

# O Conhecimento Didático de Estatística de Duas Professoras de Matemática sobre Dados Bivariados

## The Statistical Didactical Knowledge of Two Mathematics Teachers about Bivariate Data

Sandra Maria Oliveira Quintas\*

Rosa Antónia O. F. T. Ferreira\*\*

Hélia Margarida A. P. Oliveira\*\*\*

### Resumo

Este artigo debruça-se sobre o conhecimento didático de duas professoras de Matemática no ensino de dados bivariados e suas relações, no tema da Estatística na disciplina de Matemática A, do ensino secundário em Portugal. Em particular, são analisadas as dimensões do conhecimento do ensino, do conhecimento dos alunos e da aprendizagem e a forma como estas se relacionam entre si e com as outras dimensões do conhecimento didático de cada professora, principalmente a partir da observação da sua prática de sala de aula no 10.º ano. Os resultados evidenciam que, em diversos aspetos, não parece ser considerada pelas professoras a complexidade de que se reveste para os alunos o raciocínio sobre dados bivariados. Verifica-se, também, que existem noções e representações fundamentais que emergem a partir das situações que propõem nas aulas que não são exploradas com a devida profundidade.

**Palavras-chave:** Conhecimento Didático. Estatística. Dados Bivariados. Ensino Secundário.

### Abstract

This paper concerns the didactical knowledge of two mathematics teachers while teaching bivariate data and their relationships, when covering the statistics subject of the discipline of Mathematics A, at the secondary level, in Portugal. In particular, we analyze the dimensions of the teaching knowledge and knowledge of students and learning of the teachers' didactical knowledge; we also address how these dimensions relate to each other and to the others dimensions of the teachers' didactical knowledge, mainly through the observation of the teaching practice of both teachers in their 10<sup>th</sup> grade classes. The results showed that in several aspects, teachers seem not consider how complex it is for students to reason about bivariate data. There are fundamental notions and representations emerging from the situations worked in the classes that are not explored with the appropriate depth.

**Keywords:** Didactical Knowledge. Statistics. Bivariate Data. Secondary School.

---

\* Mestre em Estatística Aplicada pela Eastern Michigan University (EMU, EUA). Aluna de doutoramento do Instituto de Educação da Universidade de Lisboa (IEUL), Lisboa, Portugal. Endereço: Alameda da Universidade, 1649-013 Lisboa, Portugal. E-mail: [sandramquintas@gmail.com](mailto:sandramquintas@gmail.com).

\*\* Doutora em Educação Matemática pela Illinois State University (ISU, EUA). Docente do Departamento de Matemática da Faculdade de Ciências da Universidade do Porto (FCUP), Porto, Portugal. Endereço: Rua do Campo Alegre, n.º 687, 4169-007 Porto, Portugal. E-mail: [rferreir@fc.up.pt](mailto:rferreir@fc.up.pt).

\*\*\* Doutora em Educação Matemática pela Universidade de Lisboa (UL). Docente do Instituto de Educação da Universidade de Lisboa (IEUL), Lisboa, Portugal. Endereço para correspondência: Alameda da Universidade, 1649-013 Lisboa, Portugal. E-mail: [hmoliveira@ie.ulisboa.pt](mailto:hmoliveira@ie.ulisboa.pt).

## 1 Introdução

A análise e interpretação de relações bivariadas é uma atividade importante em várias disciplinas e, por conseguinte, a literatura acerca do raciocínio sobre estes dados aponta a relevância desta temática na investigação de diversas áreas tais como na Psicologia, Ciência, Educação Matemática e Educação Estatística. Os conceitos ligados ao estudo de dados e relações bivariados, nomeadamente, associação, correlação e regressão linear e o diagrama de dispersão são referidos no programa de Matemática A<sup>1</sup> do 10º ano do ensino secundário, em Portugal, no tópico de dados e relações bivariados.

O estudo de dados bivariados envolve análises de variáveis individuais e sua variabilidade, e da forma das distribuições bivariadas em gráficos (GARFIELD; BEN-ZVI, 2008). A compreensão da regressão e correlação exige conhecimento básico sobre funções e, sobretudo, a consideração da variação à volta de uma possível tendência (ENGEL; SEDLMEIER, 2011).

Vários autores referem a complexidade do ensino e aprendizagem sobre dados e relações bivariados e apontam a necessidade de mais investigação nessa área que informe sobre desafios a superar, em particular, sobre os aspetos do conhecimento didático a que se deve dar mais atenção na formação de professores (ESTEPA; BATANERO, 1996; GARFIELD; BEN-ZVI, 2008). É neste contexto que surge o presente estudo com o objetivo de compreender o conhecimento didático de duas professoras do ensino secundário, tendo em conta a especificidade do ensino de dados bivariados (BURRIL; BIEHLER, 2011; ENGEL; SEDLMEIER, 2011).

## 2 O conhecimento didático do professor em estatística

O *conhecimento profissional do professor de Matemática* desdobra-se por vários domínios, nomeadamente, o conhecimento na ação relativa à prática letiva, à prática não letiva, à profissão e ao seu próprio desenvolvimento profissional (PONTE; OLIVEIRA, 2002).

---

<sup>1</sup> Matemática A é uma disciplina obrigatória de cursos científico-humanísticos de ciências e tecnologias e de ciências socioeconómicas, do ensino secundário português (do 10.º ao 12.º anos – alunos dos 15 aos 17 anos), vocacionados para o prosseguimento de estudos no ensino superior.

Ponte e Oliveira (2002) designam a dimensão do conhecimento profissional chamado a intervir diretamente na prática letiva por *conhecimento didático*. É reconhecido que as práticas dos professores são contextos privilegiados que podem revelar aspetos relevantes do conhecimento didático dos professores (PONTE; CHAPMAN, 2006). Apesar de, habitualmente, a Estatística no ensino não superior ser estudada no âmbito da disciplina de Matemática vários autores reconhecem a especificidade do seu ensino (por exemplo, BATANERO et al., 2013; BURRILL; BIEHLER, 2011; SCHEAFFER, 2006). O modelo do conhecimento didático em Estatística, adotado neste trabalho, decorre do modelo do conhecimento didático do professor de Matemática de Ponte e Oliveira (2002) (com forte influência de Shulman (1986) mas assumindo tratar-se de um conhecimento fortemente orientado pelas situações da prática) e que inclui quatro domínios: conhecimento do conteúdo, conhecimento do currículo, conhecimento dos alunos e da aprendizagem e conhecimento do ensino. Estes domínios, que estão intimamente relacionados na prática letiva, são descritos a seguir tendo em conta a especificidade da Estatística, tal como apontado por vários autores da literatura nessa área.

O *conhecimento do conteúdo* refere-se, de uma forma genérica, ao conhecimento da disciplina e das interpretações dos seus conceitos, representações e procedimentos fundamentais, bem como, formas de raciocínio, argumentação e validação (PONTE; OLIVEIRA, 2002). Inclui conhecimento de Matemática e de Estatística. Integra a capacidade de reflexão epistemológica sobre a natureza do conhecimento estatístico e entendimento profundo deste conhecimento (BATANERO; GODINO, 2005). Envolve o conhecimento de ideias estatísticas essenciais (dados e contexto; variação, distribuição, representações, *transnumeração*<sup>2</sup>; associação e correlação; amostragem e inferência), suas interligações e conhecimento dos diversos dilemas e desafios que envolvem lidar com essas ideias (BURRILL; BIEHLER, 2011).

O *conhecimento do currículo* inclui o conhecimento das grandes finalidades e objetivos do currículo escolar e sua articulação vertical e horizontal (PONTE; OLIVEIRA, 2002) e, ainda, a compreensão das ideias estatísticas fundamentais expressas no currículo (BATANERO et al., 2013; BURRILL; BIEHLER, 2011). Contempla, ainda, a capacidade de analisar criticamente documentos curriculares, manuais escolares e outros recursos metodológicos disponíveis (BATANERO; GODINO, 2005).

---

<sup>2</sup> Processo de se usar representações variadas para identificar diferentes aspetos dos dados em estudo (WILD; PFANNKUCH, 1999).

O *conhecimento dos alunos e da aprendizagem* inclui o conhecimento das dificuldades, erros e obstáculos na aprendizagem dos conceitos, procedimentos e representações e das estratégias usadas pelos alunos na resolução de problemas (PONTE; OLIVEIRA, 2002). Abrange, também, o conhecimento dos diversos níveis de compreensão dos alunos (BATANERO; GODINO, 2005).

O *conhecimento do ensino* compreende a capacidade de planificação da sequência de conteúdos, nomeadamente adequando conteúdos a diferentes níveis de ensino, tendo em conta o grau de profundidade com que estes necessitam de ser tratados e relacionados. Inclui o conhecimento de estratégias de ensino e a fundamentação para as opções tomadas quanto às tarefas propostas e aos recursos usados, entre outros (BATANERO; GODINO, 2005). Esse conhecimento do ensino inclui o reconhecimento das vantagens e desvantagens do uso de diferentes representações no ensino dos diferentes tópicos (CURCIO; ARTZT, 1996). Abarca, igualmente, a capacidade de ajudar os alunos a desenvolver o raciocínio e pensamento estatísticos (BURRILL; BIEHLER, 2011; GARFIELD; BEN-ZVI, 2008).

### **3 O ensino de dados e relações bivariados**

Para ensinar Estatística o professor precisa ter experiência e familiaridade com elementos específicos do pensamento estatístico, nomeadamente, o reconhecimento da necessidade de dados e seu conhecimento contextual, e da *transnumeração*, dando atenção à variação e raciocínio com modelos, integrando-os na sua prática (WILD; PFANNKUCH, 1999). Para Ben-Zvi e Garfield (2004), o raciocínio estatístico engloba, por exemplo, ser capaz de compreender e interpretar estatísticas sumárias de dados, estabelecer ligações entre conceitos estatísticos e entre diferentes representações estatísticas ou relacionar ideias estatísticas com o acaso. Já o pensamento estatístico, para estes autores, envolve saber como e por que se usa um determinado método ou modelo, conhecer as limitações que se encontram, por exemplo, numa inferência estatística e ser capaz de criticar resultados de um estudo.

No ensino da Estatística, o raciocínio sobre dados bivariados deve suportar algo mais do que raciocinar sobre diagramas de dispersão, correlação, regressão e funções. Deve envolver o entendimento de ideias de estrutura e força na relação bivariada, a análise residual e ajuste do modelo, bem como, a compreensão do papel da relação bivariada em modelos e na previsão de eventos. Engel e Sedlmeier (2011) registam que, com frequência, os dados bivariados são trabalhados na aula como uma dependência funcional entre duas variáveis,

descurando-se a variação dos dados. Os autores apontam a necessidade de os alunos ganharem experiência com o processo de modelação, de modo a desenvolverem os seus conhecimentos de correlação e regressão. Destaca-se a importância de se propor tarefas que incluam dados reais, elementos sobre o seu contexto e questões que valorizem os dados (CURCIO; ARTZT, 1996; SCHEAFFER, 2006).

Garfield e Ben-Zvi (2008) sugerem que atividades de estabelecimento de conexões entre valores de coeficientes de correlação e diagramas de dispersão podem permitir que os alunos desenvolvam um melhor entendimento dos diferentes níveis de covariação e entendimento acerca de fatores que influenciam o maior ou menor valor do coeficiente de correlação. Por exemplo, um resultado de correlação linear alto não implica, necessariamente, a validade do modelo de regressão linear, devendo examinar-se cuidadosamente representações gráficas dos dados, como o diagrama de dispersão, medida muito influenciada por *outliers*.

Vários autores (ENGEL; SEDLMEIER, 2011; SHAUGHNESSY; CHANCE, 2005) também referem exemplos de conjuntos de dados para os quais a afirmação de que a correlação positiva entre duas variáveis traduzida como *assim que uma delas aumenta, a outra também aumenta*, nem sempre é verdadeira. Para esse caso, uma afirmação mais precisa é: *valores acima da média de uma das variáveis correspondem a valores acima da média da outra variável*. Esse conhecimento mais pormenorizado poderá contribuir para uma melhor apreciação da variação local e global dos dados e entendimento da fórmula do coeficiente de correlação (SHAUGHNESSY; CHANCE, 2005). A tecnologia pode proporcionar a visualização de representações diversificadas, o estabelecimento de conexões entre elas, e ainda a exploração e manipulação dos dados, concedendo um apoio significativo à análise (GARFIELD; BEN-ZVI, 2008).

#### 4 Metodologia

Este estudo insere-se numa investigação mais ampla, qualitativa e de índole interpretativa, sobre o conhecimento didático do professor no ensino da Estatística, no ensino secundário. O estudo foca-se, especificamente, no tópico dos dados bivariados e foi realizado com duas professoras, Estela e Lia, ambas licenciadas em ensino da matemática, possuindo a primeira também o grau de mestre em educação matemática. As suas experiências profissionais são de cerca de 20 e 15 anos, respetivamente. Na recolha dos dados que

informam este estudo foram usados diversos métodos, nomeadamente: (1) observação participante, com registo áudio e vídeo de três aulas, em duas turmas de 10.º ano com cerca de 25 alunos, uma de cada professora; (2) recolha documental dos materiais utilizados pelas professoras nessas aulas, designadamente fichas de trabalho e ainda, embora com menor expressão, (3) duas entrevistas semiestruturadas com registo áudio a cada professora (antes e após a realização das aulas) com foco no tópico de dados bivariados. A análise de dados foi efetuada de forma descritiva e interpretativa, conduzindo à identificação de situações da prática das professoras que remetiam para os aspetos centrais do conhecimento didático referente ao estudo de dados e relações bivariados, complementados com elementos das entrevistas.

No ensino do tema da Estatística, as duas professoras seguiram a sequência de conteúdos sugerida no programa de Matemática A do ensino secundário (DEPARTAMENTO DO ENSINO SECUNDÁRIO, 2001): (1) generalidades sobre Estatística (evolução histórica e objetivos); população, amostra, e amostragem; o procedimento estatístico: Estatística descritiva e Estatística inferencial; (2) organização de dados qualitativos e quantitativos (numérica e gráfica); e (3) estudo de dados bivariados (abordagem gráfica e intuitiva). Quanto ao último ponto, o programa salienta o desenvolvimento de ideias intuitivas relativamente à noção de correlação, de centro de gravidade e de reta de regressão, essencialmente a partir da visualização de diagramas de dispersão e de cálculos recorrendo à calculadora gráfica.

## **5 O estudo de dados bivariados na prática de Estela**

As aulas de Estela são habitualmente estruturadas e organizadas em torno da resolução de tarefas que, na maioria das vezes, são trabalhadas pelos alunos com o apoio da calculadora gráfica. Estela solicita a participação dos alunos, quer nos momentos de exposição dos conteúdos quer nos de correção dos trabalhos, pedindo-lhes que respondam a questões ou que vão ao quadro ou ao computador (cujo ecrã está, usualmente, projetado na tela) mostrar a sua resolução de alguma questão com a calculadora gráfica. O estudo em torno dos dados bivariados foi desenvolvido num curto espaço de tempo, no decorrer de três aulas de 90 minutos; por conseguinte, as tarefas propostas tiveram, sobretudo, o propósito de introduzir conceitos e representações, providenciando as respetivas explicações que a professora considerou necessárias. De entre estas aulas, escolhemos os seguintes três episódios

associados a diferentes tarefas realizadas, de modo a evidenciar aspetos centrais do conhecimento didático de Estela relativamente a dados bivariados.

### 5.1 Tarefa *A equipa de basquetebol do Porto*

Estela explicou, na entrevista, que a tarefa *A equipa de basquetebol do Porto* (ver anexo) foi reformulada, aproveitando uma outra mais antiga que possuía. Manteve os dados reais mas incorporou três novas questões plausíveis de serem investigadas, cada uma delas relacionando duas variáveis quantitativas. Na aula foi analisada apenas a relação entre as variáveis *minutos de jogo* e *pontos obtidos* de uma dessas questões. Depois de ter pedido aos alunos para procederem à construção do diagrama de dispersão com essas variáveis no seu caderno diário e à sua representação na calculadora, Estela passou a fazer a leitura do diagrama, chamando a atenção, em especial, para a tendência global dos dados, mesmo quando uma aluna (Joana) mencionou a existência de dados que não seguiam essa tendência. A professora deixou transparecer o entendimento que os alunos deveriam ter sobre a reta de regressão:

Professora: [...] O que é que reparam... Acontece que à medida que o tempo aumenta vocês vêem que os pontos também aumentam, ou não?

Alguns alunos: Sim!

Joana: Há exceções.

Professora: Então diz... Há exceções... Mas a maioria [dos pontos]....

Alguns alunos: Sim.

[...]

Professora: [...] Estão a ver que... Parece que se consegue fazer passar uma reta, não por todos os pontos... Ora bem... Mas pela maior parte deles ou pelo menos mais próximo deles. Agora só falta descobrir como é que se desenha essa reta [na calculadora], certo? Esta reta chama-se reta de regressão [...]. (Observação de aula, 13 de junho de 2011)

Estela realçou a proximidade da reta de regressão relativamente à tendência global dos dados que observou no diagrama de dispersão. Na aula, também definiu a reta de regressão como a *reta que melhor se ajusta aos pontos do diagrama de dispersão*. No desenvolvimento dessa tarefa, Estela orientou os alunos para uma atividade muito específica à volta da definição que forneceu para a reta de regressão. Pediu-lhes, inicialmente, que utilizassem os seus conhecimentos prévios sobre funções na determinação da expressão analítica de uma possível reta que melhor se ajustasse aos pontos do diagrama de dispersão (tendo em conta uma escolha conveniente de dois pontos quaisquer, que poderiam pertencer ou não ao diagrama de dispersão, através da observação do diagrama). Depois explicou como determinar a equação da reta de regressão na calculadora gráfica. Na perspetiva de Estela essa

atividade poderia facilitar o desenvolvimento das ideias dos alunos sobre a reta de regressão, nomeadamente, ao analisarem o quão afastada ou próxima a reta estimada se encontrava da reta de regressão. Quando os alunos confrontaram essas duas retas (a estimada e a de regressão) chegaram, de uma maneira geral, à conclusão de que eram diferentes mas que não estavam muito afastadas entre si quando visualizadas em simultâneo sobre o diagrama de dispersão na calculadora. Na sua interação com os alunos, Estela começou por lhes solicitar que representassem as duas retas na calculadora gráfica de modo a poder compará-las:

Professora: Desenham a reta que pedi... Mas sabem uma coisa, a calculadora gráfica faz isso tudo sozinha! [...] Agora quero que vocês comparem... a que obtiveram pela máquina [referindo-se à reta de regressão] com essa à mão [em que determinaram a expressão analítica]... Sabem como é que se faz na máquina? Para quem já fez na máquina, digam-me lá como é que se faz?

Isabel: Mas ó Stora, mas não dá valores iguais! [referindo-se ao facto de os declives e ordenadas na origem não serem iguais para as duas retas]

Professora: Pois não! ... É assim, eu disse-vos um ponto que tinha a certeza que a reta [de regressão] da máquina passava nele [centro de gravidade]; os vossos colegas disseram o outro [(22,4; 9,6)]... E se ele não passa por (22,4; 9,6)?

[...]

Professora: [Dá as instruções para chegarem à reta de regressão através da calculadora gráfica, explicando cuidadosamente o que cada instrução faz] E agora a equação [da reta de regressão] que a máquina me deu foi  $y=0.43x+0.135$ .

[...]

Professora: E agora façam *graph* [para aparecer no visor a reta de regressão]... Ficou muito diferente da nossa?

Vários alunos: Nem por isso. Não. (Observação de aula, 13 de junho de 2011)

Nesta interação, a professora reagiu ao facto de a aluna Isabel ter ficado surpreendida devido às duas retas construídas não apresentarem o mesmo declive e ordenada na origem. Estela referiu que, em relação aos dois pontos usados para determinar a equação inicial da reta (estimada), só havia a garantia de que um deles, o centro de gravidade, pertencesse à reta de regressão que procuravam. Acabou por ser explicitamente assumido pela professora que a reta de regressão é aquela que melhor se ajusta à nuvem de pontos considerada, é obtida através da calculadora gráfica e o centro de gravidade é um ponto que lhe pertence. Este último facto foi confirmado quando os alunos verificaram que as coordenadas do centro de gravidade satisfaziam a equação da reta de regressão determinada através da calculadora gráfica. Ainda durante o desenvolvimento desta tarefa, à semelhança do que aconteceu com Isabel, outros alunos mostraram-se intrigados com o facto de as duas retas que determinaram não coincidirem. A professora procurou explicar essa situação:

Professora: [...] Vocês viram um ponto que [achavam que] estava na reta... Como vocês não viram mais nenhum, eu sugeri um outro [ponto]... Calculámos à moda antiga a reta [ou seja, obtiveram por cálculo a expressão da reta que passava pelos dois pontos indicados], desenhei-a à mão... Depois fomos ver se o meu desenho [representação desta reta inicial, na calculadora gráfica] estava muito afastado do desenho [da reta de regressão] que a máquina fazia e vimos que quanto ao declive até nem estava muito mal [pelo facto de os seus valores estarem próximos]. Já a ordenada na origem calhou um bocado mal... Mas isto à mão! (Observação de aula, 13 de junho de 2011)

No excerto acima, Estela descreveu a sequência de passos realizados para chegarem à reta de regressão. Contudo, não incluiu explicações que ajudassem a perceber com profundidade o motivo pelo qual a reta de regressão obtida com a calculadora se ajusta melhor aos dados fornecidos do que a reta estimada inicialmente. No entanto, ainda na exploração dessa tarefa, na introdução da noção de correlação entre as variáveis em análise, Estela estabeleceu conexões entre algumas ideias: (1) a tendência observada no gráfico de dispersão para um aumento do *número de pontos* à medida que os *minutos de jogo* aumentam, com o declive positivo da reta de regressão e *correlação linear positiva*; e (2) a maior proximidade da reta de regressão aos pontos do diagrama de dispersão, o que traduz a existência de uma forte correlação linear entre as variáveis em estudo, tal como o excerto seguinte mostra:

Professora: Então, outra coisa... [observando os pontos do diagrama de dispersão] À medida que o tempo aumenta... Que os jogadores estão mais tempo em campo, em geral, eles marcam mais pontos... Certo? Então, por essa razão dizemos que há correlação linear positiva... Certo...? E quanto mais esses pontos [dados] se aproximam da reta... Quanto menor for a distância dos pontinhos [do diagrama de dispersão] à reta de regressão... For menor para todos eles, mais forte é essa correlação!

Isabel: É o que tu estavas a dizer! [diz a colega de carteira, em voz alta, para o Leonardo]

Leonardo: São aqueles quadrados...

Professora: São aqueles quadrados a ficar mais pequenos [referindo-se aos desvios que o aluno Leonardo tinha descoberto na sua calculadora gráfica de um modelo mais avançado] ... Como é que se mede essa correlação matematicamente... se é forte ou se é menos forte? À custa de cálculos que a máquina faz.... Aliás vocês têm a fórmula no livro. (Observação de aula, 13 de junho de 2011)

Apesar de Estela mencionar, de forma breve, os *quadrados* que o Leonardo tinha conseguido fazer na sua calculadora de um modelo mais avançado, parece desejar evidenciar a relação próxima entre *quadrados* eventualmente menores e a reta que melhor se ajusta à distribuição. Contudo, a maioria dos alunos não se apercebeu, efetivamente, do que tinha sido feito pelo Leonardo, dado que a professora apenas descreveu em voz alta à turma o conteúdo dos ecrãs da calculadora desse aluno, para dar uma ideia do que é possível alcançar com uma calculadora mais sofisticada. Contudo, na entrevista, Estela revela entender a ideia principal implícita na constituição da reta de regressão pelo método dos mínimos quadrados quando disse a seu respeito que: *A soma dos quadrados dos desvios é mínima.*

## 5.2 Tarefa População residente em Portugal

Em relação à tarefa que escolheu do manual adotado (ver anexo), a professora manifestou a intenção de discutir a utilidade do modelo de regressão linear. No seu desenvolvimento, solicitou aos alunos que introduzissem as listas de dados na calculadora e

através dela obtivessem o diagrama de dispersão e a reta de regressão tal como estavam exibidos no enunciado. A seguinte interação ocorreu sobre o propósito desse modelo:

Professora: Portanto... Qual o papel principal deste modelo linear, ou seja, desta reta de regressão? O que ela pede é para estimar, fazer uma estimativa... prever! [...] Porque é que este modelo não serve para eu imaginar qual será a população daqui a não sei quantos séculos, nem serve para imaginar quantas pessoas existiam há não sei quantos séculos atrás?

[...]

Mariana: Ao substituir o  $a$  por um ano [na equação da reta de regressão]... E a população dar um valor normal.

Professora: O que é a população dar um valor normal?

Leonardo: Superior a zero.

Professora: Superior a zero, pelo menos... Diz mais alto [dirigindo-se a Leonardo].

Leonardo: Ao prolongar-se a reta [à esquerda] vai passar por baixo de zero.

Professora: Exatamente. Se prolongarmos a reta [à esquerda] o que acontece?

Vários alunos: Tínhamos população negativa.

Professora: Isto é impossível. Portanto, há uns séculos atrás teríamos população negativa... Em contrapartida, se prolongássemos a reta [à direita]? O que acontecia?

Vários alunos: A população vai crescer.

Professora: A população aumentava. A população crescia infinitamente. Isto não é possível? Porquê? Está aí [no manual] uma sugestão.

(...)

Professora: “Não cabíamos cá todos”. Mais? Os recursos são...?

Professora e alunos: Limitados. (Observação de aula, 16 de junho de 2011)

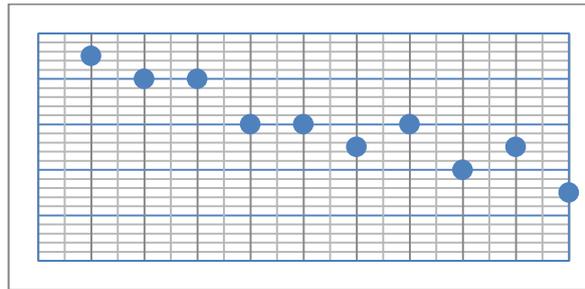
Nessa interação a professora tentou fazer com que os alunos se apercebessem da limitação do modelo linear na previsão a médio e longo prazo da evolução da população ou em estimar um valor aproximado dessa população nesses períodos. A professora acabou por fazer transparecer a ideia de que o modelo linear em causa era desapropriado para extrapolar em vários momentos no tempo, embora sem concretizar exemplos específicos.

### 5.3 Associação entre diagramas de dispersão e coeficientes de correlação

A professora escolheu do manual algumas tarefas que envolviam um conjunto de nuvens de pontos ou diagramas de dispersão aos quais se deveria fazer corresponder o respetivo valor de coeficiente de correlação linear de um conjunto de valores fornecidos. Estela referiu que, para resolver essas tarefas, os alunos deviam considerar ou imaginar a reta que melhor se ajustasse aos pontos de cada diagrama de dispersão, e aquela que melhor o fizesse era a que possuía coeficiente mais forte, caso contrário, seria a mais fraca. Estela também sugeriu o recurso à calculadora gráfica como uma primeira abordagem à questão. Para tal, os alunos teriam de atribuir uma escala à quadrícula, determinar as coordenadas de cada ponto da nuvem de pontos (ver figura 1), colocar esses dados na calculadora gráfica e determinar através dela a reta de regressão e o coeficiente de correlação. Na entrevista, quando questionada sobre essa recomendação, Estela referiu-se à oportunidade que teve de

mostrar aos alunos a utilidade da calculadora gráfica na exploração dos dados incluídos na tarefa:

Estela: [...] Para eles próprios verem que também podem meter na calculadora gráfica [os dados]... situações do manual... se tiverem dúvidas e não conseguirem ver mais ou menos de cabeça ... fazer uma estimativa [aventar uma possível resposta] [...] podem experimentar sem problemas de discutir, errar... aqui não há errar... há experimentar e concluir, colocar hipóteses e confirmá-las ou não. (Entrevista 2, 20 de junho de 2011)



**Figura 1** - Nuvem de pontos à qual correspondia o coeficiente de correlação -0.94.

Essa estratégia foi usada, sobretudo, para se alcançar o resultado do coeficiente de correlação. Quanto à realização das associações requeridas, de uma maneira geral, os alunos não revelaram dificuldades. Descreveram a correlação entre as duas variáveis observadas, em cada diagrama, em termos da *força* e do *signal* da relação. Ou seja, apoiando-se na observação do valor do coeficiente de correlação, para valores próximos de 1 ou -1 afirmavam que a correlação era forte e para valores próximos de zero afirmavam que era fraca. Embora Estela tenha confirmado cada correspondência em interação com os alunos, não lhes solicitou a justificação das suas respostas.

## 6 O estudo de dados bivariados na prática de Lia

As aulas de Lia são, habitualmente, estruturadas e organizadas em torno da exposição dos conteúdos (definições e exemplos) e da resolução de tarefas pelos alunos aos pares, ainda assim bastante conduzida por si. Essa professora mostra ter alguma preocupação em promover a participação dos alunos nas aulas, quer nos momentos de exposição dos conteúdos quer nos de correção dos trabalhos, solicitando-lhes respostas diretamente do lugar ou que apresentem as suas resoluções no quadro. A calculadora gráfica foi usada pontualmente no ensino do tema da Estatística, mas a professora deu-lhe particular destaque quando abordou o tópico dos dados bivariados. O estudo em torno dos dados bivariados foi desenvolvido por Lia no decorrer de três aulas de 90 minutos. O trabalho desenvolvido nessa temática consistiu, numa primeira fase, na apresentação de definições dos conceitos e de representações e, ainda, de

vários exemplos de aplicação desses conceitos e, numa segunda fase, na resolução de tarefas com o auxílio da calculadora gráfica. De entre as aulas, selecionámos três episódios, de modo a evidenciar aspetos centrais do conhecimento didático da professora desse tópico: o primeiro associado à introdução dos conceitos com o apoio de um *software* disponível no computador da sala de aula e os outros dois a tarefas propostas na aula (ver anexo).

### 6.1 O recurso multimédia - *escola virtual*<sup>3</sup>

Lia optou por usar o recurso multimédia *escola virtual* para apresentar conceitos e representações, além de vários exemplos de dados bivariados que queria ensinar aos alunos, do modo mais expedito possível, dado apenas lhe restaram as três últimas aulas do ano letivo para desenvolver essa temática. Através desse recurso escolheu e exibiu várias noções e representações (dados bivariados, diagrama de dispersão, coeficiente de correlação, centro de gravidade e reta de regressão) e mostrou os passos a seguir na calculadora gráfica para obter medidas estatísticas de duas variáveis em simultâneo, a equação da reta de regressão e o valor do coeficiente de correlação. Fez também alguns comentários que interligavam essas noções e representações.

Na apresentação inicial da *escola virtual* realçou alguns exemplos de conjuntos de duas variáveis estatísticas cujo estudo da existência de relação entre ambas parecia ser relevante e a representação de dados bivariados através de pares ordenados e num diagrama de dispersão. Ao abordar o diagrama de dispersão, indicou que o estudo de dados bivariados tinha em vista analisar a relação entre as duas variáveis. Quando usou a *escola virtual* sobressaíram algumas explicações que forneceu relativamente à utilidade do diagrama de dispersão, do coeficiente de correlação linear, do centro de gravidade e da reta de regressão, que analisamos de seguida.

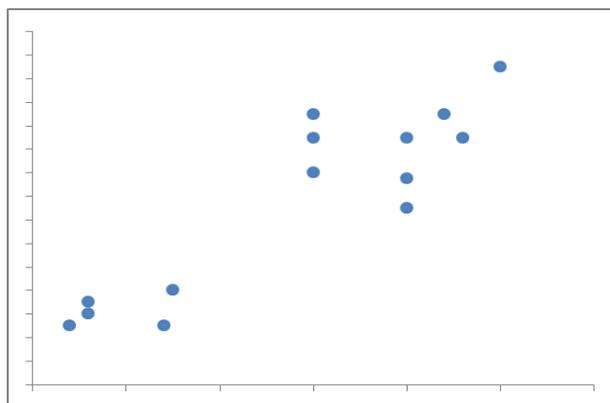
Ao solicitar aos alunos a análise da relação entre as variáveis temperatura máxima e temperatura mínima de várias cidades europeias, representadas num diagrama de dispersão (figura 2), Lia questionou os alunos sobre a relação entre as variáveis:

Professora: Será que existe alguma relação? Olhem para os valores...? O que é que acontece quando a temperatura mínima vai aumentando?

Vários alunos: Aumenta.

Professora: A temperatura máxima também vai aumentando. Há aqui uma relação entre elas. Quando uma aumenta a outra também aumenta, não de uma forma proporcional, não é?. Mas vai aumentando. [...] Há correlação positiva entre as duas variáveis, pois quando uma aumenta a outra também aumenta. (Observação de aula, 14 de junho de 2011)

<sup>3</sup> Recurso disponível *online*, preparado especialmente para a comunidade educativa, de uma editora nacional. Os conteúdos incluem aulas interativas, testes e exercícios, e outros documentos, para os diferentes níveis escolares.



**Figura 2** - Diagrama de dispersão das variáveis temperatura mínima e máxima.

Nessa interação, Lia estabeleceu uma relação entre a tendência global de crescimento que observou nos dados representados no diagrama de dispersão com uma correlação positiva. Nesse exemplo concreto existiam alguns dados que não seguiam a tendência, mas tal situação não foi mencionada.

Ao selecionar e analisar vários exemplos da *escola virtual* de diagramas de dispersão com distribuições diversificadas, a professora foi introduzindo o coeficiente de correlação linear como uma medida que toma valores numéricos entre -1 e 1 e classificou-o quanto ao sinal (positivo ou negativo) e à intensidade da correlação (perfeita, forte, moderada ou fraca). Ao sintetizar essa informação, na aula, referiu ainda:

Quando  $r=1$ , quando há correlação linear perfeita positiva, implica que os pontos estão alinhados sobre uma reta. (...) Imaginemos agora que nós temos um coeficiente de correlação  $r=-0.9$ . O que é que vai acontecer?... Os pontos não vão estar todos sobre uma reta, mas vão estar todos juntinhos [a uma reta]... vai haver uma reta próxima deles todos, dizemos [neste caso] que a correlação é forte e negativa. Quanto à intensidade da correlação, temos também a correlação moderada mas quanto mais o valor de  $r$  [coeficiente de correlação] se aproxima de zero ou a correlação é fraca ou não existe (...) à reta que se aproxima muito dos pontos nos vamos chamar reta de regressão. Ela tem uma equação. Se colocarmos os dados na calculadora [é possível] a calculadora dar-nos uma equação do tipo  $y=ax+b$ ... dá-nos o  $a$ , o  $b$  e o nosso coeficiente de correlação que é o  $r$ . (Observação de aula, 14 de junho de 2011)

No excerto transcrito, Lia deixou transparecer, também, o que entende por reta de regressão, e expressou a utilidade da calculadora gráfica na determinação da expressão analítica dessa reta, bem como na obtenção do valor do coeficiente de correlação linear.

Lia salientou da *escola virtual* um diagrama de dispersão que tinha a particularidade de ter o centro de gravidade marcado. Esse diagrama incluía, também, um eixo horizontal e outro vertical centrado no referido centro de gravidade. Na interação que teve com os alunos acerca desse gráfico, Lia realçou a utilidade do centro de gravidade na análise da tendência global dos dados e na confirmação do sinal da correlação:

Professora: Quando a correlação é negativa a maior parte dos pontos vai estar neste e neste [apontando para os quadrante pares formados com os eixos centrados no centro de gravidade da distribuição que estavam incluídos no diagrama de dispersão] mas não são todos, ok? E se a correlação for positiva?

Ivone: [aponta para os quadrantes ímpares].

Professora: Exato. Este centro de gravidade pode ajudar-nos a verificar se a correlação é positiva ou negativa... Mas olhando para o diagrama de dispersão vocês devem conseguir ver se a correlação é positiva ou não. (Observação de aula, 14 de junho de 2011)

Na referida interação, não emergiram outras explicações que ajudassem a perceber, por exemplo, a estreita relação entre o centro de gravidade e a reta de regressão. Na entrevista, Lia comentou que não relacionou o centro de gravidade com a reta de regressão na aula, sobretudo por falta de tempo. No entanto, também referiu que o entendimento da reta de regressão no ensino era prioritário.

A exposição dos conteúdos com o apoio da *escola virtual* terminou quando Lia referiu a utilidade da reta de regressão para se *estimar ou prever valores* da variável dependente ou da variável independente, dentro e fora dos intervalos de variação de cada uma delas, desde que os valores determinados fizessem sentido no contexto dos dados. Já nessa situação, Lia não deu indícios de conhecer que, dada uma reta de regressão, esta só deverá ser usada para estimar valores da variável dependente. A reta de regressão foi definida por Lia como sendo aquela *que melhor se aproxima de todos os pontos* e, de passagem, a professora referiu ainda que esta *era a [reta] dos mínimos quadrados*. No entanto, não explicou o significado da última designação, apenas tendo referido que era possível obter a equação dessa reta com a calculadora.

## 6.2 Tarefa A *evolução do crescimento do Pedro*

Lia selecionou essa tarefa do manual escolar e na sua apresentação aos alunos disse-lhes que teriam de analisar a relação entre as duas variáveis em estudo (idade e altura de uma pessoa) e recorrer à reta de regressão para estimar um valor para cada uma das variáveis. Informou, também, que teriam de utilizar a calculadora gráfica para representar esses dados num diagrama de dispersão e determinar os resultados pedidos.

Para a resolução dessa tarefa, Lia pretendia recorrer às duas calculadoras gráficas (de marcas diferentes) mais usadas na aula. Preparou o computador e começou por projetar uma das calculadoras usadas. Pediu um voluntário para trabalhar no computador a sua resolução. O aluno que se voluntariou, André, inseriu os dados nas listas e mostrou rapidamente o diagrama de dispersão com a reta de regressão incluída. Não foi feito nenhum comentário à

visualização obtida. A interação de Lia com os alunos centrou-se na equação para a reta de regressão:

Professora: [para o André] Consegues voltar para trás e mostrar os valores de  $a$ ,  $b$  da reta de regressão?

André: Sim.

Professora: ... Cá está! Tem ali [na tela] os valores  $a$  e  $b$  e o do coeficiente de correlação que é o  $r$ . Que neste caso é de 0.99. Isto quer dizer que há uma correlação de que tipo? Como podemos classificá-la?

Vários alunos: É forte.

Professora: É forte, não é? É quase perfeita! (...) Agora [nas alíneas seguintes] queremos fazer previsões. Quero utilizar a reta de regressão para dada uma idade, ver qual a possível altura e vice-versa... (...) Qual a altura do Pedro quando ele tinha 10 anos?

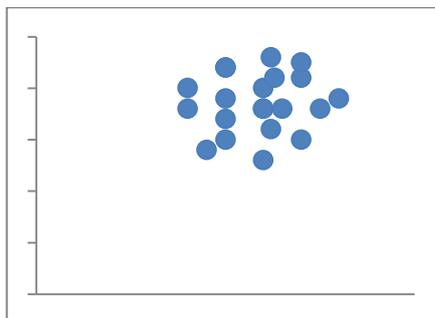
André: Não sei [determinar na calculadora]. (Observação de aula, 17 de junho de 2011)

Lia tinha a expectativa de encontrar, juntamente com os alunos, uma forma rápida de estimar o valor de uma variável sendo conhecido o valor da outra, tendo por base a reta de regressão determinada, tal como as últimas alíneas da tarefa aludiam. Contudo, ao não conseguir descobrir as instruções que pudessem conduzir aos resultados pedidos através das duas calculadoras, acabou por sugerir aos alunos que passassem para o caderno a equação da reta e que os determinassem com papel e lápis. Numa outra tarefa semelhante a professora acabou por sugerir uma outra estratégia, recorrendo à calculadora gráfica, para estimar o valor de  $x$  sendo conhecido o valor de  $y$ , através da interseção de duas retas: a de regressão e uma reta horizontal. No entanto, os alunos, de uma maneira geral, revelaram mais dificuldades em determinar com papel e lápis o valor de  $x$  sendo fornecido o valor de  $y$ , do que o valor de  $y$  conhecido o valor  $x$ , tendo por base a equação da reta de regressão em que  $x$  é a variável independente e  $y$  a dependente. Salienta-se, ainda, que nessa tarefa Lia não revelou conhecer que as retas de regressão (método dos mínimos quadrados) obtidas para estimar  $y$  em função de  $x$  ou para estimar  $x$  em função de  $y$ , não são, em geral, idênticas, servindo, cada uma delas, essencialmente para estimar certos valores da sua respetiva variável dependente.

### 6.3 Correspondência entre diagramas de dispersão e coeficientes de correlação

Na aula de Lia foram propostas algumas tarefas que solicitavam que se fizesse a correspondência entre diagramas de dispersão de distribuições diversas e valores de coeficientes de correlação. Nessas tarefas, os alunos forneceram as respostas corretas relativas às associações estabelecidas e, de uma maneira geral, os comentários justificativos foram sucintos e tiveram como base o sinal (positivo ou negativo) da tendência global dos dados e a *intensidade* da correlação observada (classificada habitualmente em: nula, fraca, moderada ou forte). Os argumentos usados pelos alunos ao comentar a *intensidade* da correlação tiveram em conta a *concentração* ou *dispersão* dos dados entre si no diagrama, não tendo sido

observado qualquer comentário em que se referissem, implícita ou explicitamente, à reta de regressão.



**Figura 3** - Diagrama de dispersão ao qual correspondia o coeficiente de correlação de valor zero.

Por exemplo, ao questionar os alunos sobre o possível coeficiente de correlação que corresponderia ao diagrama de dispersão da figura 3, Lia teve a seguinte intervenção:

Professora: Aqui há dúvidas? Qual é o coeficiente de correlação [que associam ao diagrama]?

Vários alunos: É zero!

Professora: Mas porquê? Vamos lá, digam?

Miguel: não há...

Professora: Porquê?

Rafael: Há dispersão.

Professora: Porque os pontos estão muito... muito dispersos. De facto, não existe uma nuvem ascendente...nem descendente. (Observação de aula, 17 de junho de 2011)

Nessa interação, observa-se que os alunos são muito sintéticos nas suas respostas. Quando Rafael referiu existir dispersão no gráfico, a professora aproveitou para fazer um comentário de modo a suportar a resposta desse aluno, ao invés de o questionar sobre a sua resposta.

## 7 Discussão e conclusão

As professoras deste estudo denotam um *conhecimento do currículo* que as leva a selecionar um conjunto de situações que visam os objetivos indicados pelo programa do ensino secundário vigente. No ensino de dados bivariados, ambas propõem algumas tarefas com potencial para promover o desenvolvimento do raciocínio e pensamento estatísticos dos alunos sobre dados bivariados e suas relações. No caso de Estela, as tarefas *A equipa de basquetebol do Porto* e *População residente em Portugal*, e no caso de Lia, a tarefa *A evolução do crescimento do Pedro*, podem facilitar um aprofundamento da situação real, subjacente aos dados fornecidos, com base no raciocínio sobre representações gráficas e sobre o modelo de regressão linear (SCHEAFFER, 2006). As tarefas de associação entre diagramas de dispersão/ nuvem de pontos e coeficientes de correlação, que as duas professoras

desenvolveram nas suas aulas, podem contribuir para o desenvolvimento do raciocínio dos alunos ao nível da compreensão de gráficos e aprofundamento dos conceitos (CURCIO; ARTZT, 1996; GARFIELD; BEN-ZVI, 2008).

No que diz respeito ao *conhecimento do ensino*, embora as práticas das duas professoras possam ser globalmente caracterizadas por uma abordagem de ensino direto, as diferenças que se evidenciam entre ambas parecem influenciar a percepção que detêm das dificuldades e obstáculos na aprendizagem dos dados bivariados. Lia centra a aula em si própria, optando por transmitir a informação que considera necessária, apresentar exemplos, fornecer explicações e colocar algumas questões aos alunos, o que reduz muito a possibilidade de participação destes na aula. Já Estela sustenta habitualmente a introdução e discussão dos conteúdos na realização das tarefas pelos alunos e, em simultâneo, vai colocando questões, orientando o trabalho a realizar, num ambiente que proporciona uma maior participação destes; ainda assim, nem sempre explora as situações de modo a promover o raciocínio dos alunos sobre dados e relações bivariados.

Ainda no domínio do *conhecimento do ensino*, observa-se, nas duas professoras, o desenvolvimento do tópico de dados e relações bivariados através da análise do diagrama de dispersão, do coeficiente de correlação e do uso da reta regressão com o apoio da calculadora gráfica, embora de modo superficial. Por exemplo, uma correlação positiva entre duas variáveis quantitativas foi sempre traduzida pelas duas professoras como *assim que uma aumenta a outra também aumenta*, mesmo quando se tratava de distribuições com vários dados que não seguiam essa tendência, aspeto apontado como crítico por Engel e Sedlmeier (2011). Estela optou por levar os alunos a usar a calculadora como primeira abordagem para calcular o coeficiente de correlação a partir dos dados representados em diagramas de dispersão. Essa abordagem parece prender-se mais com o uso da calculadora como um meio rápido de obtenção de resultados do que como um recurso que pode permitir trabalhar ideias intuitivas sobre correlação linear. Contudo, essa situação pode, igualmente, espelhar a importância que atribui à exploração de tarefas com essa ferramenta. No caso de Lia, o propósito principal do uso da calculadora no ensino dos dados bivariados foi a realização de cálculos ou a obtenção dos resultados.

As duas professoras definiram a reta de regressão de modo semelhante e, embora tenham feito uma abordagem diferenciada no seu ensino, não fizeram qualquer referência a ideias chave de estruturação dessa reta. Também, quando referiram o centro de gravidade, na aula, tentaram relacioná-lo com alguns conceitos mas, ainda assim, de modo pouco

aprofundado. Estela assumiu-o como um ponto pertencente à reta de regressão, mas não se deteve na exploração da sua importância e utilidade. Lia, por seu turno, não referiu que a reta de regressão passava sempre por esse ponto. O raciocínio sobre o modelo de regressão linear, interligando-o com o contexto em que se inserem os dados, nem sempre foi visível. Estela suscitou alguma discussão à volta das limitações do modelo de regressão linear embora esta pudesse ter sido mais completa, incidindo em análises de exemplos concretos em que o modelo não era conveniente para estimar a população e de outros exemplos em que era adequado, dando-se o devido valor aos dados. Lia não teve a preocupação de fazer uma análise cuidadosa dos dados em estudo, tendo focado os alunos unicamente na obtenção da equação da reta de regressão na calculadora. Embora essa reta não tenha sido referida como modelo linear, foi utilizada para estimar valores dentro do intervalo de variação de cada variável. Também não foram discutidas na aula as limitações do uso da reta de regressão na previsão de resultados.

As duas professoras revelam, de uma maneira geral, um conhecimento pouco profundo sobre o *aluno e a aprendizagem* de dados e relações bivariados. No caso das tarefas de associação entre diagramas e coeficientes de correlação, os alunos das duas turmas realizaram as associações corretas e descreveram o coeficiente de correlação em termos de *signal e força/intensidade*, tal como lhes foi ensinado. Estela não solicitou a justificação das suas respostas. Apesar de Lia ter pedido essa justificação, as respostas dos seus alunos foram sucintas e tiveram em conta, sobretudo, a dispersão/concentração dos dados no gráfico sem fazer alusão à reta de regressão. Quando os alunos das duas professoras observaram os diagramas de dispersão, na análise de cada situação, não ponderaram a forma das distribuições em termos da existência de grupos ou *outliers*, nem como estes poderiam alterar o valor do coeficiente de correlação. Essa experiência pode tê-los conduzido à ideia de que encontrar o valor do coeficiente de correlação por si só é suficiente para tirar conclusões sobre a existência de relação linear ou sobre a validade do modelo de regressão linear, tal como Engel e Sedlmeier (2011) alertam.

Na resolução da tarefa *A equipa de basquetebol do Porto* quando uma aluna associou o comentário da professora - *Quanto menor for a distância dos pontos do diagrama à reta de regressão, mais forte é a correlação* - às distâncias de cada ponto à reta de regressão representadas através de quadrados que um outro aluno conseguiu colocar no ecrã da sua calculadora mais avançada, Estela optou por não dar seguimento a esse comentário, desperdiçando, assim, uma oportunidade para aprofundar os conceitos de correlação e de reta

de regressão através das distâncias mencionadas numa análise residual (GARFIELD; BEN-ZVI, 2008). Lia ficou surpreendida com o facto de os alunos terem dificuldades em usar a calculadora para estimar um valor de  $y$  (ou de  $x$ ) de uma reta de regressão, sendo conhecido um valor de  $x$  (ou de  $y$ ), mas não conseguiu indicar, no imediato, instruções que permitissem chegar aos resultados. Essas estimativas podem ter induzido nos alunos a ideia errónea de que as retas de regressão  $y$  em função de  $x$  e  $x$  em função de  $y$  são sempre a mesma e, assim sendo, qualquer uma delas serve para estimar um valor de  $x$  ou de  $y$ . Essa situação não foi observada na prática de Estela.

Relativamente ao *conhecimento do conteúdo* destaca-se, nas duas professoras, o entendimento da relação bivariada associada, muitas vezes, à dependência funcional de duas variáveis, em que se atende sobretudo à tendência global dos dados, mesmo quando existem vários dados numa distribuição que não seguem essa tendência ou quando a observação do gráfico sugere outra relação que não a linear. Verifica-se a necessidade de as professoras promoverem análises de variação local e global e de recorrerem a descrições mais precisas sobre dados bivariados (ENGEL; SEDLMEIER, 2011; GARFIELD; BEN-ZVI, 2008). Relativamente à reta de regressão, apesar de ter havido uma referência ao método dos mínimos quadrados, de forma explícita, no caso de Lia, e de forma implícita, no caso de Estela, nenhuma das professoras sentiu necessidade de ensinar esse conceito com o devido aprofundamento. Adicionalmente, Lia revelou desconhecer propriedades estruturais da reta de regressão, ao solicitar aos alunos para prever valores não só da variável dependente mas também da independente de uma mesma reta de regressão. As duas professoras também revelaram algumas dificuldades em estabelecer relações mais profundas entre conceitos e representações no desenvolvimento desse tópico estatístico.

Em suma, em algumas instâncias, não parece ser considerada pelas professoras a complexidade de que se reveste para os alunos o raciocínio sobre dados bivariados. Foram referidas, ao longo deste texto, algumas oportunidades de ensino não aproveitadas pelas professoras para ajudar os alunos a aprofundar os seus conhecimentos e a superar dificuldades, o que também evidencia o grau de exigência requerido na aprendizagem dessa temática. Tais situações parecem estar fortemente relacionadas com a leitura que as professoras fazem do currículo escolar em Estatística e com o nível de conhecimento que detêm sobre dados bivariados. Este estudo aponta para a necessidade de um aprofundamento do conhecimento didático do professor sobre dados bivariados, que contribua mais significativamente para o desenvolvimento do raciocínio e do pensamento estatísticos dos

alunos. Uma maior compreensão do conhecimento didático do professor em Estatística nessa temática pode contribuir para desenvolver ações relacionadas com o currículo e com a formação inicial e contínua que visem melhorar o processo de ensino-aprendizagem de dados bivariados e suas relações.

### Agradecimentos

Este trabalho foi realizado no âmbito do Projeto *Desenvolver a literacia estatística: Aprendizagem do aluno e formação do professor* (contrato PTDC/CPE-CED/117933/2010) financiado por fundos nacionais através da FCT – Fundação para a Ciência e Tecnologia, de Portugal.

### Referências

- BATANERO, C.; DIAZ, C.; CONTRERAS, J. M.; ROA, R. El sentido estadístico y su desarrollo. *Números*, La Laguna, v. 83, n. 2, p. 7-18, Jul. 2013.
- BATANERO, C.; GODINO, J. D. Perspectivas de la educación estadística como área de investigación. In: LUENGO, R. (Ed.). **Líneas de investigación en didáctica de las matemáticas**. Badajoz: Universidad de Extremadura, 2005. p. 203-226.
- BEN-ZVI, D.; GARFIELD, J. Statistical literacy, reasoning and thinking: goals, definitions and challenges. In: BEN-ZVI, D.; GARFIELD, J. (Ed.). **The challenge of developing statistical literacy, reasoning and thinking**. Dordrecht, The Netherlands: Kluwer Academic Publishers, 2004. p. 3-15.
- BURRILL, G.; BIEHLER, R. Fundamental statistical ideas in the school curriculum and in training teachers. In: BATANERO C.; READING C. (Ed.). **Teaching statistics in school mathematics-challenges for teaching and teacher education: a joint ICMI/IASE study**. New York: Springer, 2011. p. 57-70.
- CURCIO, F. R.; ARTZT, A. F. Assessing students' ability to analyze data: reaching beyond computation. **The Mathematics Teacher**, Reston, v. 89, n. 8, p. 668-673, Nov. 1996.
- DEPARTAMENTO DO ENSINO SECUNDÁRIO. **Programa de Matemática A**, 10.º ano. Lisboa: ME, 2001. Disponível em: <<http://www.dgicd.min-edu.pt/ensinosecundario/>>. Acesso em: 28 Jun. de 2014.
- ENGEL, J.; SEDLMEIER, P. Correlation and regression in the training of teachers. In: BATANERO, C.; BURRILL, G.; READING, C. (Ed.). **Teaching statistics in school mathematics-challenges for teaching and teacher education**. New York: Springer, 2011. p. 247-258.
- ESTEPA, A.; BATANERO, C. Judgments of correlation in scatterplots: students' intuitive strategies and preconceptions. **Hiroshima Journal of Mathematics Education**, Hiroshima, v. 4, n. 25, p. 25-41. 1996.
- GARFIELD, J.; BEN-ZVI, D. **Developing students' statistical reasoning: connecting research and teaching practice**. Dordrecht, The Netherlands: Springer, 2008.



PONTE, J.; OLIVEIRA, H. Remar contra a maré: a construção do conhecimento e da identidade profissional na formação inicial. **Revista de Educação**, Lisboa, v. 11, n. 2, p. 145-163. 2002.

PONTE, J.; CHAPMAN, O. Mathematics teachers' knowledge and practices. In: GUTIERREZ, A.; BOERO, P. (Ed.). **Handbook of research on the psychology of mathematics education: past, present and future**. Rotterdam: Sense Publishers, 2006. p. 461-494.

SCHEAFFER, R. Statistics and mathematics: on making a happy marriage. In: BURRILL, G. (Ed.). **Thinking and reasoning with data and chance**. Reston, VA: NCTM, 2006. p. 309-321.

SHAUGHNESSY, J.; CHANCE, B. **Statistical questions from the classroom**. Reston, VA: NCTM, 2005.

SHULMAN, L. Those who understand: knowledge growth in teaching. **Educational Researcher**, USA, v. 15, n. 2, p. 4-14, Feb. 1986.

WILD, C.; PFANNKUCH, M. Statistical thinking in empirical enquiry. **International Statistical Review**, Oxford, v. 67, n. 3, p. 223-265. 1999.

**Submetido em Abril de 2014.**  
**Aprovado em Junho de 2014.**

## ANEXO

### Tarefa: A equipa de basquetebol do Porto

Na tabela abaixo estão indicados a idade, a altura e as médias por jogo dos minutos em campo e dos pontos marcados da equipa de basquetebol do Porto na época 200/2001 até à 13ª jornada segundo dados recolhidos em [www.infordesporto.pt](http://www.infordesporto.pt). Analisa as questões:

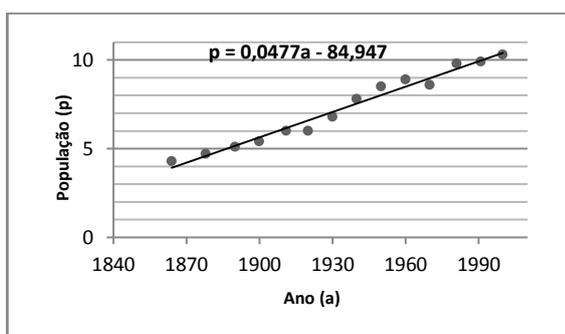
- Será que existe alguma relação entre a altura do jogador e os minutos em que ele está em campo? Dito de outra forma, será que os jogadores mais altos são solicitados mais vezes a jogar?
- E entre a idade e os pontos que marca? Será que os jogadores mais novos marcam mais pontos?
- E quanto à eficácia do jogo, será que existe relação entre minutos de jogo e pontos obtidos?

Jogador	Idade	Altura	Minutos	Pontos
A. Blackely	35	2.04	30	12.1
E. Évora	22	2.05	24.4	9.4
J. Pedrera	28	2.02	22.4	9.6
N. Marçal	25	2.05	23.8	15
N. Perdigão	22	1.92	21.8	8.1
N. Quidiongo	24	1.80	5.9	4.0
K. Vulin	25	2.04	31.2	15.4
P. Cunha	20	2.00	3.5	1.0
Rui Santos	30	1.88	30.1	8.7
J. Rocha	25	2.00	10.9	4.1
Raúl Santos	31	2.03	3.8	1.4
A. Hardad	31	1.98	28.8	14.9

### Tarefa: População residente em Portugal

Na tabela estão alguns dados sobre a população residente em Portugal desde 1864 até 2000. O diagrama de dispersão relativos a estes dados assim como a respetiva reta de regressão e a sua equação estão representados na figura abaixo.

Ano (a)	População em milhões (p)	Ano (a)	População em milhões (p)
1864	4,3	1940	7,8
1878	4,7	1950	8,5
1890	5,1	1960	8,9
1900	5,4	1970	8,6
1911	6,0	1981	9,8
1920	6,0	1991	9,9
1930	6,8	2000	10,3



Explique por que razão o modelo linear apresentado não pode ser adequado para: (a) estimar o número aproximado de habitantes, em Portugal, há uns séculos e (b) prever a evolução da população portuguesa a muito longo prazo (relacione uma tal previsão com os recursos, alimentares e outros, necessariamente limitados).

### Tarefa: A evolução do crescimento de Pedro

Na tabela seguinte estão representados alguns dados relativos à altura (cm) do Pedro. Recorre à calculadora e:

Idade	2	3	4	6	8	12	15
Altura	78	84	92	110	120	145	160



- (1) Representa a nuvem de pontos associada aos dados.
- (2) Determina a equação da reta de regressão e faz a respetiva representação gráfica.
- (3) Determina os valores aproximados: da altura do Pedro quando tinha 10 anos; da idade do Pedro quando a sua altura era de 150 cm.