

por Maria Eugénia Ferrão

[Departamento de Matemática Universidade da Beira Interior]

Ainda os *Rankings* e a Estatística

O modelo de componentes de variância aplicado aos dados da amostra portuguesa do PISA 2000 ilustra a importância da precaução na interpretação dos resultados face a eventuais limitações metodológicas. A utilização de modelos estatísticos para a criação de listas ordenadas de escolas sem ter em consideração variáveis de controlo socioeconómico constitui um "pecado" para com os alunos, professores e escolas de contextos desfavorecidos.

Introdução

Em Portugal, à semelhança de outros cinco países europeus¹, procede-se à divulgação periódica dos resultados nos exames nacionais através da publicação das listas ordenadas de escolas, usualmente conhecidas por *rankings*. A divulgação dos resultados da avaliação das escolas está prevista no artº16º do capítulo IV da Lei do Sistema de Avaliação da Educação não Superior². Os nove objectivos do sistema de avaliação enunciados na Lei podem classificar-se em dois grupos, de acordo com a sua finalidade: (a) melhoria da qualidade do sistema educativo; (b) produção de informação.

Este último propósito não se refere exclusivamente às escolas. A expectativa dos governantes e dos cidadãos é que o desempenho e a qualidade dos serviços de qualquer instituição pública sejam passíveis de ser resumidos através de alguma estatística ou indicador de desempenho.

A controvérsia em torno do assunto tem sido tratada exaustivamente em periódicos científicos internacionais, com algumas consequências na prática avaliativa dos sistemas educativos. Segundo Bird, Cox, Farewell, Goldstein e Smith [1] na República da Irlanda e na Austrália (New South Wales) há legislação contra o uso das listas ordenadas das escolas baseadas nos resultados dos exames. Aqueles autores sintetizaram a utilidade e importância dos indicadores de desempenho do seguinte modo "... quando bem feito, pode ser bastante produtivo para os envolvidos. Quando mal feito, pode ser muito custoso, e não apenas inútil mas prejudicial e até mesmo destrutivo.". Abrantes [2] referiu a forte pressão que os professores portugueses sentem como consequência dos resultados dos exames do ensino secundário ou pela má posição que Portugal ocupa nas comparações internacionais.

A temática também tem sido debatida em Portugal [4], [5]. No entanto, a ligação entre o conhecimento científico e metodológico e a concepção, planeamento e concretização da avaliação do sistema educativo ainda é ténue. Não se sabe, por exemplo, se as diferenças verificadas na série cronológica dos resultados dos exames nacionais são devidas à melhoria/pioria do processo ensino-aprendizagem, tal como sugere Justino [6], ou tão só devidas às características das provas. Há mais de 130 anos que os cientistas [7], [8] se preocupam com as propriedades métricas de escalas (e instrumentos de medida) referentes a grandezas não directamente observáveis. Na actualidade, é fundamental que as propriedades métricas dos instrumentos usados nas avaliações nacionais sejam garantidas e, assim, os resultados obtidos sejam realmente úteis e usados adequadamente.

¹Eurydice [3].

²"Os resultados da avaliação das escolas e do sistema educativo, constantes de relatórios de análise integrada, contextualizada e comparada, devem ser divulgados com o objectivo de disponibilizar aos cidadãos em geral e às comunidades educativas em particular, uma visão extensiva, actualizada, criticamente reflectiva e comparada internacionalmente do sistema educativo português".

Têm sido usados modelos estatísticos para produzir as listas ordenadas de escolas. Pestana [5] centrou as críticas às listas ordenadas decorrentes da aplicação dos modelos de regressão múltipla propostos por Grácio [9] nos valores do *coeficiente de determinação* (variam entre 0,018 e 0,213), no efeito de *confundimento* e no uso abusivo de *causalidade*. O autor exemplifica o modo de usar esta classe de modelos com dados de bioestatística (relação entre o perímetro cefálico à nascença e o comprimento do biparietal medido na 34ª semana de gravidez) onde a unidade estatística de observação é a criança.

Os dados referentes à avaliação educacional têm a estrutura hierárquica da população onde foram recolhidos, isto é, contêm variáveis que representam atributos de diferentes unidades estatísticas de observação. Nomeadamente, segundo Pestana [5], os modelos enunciados em Grácio [9] consideram:

- Variável dependente: resultados escolares definidos como a média das classificações dos alunos de cada escola nas disciplinas avaliadas no 12º ano; a unidade estatística de observação é o aluno; as respectivas classificações são agregadas para a unidade escola;
- Variáveis explicativas e de controlo: contexto social e económico aferido através do índice de poder de compra, taxa de não escolarização e número médio de anos de escolaridade da população de cada concelho; tipo de estabelecimento de ensino (público ou privado).

Os modelos de regressão múltipla não são adequados para a modelação dos dados com a estrutura hierárquica da população subjacente. Dois dos problemas decorrentes da mistura de atributos referentes a diferentes unidades estatísticas de observação são: (1) os erros padrão das estimativas dos parâmetros fixos são subestimados; (2) o modelo não leva em conta a variabilidade dos contextos socio-económicos dentro de cada concelho e associados à população discente das escolas³. Deverá, assim, usar-se o modelo de regressão multinível [10]

Concordando-se ou não com a produção das listas ordenadas de escolas, com a sua divulgação pública e/ou com a utilização das listas para a formulação de política educativa, é essencial que elas sejam válidas face aos objectivos da sua constituição.

O principal propósito deste artigo é mostrar o impacto que a especificação do modelo estatístico, no que se refere à escolha das variáveis que o modelo inclui e à respectiva medição, tem nas listas ordenadas resultantes. Para tal, aplica-se o modelo de regressão multinível aos dados do PISA 2000 e, para ilustração, usam-se critérios alternativos de ordenação baseados na definição de efeito-escola tipo A e tipo B propostos por Raudenbush e Willms [10]. São comparadas as estatísticas de posição das escolas e mostra-se que a posição ocupada depende do critério adoptado, tal como esperado. No entanto, perante o mesmo critério, a simples selecção das variáveis, ou até a forma da sua medição, tem relevância nos resultados obtidos.

Ficará fora do domínio deste trabalho, a ilustração das consequências da agregação dos dados e posterior inferência. Este assunto é conhecido na literatura por falácia ecológica [11] e está na base de análises e contribuições científicas polémicas na história da Ciência (por exemplo, Lohmoller, Falter, Link e Rijke [12]).

Metodologia

Critério de ordenação

Como é que se deve estimar a contribuição da escola (e professores) na aprendizagem e desenvolvimento dos alunos? Há uma concordância generalizada na literatura especializada de eficácia escolar de que a comparação de escolas deve ser feita através do *valor acrescentado* pela escola na aprendizagem e desenvolvimento de todos e de cada um dos seus alunos [10], [13], [14], [15]. Conceitos importados da Economia para a Educação [15] definem valor acrescentado como sendo a indicação de quão uma dada escola tem promovido o progresso de todos os alunos num conjunto de áreas (curriculares ou não curriculares) durante determinado período de tempo – por exemplo, desde a entrada na escola até ao momento da aferição dos

³Postular um contexto social e económico constante para todas as escolas de cada concelho é demasiado irrealista. Ao lidar com medidas agrupadas por escola/concelho, o modelo despreza a variabilidade dos dados, tanto dos resultados escolares como das variáveis de controlo, considerando todos os alunos como iguais. Sabe-se que esta abordagem foi adoptada, em parte, porque os dados referentes ao nível socio-económico do aluno não estavam disponíveis para análise.

resultados escolares – comparativamente com os *efeitos* de outras escolas. A operacionalização desta definição requer que a métrica usada para quantificar os resultados escolares seja a mesma nos dois momentos de aferição, o que geralmente não é possível. Assim, Goldstein e Spiegelhalter [13] propõem a designação de *comparação ajustada* por traduzir melhor a transposição daquele conceito da Economia para esta área do conhecimento.

À definição de valor acrescentado está associada a noção de *efeito-escola*. Raudenbush e Willms [10] referem que o termo tem sido usado em duas situações ligeiramente distintas: (1) a que se destina a quantificar o impacto de uma dada política/prática nos resultados escolares dos alunos, tal como o efeito de diminuir a razão nºalunos/nºprofessores ou o efeito de adoptar critérios de selecção de alunos, entre outros; (2) a que se destina a quantificar a capacidade que a escola tem em modificar os resultados escolares dos alunos, ou seja, a contribuição da escola na aprendizagem. Aqueles autores defendem que o *efeito-escola* deve ter diferentes concepções, dependendo dos sujeitos que dele farão uso. São identificados dois grupos de potenciais interessados e a cada um está associado um determinado tipo de *efeito-escola*: *tipo A* e *tipo B*.

O efeito do *tipo A* é aquele que os pais consideram quando escolhem uma escola para os seus filhos. Os pais desejam que a escola produza o maior efeito possível, independentemente de como a escola o consegue produzir – quer seja através de professores altamente qualificados, do método de ensino praticado, da influência do contexto socio-económico da comunidade na qual a escola está inserida, dos recursos materiais disponíveis etc. Deste modo, interessa-lhes saber qual é o desempenho escolar esperado condicionado às características das crianças ou jovens, tais como conhecimento prévio, idade, situação social e económica.

O tipo de efeito *B* é definido de modo a “isolar” o efeito da prática da escola, distinguindo-o do contexto da escola. O contexto da escola inclui os factores associados à escola mas que são exógenos ao trabalho lá desenvolvido – a prática educativa – por professores, directores e outros profissionais. Os factores de contexto incluem, por exemplo, as características socio-económicas da comunidade servida, assim como a composição demográfica da população discente. Estes factores podem ser catalizadores ou inibidores de ambientes de aprendizagem, independentemente do esforço educativo levado a cabo pelos profissionais na escola. O efeito-escola do *tipo B* é, geralmente, do interesse daqueles que têm a responsabilidade de promover a aprendizagem de todos os alunos, independentemente das suas características e antecedentes.

Modelo estatístico

Em estudos desta natureza, a população é constituída por alunos agrupados em turmas e estas, por sua vez, agrupadas em escolas, etc. O modelo de avaliação tem de conseguir não só dar medidas nacionais e regionais dos resultados atingidos, mas também identificar os processos intra-escolares que explicam a variabilidade dos resultados, dadas as condições de escolarização. Este assunto está amplamente registado na literatura [13], [14], [16]. O modelo estatístico adequado para, simultaneamente, acomodar atributos referentes a unidades de observação de diferentes níveis hierárquicos (aluno, turma-professor, escola, concelho) é o modelo multinível [10], [17], [18], [19].

Para estimar os efeitos do tipo *A* e *B* será usado o modelo mais simples, designado por modelo de componentes de variância⁴.

Considere-se “aluno” a unidade de nível 1, identificada pelo subscrito i , e “escola” a unidade do nível 2, identificada pelo subscrito j . Considere-se a existência de J escolas, $j=1, \dots, J$, cada uma delas com n_j alunos, $i=1, \dots, n_j$,

$$I = \sum_{j=1}^J n_j.$$
 Seja y_{ij} o desempenho escolar do i -ésimo aluno na j -ésima escola, e x_{ij} o respectivo nível socio-

económico. O modelo de componentes de variância com uma variável explicativa, especifica-se como segue:

⁴Variance components model [10].

$$\begin{aligned}
 y_{ij} &= \beta_{0j} + \beta_1 x_{ij} + \varepsilon_{ij} \\
 \beta_{0j} &= \gamma_{00} + u_{0j} \\
 \varepsilon_{ij} &\sim N(0, \sigma_\varepsilon^2) \\
 u_{0j} &\sim N(0, \sigma_{u0}^2) \\
 \text{cov}[\varepsilon_{ij}, u_{0j}] &= 0
 \end{aligned} \tag{1}$$

Com base num modelo semelhante a (1), mas em que a variável explicativa representa o desempenho prévio do aluno ij , Goldstein [145](1997, p.383) diz que os resíduos ajustados (\hat{u}_{0j}) são geralmente usados como estimativas do valor acrescentado da escola porque, ajustando (ou controlando) pelo desempenho inicial do aluno, eles quantificam o progresso relativo em diferentes escolas. Cada estimativa \hat{u}_{0j} representa o afastamento do desempenho médio da escola j à média global, $\hat{\gamma}_{00}$. Também Raudenbush e Willms e Raudenbush [101](1995, p.309) estabelecem que o efeito-escola ou valor acrescentado são estimados através do resíduo associado às escolas, mas com modelos em que o conjunto de variáveis explicativas seleccionadas se encontra em concordância com a definição dos efeitos *tipo A* e *tipo B* anteriormente expostos. Neste trabalho, para ilustrar a aplicação dos dois critérios aos dados do PISA, usou-se o desempenho em Matemática como variável dependente. A descrição da base de dados utilizada constitui o anexo A.

Estimação do efeito-escola tipo A

O modelo estatístico para estimar os efeitos do tipo A é o modelo (2) que seguidamente se apresenta⁵.

$$\begin{aligned}
 y_{ij} &= \beta_{0j} + \beta_1 \text{ano}_{-5ij} + \beta_2 \text{ano}_{-6ij} + \beta_3 \text{ano}_{-7ij} + \beta_4 \text{ano}_{-8ij} + \beta_5 \text{ano}_{-9ij} + \beta_6 \text{ano}_{-11ij} + \\
 &\quad + \beta_7 \text{Menina}_{ij} + \beta_8 \text{ISEI}_{ij} + \beta_9 \text{int}_{\text{mat}}_{ij} + \beta_{10} \text{idade}_{ij} + \varepsilon_{ij} \\
 \beta_{0j} &= \gamma_{00} + u_{0j} \\
 \varepsilon_{ij} &\sim N(0, \sigma_\varepsilon^2) \\
 u_{0j} &\sim N(0, \sigma_{u0}^2) \\
 \text{cov}[\varepsilon_{ij}, u_{0j}] &= 0
 \end{aligned} \tag{2}$$

Todas as variáveis explicativas são referentes ao aluno: sexo, nível de ensino que frequenta, idade, nível socio-económico (ISEI), interesse em Matemática.

Estimação do efeito-escola tipo B

O modelo estatístico para estimar os efeitos do tipo B é o modelo (3) que seguidamente se apresenta. As variáveis explicativas são as mesmas do modelo (2) além das que são relativas à escola: nível socio-económico contextual, existência de critério de selecção de alunos, localização da escola.

$$\begin{aligned}
 y_{ij} &= \beta_{0j} + \beta_1 \text{ano}_{-5ij} + \beta_2 \text{ano}_{-6ij} + \beta_3 \text{ano}_{-7ij} + \beta_4 \text{ano}_{-8ij} + \beta_5 \text{ano}_{-9ij} + \beta_6 \text{ano}_{-11ij} + \\
 &\quad + \beta_7 \text{Menina}_{ij} + \beta_8 \text{ISEI}_{ij} + \beta_9 \text{int}_{\text{mat}}_{ij} + \beta_{10} \text{idade}_{ij} + \varepsilon_{ij} \\
 \beta_{0j} &= \gamma_{00} + \gamma_{01} \overline{\text{ISEI}}_j + \gamma_{02} \text{HalB}_j + \dots + \gamma_{07} \text{Selecc\~ao}_j + u_{0j} \\
 \varepsilon_{ij} &\sim N(0, \sigma_\varepsilon^2) \\
 u_{0j} &\sim N(0, \sigma_{u0}^2)
 \end{aligned} \tag{3}$$

⁵Recorde-se a inexistência duma variável que represente o conhecimento/competências prévias do aluno. As variáveis indicadoras do nível de ensino que o aluno frequenta funcionam como *proxy*.

Foi usado o procedimento de estimação *Iterative Generalised Least Squares* [20] e o plano amostral considerado através de *weighted scaling method 2* [21]. Todas as análises foram realizadas no programa estatístico *MLwiN* [22].

Resultados

Com base no ajuste do modelo nulo e considerando a totalidade dos casos válidos, 2461 alunos em 142 escolas, o coeficiente de partição da variância [23], [24] tem o valor de 0,33. A tabela 1 contém as estimativas dos parâmetros fixos e respectivos erros padrão para os modelos (2) e (3).

As estimativas dos coeficientes associados às variáveis dos alunos apresentam valores muito próximos em ambos os modelos. As estimativas dos parâmetros associadas ao nível de ensino (do 5º até ao 9º ano) reflectem o défice de literacia Matemática face ao que seria esperado encontrar num jovem de 15 anos de idade. Note-se que os alunos retidos não passaram pelo processo ensino-aprendizagem dos conteúdos programáticos previstos até ao 10º ano de escolaridade, tornando-se assim impossível dar provas dessa aprendizagem. Por exemplo, o modelo 2 sugere que os jovens de 15 anos a frequentar o 5º ano têm, em média, o desempenho em Matemática reduzido em 241 unidades comparativamente aos jovens de 15 anos que frequentam o 10º ou 11º anos de escolaridade.

Tabela 1. Estimativas dos parâmetros fixos

Variável		Modelo 2 Estimativa (erro padrão)	Modelo 3 Estimativa (erro padrão)
Constante		538,51 (4,06)	536,73 (6,21)
Aluno: Nível de ensino	5º	-240,74 (19,46)	-239,74 (19,70)
	6º	-190,54 (8,88)	-186,48 (9,15)
	7º	-173,46 (4,99)	-170,59 (5,22)
	8º	-128,83 (3,96)	-126,28 (4,19)
	9º	-76,37 (3,26)	-73,66 (3,50)
	11º	19,32 (37,22)	19,18 (37,47)
Aluno: Sexo	Menina	-32,57 (2,29)	-33,04 (2,35)
Aluno: Nível socioeconómico	ISEI	0,54 (0,09)	0,47 (0,09)
Aluno: Interesse em Matemática	Int_mat	13,51 (1,18)	13,64 (1,20)
Aluno: Idade	Idade	-2,58 (0,34)	-2,60 (0,35)
Escola: Localização (milhares de habitantes)	Hab<3	---	11,66 (6,79)
	3<Hab<15	---	9,93 (4,31)
	100<Hab<1000	---	11,29 (5,36)
	Hab>1000 Centro	---	23,26 (8,43)
	Hab>1000 Periferia	---	15,72 (19,78)
Escola: Nível sócioeconómico	Média ISEI	---	0,63 (0,30)
Escola: Critério de selecção	Nunca aplica	---	-7,25 (5,09)
Nº de casos usados		2304	2202

As meninas apresentam resultados, em média, 33 unidades abaixo face aos meninos. Tanto o nível socio-económico do aluno como o interesse do jovem pela Matemática têm impacto positivo nos resultados atingidos.

O modelo 3 sugere que os resultados atingidos pelo aluno são melhores quando a escola que ele frequenta está localizada no centro de uma grande cidade (com mais de um milhão de habitantes). Não se verificou diferença estatisticamente significativa nos resultados escolares quando a escola está localizada em cidades de outra dimensão comparativamente com as de 15.000 a 100.000 habitantes. As estimativas referentes ao nível socio-económico da escola, variável de contexto, indicam que os jovens que estudam em escolas cuja população discente é mais afluenta, têm melhores resultados. Perante este modelo, não há diferença estatisticamente significativa nos resultados escolares dos alunos que frequentam escolas que adoptem critérios de selecção de alunos face às que não têm qualquer política de selecção. No entanto, modelação exploratória em torno da questão mostrou que tal fenómeno está correlacionado com o nível socio-económico da escola. Retirando do modelo esta variável contextual, aquela relação torna-se estatisticamente significativa, indicando que os jovens que estudam em escolas que nunca adoptam critérios de selecção de alunos obtêm, em média, piores resultados comparativamente com os colegas que estudam em escolas que adoptam algum critério de selecção de alunos (ver anexo B).

A tabela 2 contém as estimativas dos parâmetros aleatórios dos modelos acima mencionados, determinadas com base nos dados amostrais. O coeficiente de determinação dos modelos 2 e 3 é 0,58. A elevada capacidade explicativa dos modelos é devida à inclusão da variável “nível de ensino”⁶. A capacidade explicativa do modelo 2, mas com a exclusão dessa variável, baixa para 0,13. Adiante são comparadas as listas ordenadas com critérios de ordenação efeito do tipo B estimado com e sem a variável “nível de ensino”.

Tabela 2. Variância e coeficiente de determinação

Variância	Modelo nulo	Modelo 2	Modelo 3	Modelo 2 S/nível de ensino
Nível1 – aluno	4881,4	2842,1	2871,76	4506,87
Nível2 – escola	2425,7	250,4	191,14	1867,25
Total	7307,1	3092,5	3062,9	6374,12
R ²	---	0,58	0,58	0,13

O objectivo principal desta modelação é a obtenção dos resíduos do nível 2 para produzir listas ordenadas, considerando como critério o efeito de tipo A e de tipo B. Determinou-se o desvio absoluto da posição que cada escola ocupa em cada uma delas. Os percentis 75º, 90º e 95º desse desvio absoluto são, respectivamente, 16, 25 e 28. O mesmo é dizer que 25% das escolas distam pelo menos 16 posições nas duas listas, 10% distam pelo menos 25 posições e 5% distam pelo menos 28 posições. A associação entre as duas listas quantificada pelo coeficiente de correlação de Kendall é 0,79. O gráfico de dispersão apresentado na figura 1 ilustra tal associação.

A comparação entre a lista ordenada originada pelas estimativas do efeito tipo B (modelo 3) e a lista cujas estimativas resultam do mesmo modelo excluída a variável explicativa “nível de ensino” é mostrada no gráfico de dispersão da figura 2. A correlação de Kendall entre as duas listas é -0,22.

De modo semelhante, as figuras 3 e 4 mostram graficamente duas listas ordenadas (resíduos e respectivo intervalo de confiança de 95%) resultantes do modelo 2, com e sem a variável “nível de ensino”, respectivamente. A título de exemplo, as figuras mostram a posição que a mesma instituição ocupa em cada uma das listas. A estimativa intervalar torna evidente que a diferença entre a maioria das instituições não é estatisticamente significativa, pelo que a ordenação das escolas é desprovida de sentido.

⁶ A retenção do aluno pode ter sido originada noutra escola que não naquela onde ele se encontra no momento da aferição do PISA. Atendendo à escassez de dados relativos à trajectória escolar do aluno, torna-se impossível melhorar o modelo.

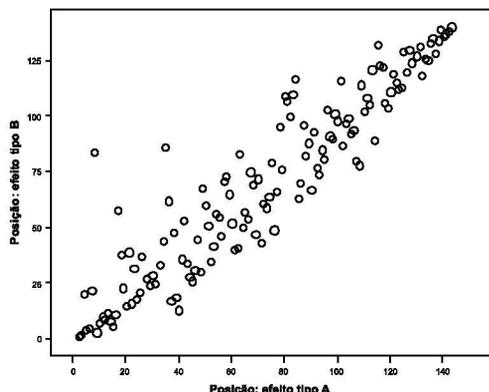


Figura 1. Gráfico de dispersão das posições das escolas em duas listas ordenadas por diferentes critérios

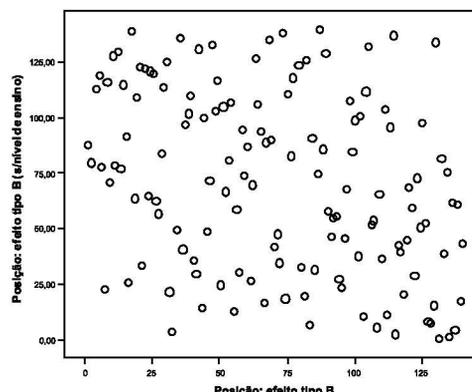


Figura 2. Gráfico de dispersão das posições das escolas em duas listas ordenadas pelo mesmo critério com diferentes variáveis

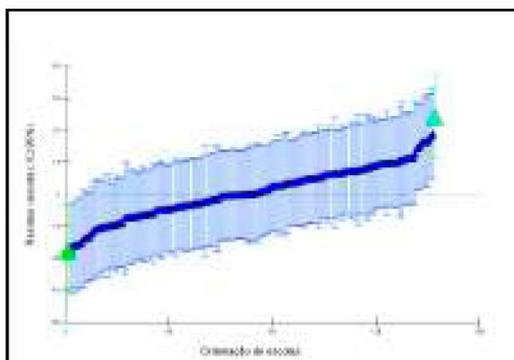


Figura 3. Estimativa pontual e intervalar do efeito tipo A (modelo 2)

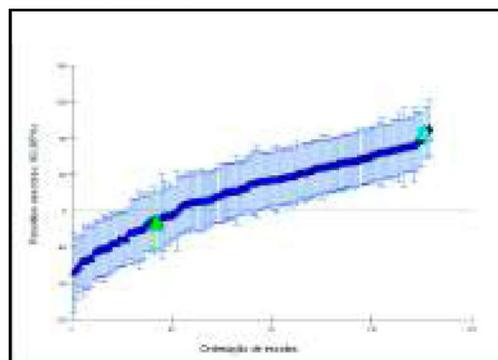


Figura 4. Estimativa pontual e intervalar do efeito tipo A (modelo 2 s/ variável "nível de ensino")

Discussão e Conclusão

A relevância da Estatística aplicada à avaliação institucional, nomeadamente em Educação, é amplamente consensual. Os cuidados metodológicos a adoptar no que se refere, entre outros, aos procedimentos de recolha dos dados, à garantia da qualidade dos dados, à escolha da classe do modelo estatístico, à selecção das variáveis, à inferência, são aspectos que têm sido tratados na literatura nacional e internacional.

Este artigo traz à discussão aspectos metodológicos relacionados com a produção de listas ordenadas das escolas, nomeadamente, mostra a discrepância dos resultados que se deve a: 1) objectivos diferentes e, por conseguinte, critérios de ordenação diferentes; 2) omissão de variáveis de controlo.

São aplicados modelos de regressão multinível, que representam a estrutura hierárquica dos alunos agrupados em escolas, aos dados portugueses do PISA 2000. A título de ilustração é usada a definição de efeito-escola proposto por Raudenbush e Willms [10].

Os resultados obtidos são elucidativos do quão arbitrário poderá ser um sistema de avaliação quando não se explicita o seu modelo conceptual, quando o modelo estatístico não especifica correctamente o modelo conceptual e não se tomam as devidas precauções face às limitações dos modelos.

Concorde-se ou não com as listas ordenadas de escolas, quaisquer que sejam os critérios estabelecidos pela tutela, decorre naturalmente dos resultados apresentados que é premente melhorar a qualidade dos dados em

Educação, garantir a disponibilidade das variáveis de controle pertinentes, adoptar procedimentos que aumentem a precisão dos resultados obtidos, interpretar resultados considerando as limitações ou eventual violação dos pressupostos dos modelos utilizados. [M](#)

Referências

- [1] **Bird, S. M., Cox, D., Farewell, V. T., Goldstein, H., Holt, T., & Smith, P.C.** (2005). "Performance indicators: good, bad, and ugly". *Journal of the Royal Statistical Society, A*, 168, p. 1-27.
- [2] **Abrantes, P.** (2001). "Mathematical competence for all: options, implications and obstacles." *Educational Studies in Mathematics*, vol 47(2), 125-143.
- [3] **Eurydice** (2004). *A avaliação dos estabelecimentos de ensino obrigatório na europa*. Disponível em <http://www.eurydice.org/documents/evals/fr/frameset.htm>
- [4] **Queiró, J. F.** (2002). "A publicação das notas dos exames do 12º ano". *Gazeta de Matemática*, 142, 28-31.
- [5] **Pestana, D. D.** (2003). "Apologia da Estatística (A Pretexto da Seriação das Escolas Secundárias)". *Gazeta de Matemática*, 144, 20-33.
- [6] **Justino** (2005). *No silêncio todos somos iguais*. Lisboa, Gradiva.
- [7] **Galton, F.** (1874). "On a proposed statistical scale". *Nature*, March, 5, 342-343.
- [8] **Edgeworth, F.** (1888). "On a new method of reducing observations relating to several quantities". *Philosophical Magazine*, 25, p. 184-191.
- [9] **Grácio, S. Franco, L. Velho, S., Sanches, E. e Rijo, S.** (2002). *Proposta de seriação das escolas secundárias segundo os resultados obtidos nos exames nacionais de 12º ano em 2001/2002*. Lisboa: FCSH, Universidade Nova de Lisboa.
- [10] **Raudenbush, S. W., & Willms, J. D.** (1995). "The estimation of school effects". *Journal of Educational and Behavioral Statistics*, 20(4), 307-335.
- [11] **Robinson, W. S.** (1950). "Ecological correlations and the behaviour of individuals". *American Sociology Review*, 15, 351-357.
- [12] **Lohmoller, J. B., Falter, J., Link, A., & de Rijke, J.** (1985). "Unemployment and the rise of national socialism: contradicting results from different regional aggregations". In P. Nijkamp (ed.) *Measuring the unmeasurable*. Martinus Nijhoof: Den Haag.
- [13] **Goldstein, H., & Spiegelhalter, D. J.** (1996). "League tables and their limitations: statistical issues in comparison of institutional performance". *Journal of the Royal Statistical Society, A*, 159(3), 385-443.
- [14] **Goldstein, H.** (1997). "Methods in school effectiveness research". *School Effectiveness and School Improvement*, 8(4), 369-395.

- [15] **Sammons, P., Thomas, S., & Mortimore, P.** (1997). *Forging links: effective schools and effective departments*. London: Chatman Publishing, Lda.
- [16] **Aitkin, M., Longford, N.** (1986). "Statistical modelling issues in school effectiveness studies". *Journal of the Royal Statistical Society, A*, 149, part 1, p. 1-43.
- [17] **Bryk, A. S., & Raudenbush, S. W.** (1992). *Hierarchical linear models*. Newbury Park, California: Sage.
- [18] **Longford, N. T.** (1993). *Random coefficient models*. Oxford: Clarendon Press.
- [19] **Ferrão, M. E.** (2003). *Introdução aos modelos de regressão multinível em educação*. Campinas: Komedi.
- [20] **Goldstein, H.** (1986). "Multilevel mixed linear models analysis using iterative generalised least squares". *Biometrika*, 73, 43-56.
- [21] **Pfeffermann, D., Skinner, C. J., Holmes, D. J., Goldstein, H., & Rasbash, J.** (1998). "Weighting for unequal selection probabilities in multilevel models". *Journal of the Royal Statistical Society, B*, 60, 23-40.
- [22] **Rasbash, J., Browne, W., Healy, M., Cameron, B., & Charlton, C.** (2001). *MLwiN. Multilevel models project*. London: University of London.
- [23] **Goldstein, H.** (2003). *Multilevel Statistical Models* (3rd ed). London: Edward Arnold.
- [24] **Goldstein, H., Browne, W., & Rasbash, J.** (2002). "Partitioning variation in multilevel models". *Understanding Statistics*, 1(4), 223-231.

Bibliografia

- Ferrão, M. E., Loureiro, M. J., Simões, M. F., & Guedes, P.** (2005). *À procura da escola eficaz – referencial teórico do projecto de investigação 3EM*. Covilhã: UBI.
- Goldstein, H., Rasbash, J., Yang, M., Woodhouse, G., Pan, H., Nuttal, D., & Thomas, S.** (1993). "A multilevel analysis of school examination results". *Oxford Review of Education*, 19(4), 425-433.
- Montgomery, D. G., & Peck, E. A.** (1982). *Introduction to linear regression analysis*. New York: John Wiley & Sons.
- Raudenbush, S. W.** (2004). "What are value-added models estimating and what does this imply for statistical practice?" *Journal of Educational and Behavioral Statistics*, 29(1), 1-129.

Anexo A

Dados do PISA 2000

PISA⁷ é um estudo da responsabilidade da OCDE sobre os conhecimentos e as competências de alunos de 15 anos dos principais países industrializados. A recolha de dados ocorre de três em três anos. No ano 2000 envolveu cerca de 265 000 alunos de 32 países, 28 dos quais membros da OCDE. A recolha de dados foi efectuada através de testes de desempenho em leitura, matemática e literacia científica e através de questionários aplicados a alunos e directores das escolas incluídas no plano amostral. Os testes de desempenho foram criados de modo a assegurar a cobertura de todos os conteúdos e a não sobrecarregar os cadernos de teste a serem aplicados aos alunos [25], [26], [27].

Neste trabalho usa-se, como medida de resultado escolar, o desempenho em Matemática. As estimativas obtidas foram transformadas numa escala com média 500 e desvio padrão 100, usando os dados dos países membros da OCDE com excepção da Holanda [26], [28].

Em Portugal foram seleccionadas aleatoriamente 156 escolas, das quais 3 se recusaram a participar no estudo e outras 4 não tinham já alunos com 15 anos. Os alunos, em cada escola, foram também aleatoriamente seleccionados. Em consequência da aplicação destes critérios, o PISA envolveu 149 escolas, 138 públicas e 11 privadas, abrangendo um total de 4604 alunos⁸. Foram incluídos na população alvo todos os alunos de 15 anos a frequentar a escola desde o 5º ano ao 11º ano de escolaridade [29].

A base de dados, bem como todos os relatórios técnicos estão disponíveis em www.pisa.oecd.org/pisa/. O ficheiro de dados `instud_math.sav` contém 2545 casos relativos aos alunos portugueses que foram submetidos aos testes de Matemática.

Para este trabalho, excluíram-se da análise os alunos que estudam em escolas privadas sem financiamento público⁹ e seleccionaram-se as seguintes variáveis:

Variáveis do aluno

- **Sexo:** variável nominal binária, “1” representa as meninas e “0” representa os meninos; os meninos constituem o grupo de referência. Segundo expansão da amostra, 52,4% da população de referência, são meninas.

- **Nível de ensino que frequenta:** variável ordinal que representa o nível de ensino, do 5º ano ao 11º ano, que os alunos da amostra frequentam. Visando a modelação, a variável foi recodificada, através de variáveis mudas, e considerou-se como grupo de referência o 10º ano (nível adequado à idade de 15 anos). Considerando os casos válidos e segundo a expansão da amostra, 52,2% da população de referência frequentam o 10º ou 11º ano de escolaridade, 28,2% frequentam o 9º ano de escolaridade e os restantes alunos distribuem-se entre o 5º e o 8º anos.

- **Idade:** medida em meses, a média é 187,6 com desvio padrão 3,4. Para a modelação a variável foi centrada em 180;

- **Nível socio-económico:** aferido pelo índice internacional socio-económico (ISEI) da situação ocupacional do pai ou mãe do aluno [30]. Usou-se o valor máximo dos dois, com média 44,0 e desvio padrão 15,9. Nos modelos 2 e 3 a variável foi centrada na média.

- **Interesse em Matemática:** índice construído segundo Baumert, Gruehn, Koller e Schnabel [31], reflectindo o grau de concordância do aluno com afirmações, tais como: “Fico totalmente absorvido quando estudo Matemática”; “A Matemática é importante para mim”; “Eu não gostaria de desistir porque a Matemática é divertida”.

Variáveis da escola

- **Nível socio-económico médio da escola:** média do nível socio-económico dos alunos da escola.

- **Localização da escola:** variável nominal que representa, em termos do número de habitantes, a dimensão

⁷ Programme for International Student Assessment.

⁸ O ficheiro de dados contém 4585 casos não tendo encontrado explicação para a diferença.

⁹ Na amostra encontram-se 2334 alunos que frequentam escolas públicas, 127 alunos que frequentam escolas particulares com financiamento público, 49 alunos que estudam em escolas particulares sem financiamento público e 35 alunos para os quais o tipo da escola que frequentam é omissa.

da vila/cidade onde a escola se localiza. As categorias (e respectiva distribuição dos alunos) são as seguintes: 1 – até 3.000 habitantes (4,8%); 2 – de 3.000 a 15.000 habitantes (36,9%); 3 – de 15.000 a 100.000 habitantes (36,9%); 4 – de 100.000 a 1.000.000 de habitantes (13,3%); 5 – centro de cidade com mais de 1.000.000 de habitantes (6,8%); 6 – periferia de cidade com mais de 1.000.000 de habitantes (1,3%). A variável foi recodificada, através de variáveis mudas, e considerou-se como grupo de referência a localização em cidade com 15.000 a 100.000 habitantes.

- Critério de selecção de alunos de acordo com o desempenho escolar: as escolas foram classificadas em dois grupos – (1) as que nunca aplicam tal critério e (0) as que o aplicam às vezes ou sempre. As últimas constituem o grupo de referência e representam 25,5% dos alunos abrangidos.

Apesar das críticas [32], a utilização secundária dos microdados do PISA constitui um avanço relativamente à utilização de qualquer outro conjunto de dados disponível referente à avaliação do sistema educativo português. No entanto, as análises aqui apresentadas ainda ficam aquém do que seria desejável no âmbito do estudo dos efeitos escolares, pois, entre outras, não se dispõe de uma variável fundamental associada ao aluno – a que mede o seu conhecimento à entrada na escola.

Referências

[25] OECD (1999). *Measuring student knowledge and skills: a new framework for assessment*. Paris: OECD.

[26] OECD (2000). *Measuring student knowledge and skills: the PISA 2000 assessment of reading, mathematical and scientific literacy*. Paris, OECD.

[27] OECD (2002a). *PISA 2000 technical report*. Paris: OECD.

[28] OECD (2002b). *Manual for the PISA 2000 database*. Paris: OECD.

[29] GAVE (2001). *Resultados do estudo internacional PISA 2000 – Primeiro relatório nacional*. Lisboa: Ministério da Educação.

[30] Gazeboom, H. B. G., De Graaf, P., & Treiman, D. J. (1992). "A standard international socio-economic index of occupational status". *Social Science Research*, 21(1), 1-56.

[31] Baumert, J., Gruehn, S., Koller, O. E Schnabel, K.U. (1997). *Bildungsverlaufe und psychosoziale entwicklung im jugendalter (BIJU): dokumentation – Band 1*. Berlin: Max-Planck-Institut fur Bildungsforschung.

[32] Collani, Elart von (2001). "OECD PISA - An example of stochastic illiteracy?" *Economic Quality Control*, 16, p. 227-253.

Anexo B

Tabela de estimativas do modelo 3

Variáveis		Modelo 3 -Final Estimativa (erro padrão)	Modelo 3 c/critério de selecção Estimativa (erro padrão)
Constante		538,05 (4,08)	546,08 (5,63)
Aluno: Nível de ensino	5º	-238,65 (19,54)	-242,76 (19,61)
	6º	-186,28 (9,03)	-188,95 (8,98)
	7º	-170,85 (5,10)	-172,50 (5,12)
	8º	-126,38 (4,10)	-127,84 (4,07)
	9º	-73,80 (3,39)	-75,12 (3,37)
	11º	---	---
Aluno: Menina		-32,62 (2,32)	-33,03 (2,35)
Aluno: Nível socio-económico		0,46 (0,09)	0,52 (0,09)
Aluno: Interesse em Matemática		13,44 (1,19)	13,59 (1,20)
Aluno: Idade		-2,59 (0,34)	-2,60 (0,35)
Escola: Localização (milhares de habitantes)	Hab<3	---	---
	3<Hab<15	---	---
	100<Hab<1000	---	---
	Hab>1000 Centro	16,14 (7,95)	18,18 (8,36)
	Hab>1000 Periferia	---	---
	Escola: Nível sócio- económico	Média ISEI	0,66 (0,26)
Escola: Critério de selecção nunca aplica		---	-10,57 (4,93)
Nº de casos usados		2202	2202