right]^2$

A expressão anterior depende de $\widehat{p}$ e de $\widehat{q} = 1 - \widehat{p}$, que são geralmente desconhecidos antes da obtenção da amostra. Nestas situações utiliza-se o valor mais conservativo $\widehat{p} = \widehat{q} = \frac{1}{2}$. Assim, a dimensão da amostra será dada pela expressão

$$n = \left[\frac{z_{\frac{\alpha}{2}}}{\omega}\right]^2$$

Exemplo 10.4.4 *Pretende-se estimar a proporção de peças defeituosas à saída de uma linha de produção será de um lote de peças constituídas para o efeito. Qual deverá ser a dimensão do lote , de modo que a amplitude do intervalo de confiança a 95% da proporção de peças defeituosas não ultrapasse 0.02?*

10.5 Intervalos de confiança a $(1-\alpha)100\%$ para a diferença dos valores esperados, $\mu_1 - \mu_2$, quando as amostras não podem ser consideradas independentes (observações emparelhadas).

Define-se uma nova variável aleatória:

$$D = X_1 - X_2.$$

Os valores da variável D são considerados como constituindo uma amostra e o intervalo de confiança determina-se como na situação em que existe apenas uma amostra (intervalo de confiança para o valor esperado).

Exemplo 10.5.1 *A reitoria de uma universidade de uma escola decidiu publicar os resultados dos inquéritos anuais de avaliação pedagógica de todos os professores. Escolhidos aleatoriamente 10 docentes, recolheram-se as pontuações obtidas antes e depois da decisão de publicação:*

$$(3.4; 3.6); (4.2; 4.0); (3.2; 3.4); (2.5; 3.1); (4.2; 4.4);$$

$$(2.9; 3.1); (2.7; 2.5); (3.5; 3.6); (2.6; 2.8); (3.1; 3.4)$$

Admitindo que a classificação (numa escala de 0-5) tem distribuição normal, determine um intervalo de confiança a 95% para a diferença entre as classificações. Pode concluir-se que a decisão de publicar os resultados dos inquéritos melhorou as pontuações dos docentes?

11 Testes de hipóteses.

objectivo: decidir, atendendo à informação fornecida por uma amostra, sobre a aceitação ou não aceitação de determinada hipótese.

Exemplo 11.0.2 *Suponha que pretende estudar se determinada moeda é equilibrada, o que corresponde a testar $p = \frac{1}{2}$, onde p é a probabilidade, por exemplo, de saída de ■face■.*

O procedimento envolvido na metodologia utilizada nos testes de hipóteses pode dividir-se em quatro fases:

1. Definição das hipóteses;
2. Identificação da Estatística de Teste e caracterização da sua distribuição amostral;
3. Definição da regra de decisão, com especificação do nível de significância do teste;
4. Cálculo da estatística de teste e tomada de decisão.

[1] Definição das hipóteses

hipótese: conjectura acerca de uma ou mais populações.
Vamos estudar as hipóteses paramétricas (quando a conjectura diz respeito apenas ao parâmetro).
O teste de hipóteses engloba duas hipóteses:

- a hipótese nula, $\mathcal{H}_0$, hipótese a testar;
- a hipótese alternativa, $\mathcal{H}_1$, que corresponde ao conjunto das alternativas.

A hipótese $\mathcal{H}_0$ corresponde a algo que se pretende testar.
A aceitação ou rejeição de $\mathcal{H}_0$ far-se-à considerando uma amostra. Note-se que a aceitação (rejeição) de $\mathcal{H}_0$ implica a rejeição (aceitação) de $\mathcal{H}_1$.

Notas importantes

1. A **hipótese alternativa** nunca contém uma igualdade.
2. A hipótese nula é considerada verdadeira ao longo do procedimento de teste até ao momento em que haja evidência estatística apontando em sentido contrário.
3. A **hipótese nula** contém sempre uma igualdade. Mesmo quando na hipótese nula faz sentido figurar o sinal $\geqslant$ ou o sinal $\leqslant$, o teste é realizado considerando apenas a situação em que $\mathcal{H}_0$ mais se aproxima de $\mathcal{H}_1$, ou seja, admitindo que é verdadeira a afirmação de $\mathcal{H}_0$ que corresponde à igualdade.
4. Quando $\mathcal{H}_1$ tiver uma desigualdade ($>$ ou $<$) o teste diz-se unilateral (à direita para o sinal $>$ e à esquerda para o sinal $<$).

   Quando $\mathcal{H}_1$ contiver uma não igualdade ($\neq$) o teste diz-se bilateral.

[2] Identificação da Estatística de Teste e caracterização da sua distribuição amostral

A estatística de teste (ET) é utilizada para verificar a plausibilidade da hipótese nula. Para que tal estatística possa cumprir a sua função é necessário conhecer a sua distribuição amostral quando se admitir que é verdadeira a hipótese nula.

Ver distribuições da tabela dos intervalos de confiança.

[3]Definição da regra de decisão, com especificação

A decisão de rejeitar ou não a hipótese nula fundamenta-se no valor que a estatística de teste toma.

A distribuição da estatística de teste é especificada na suposição que $\mathcal{H}_0$ é verdadeira.

Para que esta decisão possa ser tomada de uma forma controlada, é conveniente que, à partida, se fixe o valor a partir do qual se considera improvável a validade da hipótese nula. Tal fixação corresponde à especificação de uma **região de rejeição**.

Teste de $\mathcal{H}_0 : \theta = \theta_0$ contra $\mathcal{H}_1$

$\mathcal{H}_1$	região de rejeição
$\theta > \theta_0$	$ET > ET(\alpha)$
$\theta < \theta_0$	$ET < ET(1-\alpha)$
$\theta \neq \theta_0$	$ET < ET(1-\frac{\alpha}{2})$ ou $ET > ET(\frac{\alpha}{2})$

em que, sendo X a distribuição da Estatística de teste, $P(X > ET(\alpha)) = \alpha$.

A probabilidade α de, no caso de $\mathcal{H}_0$ ser verdadeira, a ET pertencer à região de rejeição designa-se por **nível de significância** do teste.

O nível de significância representa a probabilidade de se cometer o erro de rejeitar $\mathcal{H}_0$ quando $\mathcal{H}_0$ é verdadeira - **erro do tipo I** ou de rejeição.

[4] Cálculo da estatística de teste e tomada da decisão

A última fase do teste de hipóteses corresponde ao cálculo da estatística de teste e, face ao valor obtido, decidir pela aceitação ou rejeição de $\mathcal{H}_0$.

- Se a estatística de teste pertencer à região de rejeição, então $\mathcal{H}_0$ é rejeitada.
- Se $\mathcal{H}_0$ não fosse rejeitada com base na amostra recolhida, isso não significaria que $\mathcal{H}_0$ fosse verdadeira. Nesse caso, a acção correcta poderia ser a de recolher mais informação.

11.1 Valor de prova.

O teste de hipóteses finaliza com a decisão de *rejeição* ou *não rejeição* da hipótese nula. No entanto, esta dicotomia é artificial dado que

- a fixação do nível de significância é arbitrária e

- os dados amostrais podem contradizer a hipótese nula em maior ou menor grau.

O **valor de prova** mede a evidência que os dados fornecem em favor de $\mathcal{H}_0$. A sua definição é a seguinte:

o valor de prova é a probabilidade da ET tomar um valor igual ou mais extremo do que aquele que, de facto, é observado.

Note-se que, tal como a ET, o valor de prova é calculado admitindo que $\mathcal{H}_0$ é verdadeira.

Quanto menor for o valor de prova, maior será o grau com que a hipótese nula é contradita.

Dada a importância da informação fornecida pelo valor de prova é recomendável a sua inclusão explícita nos resultados de quaisquer testes de hipóteses.

Ver exemplo 2

Quando o teste é bilateral é necessário ter presente que, nesta situação, o conjunto dos valores assumidos por aquela estatística que são tão ou mais desfavoráveis para a hipótese nula são os que se situam nas duas caudas da distribuição da estatística de teste.

Ver exemplo 3

11.2 Erro do Tipo II- potência de teste.

Todo o caminho que vai do particular(amostra) para o todo (população) pode conduzir a erros.

Num teste de hipóteses devem considerar-se dois tipos de erros:

	$\mathcal{H}_0$ verdadeira	$\mathcal{H}_0$ falsa
$\mathcal{H}_0$ rejeitada	Erro do tipo I (probabilidade α)	Decisão Correcta (probabilidade $1-\beta$)
$\mathcal{H}_0$ não rejeitada	Decisão correcta (probabilidade $1-\alpha$)	Erro do Tipo II (probabilidade β)

A diferença $1-\beta$ traduz a probabilidade de rejeitar $\mathcal{H}_0$ quando ela é, de facto, verdadeira e designa-se por **Potência de teste**.

11.3 α *versus* β.

O teste ideal era aquele em que $\alpha = 0$ e $\beta = 0$, o que implicaria que o teste nos levaria sempre à decisão correcta, mas o teste ideal raramente existe. No entanto, é desejável que a probabilidade do erro do tipo I seja reduzida uma vez que a possibilidade de rejeitar incorrectamente a hipótese nula é considerada grave, pois esta hipótese corresponde à posição que deve ser defendida, salvo se evidência convincente apontar no sentido da hipótese alternativa.

11.4 Relação entre testes de hipóteses e intervalos de confiança.

- A relação fundamental que existe entre os testes de hipóteses e os intervalos de confiança pode ser enunciada nos termos seguintes:

uma hipótese nula $\mathcal{H}_0 : \theta = \theta_0$

pode ser rejeitada a um nível de significância α se, e somente se, o intervalo de confiança para θ não incluir o valor θ_0.
Note-se que esta condição impõe que o intervalo de confiança seja compatível com a natureza de $\mathcal{H}_1$, ou seja, que para testes bilaterais se construam intervalos de confiança bilaterais e, para testes unilaterais(num sentido), se construam intervalos de confiança unilaterais (no mesmo sentido).

- Assim pode-se proceder ao teste de hipóteses recorrendo a intervalos de confiança.

Exemplo 11.4.1 *Certo candidato às eleições da associação académica da ESTGM declarou, durante a campanha, que iria obter a maioria. Um elemento da sua comitiva resolveu recolher aleatoriamente uma amostra com 120 eleitores; 72 responderam que iriam votar nesse candidato.*

1. *A um nível de significância de 5%, verifique se os dados evidenciam que o candidato tinha razão.*

Resolução:

1. Teste de significância unilateral à direita para uma proporção.

   Para obter maioria teria de obter mais de 50% dos votos. Assim,

$$\mathcal{H}_0 : p = 0.5$$
$$\mathcal{H}_1 : p > 0.5$$

   $ET = \frac{\widehat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$; se $\mathcal{H}_0$ é verdadeira $ET \sim N(0,1)$
   $\widehat{p} = \frac{72}{120} = 0.6$

$$ET = \frac{0.6 - 0.5}{\frac{0.5\times 0.5}{120}} = 2.1908$$

   Nível de significância: 5%

$$ET(\alpha) = ET(0.05) = 1.645$$

   Região de rejeição: $ET > ET(\alpha) \Leftrightarrow ET > 1.645$.

   Como $2.1908 > 1.645$, rejeita-se $\mathcal{H}_0$ e portanto pode concluir-se que o candidato tinha razão.

$$VP = P(Z > 2.1908|\mathcal{H}_0) = 1 - P(Z \leqslant 2.1908) = 0.0143.$$

Exemplo 11.4.2 *Uma certa noite encontrava-se a jogar aos dados em casa de um amigo com outros amigos. O João estava a perder! A dada altura o João levantou a hipótese dos dados não serem perfeitos. O Simão apoiou e resolveu lançar um dos dados 200 vezes, tendo a face 4 saído 20 vezes. Diga se o dado é perfeito.*

Exemplo 11.4.3 *(Exame 2002/2003) A empresa de gelados ComGelados afirma que os seus gelados são mais baratos que os da sua principal concorrente. Um aficionado por este tipo de sobremesas decidiu testar a afirmação. Recolheu amostras de ambas as empresas e obteve o seguinte resultado.*

Amostra	n	$\overline{X}$	s'
ComGelados	*15*	*5*	*0.2*
Concorrente	*12*	*6*	*0.3*

1. *Teste se a afirmação da empresa ComGelados é verdadeira. (Para $\alpha = 5\%$)*

Exemplo 11.4.4 *(Exame 2002/2003) A Sr.ª Maria das Hortaliças, negociante de longa data, acaba de mudar o seu local de venda para uma barraca mais central. Ela considera que valeu a pena a mudança se o valor médio das vendas diárias actuais for superior em mais de 40 euros em relação ao mesmo valor obtido na barraca anterior. Para o efeito, decidiu apontar os lucros diários obtidos em 25 dias antes da mudança e em 18 dias na nova barraca. Os resultados obtidos encontram-se descritos na tabela seguinte, em euros:*

Amostra	n	$\overline{X}$	s'
Antes	*24*	*140*	*8*
Depois	*18*	*190*	*10*

1. *Ajude a Sr.ª Maria das Hortaliças a tirar a sua conclusão : valeu ou não a pena trocar para a nova barraca? (considere $\alpha = 5\%$)*

2. *Realize o mesmo teste utilizando a metodologia dos intervalos de confiança. A sua conclusão é a mesma?*