Fernandes, J. A., Martinho, M. H., Tinoco, J., & Viseu, F. (Orgs.) (2013). *Atas do XXIV Seminário de Investigação em Educação Matemática*. Braga: APM & CIEd da Universidade do Minho.

O conhecimento didático de uma professora no ensino da relação bivariada na Estatística*

Sandra Quintas[1], *Hélia Oliveira*[2], *Rosa Tomás Ferreira*[3]
[1]Unidade de Investigação do IEUL, sandramquintas@gmail.com
[2]Instituto de Educação da Universidade de Lisboa, hmoliveira@ie.ul.pt
[3]Faculdade de Ciências da Universidade do Porto & CMUP, rferreir@fc.up.pt

Resumo. *Este estudo debruça-se sobre o conhecimento didático de uma professora de Matemática no ensino da relação bivariada no tema da Estatística na disciplina de Matemática A, do ensino secundário. Em particular, foca-se no conhecimento do ensino de uma professora no âmbito deste tema e da forma como este se relaciona com outras dimensões do seu conhecimento didático. Os resultados evidenciam a complexidade de que se reveste o conhecimento do ensino da relação bivariada relativamente a como apoiar os alunos na análise e interpretação do coeficiente de correlação e no raciocínio com o modelo de regressão linear. Mostram também a estreita articulação entre o conhecimento do ensino e o conhecimento dos alunos e da aprendizagem no que diz respeito ao raciocínio sobre relações bivariadas.*

Palavras-chave: conhecimento didático; dados bivariados; raciocínio estatístico.

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 várias áreas tais como na Psicologia, na Ciência, na Educação Matemática e na Educação Estatística. Os conceitos ligados ao estudo de dados bivariados, nomeadamente, associação, correlação e regressão linear, são referidos no programa de Matemática A do 10.º ano, no tópico das distribuições bidimensionais, o tópico final do tema da Estatística. Vários autores referem a complexidade do ensino e aprendizagem sobre dados e relações bivariadas (Engel & Sedlmeier, 2011; Estepa & Batanero, 1996; Garfield & Ben-Zvi, 2008; Mugabe, Fernandes & Correia, 2012). A compreensão da regressão e correlação exige conhecimento básico sobre funções e, acima de tudo, a consideração da variação à volta de uma possível tendência (Engel & Sedlmeier, 2011). É neste contexto que surge o

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

presente estudo com o objetivo de compreender o conhecimento didático de uma professora, nomeadamente o seu conhecimento do ensino da relação bivariada e as conexões deste domínio com outros domínios do seu conhecimento didático que se evidenciam neste tema.

O conhecimento didático do professor em Estatística

O conhecimento profissional do professor de Matemática desdobra-se em várias dimensões, nomeadamente, o conhecimento na ação relativa à prática letiva, à prática não letiva, à profissão e ao seu próprio desenvolvimento profissional (Ponte & Oliveira, 2002). Ponte e Oliveira (2002) designam a dimensão do conhecimento profissional chamado a intervir diretamente na prática letiva por conhecimento didático. Apesar de, habitualmente, a Estatística ao nível escolar ser estudada na disciplina de Matemática vários autores reconhecem a especificidade do seu ensino (por exemplo, Garfield & Ben-Zvi, 2008). O modelo do conhecimento didático do professor em Estatística adotado neste trabalho tem como fonte de inspiração o modelo de Ponte e Oliveira (2002) que incorpora quatro domínios (o conhecimento de Estatística, do currículo, dos alunos e da aprendizagem, e do ensino), bem como um conjunto de aspetos apontados por Batanero e Godino (2005) que integram o conhecimento do professor que ensina Estatística. O conhecimento de Estatística refere-se ao conhecimento da disciplina e das interpretações dos seus conceitos, representações e procedimentos (Ponte & Oliveira, 2002). Integra a capacidade de reflexão sobre a natureza do conhecimento estatístico e sobre o significado de conceitos e procedimentos (Batanero & Godino, 2005). Inclui o conhecimento de ideias estatísticas essenciais (e.g. dados; variação; representações; correlação) e suas interligações (Batanero, Diaz, Contreras & Roa, 2013). O conhecimento do currículo inclui o conhecimento das grandes finalidades e objetivos do currículo escolar e sua articulação vertical e horizontal (Ponte & Oliveira, 2002). 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 e ainda o conhecimento dos diversos níveis de compreensão dos alunos (Batanero & Godino, 2005; Ponte & Oliveira, 2002). O conhecimento do ensino compreende a capacidade de planificação da sequência de conteúdos, inclui a capacidade de ajustar conteúdos a diferentes níveis de ensino, tendo em conta o grau de profundidade com que estes necessitam de ser tratados e relacionados e ainda estratégias de ensino adotadas (por exemplo, uso de tecnologia).

Abarca também a capacidade de ajudar os alunos a desenvolver o raciocínio e pensamento estatístico (Batanero et al., 2013). Este domínio do conhecimento didático é mais restrito e focado que o referido por Ponte e Oliveira (2002), na medida que se centra nos aspetos específicos do tema estatístico em estudo.

O ensino da relação bivariada

Para ensinar Estatística o professor precisa de ter experiência e familiaridade com elementos específicos do pensamento estatístico, nomeadamente, reconhecimento da necessidade de dados e seu conhecimento contextual, atenção à variação e raciocínio com modelos, integrando-os na sua prática (Wild & Pfannkuck, 1999). No ensino da Estatística, a compreensão do raciocínio sobre dados bivariados, também conhecido por raciocínio covariacional, deve suportar algo mais do que raciocinar sobre diagramas de dispersão, correlação, regressão e funções (Garfield & Ben-Zvi, 2008). 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. No entanto, podem surgir dificuldades na interpretação do coeficiente de correlação linear (Engel & Sedlmeier, 2011; Shaughnessy & Chance, 2005). Por exemplo, um resultado de correlação linear alto não implica, por si só, a validade do modelo de regressão linear, havendo necessidade de se examinar cuidadosamente representações gráficas dos dados, tais como o diagrama de dispersão, devido à possibilidade de esta medida ser muito influenciada por *outliers*. Estes autores também referem exemplos de conjuntos de dados bivariados 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 este caso, uma afirmação mais precisa é de que *valores acima da média de uma das variáveis correspondem a valores acima da média da outra variável*. Este conhecimento mais pormenorizado poderá contribuir para uma melhor apreciação da variação local e entendimento da fórmula do coeficiente de correlação.

Engel e Sedlemeier (2011) registam que, com frequência, este tipo de dados é trabalhado, na sala de aula, como uma dependência funcional de duas variáveis na Matemática, descurando-se a variação dos dados, o que consideram poder decorrer de uma falta de preparação na formação inicial sobre como lidar com esses conteúdos.

Em linhas gerais, diversos autores declaram que o tipo de tarefas propostas e a forma como são trabalhadas na sala de aula influenciam em grande medida a qualidade das aprendizagens dos alunos (Garfield & Ben-Zvi, 2008; Ponte, 2005; Scheaffer, 2006). Em particular, na Estatística, destaca-se a relevância de se propor tarefas que incluam dados reais, elementos sobre o seu contexto e questões que valorizem os dados (Curcio & Artzt, 1996). Scheaffer (2006) acrescenta que tarefas desta natureza, que estimulam, em especial a análise de dados, facilitam o desenvolvimento do raciocínio estatístico dos alunos. Garfield e Ben-Zvi (2008) sugerem que atividades de estabelecimento de conexões entre valores de coeficiente de correlação e diagramas de dispersão pode permitir que os alunos desenvolvam melhor sentido dos diferentes níveis de covariação e entendimento acerca de fatores que influenciam que o coeficiente de correlação tenha valor maior ou menor. Atividades deste cariz são relevantes no desenvolvimento da capacidade de leitura de gráficos e apreciação das representações usadas, nomeadamente, as suas vantagens e desvantagens no processo de sumarização e de análise dos dados (Curcio & Artzt, 1996), no aprofundamento dos conceitos envolvidos e desenvolvimento do raciocínio estatístico dos alunos (Garfield & Ben-Zvi, 2008). A tecnologia pode proporcionar a visualização de representações diversificadas, o estabelecimento de conexões entre elas, e ainda a exploração e manipulação dos dados, dando, assim, um apoio significativo à análise de dados (Engel & Sedlmeier, 2011).

Metodologia

Esta comunicação insere-se numa investigação mais ampla, qualitativa e de índole interpretativa sobre o conhecimento didático do professor no ensino da Estatística, no ensino secundário. Este texto refere-se a uma das três professoras participantes, Estela, que tem uma experiência profissional de mais de 20 anos.

Na recolha dos dados que informam esta comunicação foram usados diversos métodos, nomeadamente: (1) observação participante, com registo áudio e vídeo de aulas, numa turma de 10.º ano constituída por 25 alunos; (2) duas entrevistas semi-estruturadas à professora com registo áudio, antes e após a realização das aulas observadas; e (3) recolha documental dos materiais utilizados pela professora nestas aulas, designadamente fichas de trabalho. A análise de dados foi efetuada de forma descritiva e interpretativa, a partir da identificação de três episódios, cada um associado a uma tarefa realizada na aula. Estas tarefas foram escolhidas por abordarem os aspetos centrais do estudo da relação bivariada.

As aulas são habitualmente estruturadas e organizadas em torno da resolução das tarefas, que na maioria das vezes são trabalhadas pelos alunos com o apoio da calculadora gráfica. Estela tem a preocupação de envolver os alunos, quer nos momentos de exposição dos conteúdos quer nos de correção dos trabalhos, solicitando-lhes que respondam a questões ou que vão ao quadro ou ao computador (cujo ecrã é projetado na tela) mostrar a sua resolução de alguma questão com a calculadora gráfica. Todos os alunos possuem calculadoras gráficas, a maioria da mesma marca, e há alguns que possuem os modelos mais avançados.

O estudo da relação bivariada na prática de Estela

O estudo em torno das distribuições bidimensionais foi desenvolvido no decorrer de três aulas de 90 minutos, por conseguinte, as tarefas propostas nas aulas sobre esta temática foram sobretudo trabalhadas com o propósito de introduzir os conceitos e representações e ainda fornecer explicações que a professora considerou necessárias.

A equipa de basquetebol do Porto

Estela explicou que a tarefa *A equipa de basquetebol do Porto* (ver anexo) que preparou para levar à sala de aula foi reformulada aproveitando uma outra mais antiga que possuía. Manteve os dados reais e a reformulação consistiu na incorporação de três questões adicionais plausíveis de serem investigadas, em que cada uma delas relacionava duas variáveis em estudo. Na sala de 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 especialmente a atenção para a tendência global dos dados, mesmo quando uma aluna mencionou a existência de dados que não seguiam essa tendência. Adicionalmente, a professora deixou transparecer o entendimento que os alunos deveriam ter sobre a reta de regressão:

> Prof.: (…) O que é que reparam… Acontece que à medida que o tempo aumenta vocês vêm que os pontos também aumentam, ou não?
>
> Alguns alunos: sim.
>
> Joana: Há exceções.

> Prof.: Então diz…Há exceções… mas a maioria [dos pontos]…
>
> Alguns alunos: sim.
>
> (…)
>
> Prof.: (…) 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 (…)

Estela indicou que a reta de regressão servia para modelar a "tendência global dos dados" que expressa no diagrama de dispersão. Também definiu a reta de regressão como a "reta que melhor se ajusta aos pontos do diagrama de dispersão".

No desenvolvimento desta tarefa Estela orientou o trabalho dos 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 desta representação). Depois explicou como determinar a equação da reta de regressão na calculadora gráfica. Na perspetiva de Estela esta 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.

Na exploração desta tarefa, quando os alunos confrontaram estas duas retas (a estimada e a de regressão) chegaram, de uma maneira geral, à conclusão de que elas eram diferentes mas que não estavam muito afastadas entre si, quando visualizadas em simultâneo sobre o diagrama de dispersão na calculadora gráfica. Na interação que teve com os alunos, Estela começou por lhes solicitar que representassem as duas retas na calculadora gráfica de modo a poderem compará-las:

> Prof.: Desenharam 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 se determinou 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 dos declives e ordenadas na origem não serem iguais nas duas retas]
> Prof.: 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)?

(…)

Prof.: [Dá as instruções para se chegar 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.

(…)

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

Vários alunos: Nem por isso. Não.

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 só havia a garantia de que um deles, o centro de gravidade, pertencia à reta de regressão que se procurava. Acabou por ser explicitamente assumido, pela professora, que a reta de regressão é aquela que melhor se ajusta à nuvem de pontos considerada e que 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 que tinham determinado através da calculadora gráfica:

Prof.: Agora podem verificar se aquele ponto [centro de gravidade] que eu disse que estava na reta de regressão, está lá.

(…)

Jaime: Está. Certíssimo.

Prof.: Não estou aqui para enganar ninguém como veem! (…) Afinal a reta passa mesmo por esse ponto!... Que é chamado centro de gravidade.

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 esta situação:

(…) 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, obteve-se à mão 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 dos seus valores estarem próximos]. Já a ordenada na origem calhou um bocado mal…mas isto à mão!

Neste excerto, Estela descreveu a sequência de trabalho que seguiram para chegar 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 determinada inicialmente. No entanto, ainda na exploração desta tarefa, na introdução da noção de correlação entre as variáveis em análise, Estela estabeleceu conexões entre algumas ideias:

> Prof.: 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 mesa, em voz alta, para o Leonardo]
>
> Leonardo: São aqueles quadrados…
>
> Prof.: São aqueles quadrados a ficar mais pequenos [referindo-se aos resíduos que o aluno Leonardo tinha descoberto na sua calculadora… 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.

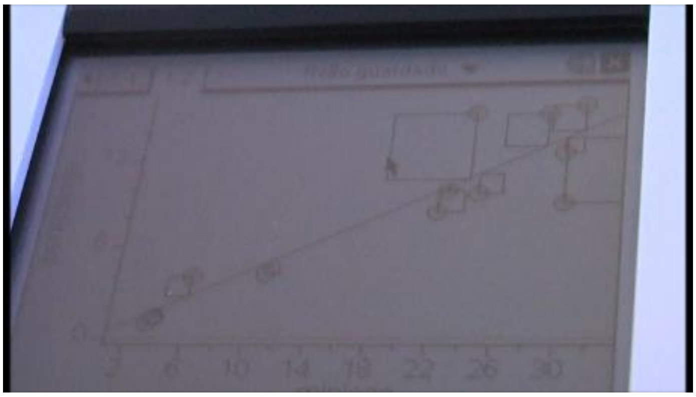


Figura1. Ecrã que inclui os resíduos obtidos pelo Leonardo na sua calculadora.

Neste excerto, a professora tentou associar 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. Apesar de Estela mencionar de forma breve os "quadradinhos" que o Leonardo tinha conseguido fazer na sua calculadora gráfica mais avançada, parece desejar evidenciar a relação próxima entre "quadrados" eventualmente mais pequenos e a reta que melhor se ajusta à distribuição. Esta ideia poderia ter sido usada para justificar o facto da reta de regressão se ajustar melhor aos dados do que a reta estimada pelos alunos. 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 do aluno, para dar uma ideia do que é possível alcançar com uma calculadora mais avançada. Não aproveitou para questionar nem explicar o que os referidos "quadradinhos", cujos tamanhos variavam, poderiam efetivamente significar.

População residente em Portugal

Nesta tarefa que Estela escolheu do manual adotado (ver anexo), a professora manifesta a intenção de discutir a utilidade do modelo de regressão linear. No seu desenvolvimento, a professora 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 deste modelo:

> Prof.: 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.
>
> Prof.: O que é a população dar um valor normal?
>
> Leonardo: Superior a zero.
>
> Prof.: Superior a zero, pelo menos… Diz mais alto [disse para o Leonardo]?
>
> Leonardo: Ao prolongar-se a reta [à esquerda] vai passar por baixo de zero.
>
> Prof.: Exatamente. Se prolongarmos a reta [à esquerda] o que acontece?
>
> Vários alunos: Tínhamos população negativa.
>
> Prof.: 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.
>
> Prof.: A população aumentava. A população crescia infinitamente. Isto não é possível? Porquê? Está aí [no manual] uma sugestão.
>
> (…)
>
> Prof.: "Não cabíamos cá todos". Mais? Os recursos são…?

Prof. e alunos: Limitados.

Nesta interação a professora tenta fazer com que os alunos se apercebam da limitação do modelo linear na previsão a médio e longo prazo da evolução da população portuguesa 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. Contudo, não foram discutidos exemplos de momentos em que este modelo poderia, eventualmente, ser útil para estimar a população.

Associação entre diagramas de dispersão/nuvem de pontos 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 estas 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, teriam de atribuir uma escala à quadrícula, determinar as coordenadas de cada ponto da nuvem de pontos (figura 2), colocar esses dados na calculadora gráfica e determinar através dela a reta de regressão e o coeficiente de correlação. Na entrevista, quando questionada sobre esta recomendação, Estela referiu-se à oportunidade que teve para mostrar a utilidade da calculadora gráfica na exploração dos dados incluídos na tarefa:

> … 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.

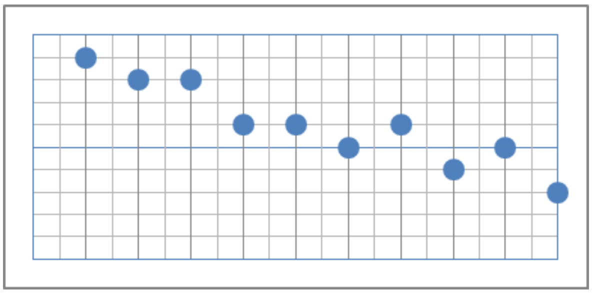


Figura 2. Nuvem de pontos à qual correspondia o coeficiente de correlação -0.94.

Esta estratégia foi usada sobretudo para se alcançar o resultado do coeficiente de correlação. Quanto à realização das associações requeridas, os alunos de uma maneira geral, não revelaram dificuldades em realizá-las. Os alunos descreveram a correlação entre as duas variáveis observadas em cada diagrama, em termos da “força” e sinal da relação. Ou seja, os alunos apoiaram-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 a correlação era fraca. Embora Estela tenha confirmado cada correspondência em interação com os alunos, não solicitou a justificação das suas respostas.

A concluir

Estela denota um *conhecimento do currículo* que a leva a selecionar um conjunto de situações que visam os objetivos indicados pelo programa do ensino secundário. No ensino da relação bivariada propôs tarefas com potencial para promover o desenvolvimento do raciocínio e pensamento estatísticos dos alunos (Curcio & Artzt, 1996; Scheaffer, 2006; Garfield & Ben-Zvi, 2008). As duas primeiras têm potencial para um aprofundamento da situação real, subjacente aos dados fornecidos, com base no raciocínio sobre o modelo de regressão linear. As tarefas de associação entre diagramas de dispersão e coeficientes de correlação linear podem contribuir para o desenvolvimento do raciocínio dos alunos ao nível da compreensão da leitura de gráficos e aprofundamento dos conceitos envolvidos.

Relativamente ao *conhecimento de Estatística* destaca-se na Estela o entendimento da relação bivariada associada, muitas vezes, à dependência funcional de duas variáveis na Matemática, em que se atende sobretudo à tendência global dos dados (Shaughnessy & Chance, 2005). Por exemplo, a correlação positiva entre duas variáveis foi traduzida como *assim que uma delas aumenta, a outra também aumenta*, mesmo quando existiam alguns dados na distribuição que não acompanhavam essa tendência. Ao nível do *conhecimento do ensino* da relação bivariada observaram-se algumas dificuldades, apontadas na literatura, sobre como proceder e apoiar os alunos na análise e interpretação do coeficiente de correlação e no raciocínio com o modelo de regressão. Nas tarefas de associação entre diagramas e coeficientes de correlação, os alunos descreveram o coeficiente de correlação em termos de sinal e “força” e não lhes foi solicitada a justificação das respostas. Na análise de cada situação não foi ponderada a forma das distribuições em termos da existência de grupos ou *outliers*, nem como estes poderiam alterar o valor do coeficiente de correlação. Esta experiência pode conduzir os alunos à ideia de que a avaliação do valor do coeficiente de correlação por si só é suficiente para se tirar conclusões sobre a validade do modelo de regressão linear. A estratégia avançada por Estela de usarem a calculadora para calcular o coeficiente de correlação dos dados representados em diagramas de dispersão pareceu ter mais um propósito de obtenção do resultado do que proporcionar desenvolvimento de ideias intuitivas sobre correlação linear. De facto, é importante que no ensino da relação bivariada se atenda a fatores que possam influenciar diferentes níveis de covariação no desenvolvimento do raciocínio covariacional nos alunos (Garfield & Ben-Zvi, 2008). Ainda no domínio do conhecimento do ensino, o raciocínio sobre o modelo de regressão linear, interligando-o com o contexto em que se inserem os dados, não foi visível na tarefa *A equipa de basquetebol do Porto*. Contudo, na tarefa *População residente em Portugal* a professora suscitou alguma discussão em torno das limitações do modelo de regressão embora pudesse ter incidindo também, por exemplo, em análises de exemplos concretos dentro e fora do intervalo de variação da variável “população”, dando-se o devido valor aos dados.

A análise da prática de Estela revela também a necessidade de um forte conhecimento dos *alunos e da aprendizagem* que considere a especificidade do raciocínio sobre relações bivariadas. Por exemplo, na tarefa *A equipa de basquetebol do Porto*, na comparação entre “a reta que melhor se ajusta aos dados” a partir de dois pontos indicados com a reta de regressão que foi determinada diretamente da calculadora gráfica, ao contrário do que a professora estava à espera, a justificação da sua proximidade não foi suficiente para que todos os alunos aceitassem que a reta de regressão obtida na calculadora era a que procuravam. E apesar de Estela revelar perceber a importância da análise residual na avaliação da associação linear, não tirou partido da tecnologia disponível para o esclarecimento dessas dúvidas. Aproveitando o trabalho exibido pelo Leonardo na sua calculadora seria possível estabelecer uma comparação entre os tamanhos dos “quadrados” de ambas as retas, dado que a reta de regressão é aquela cuja soma das áreas dos quadrados é mínima.

A análise da prática da professora relativamente ao ensino da relação bivariada permite observar a forte conexão entre o *conhecimento do ensino* e o *conhecimento dos alunos e da aprendizagem* neste tema. Efetivamente, em alguns aspetos não parece ser considerada pela professora a complexidade de que se reveste para os alunos o raciocínio sobre dados bivariados que a leve a explorar com a necessária profundidade as noções e representações fundamentais que emergem a partir das tarefas que propõe.

Referências bibliográficas

Batanero, C., Diaz, C., Contreras, J. & Roa, R. (2013). El sentido estadístico y su desarrollo. *Números*, *83*, 7-18.

Batanero, C. & Godino, J. (2005). Perspectivas de la educación estadística como área de investigación. In R. Luengo (ed.), *Líneas de investigatión en Didáctica de las Matemáticas* (pp. 203-226). Badajoz: Universidad de Extremadura.

Ben-Zvi, D., & Garfield, J. (2004). *The challenge of developing statistical literacy, reasoning and thinking*. Dordrecht, The Netherlands: Kluwer.

Curcio, F. & Artzt, A. (1996). Assessing students ability to analyze data: Reaching beyond computation. *The Mathematics Teacher*, *89*, 668-673.

Engel, J. & Sedlmeier, P. (2011). Correlation and regression in the training of teachers. In C. Batanero, G. Burril, & C. Reading (Eds.), *Teaching statistics in school mathematics-challenges for teaching and teacher education* (pp. 247-258). New York: Springer.

Estepa, A. & Batanero, C. (1996). Judgments of correlation in scatterplots: Students' intuitive strategies and preconceptions. *Hiroshima Journal of Mathematics Education*, *4*, 24-41.

Garfield, J. & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching Practice*. Dordrecht, The Netherlands: Springer.

Mugabe, D. A., Fernandes, J. A., Correia, P. F. (2012). Avaliação da associação Estatística num diagrama de dispersão por estudantes universitários. In H. Pinto, H. Jacinto, A. Henriques, A. Silvestre, C. Nunes (Orgs.), *Atas do XXIII Seminário de Investigação em Educação Matemática* (pp. 403-414). Coimbra: APM.

Ponte, J. (2005). Gestão curricular em Matemática. In GTI (Ed.), *O professor e o desenvolvimento curricular* (pp. 11-34). Lisboa: APM.

Ponte, J. & Oliveira, H. (2002). Remar contra a maré: A construção do conhecimento e da identidade profissional na formação inicial. *Revista de Educação*, *11*(2), 145-163.

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

Shaughnessy, J. & Chance, B. (2005). *Statistical questions from the classroom*. Reston, VA: NCTM.

Wild, C., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *International Statistical Review*, *67*(3), 223-265.

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 2000/2001 até à 13ª jornada segundo dados recolhidos no site da Infordesporto: www.infordesporto.pt. Analisa as questões:

(a) Será que existe alguma relação entre a altura do jogador e os minutos que este está em campo? Dito de outra forma, será que os jogadores mais altos são solicitados mais vezes a jogar?

(b) E entre a idade e os pontos que marca? Será que os jogadores mais novos marcam mais pontos?

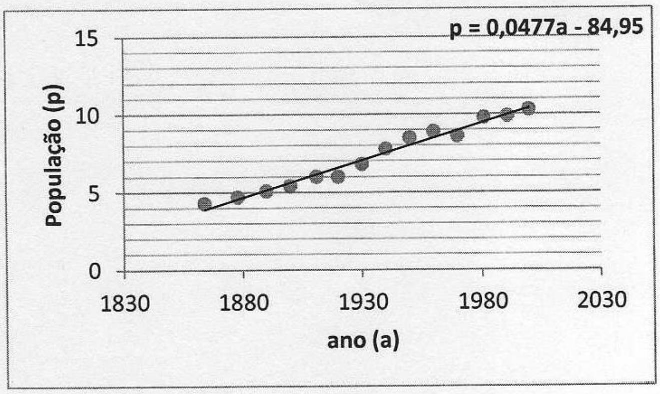
(c) E quanto à eficácia do jogo, será que existe relação entre os minutos de jogo e os pontos obtidos?

Jogador	Idade	Altura	Minutos/Jogo	Pontos/Jogo
Anthony Blackely	35	2.04	30	12.1
Elvis Évora	22	2.05	24.4	9.4
José Pedrera	28	2.02	22.4	9.6
Nuno Marçal	25	2.05	23.8	15
Nuno Perdigão	22	1.92	21.8	8.1
Nuno Quidiongo	24	1.80	5.9	4
Kevin Vulin	25	2.04	31.2	15.4
Paulo Cunha	20	2.00	3.5	1
Rui Santos	30	1.88	30.1	8.7
João Rocha	25	2.00	10.9	4.1
Raúl Santos	31	2.03	3.8	1.4
Bob Harsad	31	1.98	28.8	14.9

Tarefa ***População residente em Portugal***

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

Ano	Pop. em milhões	Ano	Pop. em milhões
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 atrás apresentado não pode ser adequado para:

(a) estimar o número aproximado de habitantes, em Portugal, há uns séculos;

(b) prever a evolução da população portuguesa, a muito longo prazo (relacione uma tal previsão com os recursos, alimentares e outros, necessariamente limitados).