

Análise do documento

METAS CURRICULARES – Ensino Básico – Matemática

Domínio – Organização e tratamento de dados (OTD)

Data are numbers, but they are not “just numbers”. Data are numbers with a context. Because data are numbers with a context, doing statistics means more than manipulating numbers”

In Moore, STATISTICS Concepts and Controversies, página XV

De seguida faremos algumas considerações sobre o documento em análise, elaborado por uma equipa constituída pelos docentes com formação matemática António Bivar (Universidade Lusíada), Carlos Grosso (Escola Superior de Educação João de Deus), Filipe Oliveira (Faculdade de Ciências e Tecnologia da UN) e Maria Clementina Timóteo (Escola Secundária Padre Alberto Neto), neste texto designados por “autores”. O facto de nenhum dos docentes ter formação Estatística, pode ser sintomático e justificativo de no tema **Organização e tratamento de dados (OTD)**, do documento colocado à discussão, se verificarem algumas situações menos claras ou até conducentes a erros. Como especialistas na área de Estatística, é somente nesta área que iremos pronunciar-nos.

Para uma melhor compreensão dos comentários ao documento em análise, inserimos a parte do documento Metas curriculares que se refere a OTD, seguido dos nossos comentários.

Lisboa, 23 de Julho de 2012

A Comissão Especializada de Educação da Sociedade Portuguesa de Estatística

Maria Eugénia Graça Martins

Luísa Canto Castro Loura

Maria Manuela Neves

1º ciclo

“No domínio Organização e Tratamento de Dados é dada ênfase a diversos processos e metodologias que permitem repertoriar e interpretar informação recolhida em contextos variados, aproveitando-se para fornecer algum vocabulário básico da Teoria dos Conjuntos, necessário à compreensão dos procedimentos efetuados. No 3.º ano é apresentada a noção de frequência absoluta e, no 4.º ano, a de frequência relativa bem como a representação de números racionais sob forma de percentagem. As questões relativas a processos aleatórios foram propositadamente deixadas de lado por se entender que apresentam um grau de complexidade demasiado elevado para este nível de ensino, por falta de critérios suficientemente simples que conduzam os alunos a utilizar adequadamente a linguagem associada à interpretação dos fenómenos regidos pelo acaso”

Comentário - A opção dos autores pela inclusão no tema de OTD da referência à Teoria dos Conjuntos, **não nos parece adequada. O termo primitivo “conjunto” e alguns dos conceitos que podem ser introduzidos a este nível de escolaridade, como os de “pertença”, “inclusão” e “cardinal”, são demasiado importantes em Matemática e poderão vir a originar confusões graves caso sejam incluídos no tema de OTD.** O problema coloca-se, principalmente, quando os dados recolhidos são de natureza quantitativa. Se se recolher, numa turma, a informação sobre o número de irmãos de cada um dos alunos, obtemos um “conjunto” de dados que, certamente, não irá verificar uma das condições básicas quanto à enumeração dos elementos de um conjunto, nomeadamente, a de não ter elementos repetidos. Aqui convém observar que em Estatística, não existem elementos repetidos já que estamos a tratar com *dados* e dados são mais do que números, são números com um contexto. Dois elementos aparentemente iguais, não o são, já que foram o resultado de observações sobre entidades diferentes. É claro que este problema poderia ser ultrapassado considerando, no “conjunto” de dados, não o número de irmãos, mas o par (aluno X, nº de irmãos do aluno X), mas seria uma complicação desnecessária. É um facto que se utilizam os diagramas de Venn, cujo suporte são conjuntos mas não vemos a necessidade da utilização da linguagem formal da Teoria dos conjuntos para a *“compreensão dos procedimentos efectuados”*.

É também uma opção dos autores terem deixado de lado as questões relativas a *“processos aleatórios...regidos pelo acaso”*. **Não é essa a nossa opinião.** Veja-se, por exemplo, no curriculum inglês - Key Stage 2 (ages 7-11) onde se diz “...explore doubt and certainty and develop an understanding of probability through classroom situations; discuss events using a vocabulary that includes the words ‘equally likely’, ‘fair’, ‘unfair’, ‘certain’. Ou ainda no Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics – NCTM no que respeita Pre-K-Grade 2 (ages 4-7) onde se lê, página 26 “Discuss events related to student’s experiences as likely or unlikely”.

1º ano

Representação de dados

1. Recolher e representar conjuntos de dados

1. Ler gráficos de pontos e pictogramas em que cada figura representa uma unidade.
2. Recolher e registar dados utilizando gráficos de pontos e pictogramas em que cada figura representa uma unidade.
3. Utilizar corretamente os termos «conjunto», «elemento» e as expressões «pertence», «não pertence» e «cardinal».
4. Representar graficamente conjuntos disjuntos e os respetivos elementos.
5. Classificar objetos de acordo com um ou dois critérios.

Comentário – Sobre o ponto 3 já nos pronunciámos no comentário geral, pelo que quanto a nós seria de excluir. Não se compreende o que é que se pretende no ponto 4, nomeadamente se se referem a algumas das representações gráficas referidas no ponto 2, para representar conjuntos de dados, ou se se trata de teoria dos conjuntos e estão a pensar na representação de conjuntos em diagramas de Venn. No ponto 5 acrescentaríamos “utilizando diagramas de Venn ou de Carroll”, pois admitimos que é isso que os autores têm em mente.

2º ano

Representação de dados

1. Recolher e representar conjuntos de dados

1. Ler tabelas de frequências absolutas, gráficos de pontos e pictogramas em diferentes escalas.
2. Recolher dados utilizando diagramas de contagem (tally charts) e representá-los em tabelas de frequências absolutas.
3. Representar dados através de gráficos de pontos e de pictogramas.

2. Interpretar representações de conjuntos de dados

1. Retirar informação de diagramas de contagem, gráficos de pontos e pictogramas identificando a característica em estudo e comparando os valores das frequências das várias categorias observadas.
2. Organizar conjuntos de dados em diagramas de Carroll.
3. Identificar a reunião e a interseção de dois conjuntos.
4. Construir e interpretar diagramas de Venn.
5. Construir e interpretar gráficos de barras.

Comentário – Quando os autores se referem no ponto 2.1 a “...categorias observadas”, pressupõem que os dados a analisar são unicamente de natureza qualitativa (ou categórica), que se organizam ou classificam em *categorias*? Efectivamente, quando temos dados de natureza quantitativa (ou numérica) os dados organizam-se em *classes* (já não se designam *categorias*). Estas *classes* apresentam-se na forma de intervalos, para o caso de dados contínuos ou na forma de um único valor, no caso de dados discretos, com um número pequeno de valores distintos. Pelo que acabámos de escrever, fica-se na dúvida se os autores estão a considerar só dados de tipo qualitativo, ou também prevêm os quantitativos discretos, já que os diagramas de pontos também se utilizam para este tipo de dados. Consideramos o ponto 2.3 inoportuno quando separado do ponto 2.4, pois construir e interpretar diagramas de Venn pressupõe que os alunos identifiquem a reunião e a interseção de dois ou mais conjuntos. Quando os autores fazem em 2.2 e 2.4, pela primeira vez,

referência aos diagramas de Venn e de Carroll, coloca-se a questão de saber a que é que se estavam a referir em 1.5, no 1º ano, quando pediam para classificar objectos segundo um ou dois critérios.

3º ano

Representação e tratamento de dados

1. Representar conjuntos de dados

1. Representar conjuntos de números naturais em diagramas de caule e folhas.

2. Tratar conjuntos de dados

1. Identificar a «frequência absoluta» de uma categoria de determinado conjunto de dados como o número de elementos da população que pertencem a essa categoria.
2. Identificar a «moda» de um conjunto de dados como a categoria com maior frequência absoluta.
3. Identificar o «máximo» e o «mínimo» de um conjunto de dados numéricos respetivamente como a categoria de maior e menor valor numérico, designando-os por «extremos».
4. Identificar a «amplitude» de um conjunto de dados numéricos como a diferença entre extremos.

3. Resolver problemas

1. Resolver problemas envolvendo a análise de dados representados em tabelas, diagramas ou gráficos e a determinação de frequências absolutas, moda, extremos e amplitude.
2. Resolver problemas envolvendo a organização de dados por categorias e a respetiva representação de uma forma adequada.

Comentário – Quando no ponto 1.1 os autores se referem a representar números naturais em diagramas de *caule-e-folhas* (e não *caule e folhas*, de acordo com Paulino, C. D., Pestana, D., Branco, J., Singer, J., Barroso, L. e Bussab, W. (2011) Glossário Inglês-Português de Estatística. 2ª ed. Sociedade Portuguesa de Estatística e Associação Brasileira de Estatística) estão a pensar em dados? Como é referido na citação com que se inicia este texto, dados são mais do que números.... Pode-se sugerir que a variável a estudar seja de natureza quantitativa discreta, o que satisfaz os objectivos dos autores. Fica ainda a questão de saber se os autores ao pedirem para representar conjuntos de números naturais em diagramas de caule-e-folhas, admitem que se possa considerar como conjunto de números naturais $\{0,0,0,1,1,2,3,4,4\}$? Quando em 2.1 se referem a *categoria* de um conjunto de dados, pensamos estarem a referir-se a *categoria* para os dados qualitativos e *classe* para os dados quantitativos discretos. Também neste ponto admitimos que seja por lapso que referem para identificar a “frequência absoluta”, o número de elementos da população, em vez de “o número de elementos do conjunto de dados”. No ponto 2.2 ao identificar a moda como a categoria de maior frequência, mais uma vez se questiona se os autores só se estão a referir a dados qualitativos, caso em que o conceito de moda apresentado está correcto (se se ressaltar o facto de poder haver mais do que uma moda se existirem 2 ou mais categorias com a mesma frequência). O mesmo não se passa se os dados forem de natureza quantitativa (veja-se GRAÇA MARTINS ET AL (2007), página 125). O ponto 2.3, por um lado, além de não estar correcto, pois se os dados são quantitativos (ou numéricos), não se organizam em categorias mas sim em classes, por outro lado, para determinar o mínimo e o máximo de um conjunto de dados, não é necessário começar por organizá-los em classes. No ponto 2.4 a identificação da amplitude de um conjunto de dados como a diferença entre extremos, não está correcta! Se pretenderem utilizar a linguagem dos *extremos* (estatísticas muito úteis para posteriormente se construírem os diagramas de extremos e quartis...) então terão de dizer

que a amplitude é a diferença entre o extremo superior e o extremos inferior. Pensamos que as coisas ficam bem mais simples se se identificar a amplitude de um conjunto de dados quantitativos (ou numéricos) como a diferença entre o máximo e o mínimo! Como nota final, abrangendo todas as alíneas do ponto 2, não se compreende porque é que não se pode exigir ao aluno a definição de frequência absoluta, frequência relativa, extremos ou amplitude! Não se compreende mesmo qual a diferença entre Identificação e Definição para aqueles conceitos tão simples.

4º ano

Tratamento de dados

1. Utilizar frequências relativas e percentagens

1. Identificar a «frequência relativa» de uma categoria de determinado conjunto de dados com a fração cujo numerador é a frequência absoluta dessa categoria e o denominador é o número total de dados.
2. Exprimir qualquer fração própria em percentagem arredondada às décimas.

2. Resolver problemas

1. Resolver problemas envolvendo o cálculo e a comparação de frequências relativas.

Comentário – Relativamente ao ponto 1.1, porquê identificar frequência relativa em vez de definir frequência relativa? Por outro lado, substituiríamos “categoria” por “categoria ou classe”.

2º ciclo

No domínio da Organização e Tratamento de Dados, retomam-se várias representações de conjuntos de dados e noções estatísticas elementares como a média, a moda, os extremos e a amplitude. Atendendo ao programa, é o momento ideal para se introduzir a noção de gráfico de pontos de uma correspondência, que será naturalmente revisitada com mais profundidade no 3.º ciclo no contexto das funções.

Comentário geral – Não se compreende o que é que os autores entendem e pretendem por “gráfico de pontos de uma correspondência”. Ao consultar o programa, deduzimos que eventualmente se estejam a referir a gráficos de linhas, que são gráficos de dispersão (*scatterplot*), utilizados para representar dados bivariados, mas numa situação especial em que uma das variáveis (a que se representa no eixo das abcissas) é o tempo ou representa uma ordenação temporal.

5º ano

Representação e tratamento de dados

1. Organizar e representar dados

1. Construir tabelas de frequências absolutas e relativas reconhecendo que a soma das frequências relativas das diversas categorias de determinado conjunto de dados é igual a 1.
2. Representar um conjunto de dados em gráfico de barras.
3. Identificar um «referencial cartesiano» como um par de retas numéricas não coincidentes que se intersectam nas respetivas origens, das quais uma é fixada como «eixo das abcissas» e a outra como «eixo das ordenadas» (os «eixos coordenados»), designar o referencial cartesiano como «ortogonal» quando os eixos são perpendiculares e por «monométrico» quando a unidade de comprimento é a mesma para ambos os eixos.
4. Identificar, dado um plano munido de um referencial cartesiano, a «abscissa» (respetivamente «ordenada») de um ponto P do plano como o número representado pela interseção com o eixo das abcissas (respetivamente ordenadas) com a reta paralela ao eixo das ordenadas (respetivamente abcissas) que passa por P e designar a abscissa e a ordenada por «coordenadas» de P.
5. Construir, num plano munido de um referencial cartesiano ortogonal, o «gráfico de pontos» referente a dois conjuntos de números tais que a todo o elemento do primeiro está associado um único elemento do segundo, representando nesse plano os pontos cujas abcissas são iguais aos valores do primeiro conjunto e as ordenadas respetivamente iguais aos valores associados às abcissas no segundo conjunto.
6. Identificar um «gráfico de linha» como o que resulta de se unirem, por segmentos de reta, os pontos de abcissas consecutivas de um gráfico de pontos.
7. Resolver problemas envolvendo a análise de dados representados em tabelas de frequência, diagramas de caule-e-folhas ou gráficos de barras e de linha.

2. Tratar conjuntos de dados

1. Identificar a «média aritmética» de um conjunto de dados numéricos como o quociente entre a soma dos respetivos valores (repetindo cada parcela um número de vezes igual à frequência absoluta da categoria em causa) e o número de dados, e representá-la por « \bar{x} ».

3. Resolver problemas

1. Resolver problemas envolvendo a média aritmética e a moda de um conjunto de dados, interpretando o respetivo significado no contexto de cada situação.

Comentário – No ponto 1.1 substituir “categorias” por “categorias ou classes”. Relativamente a este ponto, colocamos ainda a seguinte questão: se no 1º ciclo (4º ano) o aluno já sabe construir tabelas de frequências absolutas e já aprendeu a calcular a frequência relativa, porque não apresenta logo as tabelas de frequências relativas? Por outro lado porque é que se pede para reconhecer que a soma das frequências relativas é igual a 1 e não se pede para reconhecer que a soma das frequências absolutas é igual ao número de elementos do conjunto de dados em análise? Os pontos 3, 4 e 5 não têm nada a ver com Organização e tratamento de dados. Quando, em Matemática, se está a representar pontos de R^2 num sistema de eixos cartesianos, não se está a representar “dados”! A situação é exactamente a oposta! Quando temos dados quantitativos bivariados utilizamos uma representação em sistema de eixos cartesianos para nos ajudar a fazer a sua leitura e interpretação.

As ferramentas da Matemática são usadas na Estatística, mas não fazem parte da Estatística e devem, em nossa opinião, ser enquadradas num tema apropriado e não em OTD. No ponto 1.6 não está correcta a definição de “gráfico de linha”, pois como também se referiu no Comentário geral, o gráfico de linhas é um caso especial de um gráfico de dispersão em que os dados representados são os de coordenadas (t, y) em que t representa o tempo. No ponto 2.1 quando se fala em “média aritmética”, deve referir-se unicamente “média”, que é o termo utilizado em Estatística (a média em Estatística é a média aritmética em Matemática...). Por outro lado quando os autores escrevem “...soma dos respectivos valores (repetindo cada parcela um número de vezes igual à frequência absoluta da categoria em causa) não se compreende o sentido do que está entre parenteses! Qual é a categoria que está em causa? Quando se fala em soma dos respectivos valores, não se está a referir a todos os valores do conjunto de dados? Em 3.1 substituir média aritmética por média.

6º ano

Representação e tratamento de dados

1. Organizar e representar dados

1. Identificar «população estatística» ou simplesmente «população» como um conjunto de elementos, designados por «unidades estatísticas», relativamente a cada um dos quais são feitas observações e recolhidos dados.
2. Identificar «variável estatística» como uma característica que admite diferentes valores e que permite, mediante a recolha de dados (um por cada unidade estatística), dividir uma dada população em subconjuntos, designados por «categorias de dados» ou simplesmente por «categorias», de modo que cada unidade estatística pertença a uma e uma só categoria, e cada categoria é constituída exactamente pelas unidades estatísticas que, de acordo com os dados recolhidos, têm um determinado valor da característica em estudo («valor (numérico)» da categoria, no caso de ser um número ou «modalidade» da categoria, no caso de não o ser).
3. Designar Identificar «população estatística» ou simplesmente «população» como uma variável estatística por «quantitativa» ou «numérica» quando os respetivos valores são números e por «qualitativa» no caso contrário e, por abuso de linguagem, as categorias de dados qualitativos pela respetiva modalidade e as categorias de dados quantitativos pelo respetivo valor.
4. Reconhecer que a frequência absoluta de uma categoria é igual ao respetivo número de elementos e que a frequência relativa é igual ao quociente entre a frequência absoluta da categoria e o número total de unidades estatísticas da população («dimensão da população»).
5. Representar um conjunto de dados num «gráfico circular» dividindo um círculo em setores circulares sucessivamente adjacentes, associados respetivamente às diferentes categorias de

dados, de modo que as amplitudes dos setores sejam diretamente proporcionais às frequências relativas das categorias correspondentes.

6. Efetuar conversões entre diferentes representações do mesmo conjunto de dados.

7. Resolver problemas envolvendo a análise de dados representados de diferentes formas.

2. Resolver problemas

1. Resolver problemas envolvendo a análise de um conjunto de dados a partir da respetiva média, moda, extremos e amplitude.

Comentário – Os autores tentam introduzir neste ano o conceito de população. A linguagem utilizada nos pontos 1.1, 1.2, 1.3 e 1.4 é demasiado confusa para se ficar a perceber afinal o que é que se entende por população. O objectivo parece ser o de resumir a organização e tratamento de dados, ao nível do 6º ano de escolaridade, a situações em que a amostra coincide com a população mas consideramos que tal poderá trazer dificuldades, no 8º ano, quando as colecções de dados se referirem a uma amostra.. No ponto 1.3 as categorias dos dados numéricos são os próprios valores... Então porque é que lhes chamamos categorias em vez de simplesmente dados? Este ponto vai criar alguns conflitos quando se tentar organizar dados de natureza contínua..., ou mesmo de natureza discreta, mas com um número elevado de valores distintos, tal que seja necessário organizá-los em classes na forma de intervalos. Qual a vantagem de se voltar aqui (ponto 1.4) ao conceito de frequência absoluta que já havia sido introduzido no 3º ano? No ponto 1.5 quando os autores falam em representar um conjunto de dados num “gráfico circular” é pressuposto que se pode representar qualquer conjunto de dados? Mesmo para dados de natureza quantitativa? E de natureza quantitativa contínua? O ponto 1.6 é um apelo à manipulação de gráficos! Seria mais interessante pedir para representar o mesmo conjunto de dados utilizando várias representações gráficas (quando possível) e escolher as mais adequadas ou elucidativas a transmitir a informação. Quando chegamos ao fim do 2º ciclo, verifica-se que os autores introduziram (tentaram) o conceito de população. Não se percebe porquê e com que objectivo. Não se fala em amostra e continua-se a falar em conjunto de dados, sem se entender se se está a pensar em todos os dados resultantes de se terem observado todos os elementos da população ou “unidades estatísticas”, relativamente a determinada característica de interesse.

3º ciclo

Finalmente, no domínio Organização e Tratamento de Dados, são introduzidas algumas medidas de localização e dispersão de um conjunto de dados e é feita uma iniciação às probabilidades e aos fenómenos aleatórios.

Comentário geral – Como já foi dito no início deste texto, **discordamos de só agora os alunos iniciarem o contacto com as situações aleatórias**, assim como com a utilização do termo probabilidade, ainda que de forma elementar e intuitiva.

7º ano

Diagramas de extremos e quartis

1. Representar, tratar e analisar conjuntos de dados

1. Construir, considerado um conjunto de dados numéricos, uma sequência crescente em sentido lato cujos valores sejam os valores das categorias, repetidos um número de vezes igual à respetiva frequência absoluta, designando-a por «sequência ordenada dos dados».
2. Identificar, dado um conjunto de n dados numéricos, a «mediana» como o valor central no caso de n ser ímpar (valor do elemento de ordem $\frac{n+1}{2}$ da sequência ordenada dos dados), ou como a média aritmética dos dois valores centrais (valores dos elementos de ordens $\frac{n}{2}$ e $\frac{n}{2} + 1$ da sequência ordenada dos dados) no caso de n ser par e representar a mediana por « \tilde{x} » ou «Me».
3. Determinar a mediana de um conjunto de dados numéricos.
4. Identificar, dado um conjunto de n dados numéricos (sendo n ímpar), o «primeiro quartil» (respetivamente «terceiro quartil») como a mediana do subconjunto de dados de ordem inferior (respetivamente superior) a $\frac{n+1}{2}$ na sequência ordenada do conjunto inicial de dados.
5. Identificar, dado um conjunto de n dados numéricos (sendo n par), o «primeiro quartil» (respetivamente «terceiro quartil») como a mediana do subconjunto de dados de ordem inferior ou igual a $\frac{n}{2}$ (respetivamente superior ou igual a $\frac{n}{2} + 1$) na sequência ordenada do conjunto inicial de dados.
6. Identificar, dado um conjunto de dados numéricos, o «segundo quartil» como a mediana desse conjunto e representar os primeiro, segundo e terceiro quartis respetivamente por Q_1 , Q_2 e Q_3 .
7. Reconhecer, dado um conjunto de dados numéricos, que pelo menos um quarto dos dados têm valores não superiores ao primeiro quartil e que pelo menos três quartos dos dados têm valores não superiores ao terceiro quartil.
8. Representar conjuntos de dados quantitativos em diagramas de extremos e quartis.
9. Designar por «medidas de localização» a média, a moda, a mediana e os quartis de um conjunto de dados.
10. Designar um conjunto de dados como «simétrico» quando a média, a moda e a mediana são iguais.
11. Identificar a «amplitude interquartis» como a diferença entre o 3.º quartil e o 1.º quartil ($Q_3 - Q_1$) e designar por «medidas de dispersão» a amplitude e a amplitude interquartis.

2. Resolver problemas

1. Resolver problemas envolvendo a análise de dados representados em tabelas de frequência, gráficos de barras e em diagramas de extremos e quartis.

Comentário – No ponto 1.1, pressupõe-se que os dados estejam organizados em classes para se proceder à sua ordenação! Não vemos essa necessidade... Se os autores entenderem que tem de ser assim, então seria bom aconselhar o caule-e-folhas, que é uma representação que tem a vantagem de facilitar a ordenação. Por outro lado sugerimos que se chame à “sequência ordenada dos dados” proposta pelos autores, unicamente “dados ordenados”. No ponto 2.7, na mesma linha do que se faz para o 1º e 3º quartis, deveria acrescentar-se: reconhecer que, dado um conjunto de dados quantitativos ou numéricos, pelo menos metade dos dados têm valores não superiores (menores ou iguais) à mediana. No ponto 1.10, não é correcto falar em dados simétricos. Falamos em distribuições simétricas, quando o aspecto apresentado pela distribuição de frequências quando representada na forma de um gráfico (ou diagrama) de caule-e-folhas ou de barras é aproximadamente simétrico, situação em que as três características amostrais consideradas são aproximadamente iguais. No ponto 1.11 onde está amplitude interquartis deverá estar amplitude interquartil (Paulino, C. D., Pestana, D., Branco, J., Singer, J., Barroso, L. e Bussab, W. (2011) Glossário Inglês-Português de Estatística. 2ª ed. Sociedade Portuguesa de Estatística e Associação Brasileira de Estatística).

8º ano

Planeamento de um estudo estatístico

1. Identificar algumas fases do planeamento de um estudo estatístico

1. Designar por «amostra» um subconjunto de uma população na qual estão definidas uma ou mais variáveis estatísticas e por «dimensão da amostra» o número de unidades estatísticas pertencentes à amostra.
2. Saber que existem critérios para obtenção de uma amostra de forma que as medidas de localização e outras medidas estatísticas calculadas utilizando os dados da amostra sejam estimativas consideradas adequadas das correspondentes medidas da população e designar por «representativa» uma amostra que cumpre esses critérios e por «enviesada» no caso contrário.
3. Identificar alguns métodos de recolha de dados.

Comentário – No ponto 1.1 a introdução do conceito de amostra parece-nos demasiado confusa, o que naturalmente resulta do facto de a própria definição de população já ser confusa... . No ponto 1.2, quando se pretende definir uma amostra representativa como aquela que conduz a que as estatísticas calculadas utilizando os seus dados sejam estimativas adequadas das correspondentes medidas da população, coloca-se uma questão bem mais complicada, que é a de saber o que são estimativas adequadas! Veja-se a definição “*A sample that represents the characteristics of the populations as closely as possible is called a representative sample*” (MANN (1995), página 8) . Referir a seguir, que se tivermos uma amostra representativa ela pode ser utilizada para calcular medidas que servem para estimar as correspondentes medidas da população.

9º ano

Histogramas

1. Organizar e representar dados em histogramas

1. Estender a noção de variável estatística quantitativa ao caso em que se associa a cada categoria um intervalo de números fechado à esquerda e aberto à direita, sendo esses intervalos disjuntos dois a dois e de união igual a um intervalo (e estender também ao caso em que se intersesta cada um desses intervalos com um conjunto finito pré-determinado de números), designando o intervalo que se associa a cada categoria por «classe dos dados» dessa categoria.
2. Identificar uma variável estatística quantitativa como «discreta» quando se associa a cada categoria apenas um número ou um conjunto finito de números e «contínua» quando se associa a cada categoria apenas um intervalo.
3. Reagrupar as unidades de uma população em categorias com base num conjunto de dados numéricos de modo que as classes tenham uma mesma amplitude pré-fixada e designar este processo por «agrupar os dados em classes da mesma amplitude».
4. Identificar, dado um conjunto de dados agrupados em classes da mesma amplitude, «histograma» como um gráfico de barras retangulares justapostas de bases iguais e tais que a altura dos retângulos é diretamente proporcional à frequência absoluta (e portanto também à frequência relativa) de cada classe.
5. Representar, em histogramas, conjuntos de dados agrupados em classes da mesma amplitude.

2. Resolver problemas

1. Resolver problemas envolvendo a representação de dados em tabelas de frequência, diagramas de caule-e-folhas e histogramas.

Probabilidade

3. Utilizar corretamente a linguagem da probabilidade

1. Identificar uma «experiência» como um processo que conduz a um resultado pertencente a um conjunto previamente fixado designado por «universo dos resultados» ou «espaço amostral», designar os elementos do espaço amostral por «casos possíveis» e a experiência por «determinista» quando existe um único caso possível e «aleatória» em caso contrário.
2. Designar por «acontecimento» qualquer subconjunto do universo dos resultados de uma experiência aleatória e os elementos de um acontecimento por «casos favoráveis» a esse acontecimento e utilizar a expressão «o acontecimento A ocorre» para significar que o resultado da experiência aleatória pertence ao conjunto A.
3. Designar, dada uma experiência aleatória, o conjunto vazio por acontecimento «impossível», o universo dos resultados por acontecimento «certo», um acontecimento por «elementar» se existir apenas um caso que lhe seja favorável e por «composto» se existir mais do que um caso que lhe seja favorável.
4. Designar dois acontecimentos por «incompatíveis» quando forem disjuntos e por «complementares» quando forem disjuntos e a respetiva reunião for igual ao espaço amostral.
5. Descrever experiências aleatórias que possam ser repetidas mantendo um mesmo universo de resultados e construídas de modo a que se espere, num número significativo de repetições, que cada um dos casos possíveis ocorra aproximadamente com a mesma frequência e designar os acontecimentos elementares dessas experiências por «equiprováveis».
6. Designar, dada uma experiência aleatória cujos casos possíveis sejam em número finito e equiprováveis, a «probabilidade» de um acontecimento como o quociente entre o número de casos favoráveis a esse acontecimento e o número de casos possíveis, designar esta definição por «regra de Laplace» ou «definição de Laplace de probabilidade» e utilizar

corretamente os termos «mais provável», «igualmente provável», «possível», «impossível» e «certo» aplicados, neste contexto, a acontecimentos.

7. Reconhecer que a probabilidade de um acontecimento, de entre os que estão associados a uma experiência aleatória cujos casos possíveis sejam em número finito e equiprováveis, é um número entre 0 e 1 e, nesse contexto, que é igual a 1 a soma das probabilidades de acontecimentos complementares.
8. Justificar que se A e B forem acontecimentos disjuntos se tem $P(A \cup B) = P(A) + P(B)$.
9. Identificar e dar exemplos de acontecimentos possíveis, impossíveis, elementares, compostos, complementares, incompatíveis e associados a uma dada experiência aleatória.
10. Utilizar tabelas de dupla entrada e árvores de probabilidade na resolução de problemas envolvendo a noção de probabilidade e a comparação das probabilidades de diferentes acontecimentos compostos.
11. Realizar experiências envolvendo a comparação das frequências relativas com as respetivas probabilidades de acontecimentos em experiências repetíveis (aleatórias), em casos em que se presume equiprobabilidade dos casos possíveis.

Comentário – Os pontos 1.1, 1.2 e 1.3 estão pouco claros e mal se percebe que o que os autores pretendem é classificar os dados em qualitativos ou quantitativos e estes em discretos ou contínuos. Esta classificação é muito simples e basta consultar GRAÇA MARTINS ET AL (2007), página 14. Aliás esta classificação dos dados já deveria ter surgido mais cedo, pois alguma da confusão provocada por alguma terminologia incorrectamente utilizada deriva desta falha! A definição de histograma dada no ponto 1.4 não está correcta, pois o histograma é um diagrama de áreas e as áreas dos rectângulos é que são proporcionais às frequências. Obviamente que se os intervalos de classe tiverem todos a mesma amplitude, as alturas dos rectângulos vêm proporcionais às frequências. A definição dada em 3.1 de experiência aleatória não está correcta, já que é necessário acrescentar que em cada realização da experiência não se tem conhecimento suficiente de qual o resultado que se vai verificar (admitindo que o espaço de resultados tem mais de um resultado) . No ponto 3.4 os autores definem acontecimentos “incompatíveis” quando forem disjuntos, mas acontecimento disjunto é sinónimo de acontecimento incompatível. Parece-nos fundamental definir primeiro o que são conjuntos ou acontecimentos disjuntos. Mais uma vez, a nossa sugestão é de que todas as noções e definições da Teoria de Conjuntos sejam apresentadas num tópico específico, e que se apresente a conversão de linguagem, no âmbito da OTD e aqui, mais concretamente, no âmbito da Probabilidade. (Em TIAGO DE OLIVEIRA (1990), página 51, lê-se “A e B dizem-se incompatíveis se for $P(A \cap B) = 0$... Vê-se pois que a noção (absoluta) de acontecimentos disjuntos implica a noção (relativa a P) de acontecimentos incompatíveis; a incompatibilidade depende da função P em consideração”. Nesta definição de incompatibilidade, acontecimentos incompatíveis não são o mesmo que acontecimentos disjuntos, já que acontecimentos disjuntos são incompatíveis, mas acontecimentos incompatíveis não são necessariamente disjuntos. Não adoptaremos esta definição, pois em toda a bibliografia consultada, de que destacamos em português, MURTEIRA ET AL (2002), página 58 e PESTANA E VELOSA (2010), página 238, a incompatibilidade aparece como sinónimo de acontecimentos disjuntos ou mutuamente exclusivos). Em 3.5 discordamos em absoluto da necessidade de que “cada um dos casos possíveis ocorra aproximadamente com a mesma frequência...”. Não se admitem experiências realizadas com moedas viciadas ou dados truncados, para já não falar da maioria das situações decorrentes do dia a dia?! De acordo com o preconizado no ponto 3.6, os autores pensam que não se pode utilizar os termos «mais provável», «igualmente provável», «possível», «impossível» e «certo» sem se conhecer a regra de Laplace. Ficamos restritos a um mundo muito limitado onde quase nunca poderemos utilizar tais termos.

Bibliografia

GRAÇA MARTINS, M. E., LOURA, L., MENDES, F. (2007) – Análise de dados, Texto de apoio para os professores do 1º ciclo, Ministério da Educação, DGIDC. ISBN-978-972-742-261-6. Depósito legal 262674/07

MANN, P. S. (1995) – Introductory Statistics, 2nd edition. John Wiley & Sons, Inc. ISBN 0-471-31009-3

MURTEIRA, B., RIBEIRO, C. S., SILVA, J. A., PIMENTA, C. (2002) – *Introdução à Estatística*. McGraw-Hill de Portugal, Lda. ISBN: 972-773-116-3.

PESTANA, D., VELOSA, S. (2010) – *Introdução à Probabilidade e à Estatística*, Volume I, 4ª edição, Fundação Calouste Gulbenkian. ISBN 978-972-31-1150-7. Depósito Legal 311132/10.

TIAGO DE OLIVEIRA, J. (1990) – *Probabilidades e Estatística* – Conceitos, Métodos e Aplicações. Volume I. McGraw-Hill. Lisboa