Elementos de Estatística e Probabilidades I

Inês Dias

01-01-2013

O principal objetivo da deste documento é fornecer conhecimentos básicos dos diferentes tópicos da disciplina de Elementos de Estatística e Probabilidades I do curso de Educação Básica.

Introdução

A Estatística é um ramo da Matemática Aplicada. Fornece métodos para recolha, organização, descrição, análise e interpretação de dados e para a utilização dos mesmos na tomada de decisões. Divide-se em duas grandes áreas:

Estatística descritiva (conjunto de técnicas apropriadas para recolher, organizar, reduzir e apresentar dados estatísticos);

Inferência estatística - com base na informação amostral permite caracterizar uma certa população, requerendo o conhecimento das probabilidades. Esta área subdivide-se em:

- Estimação – visa determinar o valor dos parâmetros
- Testes de Hipóteses – testa suposições sobre os parâmetros da população

População e Amostra e Unidade Estatística

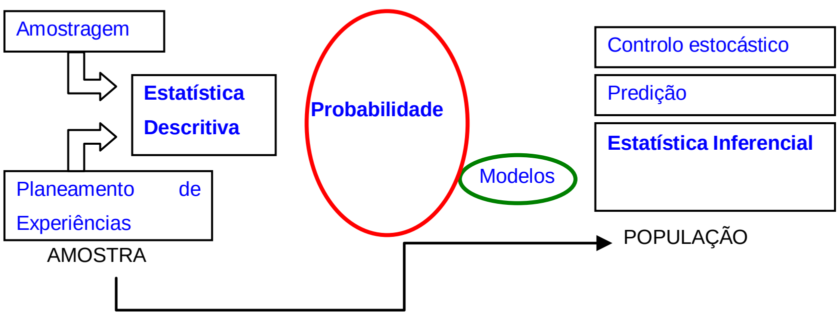
População ou **universo estatístico** é o conjunto de todos os elementos que se podem estudar e que possuem pelo menos uma característica comum.
A população pode ser finita (ex. alunos que frequentam Educação básica; Os educadores de infância da região do Alentejo, etc.) ou infinita (ex. Temperaturas em cada ponto da cidade de Évora, etc.)

Amostra é um subgrupo finito da população, selecionado para análise.

Designa-se por **unidade estatística** ou **elemento**, a cada elemento da população.

Nota: As medidas relativas à população designam-se por parâmetros, usualmente são desconhecidos mas fixos e que se pretendem conhecer.
As medidas relativas à amostra designam-se por estatísticas. Variam de amostra para amostra (são uma variável aleatória)

Podemos esquematizar a informação anterior da seguinte forma:



Técnicas de amostragem

Probabilísticas – Cada elemento da população tem hipótese de ser incluído na amostra e é possível medir a probabilidade de isso acontecer, através do cálculo de probabilidades

- *Amostragem aleatória simples* onde qualquer elemento da população tem a mesma hipótese de ser escolhido.

  Exemplo: Se pretendermos selecionar uma amostra de 30 alunos numa escola com 1200 alunos, numeramos os alunos de 1 a 1200 e seguidamente escolhemos ao acosso 30 alunos.

- *Amostragem sistemática* onde os elementos da amostra são escolhidos a partir de uma regra estabelecida.

  Exemplo: Se pretendermos selecionar uma amostra de 30 alunos numa escola com 1200 alunos, depois de numerarmos os alunos de 1 a 1200, podemos por exemplo, escolher um aluno de 30 em 30 a partir do 1º aluno selecionado, que é escolhido ao acaso de entre o primeiro grupo de 30.

- *Amostragem estratificada* que se utiliza quando a população está dividida em estratos ou grupos diferenciados.

  Exemplo: Se pretendermos selecionar uma amostra de 30 alunos numa escola secundária, considerando cada ano de escolaridade (10º, 11º e 12º) como um estrato, escolheríamos em cada um desses anos um determinado número de alunos por um dos processos anteriores. O nº de alunos a escolher em cada estrato (ano) deve ser proporcional ao número de alunos que frequenta cada ano de escolaridade.

Não Probabilísticas – não permitem definir com rigor ou calcular as probabilidades de inclusão na amostra dos diferentes elementos da população. Este método de amostragem possui muitas limitações e é inferior à amostragem probabilística em termos de precisão de resultado, pelo que os resultados obtidos por este tipo de pesquisa devem ser interpretados com cuidado.

- *Amostragem por Conveniência ou Acidental*: É utilizada quando se pretende obter informações de forma rápida e barata. É o pesquisador que define os elementos que são convenientes para a pesquisa.

  Exemplo: Numa pesquisa de opinião podemos selecionar elementos para a amostra tais como; estudantes numa sala de aula, alguns amigos e vizinhos, etc.

- *Amostragem por Julgamento:* É baseada na escolha deliberada e exclui qualquer processo aleatório. Os elementos que compõem a amostra são considerados adequados de acordo com o que o pesquisador está interessado em estudar, escolhendo-os por achar que são representativos da população.

  Exemplo: Imaginemos que se pretende estudar a aceitação de uma nova marca de roupa de um estilista famoso. So devem entrar na amostra pessoas que tenham condições financeiras de comprar esta nova marca (ficam excluídas as classes sociais mais baixas)

- *Amostragem por quotas*: A amostra por quotas nada mais é que um tipo especial de amostra intencional (Mattar, F. p. 134). No entanto, na amostragem por quotas a população deve ser conhecida, pelo menos aproximadamente, de forma que a representatividade de cada grupo dentro da população seja percebida na amostra isto é, a proporção dos elementos na amostra por quota deve ser semelhante à proporção encontrada na população de onde a amostra foi retirada. A grande diferença entre a amostragem por quotas e estratificada é que na amostragem por quotas os elementos não são selecionados através de aleatoriedade, enquanto que na estratificada a seleção dos elementos de cada estrato é feita utilizando amostragem aleatória.

Dados Estatísticos

Chama-se dado estatístico ao resultado da observação dos elementos de determinado conjunto.

Variável é uma característica comum a todos os elementos observados; algo que se pretende medir, controlar ou manipular na realização de uma análise estatística. Assim, o objecto do estudo estatístico é a análise das variáveis de forma a explorar a informação que estas podem fornecer.

Os dados estatísticos podem ser classificados quanto ao atributo, podendo sempre ser representadas numericamente. Assim temos:

Dados Qualitativos onde a escala de medida apenas indica a sua presença em categorias de classificação discreta exaustivas e mutuamente exclusivas. Estes dados podem ser medidas numa escala:

- **Nominal:** Os dados são medidos em classes discretas, mas não é possível estabelecer à partida um qualquer tipo de qualificação ou operação entre eles. Exemplos: o sexo (masculino ou feminino), a cor dos olhos (castanho, azul ou verde).
- **Ordinal:** Os dados são medidos em classes discretas entre as quais já é possível definir uma determinada ordem. Exemplos: os estratos sociais (baixo, médio ou elevado), as habilitações literárias (básico, secundário ou superior).

Dados Quantitativos dados cuja escala de medida permite a ordenação e quantificação de diferenças entre eles. Estes dados podem ser do tipo:

- **Discreto:** apenas podem tomar um número finito ou uma infinidade numerável de valores. Exemplos: o número de indivíduos por família, o número de bebidas ingeridas por pessoa e por mês, número de alunos que falta por aula, nº de erros ortográficos dados ao longo da vida…
- **Contínuas:** podem tomar qualquer valor de um intervalo de números reais. Exemplos: altura, peso, temperatura...

.

Análise Exploratória de Dados

A finalidade da Análise Exploratória de Dados é examinar os dados antes de se aplicar qualquer técnica estatística (tratamento de dados), pois consegue-se um entendimento básico dos dados e das relações existentes entre as variáveis analisada.

O tratamento de dados tem 3 vertentes:

1. Ordenar
2. Resumir
3. Evidenciar padrões através de uma representação estruturada
    a. Tabelas. Exemplos: tabelas de frequência
    b. Gráficos. Exemplos: gráficos de barras, histogramas
    c. Métodos semi-gráficos. Exemplos: Diagramas de caule e folhas

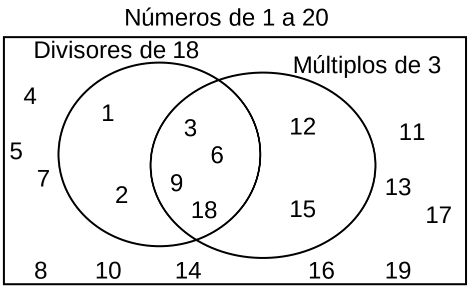
Tabelas e Gráficos

Para a realização de um estudo estatístico é fundamental saber efetuar uma correta representação de dados em tabelas e gráficos. De seguida apresenta-se alguns processos, para organizar a informação contida nos dados, de forma a realçar as características mais importantes.

Diagramas de Venn

São diagramas que utilizam, usualmente, círculos para uma classificação de objetos ou números, que partilhem características comuns. Habitualmente, considera-se um retângulo que representa o universo e dentro desse retângulo consideram-se círculos que representam os elementos com as características de interesse.

Exemplo: Suponhamos que pretendemos classificar os números de 1 a 20 considerando os divisores de 18 e os múltiplos de 3.



Diagramas de Carroll

São tabelas para organizar dados ou objetos segundo critérios de sim/não. O nome atribuído a estes diagramas, é uma homenagem a Lewis Carroll, matemático e escritor inglês, que gostava muito de problemas de lógica e de jogos matemáticos.

Exemplo: Consideremos o exemplo anterior em que pretendemos classificar os números de 1 a 20 considerando os divisores de 18 e os múltiplos de 3

Números de 1 a 20

<table>
<tr><td></td><td></td><td colspan="2">Divisores de 18</td></tr>
<tr><td></td><td></td><td>Sim</td><td>Não</td></tr>
<tr><td rowspan="2">Múltiplos de 3</td><td>Sim</td><td>3, 6, 9, 18</td><td>1, 2</td></tr>
<tr><td>Não</td><td>12, 15</td><td>4, 5, 7, 8, 10, 11, 13, 14, 16, 17, 19</td></tr>
</table>

Tabelas e Gráficos para dados qualitativos

Tally Charts

Tally Charts são um esquema de contagem gráfica.

Exemplo: Imaginemos que queremos eleger o delegado de uma turma com 25 alunos onde só quatro alunos (Ana, Miguel, Tiago e Patrícia) se mostraram disponíveis para o cargo. Assim as regras de votação serão, todos os alunos têm de votar num dos disponíveis a votação, ou caso não concordem com nenhum dos nomes poderão votar em branco.

- Depois de todos votarem vão-se contando os votos colocando um traço à frente do primeiro nome que sair.
- Se o segundo voto for para o mesmo aluno coloca-se um traço à frente do primeiro; se for para outro aluno coloca-se um traço à frente do nome correspondente;
- E assim sucessivamente, os votos vão sendo colocados à frente do nome correspondente.
- O quinto traço coloca-se de forma oblíqua a cortar os 4 traços anteriores.

No fim obtém-se um esquema do tipo:

<table>
<tr><td></td><td>Votos</td></tr>
<tr><td>Ana</td><td>|||</td></tr>
<tr><td>Miguel</td><td>𝍸 |||</td></tr>
<tr><td>Patrícia</td><td>𝍸 𝍸</td></tr>
<tr><td>Tiago</td><td>||||</td></tr>
</table>

Pela representação anterior conclui-se rapidamente que a Patrícia foi eleita delegada de turma, pois foi a que obteve maior número de votos. Também podemos verificar que não houve empates e que nenhum dos meninos votou em branco.

Este esquema de contagem gráfica permite:

- identificar todas as categorias que a variável qualitativa em estudo assume no conjunto dos dados
- organizar os dados de modo a que a contagem do número de elementos (frequências absolutas) em cada uma dessas categorias de faça de forma fácil e rápida.

Tabelas de Frequências para dados qualitativos

Seja (X_1, X_2, …, X_n) uma amostra aleatória de uma população descrita por uma variável aleatória (v.a.) X, onde X, é do tipo qualitativo.

Sejam a_1, a_2, …, a_k os valores distintos que se registam na amostra de n observações, (x_1, x_2, …, x_n), da variável X. Tem-se:

n_i - n.º de observações da amostra iguais a a_i (frequência absoluta)

$$n = \sum_{i=1}^{k} n_i = n_1 + n_2 + \ldots + n_k$$

$f_i = n_i/n$ - frequência relativa de a_i; exprime o nº de vezes que a_i é observado relativamente ao nº total de observações

A partir daqui é fácil chegar à representação dos dados quer através de tabelas.

Tabela de frequências:

valores	n_i	f_i
a_1 … a_k	n_1 … n_k	f_1 … f_k
Total	n	1

Uma tabela de frequências indica-nos as categorias que uma dada variável aleatória assume, bem como a frequência (absoluta e/ou relativa) de cada uma dessas categorias.

Para o conjunto de dados anterior, vamos construir a tabela de frequências associada:

Candidatos	n_i	f_i
Ana	3	0,12
Miguel	8	0,32
Patrícia	10	0,40
Tiago	4	0,16
Total	25	1

Representações gráfica de dados qualitativos:

DIAGRAMA DE BARRAS

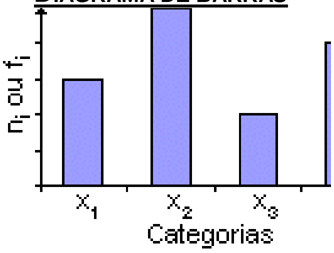
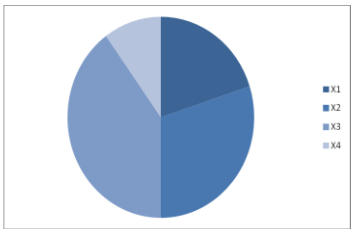


DIAGRAMA CIRCULAR



Variáveis quantitativas discretas

A variável em estudo, X, é do tipo qualitativo ou quantitativo, mas discreta.

Sejam $a_1, a_2, \ldots, a_k$ os valores distintos que se registam na amostra de n observações, $(x_1, x_2, \ldots, x_n)$, da variável X. Tem-se:

n_i - n.º de observações da amostra iguais a a_i (frequência absoluta)

$$n = \sum_{i=1}^{k} n_i = n_1 + n_2 + \ldots + n_k$$

$f_i = n_i/n$ - frequência relativa de a_i; exprime o nº de vezes que a_i é observado relativamente ao nº total de observações

$$f_1 + f_2 + \ldots + f_k = \sum_{i=1}^{k} f_i = \sum_{i=1}^{k} \frac{n_i}{n} = 1$$

$N_i = n_1 + n_2 + \ldots + n_i = \sum_{j=1}^{i} n_j$ - Frequência absoluta acumulada do valor a_i, exprime o nº de observações da amostra com valor menor ou igual a a_i.

$F_i = \frac{N_i}{n} = f_1 + f_2 + \ldots + f_i = \sum_{j=1}^{i} f_j$ - Frequência relativa acumulada do valor a_i

A partir daqui é fácil chegar à representação dos dados quer através de tabelas quer graficamente.

Tabela de frequências:

valores	n_i	f_i	N_i	F_i
a_1	n_1	f_1	N_1	F_1
…	…	…	…	…
a_k	n_k	f_k	$N_k = n$	$F_k = 1$
Total	n	1		

Representações Gráficas:

DIAGRAMA DE BARRAS

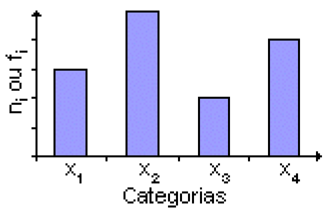


DIAGRAMA INTEGRAL

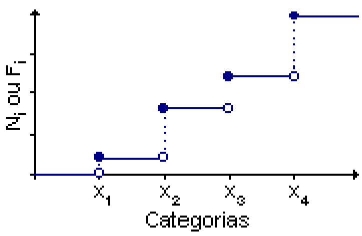
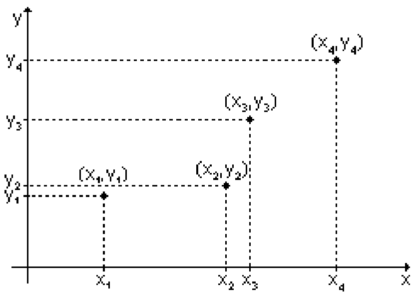


DIAGRAMA DE DISPERSÃO



Variáveis contínuas

No caso contínuo são raros os valores que se repetem na amostra, portanto o agrupamento dos dados consiste na sua classificação em classes. Logo temos duas etapas:

a) Definição das classes
b) Contagem de valores pertencentes a cada classe

Definição das classes

A classe é um intervalo de $\Re$, a que se dá o nome de intervalo de classe.

Regras que deverão ser seguidas na construção das classes:

a) Em geral o nº de classes deve estar compreendido entre 4 e 14,
b) Nenhuma classe deverá ter frequência nula,
c) As classes deverão ter, sempre que possível, amplitudes iguais,
d) Os pontos médios das classes deverão ser de cálculo fácil,
e) Classes abertas deverão ser evitadas
f) Os limites das classes são definidos de modo a que cada valor da variável é incluído num e num só intervalo.

Para determinar o nº de classes, K, por vezes é adoptada uma das seguintes estratégias

1. Escolha do nº de classes - *Regra de Sturges*
2. Escolha da amplitude da classe - *Regra (S/3, S/2)*

Nós apenas iremos falar da *Regra de Sturges*.

Escolha do nº de classes - *Regra de Sturges*

1º - Escolher o nº de classes k, que deve ser obtido por,

$$k \approx \left[\frac{\ln n}{\ln 2}\right] + 1^{(1)}$$

2º - Identificar o mínimo e o máximo na amostra

$\min(x_1, x_2, \ldots, x_n)$

$\max(x_1, x_2, \ldots, x_n)$

3º - Determinar a amplitude (range) na amostra

$R_n = \max(x_1, x_2, \ldots, x_n) - \min(x_1, x_2, \ldots, x_n)$

4º - Determinar a amplitude da classe, h,

Primeiro determinar $h^* = \dfrac{\max(x_1, x_2, \ldots, x_n) - \min(x_1, x_2, \ldots, x_n)}{k}$

Em seguida escolher $h > h^*$ para a amplitude da classe(2).

---

(1) [x] – representa a parte inteira de x

(2) Não devemos aumentar mais do que 10% do valor de h^*

5º - Determinar as classes

i) Determinar **$h * k$**

ii) Calcular o "excesso", ε **$\varepsilon = h * k - R_n$**

iii) A primeira classe a construir começará em **$\min(x_1, x_2, \ldots, x_n) - \varepsilon/2$**

Formam-se as classes para a direita somando a amplitude, h, ao limite inferior até que se encontre a última classe que termina em **$\max(x_1, x_2, \ldots, x_n) + \varepsilon/2$.**

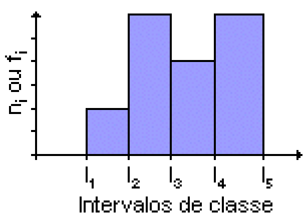
Depois de determinadas as classes há que fazer a contagem de valores pertencentes a cada classe para construir a tabela de frequências e/ou gráficos pretendidos.

Tabela de frequências:

Classes	**Ponto médio**	**Freq. Abs.**	**Freq. Rel.**	**Freq. Abs. Acum.**	**Freq. Rel. Acum.**
c_i	$x'_i = \frac{l_i + l_{i+1}}{2}$	n_i	$f_i = \frac{n_i}{n}$	$N_i = \sum_{j=1}^{i} n_j$	$F_i = \sum_{j=1}^{i} f_j = \frac{N_i}{n}$
$[l_1; l_2[$	x'_1	n_1	f_1	N_1	F_1
...	...	...	...	...	...
$[l_k; l_{k+1}[$	x'_k	n_k	f_k	$N_k = n$	$F_k = 1$
Total		***n***	***1***		

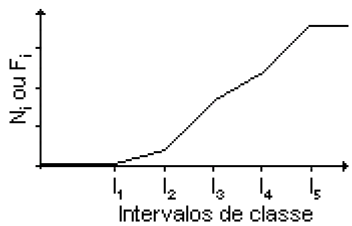
Representações Gráficas:

HISTOGRAMA



POLÍGONO DE FREQUÊNCIAS

POLÍGONO INTEGRAL



1. Distribuições de frequências

Consideremos as distribuições de frequências do tipo campanular. Dentro destas podemos ainda diferencia-las em:

- simétricas,
- assimétricas

  * assimétricas positivas ou enviesada à esquerda (têm o ramo esquerdo mais abrupto)

  * assimétricas negativas ou enviesada à direita (têm o ramo direito mais abrupto)

Características principais a estudar numa amostra

Medidas de Localização

Tendência central:

- **Média**

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{x_1 + x_2 + \ldots + x_n}{n}$$ para dados não agrupados

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{m} n_i x_i$$ para dados agrupados

- **Mediana**

O valor real tal que à sua direita e à sua esquerda se situam 50% das observações.

Se representarmos por $x_{1:n}$, $x_{2:n}$, …, $x_{n:n}$ as observações ordenadas por ordem crescente (i.e. $x_{1:n} < x_{2:n} < \ldots < x_{n:n}$), então

$$M_e = \begin{cases} x_{\frac{n+1}{2}:n}, & \text{se } n \text{ é impar} \\ \dfrac{x_{\frac{n}{2}:n} + x_{\frac{n}{2}+1:n}}{2}, & \text{se } n \text{ é par} \end{cases}$$ para dados discretos.

$$M_e = l_i + c_i \times \frac{\frac{n}{2} - N_{i-1}}{n_i}$$ para dados contínuos,

l_i – limite inferior da classe mediana

c_i – amplitude da classe mediana

N_{i-1} – freqência absoluta acumulada da classe anterior à classe mediana

N_i – frequência absoluta simples da classe mediana

- **Moda**

Representa-se por M_o e, para dados discretos, é igual ao valor a que corresponde a maior frequência.

Para dados contínuos determina-se da seguinte forma,

$$M_o = l_i + c_i \times \frac{\Delta_1}{\Delta_1 + \Delta_2},$$

l_i – limite inferior da classe modal

c_i – amplitude da classe modal

$\Delta_1 = n_i - n_{i-1}$ e $\Delta_2 = n_i - n_{i+1}$.

Comparação da média, mediana e moda

Distribuições simétricas	$\bar{x} = M_e = M_o$
Distribuições assimétricas positivas	$\bar{x} > M_e > M_o$
Distribuições assimétricas negativas	$\bar{x} < M_e < M_o$

Tendência não central:

- **Quantil de ordem α**

O valor real tal que uma percentagem α das observações se situam à sua esquerda.

$$q_\alpha = \begin{cases} x_{[n\alpha]+1:n} & \text{se } n\alpha \text{ não inteiro} \\ \dfrac{x_{n\alpha:n} + x_{n\alpha+1:n}}{2} & \text{se } n\alpha \text{ inteiro} \end{cases}$$ para dados discretos.

$$q_\alpha = l_i + c_i \times \frac{n\alpha - N_{i-1}}{n_i}$$ para dados contínuos,

l_i – limite inferior da classe do quantil

c_i – amplitude da classe do quantil

N_{i-1} – freqência absoluta acumulada da classe anterior à classe do quantil

N_i – frequência absoluta simples da classe do quantil.

- **Máximo e Mínimo**

$x_{1:n} = \min(x_1, x_2, \ldots, x_n)$

$x_{n:n} = \max(x_1, x_2, \ldots, x_n)$

Medidas de dispersão

Medidas de dispersão absoluta:

- **Amplitude amostral**

$A = x_{n:n} - x_{1:n}$

- **Amplitude ou Distância Inter-quartis**

$IQ = Q_3 - Q_1$

- **Intervalo de variação**

$P_{90} - P_{10}$

- **Desvio médio**

$$d = \frac{1}{n}\sum_{i=1}^{n} |x_i - \bar{x}|$$ para dados não agrupados

$$d = \frac{1}{n}\sum_{i=1}^{m} n_i |x_i - \bar{x}|$$ para dados agrupados

- **Variância amostral**

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \quad \text{para dados não agrupados}$$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{m} n_i\left(x_i - \overline{x}\right)^2 \quad \text{para dados agrupados}$$

- **Desvio padrão amostral**

$$s = +\sqrt{s^2}$$

Medidas de dispersão relativa:

- **Coeficiente de dispersão** $CD = \frac{s}{\overline{x}}$

- **Coeficiente de variação** $CV = \frac{s}{\overline{x}} \times 100\ (\%)$

Nota: Estes coeficientes só se utilizam quando a variável toma valores de um só sinal; todos positivos ou todos negativos

Momentos empíricos

- **Momento empírico centrado de ordem k**

$$m_k = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^k \quad k = 1, 2, \ldots \quad \text{para dados não agrupados}$$

$$m_k = \frac{1}{n}\sum_{i=1}^{m} n_i\left(x_i - \overline{x}\right)^k \quad k = 1, 2, \ldots \quad \text{para dados agrupados}$$

$m_0 = 1$

$m_1 = 0$

$m_2 = s^2$

- **Momento empírico de ordem k, relativamente a um valor arbitrário A**

$$m'_k(A) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - A\right)^k \quad k = 1, 2, \ldots \quad \text{para dados não agrupados}$$

$$m'_k(A) = \frac{1}{n}\sum_{i=1}^{m} n_i\left(x_i - A\right)^k \quad k = 1, 2, \ldots \quad \text{para dados agrupados}$$

- **Momento empírico de ordem k**

  (Caso em que A = 0)

  $m'_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k \quad k = 1, 2, \ldots$ para dados não agrupados

  $m'_k = \frac{1}{n}\sum_{i=1}^{m} n_i x_i^k \quad k = 1, 2, \ldots$ para dados agrupados

  $m'_1 = \bar{x}$

Relações existentes entre momentos empíricos centrados e momentos empíricos

$$m_k = \sum_{j=0}^{k} \binom{k}{j} m'_{k-j} \left(-m'_1\right)^j$$

Em particular tem-se

$$m_2 = m'_2 - m'^2_1$$

$$m_3 = m'_3 - 3m'_1 m'_2 + 2m'^3_1$$

$$m_4 = m'_4 - 4m'_1 m'_3 + 6m'^2_1 m'_2 - 3m'^4_1$$

Medidas de assimetria

- **Grau de assimetria de Pearson**

  $G_P = \frac{\bar{x} - M_o}{s}$

  Para distribuições moderadamente assimétricas

  $G_P = \frac{3(\bar{x} - M_o)}{s}$

  $G_P \cong 0 \Rightarrow$ a distribuição é simétrica,

  $G_P \cong 3 \Rightarrow$ a distribuição é assimétrica positiva,

  $G_P \cong -3 \Rightarrow$ a distribuição é assimétrica negativa.

- **Grau de assimetria de Bowley**

  $G_B = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$

  $G_B \cong 0 \Rightarrow$ a distribuição é simétrica,

  $G_B \cong 1 \Rightarrow$ a distribuição é assimétrica positiva,

  $G_B \cong -1 \Rightarrow$ a distribuição é assimétrica negativa.

- **Coeficiente de assimetria**

$$g_1 = \frac{m_3}{\sqrt{m_2^3}}$$

$g_1 = 0 \Rightarrow$ a distribuição é simétrica,

$g_1 > 0 \Rightarrow$ a distribuição é assimétrica positiva,

$g_1 < 0 \Rightarrow$ a distribuição é assimétrica negativa.

Medidas de achatamento

- **Coeficiente percentil de curtose (ou Kurtosis)**

$$K = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

$K = 0.263 \Rightarrow$ a curva diz-se mesocúrtica,

$K < 0.263 \Rightarrow$ a curva diz-se leptocúrtica,

$K > 0.263 \Rightarrow$ a curva diz-se platicúrtica.

- **Coeficiente de curtose (ou Kurtosis)**

$$g_2 = \frac{m_4}{m_2^2} - 3$$

$g_2 = 0 \Rightarrow$ a curva diz-se mesocúrtica,

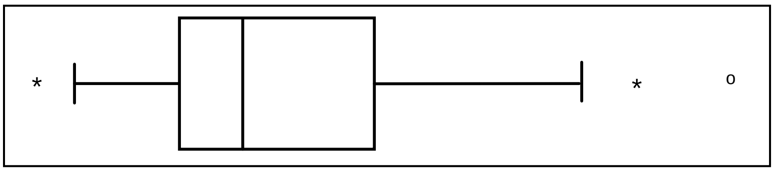
$g_2 > 0 \Rightarrow$ a curva diz-se leptocúrtica,

$g_2 < 0 \Rightarrow$ a curva diz-se platicúrtica.

- **Caixa-com-bigodes (Boxplot)**

Uma caixa-com-bigodes (ou boxplot) é uma representação gráfica que apresenta algumas das principais características descritivas de um conjunto de dados, numa imagem compacta.
A caixa com bigodes é um gráfico em que, à escala, se representam:

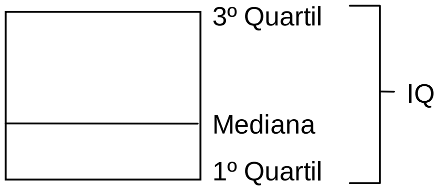
1. O 1º quartil, a mediana, o 3º quartil e a amplitude inter-quartis.
2. O menor valor não outlier e o maior valor não outlier.
3. Os outliers (moderados "*" e severos "o").



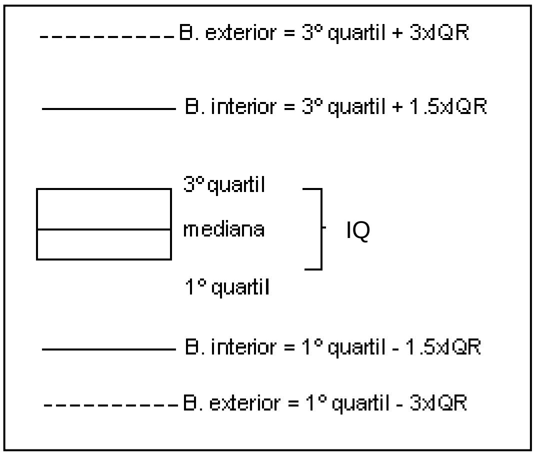
Uma caixa-com-bigodes fornece uma boa visualização da variabilidade dos dados e do tipo da assimetria e achatamento da distribuição.

A caixa central

A primeira componente da boxplot é a amplitude inter-quartis, o comprimento da caixa, que vai do primeiro quartil ao terceiro quartil. Dentro da caixa, traçamos uma linha horizontal indicando a mediana.



As barreiras



<u>barreiras interiores</u>:	1º Quartil – 1.5 x IQR
	3º Quartil + 1.5 x IQR.
<u>barreiras exteriores</u>:	1º Quartil – 3 x IQR
	3º Quartil + 3 x IQR.

Correlação e regressão linear

- **Variâncias amostrais de x e de y**

$$s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right]$$

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2\right]$$

- **Covariância amostral**

$$s_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i y_i - n\,\bar{x}\,\bar{y}\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)\right]$$

- **Coeficiente de correlação**

$$r = \frac{s_{xy}}{s_x s_y}$$

INTRODUÇÃO ÀS PROBABILIDADES

EXPERIÊNCIA ALEATÓRIA. ESPAÇO AMOSTRA. ACONTECIMENTOS

Ao processo que consiste em recolher uma observação de uma variável, ou observar um fenómeno aleatório, que se pretende estudar chamamos experiência aleatória.

Definição 1.

Uma **experiência** diz-se **aleatória** sse verificar as seguintes características:

- cada vez que é efectuada desconhecemos à partida qual o resultado que vamos obter;
- é conhecido o conjunto de todos os resultados possíveis;
- a experiência pode ser repetida em condições similares e existe regularidade quando é repetida muitas vezes.

Definição 2.

O **espaço amostra** ou **universo**, **Ω**, de uma experiência aleatória é o conjunto de todos os resultados possíveis, dessa experiência.

Exemplos:

E1: Lançamento de um dado e observação do número mostrado pela face virada para cima

$\Omega = \{1, 2, 3, 4, 5, 6\}$

E2: Número de peças defeituosas produzidas durante 24 horas por uma máquina.

$\Omega = \{0, 1, 2, \ldots, N\}$ onde N é o nº máximo de peças fabricadas em 24 horas.

E3: Número de peças fabricadas por uma máquina até se observar 10 peças defeituosas.

$\Omega = \{10, 11, 12, \ldots\}$

E4: Lançamento de dois dados e registo da soma dos números mostrados pelas faces viradas para cima

$\Omega = \{2, 3, 4, \ldots, 11,12\}$

Uma vez determinado o espaço amostra estamos interessados em associar probabilidades a subconjuntos de Ω - que designamos por acontecimentos.

Um acontecimento é, pois, um conjunto de resultados possíveis de uma experiência aleatória ou, de modo equivalente qualquer subconjunto do espaço amostra é um acontecimento definido em Ω.

Um acontecimento A respeitante a um determinado espaço amostra Ω e associado a uma experiência aleatória é simplesmente um conjunto de resultados possíveis. Dizemos que A se realizou, se o resultado da experiência aleatória, ω, é um elemento de A, i.e., $\omega \in A$.

Um acontecimento pode ser um conjunto formado por um só elemento de Ω, e nesse caso, diz-se acontecimento elementar. Chamamos aos outros acontecimentos compostos.

Ao espaço amostra, Ω, também se chama acontecimento certo.

ÁLGEBRA DOS ACONTECIMENTOS

Define-se acontecimento, como sendo um subconjunto de Ω.
Existindo um paralelismo entre conjuntos e acontecimentos há, no entanto, uma terminologia própria para acontecimentos. Assim, representando os acontecimentos por A, B, C, ..., temos:

União de acontecimentos:

Se A e B são dois acontecimentos, a união de A com B, $A \cup B$, consiste na realização de pelo menos um dos acontecimentos.

Dada uma sucessão infinita de acontecimentos A_1, A_2, ... Define-se a sua união $\bigcup_{i=1}^{\infty} A_i$ como sendo o acontecimento que se realiza sse ocorrer pelo menos um dos acontecimentos A_i.

Intersecção de acontecimentos:

Dados dois acontecimentos A e B, define-se intersecção, $A \cap B$, como o acontecimento que ocorre sse ocorrerem A e B, em simultâneo ou em sequência.

Dada uma sucessão infinita de acontecimentos A_1, A_2, ... Define-se a sua intersecção $\bigcap_{i=1}^{\infty} A_i$ como sendo o acontecimento que se realiza sse ocorrerem todos os acontecimentos A_i.

Diferença de acontecimentos:

A diferença de dois acontecimentos A e B, ou acontecimento diferença, A - B (ou A\B) é o acontecimento que se realiza quando ocorre A mas não B. Será, então, constituído por todos os elementos de A que não pertencem simultaneamente não pertencem a B
No caso particular da diferença de dois acontecimentos A e B, em que A é o próprio espaço amostra, Ω, o acontecimento diferença denomina-se **complementar** de B, e representa-se por $\overline{B}$ e é o acontecimento que ocorre quando não ocorre B.

Acontecimentos disjuntos ou **acontecimentos mutuamente exclusivos** são acontecimentos em que a realização de um deles implica a não realização do outro.

Acontecimento impossível é o acontecimento que resulta da intersecção de acontecimentos mutuamente exclusivos. Analogamente ao que se passa na teoria dos conjuntos, representa-se por $\emptyset$.

As propriedades mais importantes sobre conjuntos são válidas para os acontecimentos. Assim,

Propriedades	**União**	**Intersecção**
Comutativa	A U B = B UA	A ∩ B = B ∩ A
Associativa	A U (B U C) = (A U B) U C	A ∩ (B ∩ C) = (A ∩ B) ∩ C
Distributiva	A U (B ∩ C)=(A U B) ∩ (A U C)	A ∩ (B U C) = (A ∩ B) U (A ∩ C)
Idempotência	A U A = A	A ∩ A = A
Complementaridade	$A \cup \overline{A} = \Omega$	$A \cap \overline{A} = \emptyset$
Leis de Morgan	$\overline{A \cup B} = \overline{A} \cap \overline{B}$	$\overline{A \cap B} = \overline{A} \cup \overline{B}$
Elemento neutro	$A \cup \emptyset = A$	$A \cap \Omega = A$
Elemento absorvente	$A \cup \Omega = \Omega$	$A \cap \emptyset = \emptyset$
Absorção	$\overline{\overline{A}} = A$	
Dupla negação	Se $A \subseteq B$, então A U B = B e A ∩ B = A	

CONCEITOS DE PROBABILIDADE

"A Teoria da Probabilidade serve como meio para formular modelos de fenómenos naturais em que se supõe intervir o acaso"

in "Probabilidades e Estatística" - Vol I

J. Tiago de Oliveira

Conceito clássico de probabilidade (Probabilidade Laplaciana)

Se a uma experiência aleatória se podem associar N resultados possíveis, mutuamente exclusivos e igualmente prováveis, e se n_A desses acontecimentos tiverem o atributo A, então a probabilidade de A é a fracção $((n_A)/N)$, ou seja,

$$P[A] = \frac{\text{nº de casos favoráveis a A}}{\text{nº de casos possíveis}} = \frac{n_A}{N}.$$

Conceito frequencista de probabilidade

Se em N realizações de uma experiência aleatória, o acontecimento A ocorreu n vezes, diz-se que a frequência relativa de A nas N realizações é

$$f_A = n / N$$

a frequência relativa do acontecimento A.

Define-se probabilidade do acontecimento A, como o número para que tende a frequência relativa f_A, quando se aumenta o número de provas, ou seja,

$$P[A] = \lim_{N\to\infty}(f_A).$$

Conceito subjectivo de probabilidade

Existem experiências aleatórias que não se podem repetir várias vezes nas mesmas condições e cujos resultados não são equiprováveis. Nestes casos, os conceitos frequencista e clássico não se podem aplicar, e a probabilidade de um acontecimento é dada pelo grau de credibilidade que cada pessoa dá à realização de um acontecimento.

AXIOMAS DA TEORIA DAS PROBABILIDADES

A probabilidade, $P[\cdot]$, é uma função que associa a todo o acontecimento A definido em Ω um número pertencente ao intervalo [0,1], e que satisfaz os seguintes axiomas

Axiomas:

1. $0 \le P[A] \le 1, \ \forall A \subseteq \Omega$;
2. $P[\Omega] = 1$ (Ω é o acontecimento certo);
3. Sendo A e B acontecimentos <u>mutuamente exclusivos</u> definidos em Ω, ou seja, $A \cap B = \emptyset$, então

$$P[A \cup B] = P[A] + P[B].$$

Generalizando axioma 3, sendo $A_1, A_2, \ldots, A_n$ acontecimentos <u>mutuamente exclusivos</u> definidos em Ω,

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P[A_i]$$

Estes axiomas são válidos para as definições de probabilidade frequencista e clássica.

Como consequências imediatas dos axiomas temos as seguintes propriedades:

Propriedade 1

Dado um acontecimento A, com probabilidade P[A], a probabilidade do seu complementar, $\overline{\mathrm{A}}$, é

$$P\left[\overline{\mathrm{A}}\right]=1-P\left[\mathrm{A}\right].$$

Propriedade 2

A probabilidade do acontecimento ímpossível, $\varnothing$, é dada por

$$P[\varnothing]=0.$$

Propriedade 3

Sejam A e B dois acontecimentos tais que A $\subseteq$ B, então

$$P[\mathrm{A}] \le P[\mathrm{B}]$$

Propriedade 4

Sejam A e B dois acontecimentos quaisquer, a probabilidade da diferença, B - A, é dada por

$$P[\mathrm{B} - \mathrm{A}] = P[\mathrm{B}] - P[\mathrm{A} \cap \mathrm{B}].$$

Propriedade 5

Sejam A e B dois acontecimentos quaisquer, a probabilidade da união A U B é

$$P[\mathrm{A} \cup \mathrm{B}] = P[\mathrm{A}] + P[\mathrm{B}] - P[\mathrm{A} \cap \mathrm{B}].$$

Propriedade 6

Sejam A e B dois acontecimentos quaisquer,

$$P[\mathrm{A} \cup \mathrm{B}] \le P[\mathrm{A}] + P[\mathrm{B}].$$

Propriedade 7

Dados $A_1, A_2, \ldots, A_n$ acontecimentos quaisquer definidos em Ω, então

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P[A_i] - \sum_{i<j} P\left[A_i \cap A_j\right] + \sum_{i<j<k} P\left[A_i \cap A_j \cap A_k\right] + \ldots + (-1)^{n+1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

PROBABILIDADE CONDICIONAL. ACONTECIMENTOS INDEPENDENTES

Definição 3

A probabilidade de um acontecimento A condicionado pela ocorrência de outro acontecimento B, em que $P[\mathrm{B}] > 0$, é dada por

$$P[\mathrm{A}|\mathrm{B}] = \frac{P[\mathrm{A} \cap \mathrm{B}]}{P[\mathrm{B}]}.$$

Consequências imediatas:

- $P[\overline{\mathrm{A}}|\mathrm{B}] = 1 - P[\mathrm{A}|\mathrm{B}]$;
- $P[\mathrm{A}|\mathrm{A}] = 1$ e $P[\overline{\mathrm{A}}|\mathrm{A}] = 0$;
- Se $\mathrm{A} \cap \mathrm{B} = \varnothing \Rightarrow P[\mathrm{A}|\mathrm{B}] = 0$.

Teorema 1 (*Teorema da probabilidade composta*)

A probabilidade de intersecção de dois acontecimentos A e B, que se designa de probabilidade composta, é igual ao produto da probabilidade de um deles pela probabilidade do outro suposto aquele realizado, ou seja,

$$P[\mathrm{A} \cap \mathrm{B}] = P[\mathrm{A}].P[\mathrm{B}|\mathrm{A}] = P[\mathrm{B}].P[\mathrm{A}|\mathrm{B}].$$

Generalização para o caso de n acontecimentos $A_1, A_2, ..., A_n$

$$P\left[\bigcap_{i=1}^{n} A_i\right] = P[A_1] \cdot P[A_2|A_1] \cdot P[A_3|A_1 \cap A_2] \cdot \ldots \cdot P\left(A_n \middle| \bigcap_{i=1}^{n-1} A_i\right)$$

Definição 4

Dois acontecimentos A e B dizem-se **independentes** sse

$$P[\mathrm{A} \cap \mathrm{B}] = P[\mathrm{A}] \cdot P[\mathrm{B}].$$

Se A e B são acontecimentos independentes, então,

$$P[\mathrm{A}|\mathrm{B}] = P[\mathrm{A}] \quad \text{se } P[\mathrm{B}]>0,$$

$$P[\mathrm{B}|\mathrm{A}] = P[\mathrm{B}] \quad \text{se } P[\mathrm{A}]>0.$$

Generalizando, $A_1, A_2, ..., A_n$ dizem-se acontecimentos independentes se

$$P\left[\bigcap_{i=1}^{n} A_i\right] = \prod_{i=1}^{n} P[A_i].$$

Acontecimentos independentes versus acontecimentos mutuamente exclusivos

Dados dois acontecimentos, A e B, tais que $P[A] > 0$ e $P[B] > 0$,

- se os acontecimentos são mutuamente exclusivos, então

$$A \cap B = \emptyset \text{ e } P[A \cap B] = 0,$$

pelo que não podem ser independentes, pois para tal, $P[A \cap B] = P[A].P[B] > 0$;

- se os acontecimentos são independentes, então

$$P[A \cap B] = P[A].P[B] > 0,$$

não podendo ser mutuamente exclusivos, pois para tal, $P[A \cap B] = 0$.

Assim em geral, dois acontecimentos não podem ser simultaneamente independentes e mutuamente exclusivos. Existe, no entanto, um caso particular em que tal pode ocorrer: é o caso em que um dos acontecimentos é impossível, porque este é sempre independente e mutuamente exclusivo de todo e qualquer outro acontecimento possível.

TEOREMA DA PROBABILIDADE TOTAL. FÓRMULA DE BAYES

Definição 5

Os acontecimentos $A_1, A_2, ..., A_n$ definem uma **partição** em Ω, quando se verificam simultaneamente as seguintes condições:

(i) a união de todos os acontecimentos é o próprio espaço amostra, Ω:

$$\bigcup_{i=1}^{n} A_i = \Omega\,;$$

(ii) os acontecimentos são mutuamente exclusivos, dois a dois.

$$A_i \cap A_j = \emptyset, \qquad i \neq j, \;\; i, j = 1, 2, \ldots, n;$$

(iii) todos os acontecimentos têm probabilidade não nula:

$$P[A_i] > 0, \qquad i = 1, 2, \ldots, n.$$

Teorema 2 *(Teorema da probabilidade total)*

Se os acontecimentos A_1, A_2, ..., A_n definem uma partição sobre Ω, então para qualquer acontecimento B definido em Ω, tem-se que

$$P[B] = \sum_{i=1}^{n} P[B|A_i] \cdot P[A_i] =$$
$$= P[B|A_1] \cdot P[A_1] + P[B|A_2] \cdot P[A_2] + \ldots + P[B|A_n] \cdot P[A_n]$$

Teorema 3 *(Fórmula de Bayes)*

Se A_1, A_2, ..., A_n definem uma partição sobre Ω, então para qualquer acontecimento B definido em Ω, com P[B] > 0,

$$P[A_j|B] = \frac{P[A_j] \cdot P[B|A_j]}{\sum_{i=1}^{n} P[A_i] \cdot P[B|A_i]}, \qquad j = 1, 2, \ldots, n$$

Bibliografia:

Afonso, A., Nunes, C., (2010) *Estatística e Probabilidades - Aplicações e Soluções em SPSS*. Escolar Editora

Silva, M. C. M. (1993) *Estatística Aplicada à Psicologia e Ciências Sociais*, Lisboa McGraw-Hill.

Pestana D., Velosa S. (2002). *Introdução à probabilidade e estatística*. Volume 1. Fundação Calouste Gulbenkian.

Fontes:

http://area.dgidc.min-edu.pt/materiais_npmeb/matematicaOTD_Final.pdf