

Dificuldades de Alunos (Futuros Professores) em Conceitos de Estatística e Probabilidades

Paula Maria Barros, Escola Superior de Educação de Bragança
José António Fernandes, Universidade do Minho

Introdução

Neste texto apresentam-se alguns resultados de um estudo sobre dificuldades sentidas por alunos (futuros professores do 1º e 2º ciclos do ensino básico) em conceitos elementares de estatística e probabilidades.

Muito embora tenham estado envolvidos no estudo alunos do ensino superior, o carácter básico das tarefas exploradas faz com que os resultados obtidos sejam igualmente pertinentes para alunos dos ensinos básico e secundário.

Dificuldades em conceitos de estatística e probabilidades

São muitos os estudos que revelaram dificuldades dos alunos na aprendizagem de conceitos estocásticos, isto é, de estatística e probabilidades. Nesta secção faz-se uma breve revisão de estudos centrados nas medidas de tendência central e em conceitos elementares de probabilidades.

Muito embora as medidas de tendência central pareçam aparentemente simples, há, no entanto, investigações que mostram dificuldades conceptuais e procedimentais relacionadas com estes conceitos em alunos de diferentes idades.

No caso do cálculo de uma média ponderada em várias situações, Pollatsek, Lima e Well (1981) verificaram que muitos estudantes universitários sentiram dificuldades no cálculo de uma média global a partir do conhecimento de duas médias parciais. Nestas situações, a maioria dos estudantes determinou simplesmente a média simples dos valores das duas médias dadas, não afectando os seus valores com os pesos adequados.

A não ponderação no cálculo da média foi também observada por Li e Shen (1994) quando foi pedido aos estudantes para calcularem a média a partir de uma tabela de frequências com dados agrupados em classes.

No sentido de ampliar o estudo de Pollatsek, Lima e Well (1981), Mevarech (1983) realizou um estudo com estudantes universitários que possuíam os conhecimentos pré-requisito de cálculo. Muito embora todos reconhecessem as fórmulas de cálculo da média simples e ponderada e da variância, os conhecimentos de cálculo não se revelaram suficientes para a maioria dos estudantes adquirirem o esquema estatístico

apropriado, ou seja, apenas poucos estudantes possuíam o conjunto das estruturas conceituais necessárias para a resolução dos problemas propostos.

No caso do cálculo da média, aproximadamente 65% dos estudantes usaram a lei do fecho no cálculo da média total de grupos de tamanho diferente (propriedade estudada por Pollatsek, Lima e Well, 1981), cerca de 80% aplicaram a propriedade associativa no cálculo da média total de três grupos, 60% não compreenderam a não existência de elemento inverso no cálculo de uma média aritmética e aproximadamente 30% consideraram o valor 0 (zero) como elemento neutro, isto é, o valor da média não se altera quando se acrescenta o valor zero ao conjunto de dados. No caso do cálculo da variância, observaram-se concepções erradas semelhantes.

Para Mevarech (1983), os resultados obtidos sustentam a hipótese de que os estudantes conceptualizaram erradamente as operações de cálculo da média e da variância como duas operações binárias satisfazendo as quatro leis dos grupos aditivos. Este modelo é perfeitamente plausível quando se considera os valores da média e da variância como simples números, esquecendo que eles são medidas de tendência central e de dispersão.

Num estudo com estudantes do 6º ano de escolaridade, em que a maioria dos alunos evidenciou conhecer o algoritmo de cálculo da média, Cai (1995) observou que apenas 50% dos alunos foram capazes de determinar um valor desconhecido num pequeno conjunto de dados, apresentado sob a forma de pictograma, para se obter um dado valor da média. Este resultado agravou-se ainda mais, pois, dos alunos que encontraram o valor desconhecido, apenas 59% o determinaram através de uma utilização compreensiva do algoritmo (multiplicar o valor da média pelo número total de valores e subtrair a soma dos valores dados), e, destes, 35% recorreram a uma estratégia de tentativa e erro.

Acerca dos conceitos de média, moda e mediana, Carvalho (1996) analisou as realizações de dois grupos de alunos do 7º ano, cada um numa tarefa distinta. Num dos grupos, foi dado um conjunto de dados que os alunos deviam organizar numa tabela de frequências, no outro, os dados foram apresentados através de um gráfico de barras.

No caso da construção da tabela, os alunos não tiveram dificuldades em determinar as frequências absolutas, mas só 21% dos alunos calcularam correctamente as frequências relativas. Neste último caso, os erros dos alunos resultaram de considerar no denominador da fracção a frequência absoluta em vez do efectivo total dos dados.

Quanto às medidas de tendência central, no conjunto das duas tarefas, determinaram correctamente a moda 57% dos alunos, a média 20% dos alunos e a mediana 12% dos alunos. Entre as duas tarefas, foi na tarefa estabelecida através do gráfico que os alunos

revelaram mais dificuldades e foi também nessa tarefa que mais alunos tentarem responder.

Em relação às dificuldades dos alunos, Carvalho (1996) salienta que, no caso da tarefa em que se partiu dos dados, eles não tiveram em conta a frequência absoluta de cada valor no cálculo da mediana, quando usaram a tabela construída antes, ou não ordenaram os dados previamente à sua localização. No caso da tarefa em que se partiu do gráfico, os alunos não consideraram as frequências absolutas dos diferentes valores no cálculo da média e adicionaram as frequências absolutas e dividiram por 2 no cálculo da mediana. No caso da moda, em que se obteve o maior número de respostas correctas, a autora destaca a facilidade com ela é visualizada no gráfico de barras, correspondendo ao valor com a 'barra mais alta'.

Com o fim de avaliar as competências computacionais do conceito de média e de aplicação das suas propriedades, comparativamente com a escolha entre a média e a mediana, enquanto estatística mais adequada à caracterização de um conjunto de dados, Carvalho e César (2000) analisaram o desempenho de alunos do 7º ano trabalhando em 84 díades. Em termos de resultados, verificou-se que a maioria das díades aplicou, com sucesso, o procedimento de cálculo da média (94%) e da mediana (89%). Tal como observou Carvalho (1996), também neste estudo, comparativamente com a média, a mediana levantou mais dificuldades.

Quando se tratou de seleccionar entre a média e a mediana, envolvendo uma compreensão relacional, em contraste com uma compreensão instrumental implicada na aplicação dos respectivos algoritmos de cálculo, acentuaram-se as discrepâncias anteriores. No caso da média, apenas 27% dos alunos apresentaram argumentos que não apelam para um significado matemático, e, no caso da mediana, 44% dos alunos não foram capazes de usar argumentos matemáticos. Para estas autoras, a compreensão mais profunda do conceito de média explica-se pela sua frequente utilização nos mais variados contextos sociais, o que não acontece com o conceito de mediana.

A importância de uma compreensão relacional, face a uma compreensão apenas instrumental, é reforçada por Batanero (2000) ao afirmar que a compreensão de um conceito, para além do conhecimento das definições e propriedades, inclui também o reconhecimento dos problemas em que ele deve ser utilizado.

Eisenbach (citado em Batanero, 2000) questionou estudantes universitários de um curso introdutório de estatística sobre o significado da afirmação 'Que quer dizer que o salário médio de um empregado é de 3600 dólares?', tendo obtido respostas do tipo: 'a maioria dos empregados ganha cerca de 3600 dólares', 'é o salário central' e 'os outros trabalhadores ganham mais ou menos 3600 dólares'. Estas respostas mostram a confusão terminológica entre as palavras 'média', 'mediana' e 'moda'.

Numa perspectiva semelhante, Dreyfus e Levy (1996), num estudo com alunos de 11 e 12 anos, identificaram algumas concepções erradas a respeito da média. Uma das concepções frequentes consistiu em considerar a média como o valor central, o que denota uma confusão da média com a mediana. Outra concepção, embora menos frequente, consistiu em determinar a média adicionando os valores da variável e dividir por dois.

São também muitos os estudos que mostraram dificuldades dos alunos em probabilidades. Fischbein, Nello e Marino (1991) observaram, relativamente a alunos do 4º e 5º anos (9-11 anos) e do 6º, 7º e 8º anos (11-14 anos), que a maioria dos alunos identificou acontecimentos certos, possíveis e impossíveis e reconheceram situações com mesma estrutura estocástica. Já no caso da comparação de probabilidades em experiências compostas, os alunos sentiram muitas dificuldades. De entre os vários tipos de acontecimentos, os alunos revelaram mais dificuldades na categoria dos acontecimentos certos e na formulação de acontecimentos relativamente à sua classificação.

Green (1983) desenvolveu um estudo em larga escala em Inglaterra envolvendo 2930 alunos do 1º ao 5º ano de escolas secundárias (11-16 anos). Nos itens que requeriam apenas o conceito de contagem todos os alunos obtiveram um nível de realização satisfatório. Diferentemente, os itens que requeriam o conceito de razão revelaram-se particularmente difíceis, especialmente entre os alunos dos três primeiros anos. A ideia de esperança matemática foi acessível à maior parte dos alunos, apesar de ser raramente tratada nas aulas de matemática. Já nos itens sobre aleatoriedade, multiplicação de probabilidades, inferência amostral e estabilidade de frequências, os alunos sentiram maiores dificuldades. Globalmente, estes resultados foram corroborados por um estudo de Munisamy e Doraisamy (1998), envolvendo alunos malaios do 4º e 6º anos do ensino secundário. Nos estudos de Green (1983) e de Munisamy e Doraisamy (1998) verificou-se um aumento claro das respostas correctas com o ano escolar e com o desempenho em matemática.

No nosso país, Fernandes (1999) verificou que alunos do 8º e 11º anos de escolaridade revelaram mais dificuldades em identificar acontecimentos em certos e/ou que envolviam conectivos lógicos, na comparação de probabilidades em experiências simples que envolviam o conceito de razão e, mais acentuadas, na comparação de probabilidades em experiências compostas. Em termos de respostas correctas, observou-se um aumento sistemático com o ano escolar e com o desempenho em matemática.

Metodologia

Este estudo teve como objectivo identificar dificuldades e processos de raciocínio de futuros professores em conceitos elementares de estatística e probabilidades. Com esse intuito, passou-se um questionário a 37 alunos de uma turma de 4º ano do Curso de Matemática/Ciências da Natureza de uma Escola Superior de Educação, futuros professores do 1º e 2º ciclos do ensino básico. Todos estes alunos tinham frequentado uma disciplina de Probabilidades e Estatística no ensino superior.

O questionário era constituído por 13 questões, algumas delas com sub-questões, que envolviam conceitos ligados às medidas de tendência central, à classificação de acontecimentos e à comparação de probabilidades de acontecimentos simples. Nesta comunicação analisam-se apenas três dessas questões.

Apresentação e discussão de resultados

Questão 1. Relativamente às preferências de clube desportivo dos 200 alunos de uma escola, cuja distribuição era apresentada sob a forma de gráfico circular, pretendia-se identificar a moda e averiguar da possibilidade de calcular a média.

Quanto à identificação da moda, a maioria dos inquiridos respondeu correctamente que a moda era o Porto e 22% dos alunos identificaram a moda com a frequência relativa.

Em termos de raciocínios, 95% dos alunos fizeram *Referência à maior frequência*, apresentando argumentos do tipo: “é o que tem maior percentagem”, “é o que se repete mais” ou “é o valor que se repete com maior frequência”. Este raciocínio, além de justificar a resposta correcta, justificou também a resposta dos alunos que identificaram a moda como a frequência relativa.

Destes resultados depreende-se que, quando a variável em estudo é qualitativa, a identificação da moda apresenta ainda dificuldades para alguns alunos, já que, embora pareçam ter uma noção do seu significado, prevaleceu a ideia de que a moda tem que ser um número.

No que concerne ao cálculo da média, apenas 38% dos alunos responderam correctamente que não era possível calcular a média, tendo uma percentagem considerável (57%) concluído que a média era 20% ou a correspondente frequência absoluta.

Quando se passa à análise dos raciocínios, as dificuldades acentuam-se, verificando-se que apenas 13% dos alunos justificaram a sua resposta no raciocínio *Alusão a variáveis qualitativas*, referindo explicitamente que se tratava de uma variável

qualitativa ou afirmando esse facto de forma menos directa, alegando que “estamos em presença de nomes e não de números” ou que “não é possível atingir um valor numérico que expresse a média”.

No raciocínio *Cálculo da média das frequências*, referido pela maioria dos alunos (60%), a resposta baseou-se na aplicação do algoritmo da média. Neste caso, 57% dos alunos calcularam o quociente da soma das frequências relativas (ou absolutas) pelo número de valores que tomava a variável e chegaram à conclusão de que a média era 20% (ou 40) e 3% adicionaram as frequências relativas e dividiram a soma obtida pelo número total de inquiridos. Este raciocínio denota uma tendência para manipular os dados disponíveis de forma a aplicar uma fórmula, sem reflectir sobre o seu sentido no contexto apresentado. Como no caso da moda, parece que o fundamental é tentar encontrar um valor numérico que responda ao problema.

Finalmente, o raciocínio *Argumentos não matemáticos*, referido por 8% dos alunos, conduziu também à escolha da resposta correcta, mas neste caso os alunos apresentaram razões que tinham a ver com o clube ou com a falta de dados: “estão vários clubes envolvidos”, “foram os alunos a escolher o clube” e “não estão especificados os clubes representados por outros”.

Questão 2. Há 10 pessoas num elevador, 6 mulheres e 4 homens. A média do peso das mulheres é 60 quilos, e a média do peso dos homens é 80 quilos. Determine a média das 10 pessoas que se encontram no elevador.

Nesta questão verificou-se que mais de metade dos alunos (60%) responderam correctamente que a média era 68 quilos e uma percentagem considerável (30%) afirmou que a média era 70 quilos.

De entre os raciocínios referidos, o *Algoritmo da média* (ponderada) foi o mais utilizado (60%) e conduziu sempre à selecção da resposta correcta.

Destaca-se, ainda, 30% de alunos que recorreram à *Lei do fecho*, ou seja, calcularam a média das duas médias dadas, tendo concluído que a média era 70 quilos. De modo semelhante, Pollatsek, Lima e Well (1981) e Mevarech (1983) observaram que muitos estudantes universitários atribuem à média tal propriedade, que de facto não tem.

Por último, no raciocínio *Adição das médias* 5% de alunos adicionaram as médias dadas para obter a média final pretendida. Observando que $x/6=60$ e que $x/4=80$, concluíram que $x/10=140$ e que a média era 140 quilos, parecendo denotar algumas dificuldades na compreensão do significado das letras e da adição de fracções.

Questão 3. Num saco há 5 bolas vermelhas, 2 verdes e 4 brancas. As bolas são todas iguais excepto na cor. Sem ver, tira-se do saco um bola de cada vez, sem a voltar a repor. Quantas bolas se tem de tirar do saco para ter a certeza de obter, pelo menos, uma bola de cada cor?

Nesta questão apenas 24% dos alunos deram a resposta correcta, afirmando que é necessário tirar pelo menos 10 bolas do saco, e as restantes respostas foram muito variadas.

Tal como as respostas, os raciocínios referidos foram também variados. No raciocínio *Bolas mais numerosas*, que conduziu sempre à resposta correcta, 19% de alunos adicionaram o número de bolas vermelhas e brancas (bolas mais numerosas) e concluíram que no fim ficavam duas verdes, pelo que bastava tirar mais uma bola.

No raciocínio *Bolas menos numerosas*, 3% de alunos adicionaram o número de bolas verdes e brancas (bolas menos numerosas) e concluíram que a seguinte já seria vermelha. Este raciocínio originou a resposta incorrecta de que era necessário tirar 7 bolas.

Relativamente ao raciocínio *Dois extracções para cada cor*, 3% de alunos consideraram que, como existiam bastantes bolas, as duas tentativas para cada cor talvez fossem suficientes, levando a afirmar que era necessário retirar 6 bolas do saco.

No raciocínio *Cálculo combinatório*, 16 % de alunos utilizaram a multiplicação ou a adição de combinações. Resultados absurdos obtidos a partir do cálculo incorrecto das combinações (por exemplo, o resultado de 27 bolas) não suscitou qualquer reserva por parte dos alunos.

Já no raciocínio *Cálculo de probabilidades*, 27% de alunos adicionaram ou multiplicaram as probabilidades dos acontecimentos “sair bola branca”, “sair bola vermelha” e “sair bola verde”. Relativamente à multiplicação, no caso da extracção com reposição, obteve-se o resultado 0,03 e, no caso da extracção sem reposição, o resultado 0,04.

Como se pode ver pelos resultados apresentados, esta questão originou bastantes dificuldades, reforçando a tendência, já observada antes, para recorrer a fórmulas e manipular dados sem ter em conta o contexto, e parece sugerir a ideia de que os conhecimentos que os alunos têm prejudicam o seu raciocínio em situações mais simples.

Considerações didácticas

Visto que os inquiridos serão futuramente professores do 1º e 2º ciclos, e que fazem parte destes níveis de ensino alguns conceitos elementares de estatística e probabilidades, inclusivamente no 2º ciclo questões semelhantes às apresentadas poderão ser tratadas, e que, mesmo estando no último ano da sua formação inicial, há dificuldades que se mantêm, torna-se imprescindível confrontar os estudantes com estes e outros tipos de problemas, discutindo as suas respostas e fazendo-os reflectir sobre

elas. Esta atitude de reflexão, para além de esclarecer dúvidas e promover um debate sobre as dificuldades sentidas, abre horizontes para um conhecimento mais amplo da diversidade de problemas estocásticos que se podem debater na sala de aula, permitindo-lhes sair da rotina dos problemas típicos de aplicação directa de fórmulas e orientar os seus alunos na superação de dificuldades que parecem persistir ao longo de toda uma escolaridade.

Referências

- Batanero, C. (2000) Dificultades de los estudiantes en los conceptos estadísticos elementales: el caso de las medidas de posición central. Em C. Loureiro, O. Oliveira. e L. Brunheira. (Orgs.) *Ensino e aprendizagem da estatística*.(pp.31-48). Lisboa: Sociedade Portuguesa de Estatística, Associação de Professores de Matemática, Departamento de Educação e de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa.
- Cai, J. (1995). Beyond the computational algorithm: student's understanding of the arithmetic average concept. Em L. Meira e D. Carraher (Eds.), *Proceedings of the 19th PME conference* (Vol. 3, pp. 144-151). Universidade Federal de Pernambuco.
- Carvalho, C. (1996). Algumas questões em torno de tarefas estatísticas com alunos do 7º ano. Em António Roque e Maria João Lagarto(Orgs.), *Actas do ProfMat 96*. Lisboa: Associação de Professores de Matemática.
- Carvalho, C. e César, M. (2000) As aparências iludem: reflexões em torno do ensino da estatística no ensino básico. Em C. Loureiro, O. Oliveira. e L. Brunheira. (Orgs.), *Ensino e aprendizagem da estatística*, (pp.212 - 225). Lisboa: Sociedade Portuguesa de Estatística, Associação de Professores de Matemática, Departamento de Educação e de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa.
- Dreyfus, A. e Levy, O. (1996). Are the notion of mean and related concepts too difficult for 6th e 7th grade biology students? *European Journal of Teacher Education*, 19(2), 137-152.
- Fernandes, J. A. (1999). *Intuições e aprendizagem de probabilidades: Uma proposta de ensino de probabilidades no 9º ano de escolaridade*. Tese de doutoramento não publicada, Universidade do Minho, Braga.
- Fischbein, E., Nello, M. S. e Marino, M. S. (1991). Factors affecting probabilistic judgments in children and adolescents. *Educational Studies in Mathematics*, 22, 523-549.
- Green., D. (1983). A survey of probability concepts in 3000 pupils aged 11-16 years. Em D. R. Grey, P. Holmes, V. Barnett e G. M. Constable (eds.), *Proceedings of the First International Conference on Teaching Statistics* (Vol. 2, pp. 766-783). Sheffield, UK: Teaching Statistics Trust.
- Li, K. Y. e Shen, S. M. (1994). Students' weaknesses in statistical projects. Em D. Green (Ed.), *Teaching statistics at its best* (pp. 42-48). Sheffield, UK: Teaching Statistics Trust.
- Mevarech, Z. R. (1983). A deep structure model of students' statistical misconceptions. *Educational Studies in Mathematics*, 14, 415-429.
- Munisamy, S. e Doraisamy, L. (1998). Levels of understanding of probability concepts among secondary school pupils. *International Journal for Mathematical Education in Science and Technology*, 29(1), 39-45.
- Pollatsek, A., Lima, S. e Well, A. D. (1981). Concept or computation: students' understanding of the mean. *Educational Studies in Mathematics*,12, 191-204.