



Escolha e aplicação de ferramentas estatísticas gráficas e numéricas por futuros professores dos primeiros anos

José António Fernandes ^a
 Adelaide Freitas ^b

^a Universidade do Minho, Instituto de Educação, Braga, Portugal

^b Universidade de Aveiro, Departamento de Matemática, Aveiro, Portugal

Recebido para publicação em 29 jul. 2019. Aceito, após revisão, em 21 out. 2019.

Editor designado: Claudia Lisete Oliveira Groenwald

RESUMO

O ensino da Estatística nos primeiros anos de escolaridade requer que os professores desse nível escolar desenvolvam competências de análise de pequenas coleções de dados. Neste artigo, estudamos a habilidade de, perante uma coleção de 12 observações de tipo quantitativo, estudantes, futuros professores dos primeiros anos escolares, selecionarem e construirão gráficos adequados e identificarem e determinarem medidas estatísticas convenientes à sumarização dos dados, incluindo a interpretação do 3.º quartil. No estudo participaram 50 estudantes que se encontravam a frequentar o 2.º ano da Licenciatura em Educação Básica numa universidade do norte de Portugal. Os dados recolhidos correspondem às respostas dadas pelos estudantes numa prova de avaliação formal da disciplina de Probabilidades e Estatística. Em termos de resultados, uma análise das respostas permitiu verificar que os estudantes revelam dificuldades tanto ao nível da escolha como da aplicação dos métodos estatísticos, mais acentuadas quando se tratava de identificar os gráficos adequados para representar os dados, e de determinar os quartis e interpretar o 3.º quartil, e menos acentuadas no caso da determinação de outras medidas estatísticas.

Palavras-chave: gráficos estatísticos; medidas estatísticas; futuros professores dos primeiros anos.

Selecting and applying graphical and numerical statistical tools by prospective primary school teachers

ABSTRACT

Teaching statistics in the early years requires that teachers at this school level develop skills in analyzing small collections of data. Given a collection of quantitative data (12 observations), in this paper, we study the ability of students, prospective primary school teachers, to select and construct appropriate graphs and to identify and determine statistical measures suitable for summarizing the data, including the interpretation of the third quartile. The study involved 50 students who were attending the 2nd year of the Bachelor in Basic Education at a university in the north of Portugal. The collected data correspond to the responses given by the students in a formal evaluation test

Autor correspondente: Adelaide Freitas. Email: adelaide@ua.pt

of a course of Probability and Statistics. In terms of results, an analysis of the responses revealed that the students showed difficulties in both the choice and the application of statistical methods, which were more pronounced when it came to identifying the appropriate graphs to represent the data and to determine the quartiles and to interpret the third quartile, and less pronounced in the case of determining other statistical measures.

Keywords: statistical graphs; statistical measures; prospective primary school teachers.

INTRODUÇÃO

Atualmente vem-se assistindo a um crescendo da utilização da Estatística nos mais variados setores da sociedade e a nível individual e profissional, o que explica o aprofundamento do seu ensino nas escolas. Em Portugal, conteúdos de Probabilidades e Estatística, integrados no tema Organização e Tratamento de Dados, incluem-se nos programas da disciplina de Matemática de todos os anos escolares, tanto ao nível do ensino básico (Ministério da Educação e Ciência, 2013) como do ensino secundário (Ministério da Educação e Ciência, 2014).

Ora, em consonância com os programas escolares, o aprofundamento do estudo de conteúdos de Probabilidades e Estatística na escola requer que os professores também adquiram formação matemática que lhes permita implementar um ensino dessas temáticas adequado às necessidades dos alunos e do sistema de ensino. Em particular, para professores dos primeiros anos, importa que eles adquiram, na sua formação, conhecimentos básicos de Estatística Descritiva para o ensino e competências essenciais para a realização ou criação de atividades investigativas simples envolvendo ferramentas estatísticas clássicas, gráficas e numéricas, de sumarização de dados qualitativos e quantitativos (de natureza discreta ou contínua).

Cerca de uma década atrás, Batanero (2009) alertava para o facto da componente de formação para ensinar Probabilidades e Estatística não ser geralmente incluída na formação inicial de todos os professores. No caso dos professores dos primeiros anos, possivelmente, esta questão é ainda mais grave uma vez que só muito recentemente, concretamente no início do século (Ministério da Educação, 2007), esses conteúdos foram incluídos nos programas escolares e, em consequência, na formação inicial desses professores.

Assim, no âmbito da formação inicial dos futuros professores dos primeiros anos, neste artigo estuda-se o conhecimento estatístico e as habilidades destes futuros professores em Estatística Descritiva. Mais especificamente, avalia-se a adequação dos métodos estatísticos exploratórios escolhidos pelos futuros professores para uma análise de dados de tipo quantitativo, envolvendo gráficos e medidas estatísticas, bem como a correta implementação desses métodos.

A correta escolha de métodos exploratórios a usar numa análise preliminar de dados é da maior importância pois, por um lado, constitui a primeira etapa que deve ser superada na exploração de tarefas mais abertas, como acontece com os projetos investigativos em Estatística, e, por outro lado, está subjacente à compreensão de estratégias de resolução

de exercícios e problemas, mesmo simples, os quais podem conduzir ao aprofundamento de conceitos ou de relações entre conceitos. Num estudo anterior, Fernandes, Batanero e Gea (2019) estudaram a adequação dos métodos estatísticos escolhidos por futuros professores dos primeiros anos para analisar um conjunto de dados relativos a uma variável qualitativa (nominal), tendo concluído que é na seleção dos métodos estatísticos onde são sentidas mais dificuldades.

Em continuação do estudo referido, numa perspetiva de complementaridade, no presente estudo analisa-se o conhecimento e competência de futuros professores dos primeiros anos na seleção de métodos estatísticos (gráficos e medidas) para analisar um conjunto de dados relativos a uma variável quantitativa (de natureza contínua). A questão do tipo de variável estatística, isto é, tratar-se de uma variável qualitativa (nominal ou ordinal) ou de uma variável quantitativa (discreta ou contínua), é da maior relevância num estudo estatístico pois condiciona os métodos estatísticos que podem ser aplicados.

Nas próximas secções apresentamos o marco teórico e antecedentes do estudo, onde se enquadra o estudo e se relatam e discutem alguns estudos relacionados com a problemática do presente estudo; a metodologia de investigação, em que se explicita o tipo de estudo, os participantes e os métodos de recolha e análise de dados; a apresentação de resultados segundo os conteúdos gráficos e medidas estatísticas; e, por último, sintetizam-se e discutem-se as conclusões do estudo.

MARCO TEÓRICO E ANTECEDENTES DO ESTUDO

Na investigação sobre o conhecimento que um professor deve ter para tratar com êxito o ensino da Matemática, diversos quadros de referência têm sido disseminados. Desses modelos teóricos mais referenciados na literatura (Gonzalez & Eudave, 2018), neste trabalho optámos como marco teórico por considerar o Modelo do Conhecimento Matemático para o Ensino (*Mathematical Knowledge for Teaching*, MKT), amplamente referenciado e originalmente proposto por Ball (2000). Este modelo, que pretende caracterizar o conhecimento matemático que um professor necessita possuir para o desenvolvimento da sua prática docente, e assim favorecer a aprendizagem da Matemática dos seus alunos, classifica o conhecimento matemático dos professores em duas categorias: Conhecimento do Conteúdo e Conhecimento Pedagógico do Conteúdo, e é baseado no modelo de Shulman (1986), o qual classifica o conhecimento do professor em três categorias: Conhecimento do conteúdo, Conhecimento Pedagógico do conteúdo e Conhecimento do currículo. No Conhecimento do Conteúdo está implicado o conhecimento matemático e a transformação desse conhecimento de modo a torná-lo acessível aos alunos, enquanto no Conhecimento Pedagógico do Conteúdo se estabelecem relações entre o conhecimento matemático e vários fatores, como sejam os estudantes, o ensino e o currículo.

Uma vez que o presente trabalho incide sobre o conhecimento matemático de futuros professores, restringimos a investigação à categoria do Conhecimento do Conteúdo, a qual se compõe das três subcategorias seguintes (Hill, Ball & Schilling, 2008):

— Conhecimento Comum do Conteúdo (*Common Content Knowledge, CCK*), que consiste no conhecimento que um adulto mobiliza para resolver problemas matemáticos;

— Conhecimento Especializado do Conteúdo (*Specialized Content Knowledge, SCK*), que diz respeito ao conhecimento especial do professor que o habilita a planificar e desenvolver sequências de ensino;

— Conhecimento Matemático Avançado (*Knowledge at the mathematical horizon*), que se refere ao conhecimento da sequencialização dos tópicos matemáticos no currículo, devendo o professor relacionar o que os seus alunos estão a aprender com o que aprenderão futuramente.

Destas três subcategorias, neste estudo destacam-se o Conhecimento Comum do Conteúdo e o Conhecimento Matemático Avançado, já que no estudo realizado não se salientam aspetos do Conhecimento Especializado do Conteúdo.

Tratando-se de futuros professores dos primeiros anos, é de suma importância que estes adquiram na sua formação competências para decidir quais os métodos a usar em análises estatísticas (simples) de dados, como surgem, por exemplo, nos projetos investigativos. Segundo Batanero, Díaz, Contreras e Arteaga (2011), atualmente, a exploração de tarefas abertas é muito valorizada em virtude de aumentar a motivação dos alunos e relevar o contexto e a sua natureza realística. Por outro lado, estas tarefas permitem aos alunos adquirir conhecimentos estratégicos, que se acrescentam aos conhecimentos técnicos exercitados em tarefas convencionais.

Também, na realização de um estudo estatístico, como sejam os projetos investigativos, que envolvem a exploração de tarefas mais abertas, podem delimitar-se várias etapas, tal como advogam Wild e Pfannkuch (1999) ao proporem o modelo PPDAC, que denominam de ciclo investigativo, no qual se distinguem cinco etapas: problema (P); plano (P); dados (D); análise (A) e conclusões (C). No presente estudo salientam-se, sobretudo, a segunda e quarta etapas, relativas à planificação e à implementação das tarefas a realizar com vista a resolver o problema ou dar resposta à questão de investigação definida na etapa anterior. Na segunda etapa devem ser decididos quais os métodos/procedimentos de análise estatística ou medidas estatísticas que são adequados para o estudo, enquanto na quarta etapa se trata de aplicar os métodos/procedimentos antes selecionados.

Ora, o tipo de variável estatística em estudo determina, em grande parte, os métodos estatísticos que podem ser usados na análise dos dados, seja em relação às frequências, às medidas estatísticas ou aos gráficos estatísticos, o que, portanto, tem implicações nas etapas de planificação e análise. Relativamente às frequências, é essencial para o seu ensino que os futuros professores saibam relacionar o conceito com o procedimento. Enquanto as

frequências absolutas e relativas podem ser determinadas em qualquer tipo de variável, já as frequências absolutas e relativas acumuladas não podem ser determinadas em variáveis estatísticas nominais, pois essas frequências requerem o estabelecimento de uma relação de ordem. Desse modo, as frequências acumuladas só podem ser determinadas no caso de uma variável qualitativa ordinal ou quantitativa. Contudo, é conhecido que na impossibilidade de aplicar alguns métodos estatísticos em certas situações, porque requerem dados numéricos, os alunos recorrem incorretamente às frequências de modo a ultrapassar o problema dos valores das variáveis estatísticas qualitativas (Fernandes & Barros, 2005; Fernandes, Carvalho & Ribeiro, 2007; Fernandes, Carvalho & Correia, 2011).

Também a seleção do tipo de gráfico a usar na representação dos dados é muito importante pois pode influenciar o sucesso das outras etapas que se seguem no PPDAC. No caso dos *media*, é expectável que qualquer indivíduo, com formação matemática, interprete os diversos tipos de gráficos; já a sua construção ou decidir qual o mais adequado face a um conjunto de dados (qualitativos ou quantitativos, discretos ou contínuos, poucos ou muitos dados) requer um conhecimento mais profundo e nem sempre fácil de perceber. Por exemplo, Morais e Fernandes (2011) e Fernandes, Morais e Lacaz (2011) constataram que alunos do 9.º ano de escolaridade revelam dificuldades na escolha dos gráficos adequados para a representação dos dados. Nos três itens propostos aos alunos, verificou-se um melhor desempenho naquele que envolvia a representação gráfica de uma variável quantitativa discreta (idade), seguindo-se aquele em que se pedia a representação gráfica para comparar os valores da variável anterior segundo as categorias género masculino e feminino e, por último, um desempenho muito fraco no item em que se requeria a representação gráfica de uma variável quantitativa contínua.

Nos gráficos construídos pelos alunos destacou-se claramente a utilização do gráfico de barras simples. Quando este gráfico se revelou adequado para representar a variável, verificou-se uma elevada percentagem de respostas corretas ou parcialmente corretas; quando este gráfico se revelou não apropriado para representar a variável, verificou-se uma redução considerável das respostas corretas e parcialmente corretas. Essa redução foi particularmente acentuada no caso do item que envolvia a construção de um histograma, no qual quase metade dos alunos não responderam e, à exceção de um aluno com resposta correta e outro com resposta parcialmente correta, todos os restantes construíram gráficos de barras, gráficos circulares, gráficos de linhas e gráficos cartesianos. Segundo Fernandes et al. (2011),

a prevalência do gráfico de barras simples pode ter sido consequência de terem sido mais trabalhados nas aulas destes alunos, além de que se trata de um gráfico de mais fácil construção do que outros tipos de gráficos, como, por exemplo, os gráficos de barras agrupados ou empilhados e os histogramas. (p. 11)

Para além da escolha de um gráfico não adequado para representar os dados, os gráficos construídos pelos alunos também apresentavam várias falhas que são documentadas na

literatura (e.g., Espinel, González, Bruno & Pinto, 2009; Ruiz, Arteaga & Batanero, 2009), como seja a ausência de título e de rótulos nos eixos, o estabelecimento de escalas não adequadas e a falta de rigor na construção do gráfico. Para Friel, Curcio e Bright (2001), esses aspetos, que eles designam por estrutura do gráfico, são importantes na medida em que nos dão informação sobre o tipo de medições que estão a ser utilizadas e os dados que estão a ser medidos. Saliente-se que a mudança de escalas pode distorcer a forma de um gráfico de barras e de um histograma, podendo conduzir a interpretações erróneas em indivíduos com menor formação matemática. Também na representação dos dados por meio de um diagrama de extremos e quartis, gráfico menos convencional, mas muito utilizado em estudos exploratórios em diversas áreas do conhecimento, são conhecidas dificuldades, mais acentuadas na sua interpretação (Carvalho, Fernandes & Freitas, 2019).

Analogamente às frequências e aos gráficos, também a escolha das medidas estatísticas para resumir os dados está dependente do tipo de variável estatística em estudo. Excetuando a moda, que é uma estatística que pode ser aplicada a qualquer tipo de variável, nenhuma das outras medidas de localização (mediana, quartis e média) e de dispersão (amplitude, variância e desvio padrão) se podem aplicar a todo o tipo de variável estatística. Contudo, como se mostra na literatura (e.g., Boaventura & Fernandes, 2004; Fernandes et al., 2007), na impossibilidade de determinar a mediana ou a média, por se tratar de uma variável qualitativa nominal ou de uma variável não quantitativa, respetivamente, os alunos recorrem às frequências para obter valores numéricos que lhes permita aplicar as fórmulas de cálculo de tais estatísticas.

No estudo de Fernandes et al. (2007), tendo a professora pedido a alunos do 7.º ano que calculassem a média, a moda e a mediana da variável cor dos olhos dos alunos da turma, verificou-se que os alunos se sentiram confusos e consideraram difícil o cálculo da mediana e da média, tendo um aluno sugerido: “Podemos passar isso a números: púnhamos azul 1, verde 2 e por aí fora” (p. 49). De seguida, tendo a professora recordado que no cálculo da mediana é necessário ordenar os dados, outro aluno acrescentou: “Só se ordenássemos por ordem alfabética”. Nestes extratos constata-se a persistência dos alunos para contornarem, erradamente, a inexistência de valores numéricos, que sabem ser necessários para a determinação das estatísticas.

Já no estudo de Boaventura e Fernandes (2004), dada a distribuição de uma variável qualitativa nominal (bebida preferida dos alunos), foi pedido a alunos do 12.º ano a determinação, caso fosse possível, da moda, mediana e média, tendo-se verificado que muitos alunos sentiram muitas dificuldades na resolução da tarefa, mais acentuadas no caso da mediana e da média. No caso da mediana e da média, os alunos recorreram às frequências dos valores da variável para determinarem essas estatísticas e, no caso da moda, muitos alunos indicaram a frequência em vez do valor da variável. Também Fernandes e Barros (2005) observaram resultados semelhantes, agora com futuros professores dos primeiros anos.

Num estudo mais recente, envolvendo a mesma variável qualitativa nominal (bebida preferida dos alunos), Fernandes, Batanero e Gea (2019) verificaram que cerca de metade dos estudantes, futuros professores dos primeiros anos, determinaram

frequências absolutas e relativas acumuladas, o que não é adequado, cerca de três em cada quatro estudantes construíram um gráfico de barras ou circular, o que é adequado, quase todos (82%) indicaram o valor da moda, o que é igualmente adequado, e também muitos (78%) determinaram o valor da média, mediana, desvio padrão, quartis, amplitude ou variância, o que manifestamente não são medidas estatísticas adequadas ao tipo de variável em estudo. Estas dificuldades foram mais acentuadas na escolha dos métodos estatísticos do que na aplicação desses métodos, especialmente no caso das frequências e das medidas estatísticas.

Na presente investigação estudamos em que medida futuros professores dos primeiros anos escolhem e aplicam métodos de análise estatística adequados a uma distribuição de dados envolvendo uma variável quantitativa (de natureza contínua), relativamente a gráficos e medidas estatísticas.

METODOLOGIA

Neste estudo, de natureza fundamentalmente descritiva, inspecionam-se as escolhas de métodos estatísticos, no que respeita a gráficos e medidas estatísticas, realizadas por estudantes, futuros professores dos primeiros anos, e posterior aplicação desses métodos na análise de uma tarefa estatística.

Participaram no estudo 50 estudantes (E_i , com $i = 1, 2, \dots, 50$) de uma universidade do norte de Portugal, que estavam a frequentar a disciplina de Probabilidades e Estatística, integrada no 2.º ano do curso de Licenciatura em Educação Básica. À entrada na universidade, estes estudantes tinham uma formação matemática muito variada, o que explica também perceções diversas sobre as suas dificuldades nas disciplinas de Matemática que tinham frequentado, até então, na universidade.

A recolha de dados foi efetuada através da aplicação de um teste de avaliação formal, que os estudantes realizaram depois de concluída a lecionação da disciplina de Probabilidades e Estatística. Esta unidade curricular está focada no conhecimento do conteúdo e nela não se exploram aspetos didáticos. Das várias tarefas que faziam parte do teste, estudamos aqui apenas uma (Figura 1), focada numa coleção de 12 observações de uma variável quantitativa de natureza contínua. Para a resolução desta tarefa, os estudantes puderam usar uma calculadora ou uma folha de cálculo de computador.

No quadro seguinte apresentam-se as classificações, numa escala de 0 a 100, obtidas num teste Matemática por 12 alunos de uma turma do 5.º ano.

55	60	84	45	48	64
90	65	55	58	70	50

- a) Que tipos de gráficos são adequados para representar os dados? Desses gráficos, escolher um e representar os dados.
- b) Caso seja possível, identificar a moda e determinar os valores da média, dos quartis e do desvio padrão. Em geral, qual o significado do terceiro quartil?

Figura 1. Enunciado da tarefa proposta aos estudantes

Em termos da análise e organização de dados, começamos por listar os gráficos estatísticos escolhidos e seguidamente identificam-se falhas nos gráficos construídos pelos estudantes, na questão a); e, depois, avaliamos a determinação das seis medidas estatísticas requeridas (moda, média, 1.º quartil, mediana, 3.º quartil e desvio padrão) segundo o tipo de resposta (correta e incorreta) e a interpretação dada ao terceiro quartil, na questão b). Quando pertinente, determinam-se frequências de estudantes segundo os diferentes tipos de resposta (correta e incorreta) e não resposta, recorre-se a tabelas para resumir a informação e apresentam-se exemplos de resoluções dos estudantes tendo em vista clarificar os seus raciocínios.

RESULTADOS E ANÁLISE

Nesta secção apresentam-se os resultados obtidos segundo os conteúdos contemplados em cada uma das questões da tarefa: a) gráficos; e b) medidas estatísticas.

Gráficos

Na questão a) questionam-se os estudantes sobre os gráficos que são adequados para representar os dados fornecidos e, de seguida, pede-se para construírem um desses gráficos para representar os dados. Tendo em conta a natureza contínua dos dados e as aprendizagens dos estudantes, os gráficos mais adequados seriam um histograma ou um diagrama de extremos e quartis, e dado o número reduzido de observações (12), um diagrama de caule-e-folhas. No contexto do problema, os estudantes também poderiam efetuar uma agregação dos dados em classes/categorias e, deste modo, selecionar gráficos mais adequados a variáveis qualitativas, nomeadamente um gráfico de barras ou um gráfico circular. Embora tenham sido assinalados estes dois últimos gráficos, os estudantes não indicaram qualquer tipo de categorização prévia dos dados. Para além destes gráficos, os estudantes indicaram e/ou construíram outros gráficos não adequados, como se constata pela Tabela 1.

Tabela 1
Frequências dos tipos de gráficos indicados e construídos pelos estudantes

Tipo de gráfico	N.º de estudantes (%)	
	Gráficos indicados	Gráficos construídos
Histograma	35(70)	24(48)
Diagrama de caule-e-folhas	3(6)	2(4)
Gráfico de barras	36(72)	12(24)
Gráfico de linhas	7(14)	3(6)
Gráfico circular	5(10)	—
Diagrama de dispersão	5(10)	—
Tabela de frequências	2(4)	1(2)
Gráfico de barras agrupadas	1(2)	—
Não responde	1(2)	8(16)

Analisando a seleção de gráficos efetuada pelos estudantes, conclui-se que, em média, cada estudante assinalou dois tipos distintos de gráficos para representar os dados. Por outro lado, foram indicados muitos mais gráficos não adequados (56) do que gráficos adequados (38) e não foi referido o diagrama de extremos e quartis como gráfico adequado para representar os dados.

No caso dos gráficos considerados adequados, salienta-se o histograma, que foi referido por mais de metade dos estudantes (70%), seguindo-se o diagrama de caule-e-folhas, embora com uma frequência muito reduzida (6%).

Já no caso dos gráficos considerados não adequados, destaca-se o gráfico de barras com mais de metade dos estudantes a referi-lo (72%), tendo sido mesmo o mais selecionado, seguindo-se o gráfico de linhas (14%), o gráfico circular e o diagrama de dispersão (ambos com 10%), a tabela de frequências (4%) e o gráfico de barras agrupadas (2%).

Assim, globalmente, quando foi pedido aos estudantes para indicarem os gráficos adequados para representar os dados, eles revelaram um fraco desempenho. Para além do gráfico de barras, de linhas e circular, que não se adequam porque os dados são quase todos distintos, ainda menos se compreende a referência ao diagrama de dispersão, que se aplica a distribuições bidimensionais, portanto envolvendo duas variáveis estatísticas, e a tabelas de frequências.

Nos gráficos construídos pelos estudantes para representar os dados salienta-se o uso do histograma por cerca de metade dos estudantes (48%), que é um gráfico adequado à situação apresentada. Contudo, muitos estudantes usaram um número muito variado de classes, como seja 10 classes (8%), 6 classes (14%), 5 classes (16%), 4 classes (2%) e 3 classes (8%), quando, para um conjunto de 12 dados, seria adequado usar 3 ou 4 classes (recorrendo à Tabela de Truman L. Kelly ou à regra da raiz quadrada da dimensão da amostra). Destes gráficos, em três consideraram-se as barras separadas. Na Figura 2 apresenta-se um exemplo de histograma com 10 classes. No caso dos diagramas de caule-e-folhas os estudantes não especificaram as unidades do tronco nem das folhas.

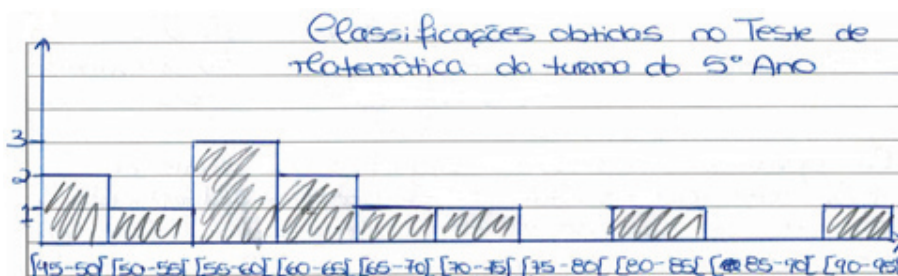


Figura 2. Gráfico construído pelo estudante E41 na questão a)

Embora não fosse adequado, uma percentagem razoável de estudantes (24%) construiu um gráfico de barras para representar os dados, tal como se exemplifica na Figura 3. Podemos constatar que o estudante E4 não tem em atenção a escala das observações e considerou apenas 11 dados, esquecendo-se do dado 58.

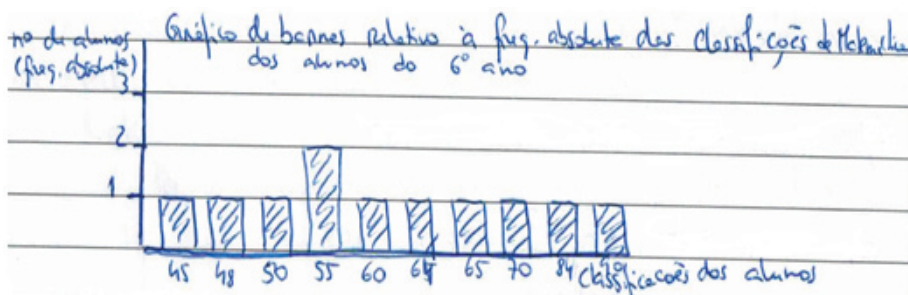


Figura 3. Gráfico construído pelo estudante E4 na questão a)

O facto de não haver quase repetição de dados, como acontece nesta situação, tem por consequência que a distribuição dada pelo gráfico de barras não é muito distinta dos dados brutos fornecidos, sendo mínima a redução de dados obtida com a representação gráfica. Destes estudantes, alguns (12%) consideraram no eixo horizontal os índices dos dados (1 a 12) e no eixo vertical as suas classificações, como se mostra na Figura 4.

Nesta representação não há mesmo qualquer redução dos dados, tratando-se tão somente de representar os dados fornecidos no enunciado da tarefa através de um gráfico de barras. Este tipo de dificuldade também foi observado por Fernandes e Correia (2009) em alguns professores dos primeiros anos em exercício.

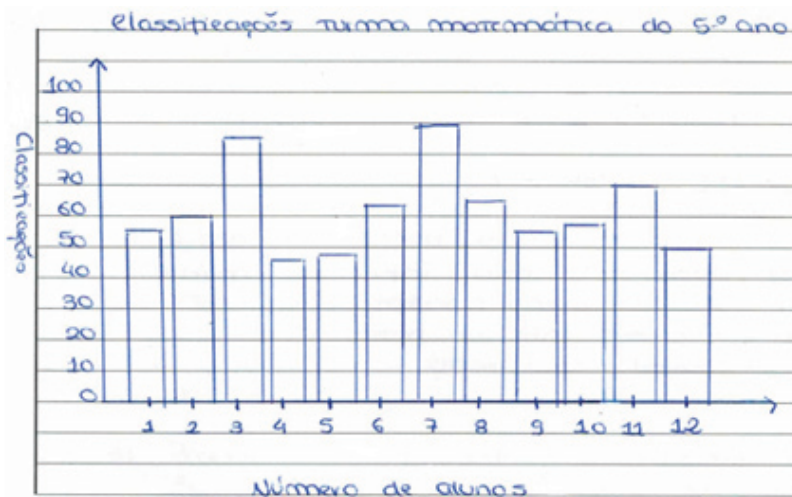


Figura 4. Gráfico construído pelo estudante E4 na questão a)

Verificou-se ainda que, frequentemente, os estudantes não estabeleceram um título para o gráfico nem nomearam os respetivos eixos coordenados. Na globalidade dos gráficos, 32% não tinham título, 26% não nomeavam o eixo horizontal e 24% não nomeavam o eixo vertical.

Medidas estatísticas

Em geral, na questão b), os estudantes consideraram que seria possível determinar a moda, a média, os quartis e o desvio padrão. Na Tabela 2 apresentam-se as frequências de estudantes segundo o tipo de resposta (correta e incorreta) e de não resposta.

Tabela 2
Frequências dos tipos de resposta nas diferentes estatísticas

Tipo de resposta	N.º de estudantes (em %)			
	Moda	Média	Quartis	Desvio padrão
Correta	50(100)	43(86)	14(28)	26(52)
Incorreta	—	6(12)	29(58)	16(32)
Não resposta	—	1(2)	7(14)	8(16)

À exceção de um estudante, que indicou a classe modal, todos os outros indicaram o valor da moda dos dados simples. Analogamente, também quase todos os estudantes (86%) determinaram corretamente o valor da média, dos quais dois (4%) usaram as marcas de classe em vez dos valores da variável estatística, salientando-se ainda as respostas incorretas devidas a erros de cálculo (12%).

Já a determinação dos quartis revelou-se uma tarefa difícil para os estudantes, com poucos estudantes (28%) a determinarem corretamente os três quartis. A maioria (58%) determinou pelo menos um quartil incorretamente e alguns nem sequer responderam (14%). Dessas respostas, salienta-se o número de estudantes que não determinaram corretamente nenhum dos quartis (16%) e os que determinaram corretamente apenas o segundo quartil (38%). Neste último caso, os estudantes começaram por determinar corretamente o segundo quartil (ou mediana), correspondente à média dos dois valores centrais e, seguidamente, eliminaram esses valores na determinação do primeiro e terceiro quartil, como se mostra na Figura 5. Deste modo, estes estudantes determinam o primeiro e terceiro quartis como sendo o valor central de séries de dados ímpar quando deviam determinar a média aritmética dos dois valores centrais de séries de dados par.

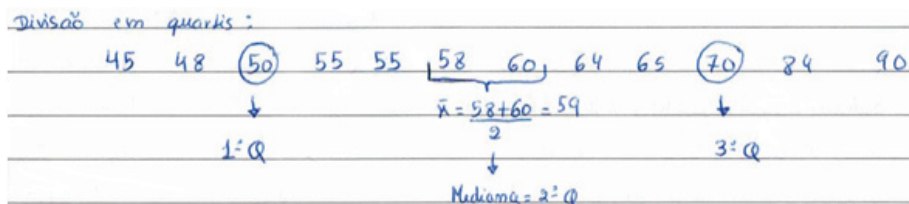


Figura 5. Determinação dos quartis pelo estudante E7 na questão b)

Aproximadamente metade dos estudantes (52%) determinaram corretamente o valor do desvio padrão, dos quais dois (4%) recorreram às marcas de classe, enquanto os restantes cometeram erros na sua determinação (32%) ou não responderam (16%). No caso das respostas erradas, salientam-se erros de cálculo (12%) e erros resultantes do uso de uma fórmula inadequada para a determinação do desvio padrão, como seja: confundir desvio padrão com variância (4%); não dividir pela dimensão da amostra (4%); determinar quadrados das diferenças entre as frequências absolutas e os valores da variável (4%); usar o valor da moda em vez do valor da média (2%); multiplicar os valores da variável pela média (2%); dividir pelo valor da média em vez de pela dimensão da amostra (2%) e considerar 11 dados em vez de 12 (2%).

Por último, na interpretação do terceiro quartil, menos de metade dos estudantes (38%) recorreram à divisão da série dos dados ordenados em quatro partes, cada uma com cerca de 25% de dados, para concluir o seu significado. Alguns estudantes (22%) responderam corretamente referindo-se, simultaneamente, a cerca de 75% de dados inferiores ou iguais e 25% de dados superiores ou iguais ao terceiro quartil, como se mostra na Figura 5.

O 3:Q significa que cerca de 75% dos dados são inferiores ou iguais ao valor do 3:Q e cerca de 25% dos dados são superiores ou iguais ao valor do 3:Q.
Neste caso, 75% dos dados são inferiores ou iguais a 67,5% e 25% dos dados são superiores ou iguais a 67,5%.

Figura 5. Determinação dos quartis pelo estudante E7 na questão b)

Menos estudantes (16%) referiram apenas que cerca de 75% dos dados são inferiores ou iguais ao terceiro quartil. Tal como o significado anterior, também este foi considerado correto pois o valor do terceiro quartil não é um valor observado na coleção de dados pelo que, nessas circunstâncias, se terá que a percentagem de valores superiores ou iguais ao valor do terceiro quartil é o complemento para 100% da percentagem de 75% (para maior detalhe sobre esta problemática da correta interpretação a dar aos quartis, associado à existência de observações iguais ao valor do quartil na coleção de dados veja-se, por exemplo, Freitas, Cruz & Silva (2017) sobre o caso particular do segundo quartil).

Para além da elevada percentagem de não respostas (34%), a atribuição de significados considerados incorretos deveu-se, sobretudo, a afirmações imprecisas dos estudantes (14%), à confusão do terceiro quartil com o primeiro quartil ou a mediana (8%), à confusão do terceiro quartil com a média ou moda (4%) e à confusão do terceiro quartil com um valor extremo (2%). Estes resultados mostram que, tal como na determinação dos valores dos quartis, também a interpretação do 3.º quartil levantou muitas dificuldades aos estudantes.

CONCLUSÕES

No presente estudo questionaram-se estudantes, futuros professores dos primeiros anos, sobre as duas etapas fundamentais de um estudo estatístico: a escolha de métodos (gráficos e numéricos) a usar numa análise estatística, seguida da aplicação desses métodos.

Considerando o desempenho global dos estudantes na questão a), conclui-se que 76% indicaram um gráfico adequado, mas destes apenas 22% indicaram só esse gráfico adequado, o que significa que a maior parte destes estudantes apresentou, além do gráfico adequado, pelo menos um gráfico não adequado à situação dada; já 52% construíram um tipo de gráfico adequado, embora podendo apresentar falhas na sua estrutura. Na questão b), considerando a determinação das medidas estatísticas e a interpretação do 3.º quartil, em média, 60,8% respondeu corretamente.

Estes resultados revelam que a questão envolvendo a indicação dos gráficos adequados à representação dos dados e a construção de um desses gráficos revelou-se ser a que mais dificuldades suscitou aos estudantes, sendo que a seleção de gráficos não adequados e a omissão de aspetos estruturais dos gráficos, como seja o título, a escala e os rótulos dos eixos (Friel et al., 2001), têm sido observados em estudos com alunos do 9.º ano (Fernandes et al., 2011; Morais & Fernandes, 2011) e com futuros professores dos primeiros anos (Fernandes et al., 2019). Na construção do gráfico salienta-se o recurso ao gráfico de barras em que, devido à quase ausência de repetição, a representação gráfica não se distingue dos dados fornecidos. Segundo Fernandes e Correia (2009), a representação gráfica de cada dado individual, como aconteceu nesta situação, é o nível mais básico de representação dos dados. Neste caso está-se perante uma representação gráfica que não produz qualquer redução de dados.

Foi na questão envolvendo a determinação de medidas estatísticas e a interpretação do 3.º quartil que se verificou a maior percentagem de respostas corretas. Geralmente, na determinação de medidas estatísticas, os estudantes revelam um melhor desempenho do que na escolha e interpretação dessas estatísticas (Boaventura & Fernandes, 2004; Fernandes et al., 2007), o que, de entre outras possíveis razões, se deve a um ensino mais orientado para as técnicas e procedimentos (Fernandes et al., 2011). Comparativamente com outras estatísticas, a maior dificuldade dos estudantes na determinação dos quartis generaliza as dificuldades sentidas pelos estudantes na determinação da mediana, conforme tem sido confirmado em vários estudos (e.g., Carvalho, Fernandes & Freitas, 2019; Fernandes & Barros, 2005).

Confrontando os resultados do presente estudo, cujos dados se referem a uma variável quantitativa, com aqueles que foram obtidos no estudo de Fernandes et al. (2019), cujos dados se referem a uma variável qualitativa nominal, estudos nos quais participaram futuros professores dos primeiros anos, não se verifica que os estudantes tenham escolhido métodos estatísticos muito distintos. Contudo, o facto de os dados se reportarem a uma variável qualitativa nominal ou a uma variável quantitativa implica a adoção de métodos estatísticos frequentemente distintos, o que não parece ter sido percebido por muitos estudantes. Então, esses estudantes determinam frequências acumuladas e várias medidas estatísticas em variáveis qualitativas nominais, recorrem às frequências em vez dos valores da variável no caso das estatísticas, e representam os dados através de gráficos que não reduzem significativamente os dados, de modo a permitir reconhecer padrões e tendências.

Portanto, na identificação e aplicação de métodos estatísticos, parece que os estudantes consideram poder aplicar tais métodos a qualquer tipo de variável estatística, sobregeneralizando as possibilidades de uso desses métodos.

Assim, decorre do presente estudo a necessidade e conveniência dos futuros professores dos primeiros anos explorarem ao longo da sua formação inicial tarefas do tipo daquelas que foram propostas neste estudo, destacando-se aquelas que envolvem a tomada de decisão acerca dos métodos estatísticos a adotar e a interpretações de resultados estatísticos. A experimentação, por parte dos futuros professores, destas tarefas é da maior relevância se pretendemos que eles também as explorem com os seus futuros alunos, pois essas experiências formativas influenciam as suas futuras práticas pedagógicas (Almeida & Fernandes, 2010).

Perante as dificuldades sentidas pelos estudantes, a solução é mesmo enfrentá-las, não as ignorar, pois a escolha dos métodos estatísticos e a interpretação de resultados são atividades intrinsecamente associadas à exploração de tarefas mais abertas, de aplicação e a projetos investigativos (Batanero et al. 2011). Dispensar os estudantes de decidir sobre quais métodos estatísticos devem usar, porque já são especificados nos enunciados das tarefas, e não incentivar a interpretação realça um ensino e aprendizagem da Estatística centrados nas fórmulas e nos cálculos, o que é uma abordagem limitada face às recomendações atuais para o ensino da Estatística (Fernandes et al., 2007; MacGillivray & Pereira-Mendoza, 2011).

AGRADECIMENTOS

Este trabalho é financiado pelo CIED — Centro de Investigação em Educação, UID/CED/01661/, Instituto de Educação, Universidade do Minho, através de fundos nacionais da FCT/MCTES-PT; EDU2016-74848-P (AEI, FEDER) e também pelo CIDMA (Centro de Investigação e Desenvolvimento em Matemática e Aplicações) da Universidade de Aveiro e FCT (Fundação para a Ciência e a Tecnologia), dentro do projeto UID/MAT/04106/2019.

DECLARAÇÕES DE CONTRIBUIÇÃO DOS AUTORES

JAF foi mais responsável pela planificação do estudo e organização do artigo, enquanto AF assumiu um papel mais proeminente no âmbito dos aspetos estatísticos. Ambos os autores participaram igualmente na elaboração do artigo, designadamente ao nível do enquadramento teórico e antecedentes, da metodologia, da análise de dados e das conclusões.

DECLARAÇÃO DE DISPONIBILIDADE DE DADOS

Os autores concordam em disponibilizar os dados que suportam os resultados deste estudo mediante solicitação razoável de um leitor, cabendo aos autores determinar se uma solicitação é razoável ou não.

REFERÊNCIAS

- Almeida, M. G., & Fernandes, J. A. (2010). A comunicação promovida por futuros professores na aula de Matemática. *Zetetiké*, 18(34), 109-154.
- Ball, D. L. (2000). Bridging Practices. Intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education*, 51(3), 241–247.
- Batanero, C. (2009). Retos para la formación estadística de los profesores. In J. A. Fernandes, M. H. Martinho, F. Viseu & P. F. Correia (Eds.), *Actas do II Encontro de Probabilidades e Estatística na Escola* (pp. 52-71). Braga: Centro de Investigação em Educação da Universidade do Minho.
- Batanero, C., Díaz, C., Contreras, J. M., & Arteaga, P. (2011). Enseñanza de la Estadística a través de proyectos. In C. Batanero & C. Díaz (Eds.), *Estatística con proyectos* (pp. 9-46). Granada: Universidad de Granada.
- Boaventura, M. G., & Fernandes, J. A. (2004). Dificuldades de alunos do 12.º ano nas medidas de tendência central: O contributo dos manuais escolares. In J. A. Fernandes, M. V. Sousa & S. A. Ribeiro (Eds.), *Actas do I Encontro de Probabilidades e Estatística na Escola* (pp. 103-126). Braga: Centro de Investigação em Educação da Universidade do Minho.
- Carvalho, M. J., Fernandes, J. A., & Freitas, A. (2019). Dificuldades de alunos do 8.º ano na construção de diagramas de extremos e quartis. In J. M. Contreras, M. M. Gea, M. M. López-Martín, & E. Molina- Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. Granada: Universidad de Granada.
- Espinel, M. C., González, M. T., Bruno, A., & Pinto, J. (2009). Las gráficas estadísticas. In L. Serrano (Ed.), *Tendencias actuales de la investigación en educación estocástica* (pp.57-74). Málaga: Gráficas San Pancrácio.
- Fernandes, J. A., & Barros, P. M. (2005). Dificuldades de futuros professores do 1.º e 2.º ciclos em estocástica. In *Actas do V Congresso Ibero-Americano de Educação Matemática (CIBEM)* (13 pp.), Faculdade de Ciências da Universidade do Porto, 17-22 de julho.
- Fernandes, J. A., & Correia, P. F. (2009). Variação em contexto de probabilidades na perspectiva de professores de Matemática. In A. Gomes (Ed.), *EME 2008 – Elementary*

Mathematics Education (pp. 197-208). Braga: Universidade do Minho e Associação para a Educação Matemática Elementar.

Fernandes, J. A., Batanero, C., & Gea, M. M. (2019). Escolha e aplicação de métodos estatísticos por futuros professores dos primeiros anos. In J. M. Contreras, M. M. Gea, M. M. López-Martín, & E. Molina-Portillo (Eds.), *Actas del Tercer Congreso Internacional Virtual de Educación Estadística*. Granada: Universidade de Granada.

Fernandes, J. A., Carvalho, C., & Correia, P. F. (2011). Contributos para a caracterização do ensino da Estatística nas escolas. *Boletim de Educação Matemática (BOLEMA)*, 24(39), 585-606.

Fernandes, J. A., Carvalho, C., & Ribeiro, S. A. (2007). Caracterização e implementação de tarefas de Estatística: um exemplo no 7.º ano de escolaridade. *Zetetiké*, 15(28), 27-61.

Fernandes, J. A., Morais, P. C., & Lacaz, T. V. S. (2011). Representação de dados através de gráficos estatísticos por alunos do 9º ano de escolaridade. *Anais da XIII Conferência Interamericana de Educação Matemática*, Recife, Brasil, 26-30 junho de 2011.

Freitas, A., Cruz, J. P., & Silva, N. (2017) Mediana de dados não agrupados: a questão de ser pelo menos 50%. *Educação e Matemática*, 143, 18-21.

Friel, S., Curcio, F., & Bright, G. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education*, 32(2), 124-158.

Gonzalez, J. F., & Eudave, D. (2018). Modelos de análisis del conocimiento matemático y didáctico para la enseñanza de los profesores. *Unión: Revista Iberoamericana de Educación Matemática*, 54, 25-45.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 39(4), 372-400.

MacGillivray, H., & Pereira-Mendoza, L. (2011). Teaching statistical thinking through investigative projects. In C. Batanero, G. Burril & C. Reading (Eds.), *Teaching statistics in school mathematics – Challenges for teaching and teacher education: A joint ICMI/IASE study* (pp. 109-120). New York: Springer.

Ministério da Educação (2007). *Programa de matemática do ensino básico*. Lisboa: Autor.

Ministério da Educação e Ciência (2014). *Programa de Matemática A – Ensino Secundário*. Lisboa: Autor.

Ministério da Educação e Ciência. (2013). *Programa de matemática para o ensino básico*. Lisboa: Autor.

Morais, P. C., & Fernandes, J. A. (2011). Realização de duas tarefas sobre construção, leitura e interpretação de gráficos estatísticos por alunos do 9º ano. In *Actas do XXII Seminário de Investigação em Educação Matemática (XXII SIEM)*. Lisboa: Associação de Professores de Matemática.

Ruiz, B., Arteaga, P., & Batanero, C. (2009). Competencias de futuros profesores en la comparación de datos. In L. Serrano (Ed.), *Tendencias actuales de la investigación en educación estocástica* (pp. 57-74). Melilla: Grupo de Investigación en Educación Estadística.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, 67(3), 223-248.