6
Metodologia para a associação entre as práticas de Língua Portuguesa e a proficiência nas diferentes habilidades de leitura

Neste capítulo, inicialmente, discuto a pertinência do uso da Modelagem Hierárquica Multinível em pesquisas que lidam com dados educacionais. Na sequência, teço alguns esclarecimentos sobre decisões tomadas quanto à definição dos dados, decisões essas decorrentes da questão central investigada ou de limitações encontradas no decorrer do estudo. Logo após, apresento a descrição das medidas e os procedimentos utilizados. Ainda nesta seção, descrevo os modelos construídos, que serão utilizados para a investigação dos efeitos das práticas dos professores na proficiência das habilidades de leitura de seus alunos.

PUC-Rio - Certificação Digital Nº 0811296/CA

6.1
Pertinência da metodologia utilizada e delimitações dos dados

Diante da impossibilidade de manter as causas constantes, como nas investigações experimentais, as ciências sociais admitem diversos fatores passíveis de interferência na variável que se pretende explicar, observando-os, registrando suas variações e procurando determinar, no resultado final, que influências cabem a cada um deles. Entre as várias metodologias de que se pode lançar mão para a investigação dessas influências está a Modelagem Hierárquica Multinível, resumidamente denominada como HLM (Hierarchical Linear Models).

O uso de modelos lineares hierárquicos encontra-se amplamente difundido em contextos de educação devido à própria natureza das estruturas de dados educacionais, que são, em geral, fortemente hierarquizadas. Tal fato se evidencia quando se observa, por exemplo, o desempenho, a atitude frente aos estudos ou outra variável individual no nível dos alunos, os quais se agrupam em turmas, que se agrupam em escolas, que, por sua vez, podem se reunir em níveis ainda maiores de agregação, como municípios, estados e países.

Acontece, porém, que a distribuição dos alunos por turmas e por escolas, por exemplo, não ocorre de forma aleatória e existem características próprias desses contextos (das escolas e das turmas, e não só dos alunos) que influenciam

na forma como os alunos aprendem. A título de exemplificação, o fato de alunos pertencerem a uma determinada turma ou a uma determinada escola faz com que eles tendam a se assemelhar mais entre si do que com alunos de outras turmas ou de outras escolas. Tais semelhanças costumam ocorrer no próprio nível dos alunos: é comum que colegas de uma mesma escola, por exemplo, provenham de famílias com um nível socioeconômico e cultural semelhante. Entretanto, as semelhanças obviamente se tornam ainda mais fortes quando se consideram níveis maiores de agregação, como turmas e escolas. No caso das turmas, por exemplo, os alunos compartilham entre si o fato de estarem expostos ao ensino ministrado pelo mesmo professor, que tem, por sua vez, características didáticas e de formação específicas, entre outras. Algo semelhante se passa no nível das escolas, onde os alunos de um mesmo estabelecimento educacional normalmente estão submetidos a uma mesma direção escolar, com suas dinâmicas próprias e as mesmas condições de infraestrutura. Diante dessa característica normalmente hierarquizada dos dados, quando se realiza a regressão da proficiência dos alunos em relação a uma ou mais variáveis explicativas, é natural que os resíduos dessa regressão dependam fortemente da turma ou da escola especificamente considerada e que tais resíduos se distribuam em torno das retas de regressão com diferentes valores de variância, conforme a turma ou escola analisada[30].

Conforme Tufi M. Soares (2003, p. 105) a estruturação hierárquica constitui "uma sequência natural de agrupamentos aninhados, de tal forma, que as variáveis representativas das características nos diversos níveis podem interagir com outras variáveis dentro do mesmo nível hierárquico e, também, com variáveis de outro nível. Tendo em vista, principalmente, essa natureza dos dados, é frequente a utilização de modelos de regressão hierárquicos[31], que permitem investigar a influência das características de cada nível da hierarquia no desempenho escolar dos alunos e na diferenciação entre as escolas e, ainda, separar a variabilidade nos resultados – associada às escolas – da variabilidade dentro de cada escola – associada aos alunos ou turmas de alunos".

As técnicas de multinível trabalham com uma ideia relativamente simples e engenhosa para diminuir os obstáculos característicos dos dados hierarquizados:

[30]A ocorrência dessas relações é denominada no meio estatístico como violação da homocedastidicidade e da independência das observações. Para detalhes ver BRYK et al. (2001).
[31] O autor está citando trabalhos de Lee, 2001e de Goldestein, 1995.

PUC-Rio - Certificação Digital Nº 0811296/CA

a determinação simultânea de um grande número de modelos lineares, um para cada unidade de interesse. Desta forma, se o interesse do pesquisador for analisar escolas diferentes ou turmas diferentes estabelece-se uma equação de regressão para cada escola ou para cada turma, conforme o caso. Além disso, ainda segundo Tufi M. Soares (*ibidem*, 113) a incorporação da estrutura de agrupamento dos dados reflete concretamente na modelagem multinível, pois enquanto para o modelo de regressão clássico o intercepto e o coeficiente de inclinação são parâmetros fixos para o modelo multinível eles são considerados parâmetros aleatórios, dependentes da influência do nível hierárquico mais alto.

Retomando a minha questão de pesquisa, busco averiguar uma possível existência de associações significativas entre a proficiência dos alunos e determinadas características didáticas dos seus respectivos professores. Como os professores, mesmo dentro de uma mesma escola, geralmente variam entre as diferentes turmas, as variáveis que expressam o seu fazer pedagógico influenciam, prioritariamente, o nível da turma. Além disso, variáveis identificadas como importantes para a explicação da proficiência estão associadas aos alunos. O nível socioeconômico, a proficiência anterior, o sexo, são variáveis normalmente empregadas como controles em estudos de eficácia escolar. Assim sendo, para este estudo foram considerados os níveis referentes ao aluno (N1) e à turma (N2).

Devo comentar ainda sobre a não utilização da variável referente aos pesos amostrais, que têm a função de quantificar e qualificar a informação acarretada por cada caso nas estimativas. O GERES possui a variável peso calculada para os alunos que participaram de todas as avaliações, ou seja, para os casos em que não há dados faltantes. No entanto, optei por trabalhar com os dados obtidos para todos os alunos que possuem duas avaliações subsequentes, ainda que não participantes dos testes de todos os anos da pesquisa. Essa decisão está relacionada com o fato de a quantidade de casos ser de extrema importância para a obtenção de confiabilidade em estudos que adotam a metodologia multinível, o que implicou a não utilização dos pesos. Devo mencionar que a determinação de pesos em casos como esse, de dados longitudinais com observações faltantes, é um problema ainda não completamente solucionado na literatura e diferentes autores recomendam diferentes procedimentos para se lidar com ele. Uma das recomendações feitas pelos pesquisadores é de se calcular os pesos transversalmente, isto é, por ano. O problema dos pesos transversais é que eles só

PUC-Rio - Certificação Digital Nº 0811296/CA

podem ser usados em análises também transversais, o que foge ao propósito das pesquisas que construíram bancos longitudinais.

Além disso, algumas simulações mostram que, para efeito de uma análise exploratória, as conclusões obtidas por meio de modelos que levam em consideração pesos amostrais são muito similares às obtidas quando os pesos não são considerados. De fato, se todas as variáveis relevantes estão presentes no modelo é de se esperar que elas produzam o controle e o ajustamento necessário às estimativas. Deve-se considerar ainda que variáveis não observadas farão falta tanto na modelagem em si quanto no cálculo dos pesos amostrais.

De toda forma, a não utilização dos pesos faz com que as incertezas sobre as conclusões do modelo sejam maiores do que se fosse possível utilizá-los, situação em que a representatividade de cada caso seria obtida com precisão ainda maior.

PUC-Rio - Certificação Digital Nº 0811296/CA

6.2
Medidas e procedimentos adotados

Quando se fazem análises lineares hierárquicas, é conveniente, antes de se introduzirem modelos mais elaborados, com variáveis explicativas nos diversos níveis, considerar primeiro o chamado modelo plenamente incondicional, que também pode ser denominado como modelo nulo ou FUM (Fully Unconditional Model, na literatura internacional). Esse procedimento preliminar consiste no seguinte modelo:

Nível 1:
$Y_{ij} = \beta_{0j} + r_{ij}$

Nível 2:
$\beta_{0j} = \gamma_{00} + u_{0j}$

Conforme se pode observar, não há qualquer variável explicativa nas equações do modelo acima. A interpretação para o referido modelo é a seguinte: no nível 1, a nota de cada aluno i da turma j (Yij) corresponde, neste caso, à média de sua respectiva turma (β0j) mais um valor referente ao erro (rij) associado a cada aluno na regressão. Por sua vez, no nível 2, a média de cada turma j (β0j) corresponde à grande média populacional (γ00) mais um valor referente ao erro u0j associado a cada turma.

A consideração desse modelo permite responder a duas questões de grande interesse para estudos que buscam associações entre variáveis. Uma delas é saber se a proficiência média de uma dada disciplina, no final de uma determinada etapa de ensino, varia significativamente entre as turmas. A segunda, em caso afirmativo para a questão anterior, diz respeito a qual é o percentual da variação total dos resultados entre as turmas.

Após a verificação da existência de variação que justificasse a investigação, foram construídos diversos modelos lineares hierárquicos de dois níveis com o intuito de se verificar uma possível associação entre a proficiência dos alunos e algumas modalidades de práticas docentes com as quais eles tiveram contato durante o 2º e o 3º anos do ensino fundamental. Em todos os casos considerados, no nível 1 foram incluídas as variáveis mensuradas para o nível dos alunos, ao passo que, no nível 2, foram inseridas as variáveis características de suas respectivas turmas. Passo, em seguida, a descrever mais detalhadamente as variáveis consideradas nesses vários modelos.

As variáveis dependentes nos modelos aqui considerados correspondem aos resultados obtidos pelos alunos no teste de leitura da pesquisa GERES. Como explicado no capítulo 5, a proficiência estimada com base na totalidade de itens presentes no teste foi subdividida em quatro proficiências: “Processamento”, “Localização”, “Integração” e “Aspectos Discursivos”. Foram analisados os resultados alcançados pelos alunos nessas diferentes subdimensões, que se tornaram, cada uma delas, as variáveis dependentes das regressões.

Para efeitos de manipulação dos dados no HLM, todas essas variáveis, quantitativas e contínuas, foram padronizadas, de modo a possuírem uma média igual a zero e desvio-padrão de uma unidade. Tal procedimento apresenta, entre outras vantagens, a de padronizar a interpretação dos resultados das diversas regressões.

As variáveis independentes, utilizadas como hipótese explicativa nos modelos, são todas do âmbito da sala de aula e correspondentes às práticas declaradas pelos professores. Conforme também detalhado no capítulo 5, foi possível construir escalas capazes de mensurar a ênfase, declarada pelos docentes, na adoção das seguintes práticas: leitura para os alunos; leitura silenciosa pelos alunos; leitura em voz alta pelos alunos; cópia, ditado e caligrafia; escrita de redação e adoção de práticas menos contextualizadas ou mais contextualizadas de

PUC-Rio - Certificação Digital Nº 0811296/CA

alfabetizar. Cada uma dessas escalas, nos modelos de regressão, são variáveis explicativas que representam a intensidade de utilização das práticas, ou seja, quanto maior o índice obtido pelo professor, maior a sua prioridade em relação à determinada prática e vice-versa.

Com o objetivo de permitir a interpretação comparativa dos resultados, assim como o procedimento realizado para as variáveis dependentes, todas as variáveis independentes também foram padronizadas, de modo a possuírem uma média igual a zero e desvio padrão igual a um. No caso da minha pesquisa, não poderia ser diferente, já que cada escala elaborada para as variáveis explicativas atinge valores numéricos diferenciados e, assim sendo, não podem ser comparadas entre si. A uniformização dos dados permite que os resultados obtidos a partir dessas medidas possam ser compreendidos em relação a quanto representam em termos de desvio padrão.

O estudo incluiu, também, quatro covariáveis, conforme mencionado brevemente na definição sobre os níveis de análise adotados. Esse é um tipo de variável que está associada a fatores comumente correlacionados com a proficiência acadêmica. O objetivo de utilizá-las é o de se obter resultados menos "contaminados" por influências de variáveis que não as explicativas. Entre essas variáveis, adiante denominadas como "controles", está o nível socioeconômico (NSE) do aluno. Conforme ressaltam Maria Teresa Alves e Francisco Soares (2007) "talvez uma das teses mais importantes da Sociologia da Educação seja o argumento de que o desempenho escolar é fortemente associado à origem social dos alunos. O peso explicativo dos fatores extraescolares, associados ao nível socioeconômico das famílias dos alunos, foi comprovado empiricamente através de grandes *surveys* educacionais".

Também foi incluída nos modelos uma variável de controle indicadora do sexo feminino contrastando com o masculino, tendo em vista que os efeitos dessa variável na proficiência escolar, com piores resultados para os meninos, em Língua Portuguesa, são universalmente conhecidos (cf. Tufi M. Soares, 2005, p. 82).

Outra variável de grande importância utilizada nos modelos foi a proficiência prévia dos alunos. Essa medida permite "descontar" o efeito da proficiência adquirida pelos alunos em anos anteriores de modo a se obter informações mais precisas sobre os ganhos individuais de aprendizagem ocorridos

PUC-Rio - Certificação Digital Nº 0811296/CA

durante um período de escolarização específico no qual o aluno esteve submetido a práticas pedagógicas também específicas.

Foi adicionada, ainda, aos modelos, a variável "dependência administrativa" (rede) utilizada como um controle adicional. Normalmente alocada no nível "escola", esta variável foi trazida para o nível imediatamente anterior, o da turma, tendo em vista a sua grande influência em contextos de desigualdade, como é o caso brasileiro, e levando-se em conta os estudos já realizados com os dados GERES, que evidenciam uma acentuada discrepância nos resultados escolares associada ao fato de os alunos pertencerem a escolas públicas ou a escolas privadas.

Devo mencionar nesse ponto que é de praxe em pesquisas com HLM serem utilizadas, como covariáveis explicativas do coeficiente aleatório, a média das variáveis de controle agregadas no nível hierárquico superior. Neste trabalho, optei por não utilizar esse tipo de controle devido ao fato de a unidade turma sofrer alterações de um ano para o outro.

O quadro 9 contém a descrição das variáveis usadas nos modelos estimados de acordo com o ano de escolaridade.

PUC-Rio - Certificação Digital Nº 0811296/CA

PUC-Rio - Certificação Digital Nº 0811296/CA

Quadro 9: Descrição das variáveis

		Variáveis	Tipo de codificação	Descrição	2º ano	3º ano
Nível Aluno (N1)	Dependentes	Proficiência em processamento	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente à habilidade de processamento, no **final** do ano escolar indicado ao lado.	X	
		Proficiência em localização	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente à habilidade em localização, no **final** de cada ano escolar indicado ao lado.	X	X
		Proficiência em integração	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente à habilidade de integração, no final do ano escolar indicado ao lado.		X
		Proficiência em aspectos discursivos	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente ao domínio de aspectos discursivos, no **final** do ano escolar indicado ao lado.		X
	Controles	Proficiência prévia em processamento	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente à habilidade de processamento, no **início** do ano escolar indicado ao lado.	X	
		Proficiência prévia em localização	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente à habilidade em localização, no **início** de cada ano escolar indicado ao lado.	X	X
		Proficiência prévia em integração	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente à habilidade de integração, no **início** do ano escolar indicado ao lado.		X
		Proficiência prévia aspectos discursivos	Contínua	Variável obtida a partir do conjunto de itens dos testes GERES referente ao domínio de aspectos discursivos, no **início** do ano escolar indicado ao lado.		X
		Nível socioeconômico	Contínua	Variável original do GERES construída a partir do *status* ocupacional e da escolaridade dos pais e da posse de bens familiar.	X	X
		Sexo (feminino)	Dicotômica	Variável original do GERES construída a partir dos questionários aplicados aos pais.	X	X
Nível Turma (N2)	Explicativas	Rede (privada) Controle adicional	Dicotômica	Variável original do GERES construída a partir do plano amostral. Rede pública inclui escoas municipais e estaduais e rede privada inclui escolas privadas e federias.	X	X
		Práticas menos contextualizadas de alfabetização	Contínua	Variável obtida a partir do conjunto de resposta dos professores às questões 94, 95, 96 e 99 do questionário GERES.	X	X
		Leitura realizada pelo professor	Contínua	Variável obtida a partir do conjunto de resposta dos professores às questões 67 e 68 do questionário GERES.	X	X
		Leitura silenciosa realizada pelo aluno	Contínua	Variável obtida a partir do conjunto de resposta dos professores às questões 69, 70, 71 e 79 do questionário GERES.	X	X
		Leitura em voz alta realizada pelo aluno	Contínua	Variável obtida a partir do conjunto de resposta dos professores às questões 72 e 73 do questionário GERES.	X	X
		Escrita: cópia, ditado e caligrafia.	Contínua	Variável obtida a partir do conjunto de resposta dos professores às questões 74, 75 e 76 do questionário GERES.	X	X

Na tabela 13, apresento a estatística descritiva das variáveis utilizadas nos modelos correspondentes ao 2º ano.

Tabela 1: Estatística descritiva das variáveis (2º ano)

2º ano						
		Variáveis	Mínimo	Máximo	Média	Desvio padrão
Nível Aluno (N1)	Desfecho	Proficiência em processamento	-1,84	2,37	0	1
		Proficiência em localização	-1,45	3,12	0	1
	Controles	Proficiência prévia em processamento	-1,57	2,01	0	1
		Proficiência prévia em localização	-1,00	2,08	0	1
		Nível socioeconômico	-2,93	2,85	0	1
		Sexo (feminino)	0	1	0	1
Nível Turma (N2)	Explicativas	Rede (privada) Controle adicional	0	1	0	1
		Média da proficiência da turma em processamento	-2,20	2,51	0	1
		Média da proficiência da turma em localização	-1,86	2,46	0	1
		Média da proficiência prévia da turma em processamento	-1,80	2,29	0	1
		Média da proficiência prévia da turma em localização	-1,25	2,47	0	1
		Práticas menos contextualizadas de alfabetização	-0,60	3,01	0	1
		Leitura realizada pelo professor	-0,72	5,57	0	1
		Leitura silenciosa realizada pelo aluno	-5,12	2,51	0	1
		Leitura em voz alta realizada pelo aluno	-4,14	1,21	0	1
		Escrita: cópia, ditado e caligrafia.	-1,61	2,81	0	1

PUC-Rio - Certificação Digital Nº 0811296/CA

Na tabela 14, apresento a estatística descritiva das variáveis utilizadas nos modelos correspondentes ao 3º ano.

Tabela 2: Estatística descritiva das variáveis (3º ano)

3º ano						
		Variáveis	Mínimo	Máximo	Média	Desvio padrão
Nível Aluno (N1)	Desfecho	Proficiência em localização	-1,37	3,79	0	1
		Proficiência em integração	-1,33	3,38	0	1
		Proficiência em aspectos discursivos	-1,23	3,01	0	1
	Controles	Proficiência prévia em localização	-1,48	3,11	0	1
		Proficiência prévia em integração	-1,16	3,23	0	1
		Proficiência prévia em aspectos discursivos	-1,29	2,99	0	1
		Nível socioeconômico	-2,95	2,87	0	1
		Sexo (feminino)	0	1	0	1
Nível Turma (N2)	Explicativas	Rede (privada) Controle adicional	0	1	0	1
		Média da proficiência da turma em localização	-2,02	2,66	0	1
		Média da proficiência da turma em integração	-1,84	2,75	0	1
		Média da proficiência da turma em aspectos discursivos	-1,69	2,69	0	1
		Práticas menos contextualizadas de alfabetização	-0,59	3,30	0	1
		Leitura realizada pelo professor	-5,02	0,86	0	1
		Leitura silenciosa realizada pelo aluno	-3,53	3,84	0	1
		Leitura em voz alta realizada pelo aluno	-1,83	4,13	0	1
		Escrita: cópia, ditado e caligrafia.	-2,35	1,87	0	1

6.3 Abordagem de análise

A seguir, apresento um exemplo genérico de um modelo de dois níveis, sendo, o primeiro, o do aluno e, o segundo, o da turma.

Nível 1:
$Y_{ij} = \beta_{0j} + \Sigma\, \beta_{nj} X_{ij} + r_{ij}$

Nível 2:
$\beta_{0j} = \gamma_{00} + \Sigma\, \gamma_{0k} W_j + u_{0j}$
$\beta_{nj} = \gamma_{n0} + \Sigma\, \gamma_{mk} W_j + u_{nj}$

PUC-Rio - Certificação Digital Nº 0811296/CA

A variável Y corresponde à proficiência do aluno numa disciplina específica.

No nível 1, há n variáveis explicativas X, (n+1), coeficientes de regressão β e um r corresponde ao erro associado à proficiência observada de cada aluno.

No nível 2, cada coeficiente da turma β pode, por sua vez, também ser modelado analogamente, por meio de m variáveis W, (m+1), coeficientes γ, e pelo erro associado à proficiência média de cada turma (unj).

Dessa forma, é possível obter diferentes parâmetros de regressão, os coeficientes lineares e angulares das equações, para as diferentes turmas consideradas no estudo. Assim sendo, as turmas de maior proficiência média devem apresentar maiores coeficientes lineares da equação (maiores valores de β0j).

Um ponto de grande interesse na análise é a consideração dos coeficientes angulares βnj que estejam significativamente associados à proficiência Y. Desta forma, um coeficiente positivo indica que a ênfase numa determinada prática pedagógica está associada a um aumento médio da proficiência dos alunos e um coeficiente negativo corresponde à interpretação contrária. Além da direção positiva ou negativa dessa associação, a consideração do valor numérico dos coeficientes também permitirá mensurar o efeito quantitativo da influência das variáveis explicativas.

Abaixo, está a equação do modelo estimado para medir a associação entre a proficiência em processamento apurada no final do 2º ano (variável dependente) e as práticas declaradas pelos professores, com os devidos controles (variáveis explicativas). Para todas as variáveis dependentes (processamento, localização, integração e aspectos discursivos) os modelos estimados seguem o mesmo padrão. Sendo assim, não será necessário apresentar as equações para todos os modelos.

Level-1 Model

$$PRFP052Z_{ij} = \beta_{0j} + \beta_{1j}*(PRFP051Z_{ij}) + \beta_{2j}*(ZNSE_{ij}) + \beta_{3j}*(FEM_{ij}) + r_{ij}$$

Level-2 Model

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(PARTIC_j) + \gamma_{02}*(ZALFTR05_j) + \gamma_{03}*(ZLERP05_j) + \gamma_{04}*(ZLERS05_j) + \gamma_{05}*(ZLERV05_j) + \gamma_{06}*(ZESCPT05_j) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + u_{1j}$$
$$\beta_{2j} = \gamma_{20}$$
$$\beta_{3j} = \gamma_{30}$$

PUC-Rio - Certificação Digital Nº 0811296/CA

No nível 1,

$PRFP052Z_{ij}$ corresponde à nota, padronizada, em processamento, no final do 2º ano, referente ao aluno *i* da turma *j*;

$PRFP051Z_{ij}$ corresponde à nota padronizada do mesmo aluno no teste realizado no início do 2º ano (controle);

$ZNSE_{ij}$ é o índice socioeconômico padronizado do aluno;

FEM é uma variável dicotômica indicadora do sexo feminino (1=feminino e 0=masculino)

r é o erro da regressão, ou seja, a diferença entre o valor real da variável dependente e o seu respectivo valor previsto pelo modelo.

No nível 2,

$PARTIC_j$: é uma variável dicotômica indicadora da rede em que o aluno está matriculado (0=rede pública;1=rede privada);

$ZALFTR05_j$ é a variável padronizada que representa "práticas menos contextualizadas" de alfabetizar do professor da turma *j* no 2º ano;

$ZLERP05_j$ é a variável padronizada relacionada à prática de leitura em voz alta na turma *j* no 2º ano;

$ZLERS05_j$ é a variável padronizada relacionada à prática de leitura silenciosa na turma *j* no 2º ano;

$ZLERV05_j$ é a variável padronizada relacionada à prática de leitura em voz alta na turma *j* no 2º ano;

$ZESCPT05_j$ é a variável padronizada relacionada à prática de cópia, ditado e caligrafia na turma *j* no 2º ano;

u é o erro da regressão associado à turma *j*.

Por fim, resta mencionar como as variáveis foram centralizadas. A variável referente à proficiência prévia foi centralizada na média do grupo. Explicando melhor, um aluno que, no início do 2º ano, teve uma nota igual à média de sua respectiva turma, ficou com um valor nulo para essa variável. Em decorrência desse tipo de centralização, no nível 2, o respectivo coeficiente β_{ij} teve a sua variação permitida. As demais variáveis foram todas centralizadas na grande média, e seus respectivos coeficientes de regressão foram conservados fixos, como normalmente ocorre nesse tipo de modelagem.

PUC-Rio - Certificação Digital Nº 0811296/CA