



Universidade de Brasília
Departamento de Estatística

Avaliação da influência de fatores associados ao clima em sala de aula a partir da percepção de estudantes do 1º ano do Ensino Médio sobre o desempenho em matemática

Thalyta Brito dos Santos

Brasília
2017

Thalyta Brito dos Santos

Avaliação da influência de fatores associados ao clima em sala de aula a partir da percepção de estudantes do 1º ano do Ensino Médio sobre o desempenho em matemática

Orientador:

Prof. Dr. **Luís Gustavo do Amaral Vinha**

Projeto apresentado para obtenção do título de Bacharel em Estatística.

Brasília

2017

RESUMO

O presente estudo tem como objetivo avaliar o impacto dos fatores associados ao clima em sala de aula sobre o desempenho dos alunos na prova de proficiência em matemática, utilizando análise multinível. Foram avaliados 10.868 alunos do 1º ano do Ensino Médio, distribuídos em 145 escolas públicas da cidade de Fortaleza, Ceará. Na análise multinível foram considerados dois níveis de hierarquia, alunos (nível 1) e escola (nível 2), sendo que as variáveis consideradas no modelo final concernentes aos alunos foram sexo, reprovação, bolsa família, fator relação professor-aluno e fator indisciplina. Quanto à escola, as variáveis consideradas foram número de salas de aula e número de funcionários. O modelo final também apresentou efeito randômico para a variável relacionada a afinidade do aluno com o estudo de matemática e interação entre o fator relação professor-aluno e o número de funcionários nas escolas. Os resultados encontrados mostraram que os fatores associados ao clima em sala de aula, relação professor-aluno e a indisciplina, afetam significativamente o desempenho dos estudantes.

Palavras-chave: Análise multinível. Avaliação educacional. Desempenho em matemática. Clima em sala de aula. Ensino Médio. Fatores associados ao desempenho escolar.

SUMÁRIO

1 INTRODUÇÃO	7
2 REFERENCIAL TEÓRICO	11
2.1 Formulação do Modelo de Regressão Multinível	11
2.2 Agregação e Desagregação	12
2.3 Construção do modelo	12
2.4 Estimação dos parâmetros	14
2.5 Comparação de modelos	15
2.6 Análise de resíduos	15
3 RESULTADOS	17
3.1 Análise Descritiva	17
3.1.1 Desempenho em matemática	18
3.1.2 Relação do desempenho em matemática com as demais variáveis	19
3.2 Análise Fatorial	27
3.3 Análise Multinível	31
4 CONCLUSÃO	41
REFERÊNCIAS	43

1 INTRODUÇÃO

A Educação é um dos principais fatores envolvidos no desenvolvimento de qualquer país, portanto, conhecer a realidade educacional nos contextos municipal, estadual e nacional é de suma importância. A avaliação educacional é uma das formas que permite a verificação dessa realidade através dos resultados obtidos pelos alunos, assumindo que o desempenho dos estudantes reflete a qualidade de ensino nas escolas. No entanto, para o conhecimento pleno da realidade educacional não basta apenas conhecer o rendimento dos alunos, mas também as características que exercem influência sobre o desempenho desses.

A análise estatística adequada para a identificação das variáveis que influenciam o resultado dos alunos é a análise de regressão multinível, pois leva em consideração a estrutura hierárquica do sistema educacional. Por exemplo, o desempenho dos alunos pode ser melhor compreendido se consideradas tanto as características pessoais como as do contexto em que estão inseridos, onde essas características pertencem a níveis distintos de um sistema hierárquico. Em outras palavras, alunos se agrupam em turmas, turmas em escolas e assim sucessivamente, de forma que cada grupo possui características próprias ou agregadas de um grupo distinto (Hox, 2010).

Existem inúmeros sistemas de avaliação da educação brasileira em diversos níveis e modalidades de ensino, como o Sistema Nacional de Avaliação Básica - SAEB, Exame Nacional do Ensino Médio - ENEM, o Sistema Nacional de Avaliação do Ensino Superior - SINAES e outros. Considerando a educação básica, diversos estudos já foram realizados a fim de identificar fatores que afetam o desempenho dos alunos através dos dados obtidos destes sistemas de avaliações educacionais, utilizando análise multinível. São exemplos: Fletcher (1998), Soares e Alves (2003), Andrade e Laros (2007) e Laros, Marciano e Andrade (2012), citados a seguir.

Fletcher (1998) foi um dos pioneiros a utilizar análise de regressão multinível para dados de avaliação educacional em larga escala no Brasil. Utilizando os dados do SAEB de 1995, esse avaliou o desempenho escolar dos alunos da 8ª série do Ensino Fundamental, levando em consideração as características relativas ao ambiente familiar e ao nível socioeconômico da comunidade em torno da escola. Para o autor, avaliar o resultado dos alunos sem considerar as diferenças na composição dos estudantes nas escolas, distorcem os re-

sultados das análises. Nesse estudo foi observado que 63% da variância no desempenho em matemática entre escolas eram atribuídas à seletividade de ordem socioeconômica.

Soares e Alves (2003) analisaram a diferença de desempenho entre alunos segundo a sua raça, com ênfase no impacto de algumas políticas e práticas escolares na produção de equidade entre esses grupos. Os resultados indicaram que melhor qualificação e remuneração docente e escolas mais bem equipadas com diretores mais envolvidos, são variáveis que tiveram impacto negativo na equidade entre brancos e pardos e entre brancos e negros de forma que elevaram as diferenças de desempenhos, principalmente entre negros e brancos. Em outras palavras, a melhoria nas condições de ensino contribui de forma positiva para o desempenho escolar, porém, pode aumentar as diferenças entre os grupos raciais. Os dados são referentes ao desempenho dos alunos da 8ª série do Ensino Fundamental na prova de matemática, do SAEB 2001.

Andrade e Laros (2007) buscaram avaliar quais fatores estão associados ao desempenho dos alunos em língua portuguesa e matemática. Neste estudo, foi utilizada a regressão multinível para análise dos dados e observou-se que a variável que mais afetou o desempenho no nível do aluno foi o atraso escolar seguida da comparação do aluno com os colegas, no nível da escola foram o atraso escolar e os recursos culturais, ambos agregados para a escola. Em outro estudo Laros, Marciano e Andrade (2012) avaliaram o desempenho dos alunos em língua portuguesa, discriminado por regiões geográficas brasileiras. Observou-se que as variáveis que mais afetaram o desempenho escolar no nível dos alunos foram recursos culturais seguido do atraso escolar, neste estudo não foi considerada a comparação dos alunos com os colegas. E no nível da escola, a variável com maior efeito foi o percentual de repetência na turma, seguido do percentual de alunos que trabalham. Ressalta-se que ambos os trabalhos são relativos aos alunos da 3ª série do Ensino Médio avaliados no SAEB 2001, e obtiveram como resultado, no nível do aluno, o atraso escolar como uma característica significativa no desempenho. Já no nível da escola, não se observa um efeito em comum, pois as variáveis consideradas nos estudos foram distintas, obtendo resultados diferenciados. No entanto, os modelos construídos em ambos os trabalhos foram capazes de explicar 13% e 14,1% da variância do desempenho no nível do aluno e no mínimo 37,2% e 40,6% da variância no nível da escola, respectivamente. Além de possuírem resultados similares com relação a variabilidade do desempenho nos testes entre as escolas, resultante do efeito de seletividade e composição escolar, sendo 76% e

74,6% respectivamente.

Nesse contexto, o presente estudo visa contribuir com a literatura existente através da avaliação de fatores associados ao clima em sala de aula como indisciplina e a relação entre alunos e professores, na percepção de alunos, sobre o desempenho escolar, considerando os fatores contextuais que influenciam na aprendizagem. O desempenho dos alunos são obtidos por meio do Sistema Permanente de Avaliação da Educação Básica do Ceará (SPAECE), onde são avaliados por uma prova de matemática além de responderem questionários contextuais. Foram considerados no estudo alunos do 1º ano do Ensino Médio oriundos da rede pública da cidade de Fortaleza.

2 REFERENCIAL TEÓRICO

2.1 Formulação do Modelo de Regressão Multinível

No estudo de dados educacionais, nota-se uma estrutura de agrupamento que ocorre de forma natural, isto é, alunos se agrupam em turmas, as turmas em escolas, as escolas em localidades e assim por diante. Nesse sentido, faz-se necessário a utilização de técnicas estatísticas que consideram as diferenças das formações dos grupos, levando em consideração o contexto em que os indivíduos estão inseridos e a influência que recebem desse. Portanto, o modelo de regressão multinível, também chamado de modelo linear hierárquico, é a técnica estatística mais adequada.

Considere, por exemplo, um estudo com apenas dois níveis, isto é, alunos agrupados em escolas, em que cada aluno é representado pelo índice i , cada escola pelo índice j e a variável resposta dada pelo desempenho escolar do aluno. O modelo multinível terá, então, a seguinte expressão geral

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + e_{ij}, \quad (1)$$

em que

- Y_{ij} é o desempenho escolar do estudante i na escola j ;
- β_{0j} é o intercepto, ou seja, valor médio do desempenho dos alunos na escola j ;
- β_{1j} é o coeficiente de regressão da escola j relativo a variável explicativa X_1 ;
- e_{ij} é o erro aleatório.

Como pode ser observado na equação 1, na regressão multinível, tanto intercepto como o coeficiente de inclinação podem variar de escola para escola. Por isso, β_{0j} e β_{1j} têm o índice j indicando a existência de um intercepto e um coeficiente para cada escola.

Outro aspecto a ser observado é a variação nos coeficientes de regressão com a inclusão de variáveis explicativas no nível da escola. Ao introduzir uma variável explicativa Z associada a escola, os coeficientes de regressão apresentados na equação 1 são dados por

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j}, \quad (2)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j}, \quad (3)$$

em que, os γ 's são parâmetros fixos, o Z é a variável referente a escola que, segundo Hox (2010), age como variável moderadora na relação entre a variável resposta e a variável do nível mais baixo, e por fim os u 's são erros aleatórios do nível da escola.

Substituindo então, (2) e (3) em (1) o modelo pode ser reescrito como

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + \gamma_{01}Z_j + \gamma_{11}X_{1ij}Z_j + u_{1j}X_{1ij} + u_{0j} + e_{ij}. \quad (4)$$

As suposições desse modelo são que o erro do nível 1 (e_{ij}) segue uma distribuição normal com média zero e variância constante e, é independente dos erros do nível 2. Já os erros do nível 2 (u_{0j} e u_{1j}) possuem distribuição normal multivariada com médias iguais a zero e variâncias dadas respectivamente por $\sigma_{u_0}^2$ e $\sigma_{u_1}^2$, além de covariância não nula.

No modelo apresentado foi considerado apenas uma variável explicativa no nível do aluno (X_{1ij}) e uma no nível da escola (Z_j). Porém, poderão ser adicionadas outras variáveis em ambos os níveis. Sendo assim, a extensão do modelo multinível considerando as novas variáveis incluídas, permitirá obter o impacto dessas no desempenho escolar, bem como melhorar o poder explicativo do modelo, além de analisar como as diversas variáveis se interagem.

2.2 Agregação e Desagregação

Segundo Hox (2010), as variáveis explicativas podem ser definidas em qualquer nível da hierarquia. Ou seja, podem ser medidas diretamente no seu próprio nível ou podem ser movidas de um nível para outro por agregação ou desagregação.

A agregação ocorre quando a variável do nível inferior é realocada para o nível superior. Por exemplo, obter o nível socioeconômico da escola a partir da média do nível socioeconômico dos alunos. Já a desagregação ocorre quando a variável do nível superior é realocada para o nível inferior. Por exemplo, atribuir aos alunos uma característica da escola onde estudam.

2.3 Construção do modelo

Para a construção do modelo multinível, Hox (2010) sugere cinco passos que podem ser seguidos na análise.

- Primeiro passo

O primeiro passo consiste no ajuste de um modelo sem qualquer variável explicativa, formado apenas pelo intercepto e pelo erro do nível do grupo e do nível do indivíduo. Esse modelo é chamado de modelo nulo ou vazio e é dado por

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}. \quad (5)$$

O modelo nulo é útil porque possibilita o cálculo do coeficiente de correlação intraclasse, que indica o quanto da variância total decorre das diferenças entre os grupos. Este coeficiente é dado pela seguinte fórmula

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2}. \quad (6)$$

O valor de ρ varia de zero a um, sendo que quanto mais próximo de um maior a proporção da variância que é devida ao nível mais alto, indicando a necessidade da utilização de modelos multiníveis.

- Segundo passo

No segundo passo são introduzidas as variáveis explicativas do nível mais baixo, porém com a especificação de efeitos fixos. O modelo é dado por

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + u_{0j} + e_{ij}. \quad (7)$$

- Terceiro passo

No terceiro passo são incluídas as variáveis explicativas do nível mais alto. O modelo obtido é

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + \gamma_{01}Z_{1j} + u_{0j} + e_{ij}. \quad (8)$$

- Quarto passo

O quarto passo consiste em investigar quais variáveis explicativas do nível mais baixo possuem efeitos aleatórios. Este modelo é dado por

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + \gamma_{01}Z_{1j} + u_{1j}X_{1ij} + u_{0j} + e_{ij}. \quad (9)$$

- Quinto passo

No último passo, adicionam-se as interações entre as variáveis explicativas do nível mais alto e aquelas variáveis explicativas do nível mais baixo (efeito *cross-level*). Assim, o modelo completo é dado por

$$Y_{ij} = \gamma_{00} + \gamma_{10}X_{1ij} + \gamma_{01}Z_{1j} + \gamma_{11}X_{1ij}Z_j + u_{1j}X_{1ij} + u_{0j} + e_{ij}. \quad (10)$$

Neste estudo, foi aplicado o procedimento proposto por Hox (2010), apresentado acima, a fim de obter um modelo multinível. Sendo que, a variável resposta considerada foi o desempenho dos alunos em matemática e as variáveis explicativas foram sexo, etnia, escolaridade da mãe ou responsável, variáveis relacionadas ao clima em sala de aula e variáveis indicadoras do nível socioeconômico. Estas variáveis foram medidas no nível do aluno, e algumas agregadas para o nível da escola.

2.4 Estimação dos parâmetros

A estimação dos parâmetros em modelos multiníveis pode ser realizada através dos métodos de máxima verossimilhança, mínimos quadrados generalizados, equações de estimação generalizadas ou por métodos bayesianos. Sendo que, o mais utilizado é o de máxima verossimilhança pela vantagem de produzir estimativas que são assintoticamente eficientes e consistentes (Hox, 2010).

O método de máxima verossimilhança consiste em maximizar a função de verossimilhança. Na análise multinível existem duas funções: máxima verossimilhança completa e a máxima verossimilhança restrita. Na máxima verossimilhança completa tanto os coeficientes de regressão como os componentes de variância são incluídos na função de verossimilhança. Enquanto que na máxima verossimilhança restrita apenas os componentes de variância são incluídos na função de verossimilhança e os coeficientes de regressão

são estimados numa segunda etapa. Na prática, embora o método de máxima verossimilhança restrita apresente estimativas com menos viés, os métodos possuem resultados bem parecidos.

2.5 Comparação de modelos

A partir da função de verossimilhança pode-se calcular uma estatística chamada *deviance*, utilizada para comparar um modelo mais simples com um modelo mais geral. Essa estatística é dada por

$$Deviance = -2 \times \ln(\text{Verossimilhança}). \quad (11)$$

A *deviance* indica quão bem o modelo se ajusta aos dados, sendo que modelos com valores menores da *deviance* normalmente estão mais ajustados. No entanto, essa medida não pode ser interpretada sozinha, mas deve-se comparar a diferença do seu valor com o valor encontrado no modelo seguinte. Então, para testar se a diferença entre as *deviances* de dois modelos é significativa, utiliza-se o teste χ^2 para diferença das *deviances* com graus de liberdade igual à diferença do número de parâmetros estimados nos dois modelos.

2.6 Análise de resíduos

Após a construção do modelo, o próximo passo fundamental é a análise de resíduos. Esta é utilizada para verificar se os pressupostos do modelo são atendidos, como linearidade e normalidade. Sem esses pressupostos pode-se chegar a resultados pouco confiáveis e a qualidade do modelo ficar comprometida.

Na regressão multinível, diferente da regressão clássica, tem-se resíduos distintos para cada efeito aleatório do modelo, sendo necessário então uma análise para cada um destes a fim de verificar os pressupostos. A título de exemplificação, ao considerar o modelo apresentado anteriormente, observa-se três resíduos (u_{1j}, u_{0j}, e_{ij}) , que levaria a uma verificação de cada um separadamente.

Uma forma de verificação subjetiva mas bastante eficiente e utilizada é através de gráficos. Com estes, a linearidade, a normalidade e a homocedasticidade podem ser observadas. Para verificar a suposição de normalidade, o gráfico dos resíduos padronizados *versus* os escores normais deverá apresentar uma linha diagonal à direita para que a suposição seja atendida.

Outro gráfico utilizado para verificação dos pressupostos é o gráfico de dispersão. Ao plotar os valores dos resíduos *versus* os valores preditos da variável resposta, os pontos devem se distribuir de forma aleatória em torno do zero, sem apresentar nenhuma tendência, para que os pressupostos de linearidade e homocedasticidade sejam válidos.

3 RESULTADOS

O banco de dados utilizado neste estudo é referente a 10.868 alunos do 1º ano do Ensino Médio distribuídos em 145 escolas públicas da cidade de Fortaleza. Este banco de dados contém informações sobre o desempenho dos alunos na prova de matemática, perfil socioeconômico e trajetória escolar dos estudantes, além de informações relacionadas ao professor e características da turma.

Para a análise desses dados a primeira etapa consistiu na Análise Descritiva, que foi utilizada como ferramenta para o conhecimento do comportamento das variáveis e das relações existentes entre elas, contribuindo com as etapas posteriores.

A segunda etapa consistiu na criação de fatores (indicadores), por meio da Análise Fatorial Exploratória (Mingoti, 2005). Esse método permite que os fatores sejam construídos de modo a agregar informações correlacionadas sem perder grande parte do conteúdo das mesmas.

Em seguida realizou-se a modelagem dos dados através da Análise de Regressão Multinível, a fim de obter o efeito da indisciplina e da relação entre professor-aluno no desempenho dos estudantes. O modelo proposto considerou dois níveis hierárquicos, sendo o primeiro aluno e o segundo a escola, contemplando, portanto, características dos alunos e características das unidades escolares e possíveis interações entre essas.

Por fim, as etapas citadas anteriormente foram realizadas com o auxílio do software SAS - *Statistical Analysis System*.

3.1 Análise Descritiva

Inicialmente foi feita uma análise descritiva de todas as variáveis utilizadas neste estudo. Essas variáveis, apresentadas na Tabela 1, foram obtidas através de um questionário contextual aplicado aos alunos do 1º ano do Ensino Médio, em 2009 e do teste de proficiência em matemática.

Tabela 1 – Variáveis utilizadas no estudo

Variáveis	Descrição/Codificação
MAT	Desempenho em Matemática em 2009.
Sexo	1: masculino, 2: feminino.
Etnia	Etnia autodeclarada: branco, pardo, negro, amarelo e indígena.
Idade	Idade reportada em 2009 (em anos).
Repetência	Número de reprovações: 0, 1, 2, 3 ou mais .
Escolaridade da mãe ou mulher responsável	A escolaridade da mãe ou mulher responsável é classificada em: -Nunca estudou -Entre 1ª e a 4ª série do ensino fundamental -Entre 5ª e a 8ª série do ensino fundamental -Entre 1ª e a 3ª série do ensino médio -Ensino superior -Não sabe
Bolsa Família	Assume o valor 1 se a família do estudante recebe Bolsa Família e 0 caso contrário.
-Banheiro -Rádio -Geladeira -Televisão a cores -Máquina de lavar roupa -Aparelho de DVD -Automóvel	Quantidade de itens existentes na casa do aluno: nenhum; um; dois; três ou mais.
Computador	Sim, com acesso a internet; Sim, sem acesso a internet; Não.
Disciplina que mais gosta	Língua Portuguesa; Matemática; Educação Física; Outra.
O professor exige atenção dos alunos nas aulas	Frequentemente; Raramente; Às vezes; Nunca.
O professore mostra interesse no aprendizado dos alunos	Frequentemente; Raramente; Às vezes; Nunca.
O professor está disponível para esclarecer as dúvidas dos alunos	Frequentemente; Raramente; Às vezes; Nunca.
O professor dá as notas de maneira justa	Frequentemente; Raramente; Às vezes; Nunca.
O professor espera a turma silenciar	Frequentemente; Raramente; Às vezes; Nunca.
Há bagunça e desordem na sala	Frequentemente; Raramente; Às vezes; Nunca.
O professor sai da sala antes do término da aula	Frequentemente; Raramente; Às vezes; Nunca.
O professor falta às aulas	Frequentemente; Raramente; Às vezes; Nunca.
O professor dá oportunidade dos alunos se expressarem	Frequentemente; Raramente; Às vezes; Nunca.
Os alunos prestam atenção ao que o professor fala	Frequentemente; Raramente; Às vezes; Nunca.

3.1.1 Desempenho em matemática

A Tabela 2 e a Figura 1 apresentadas a seguir proporcionam uma visão geral do desempenho obtido pelos estudantes de escolas públicas da cidade de Fortaleza.

Tabela 2 – Medidas resumo do desempenho geral dos alunos na prova de matemática

n	Média	Mediana	Mínimo	Máximo	Desvio Padrão
10.832	246,52	241,14	116,93	414,24	48,00

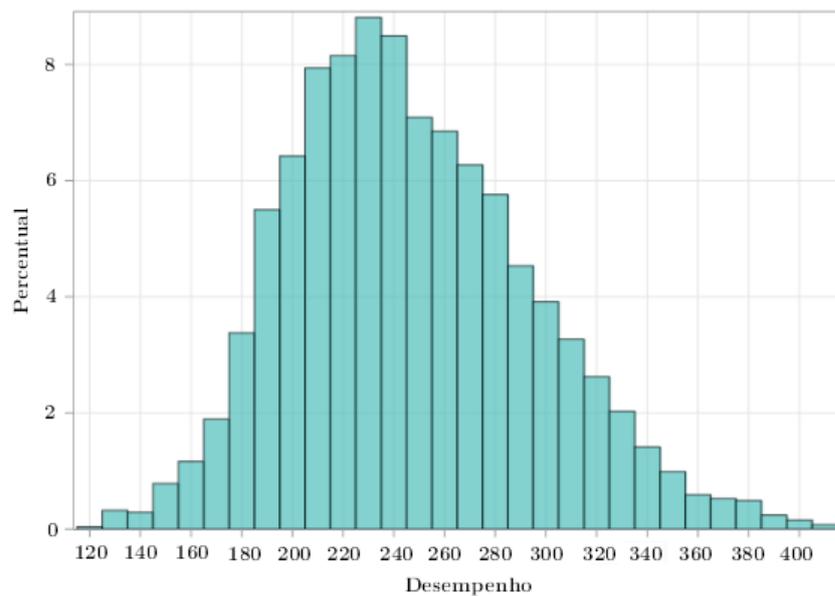


Figura 1 – Distribuição do desempenho dos alunos na prova de matemática.

De forma geral, observa-se que o desempenho médio dos alunos foi de 246,52 pontos, com uma variação de 48 pontos em torno da média (Tabela 2). De acordo com a Figura 1 observa-se também que o desempenho apresenta uma leve assimetria positiva. As notas mais frequentes entre os alunos estão entre 200 e 280 pontos, já à medida que as notas aumentam, o percentual de alunos que alcançam essas notas diminui gradativamente, menos de 2% dos alunos atingiram nota acima de 330 pontos.

3.1.2 Relação do desempenho em matemática com as demais variáveis

A fim de verificar de forma exploratória quais variáveis influenciam o desempenho dos alunos e analisar o que acontece com o desempenho dentro de cada nível das variáveis consideradas no estudo, foi feita uma análise bivariada por meio de medidas resumo e gráficos.

Sexo

Observa-se através da Tabela 3 e Figura 2 que os alunos do sexo masculino alcançaram desempenho médio maior que os do sexo feminino, além de possuírem também uma maior variabilidade de notas.

Tabela 3 – Medidas resumo da variável desempenho em matemática segundo sexo do aluno

Sexo	n	%	Média	Desvio Padrão
Masculino	4.939	45,82	255,43	50,36
Feminino	5.839	54,18	239,31	44,50

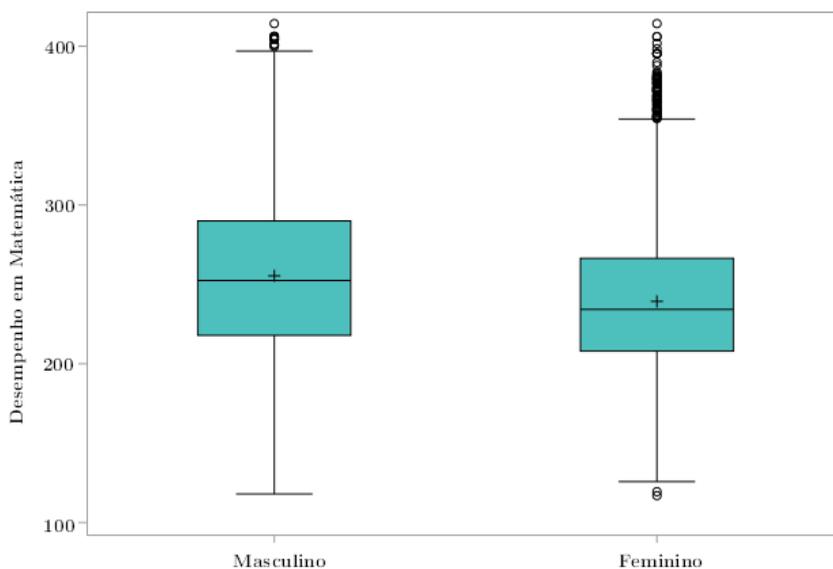


Figura 2 – Desempenho em matemática dos alunos por sexo.

Etnia autodeclarada

Pela Tabela 4 e Figura 3 é possível notar que, apesar da diferença ser pequena, os brancos e amarelos possuem notas maiores que os demais grupos. Também é possível notar que os alunos que se autodeclararam pardos e brancos correspondem a 74% dos avaliados, são a maioria no estudo.

Tabela 4 – Medidas resumo da variável desempenho em matemática segundo a etnia do aluno

Etnia	n	%	Média	Desvio Padrão
Branco	2.074	19,35	250,11	51,42
Pardo	5.843	54,53	246,29	46,05
Negro	1.532	14,30	243,99	47,94
Amarelo	817	7,64	249,44	51,25
Indígena	446	4,18	242,53	48,78

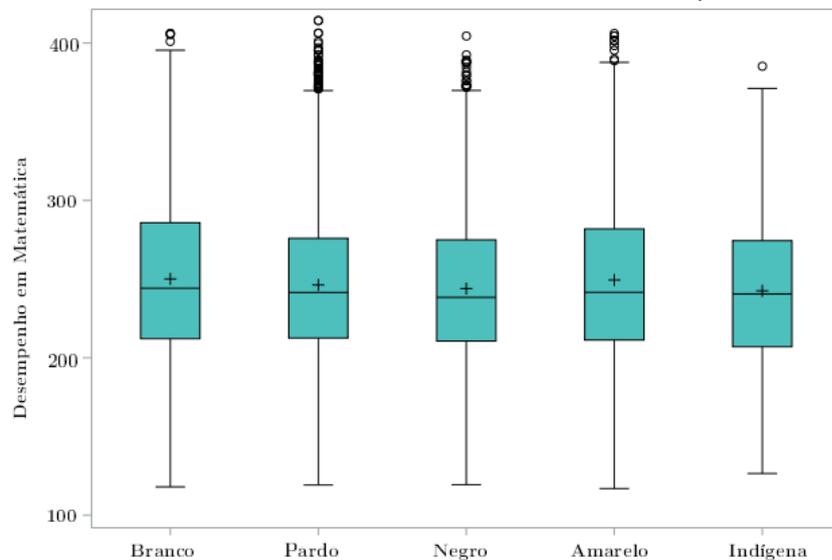


Figura 3 – Desempenho em matemática dos alunos por etnia.

Idade

Com relação a idade, observa-se um desempenho maior entre os estudantes de 14 a 16 anos, faixa etária esperada para alunos do 1º ano do Ensino Médio (Tabela 5 e Figura 4). Provavelmente alunos acima de 17 anos repetiram alguma série e alunos com idade abaixo de 14 anos são possíveis erros de coleta de informação. Assim, pode-se concluir que alunos que estão com a idade adequada para essa etapa de escolarização possuem desempenho superior aos demais.

Tabela 5 – Medidas resumo da variável desempenho em matemática segundo a idade do aluno

Idade	n	%	Média	Desvio Padrão
Menor que 14 anos	209	2,17	220,99	47,15
14 anos	239	2,49	269,30	54,97
15 anos	2.968	30,86	265,20	48,97
16 anos	3.152	32,78	250,52	45,73
17 anos	1.804	18,76	238,55	42,19
18 anos	872	9,06	227,78	40,85
19 anos	373	3,88	226,20	39,47

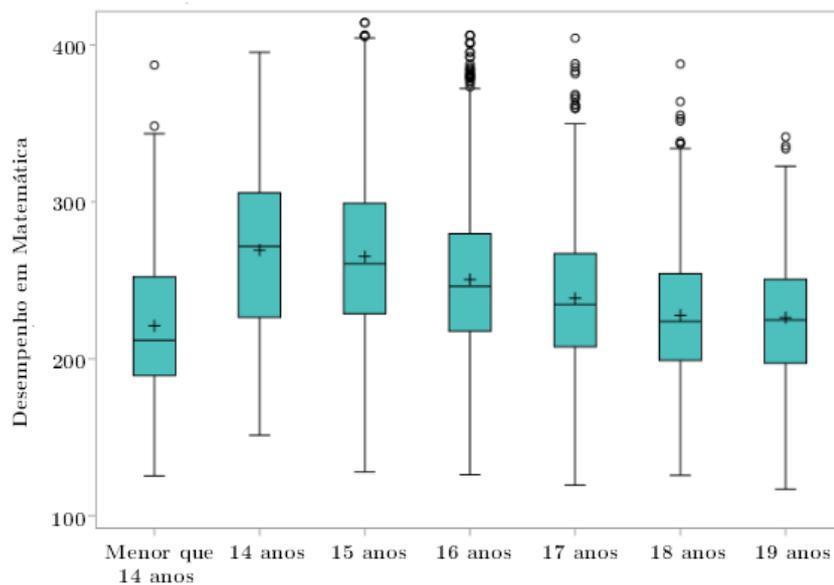


Figura 4 – Desempenho em matemática dos alunos por idade.

Número de reprovações

Através da Tabela 6 e Figura 5, nota-se que os alunos que nunca reprovaram possuem desempenho notavelmente maior em comparação com os alunos que já reprovaram pelo menos uma vez. De maneira geral, conforme aumenta o número de reprovações, mais baixo é o desempenho dos alunos.

Tabela 6 – Medidas resumo da variável desempenho em matemática segundo o número de reprovações dos alunos

Número de reprovações	n	%	Média	Desvio Padrão
Nenhuma vez	5.881	54,81	257,97	49,30
Uma vez	2.867	26,73	236,34	42,55
Duas vezes	1.466	13,67	229,68	41,70
Três vezes ou mais	513	4,79	225,69	42,97

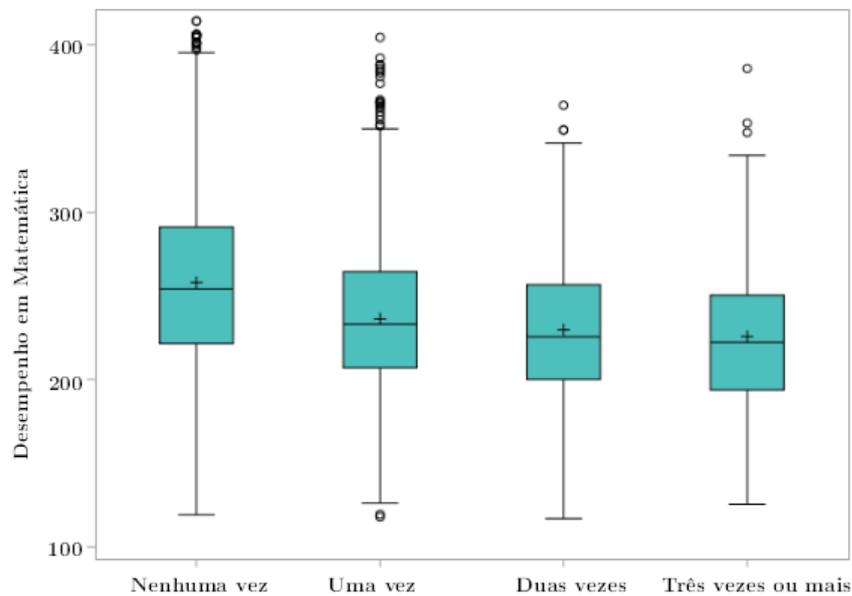


Figura 5 – Desempenho em matemática dos alunos segundo o número de reprovações.

Escolaridade da mãe ou mulher responsável

Pode-se observar que alunos cuja mãe possui escolaridade entre 1^a e a 3^a série do Ensino Médio se sobressaem na média obtida na prova de matemática, seguido daqueles cuja mãe possui Ensino Superior (Tabela 7 e Figura 6). Apesar da pequena diferença entre os dois grupos citados, a situação apresentada é inesperada, já que é suposto um maior desempenho para aqueles cuja mãe possui maior grau de instrução. Vale ressaltar também que a quantidade de alunos que possuem mãe que nunca estudou é muito baixa, apenas 4,56%, sendo que esses obtiveram o pior desempenho. Além disso, um dado que chama bastante atenção é o percentual relativamente alto de alunos que declararam não saber o nível de escolaridade da mãe (13,38%).

Tabela 7 – Medidas resumo do desempenho dos alunos em matemática segundo a escolaridade da mãe ou mulher responsável

Escolaridade da mãe ou mulher responsável	n	%	Média	Desvio Padrão
Nunca estudou	487	4,56	229,62	45,79
Entre 1ª e a 4ª série do Ensino Fundamental	2.175	20,31	240,08	44,75
Entre 5ª e a 8ª série do Ensino Fundamental	3.019	28,22	248,42	46,93
Entre 1ª e a 3ª série do Ensino Médio	2.892	27,01	256,74	49,76
Ensino Superior	696	6,51	251,34	53,72
Não sabe	1.431	13,38	236,03	43,46

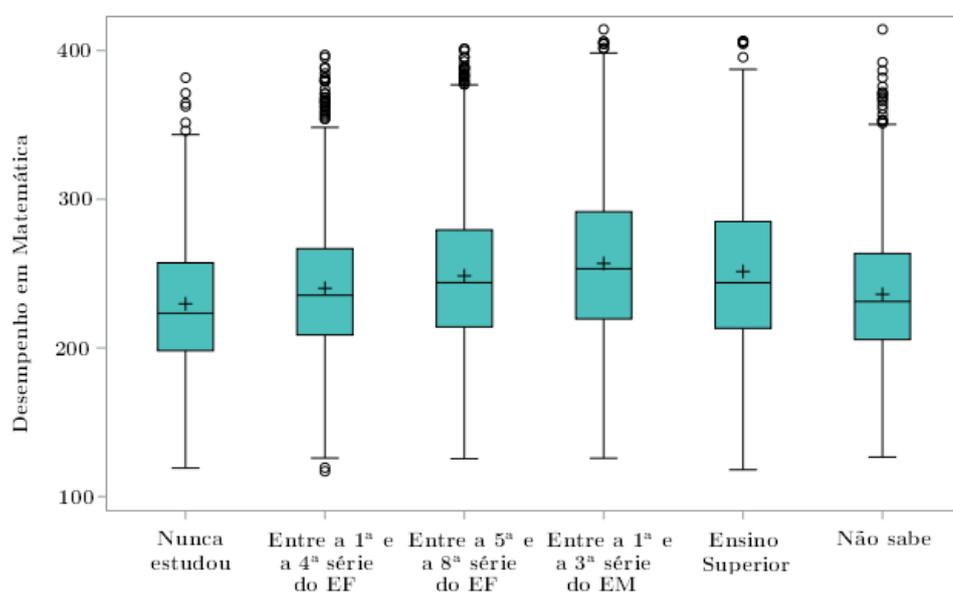


Figura 6 – Desempenho em matemática dos alunos segundo escolaridade da mãe ou mulher responsável.

Bolsa família

É possível perceber, por meio da Tabela 8 e Figura 7, que os estudantes cujas famílias não recebem o benefício do governo possuem desempenho médio maior. Além disso, 50% dos alunos que não recebem o benefício possuem desempenho de até 245,05 pontos na prova de matemática, valor que cai para 238,26 pontos para aqueles que recebem.

Tabela 8 – Medidas resumo do desempenho dos alunos em matemática de acordo com a bolsa família

Bolsa família	n	%	Média	Desvio Padrão
Sim	5.558	52	243,14	45,49
Não	5.133	48	250,68	50,19

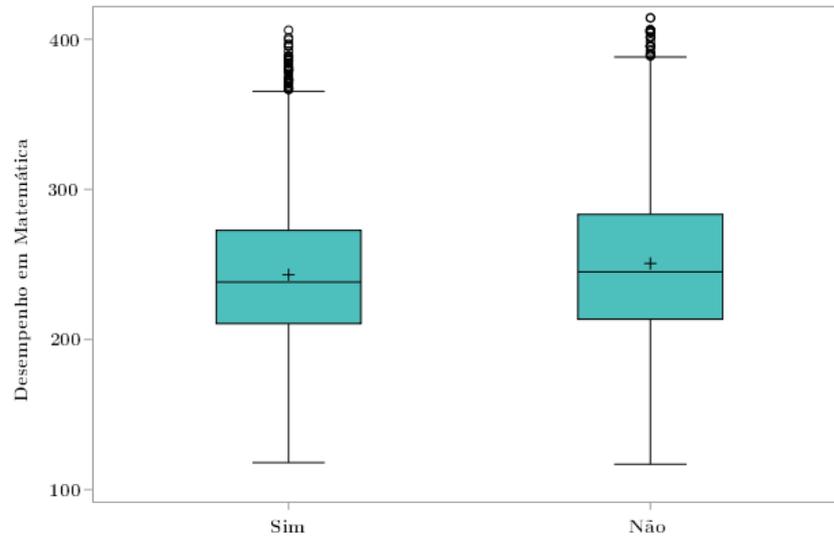


Figura 7 – Desempenho em matemática dos alunos de acordo com a bolsa família.

Disciplina que mais gosta

Somente 19,57% dos estudantes preferem matemática às demais matérias, sendo que as notas destes estudantes são maiores que as dos demais, como mostra a Tabela 9 e a Figura 8. Observa-se também que os estudantes que preferem língua portuguesa às demais matérias foram os que tiveram pior desempenho na prova de proficiência em matemática.

Tabela 9 – Medidas resumo do desempenho em matemática segundo a disciplina que o aluno mais gosta

Disciplina que mais gosta	n	%	Média	Desvio Padrão
Língua portuguesa	2.600	24,52	235,72	43,42
Matemática	2.075	19,57	272,62	53,41
Educação física	1.533	14,48	238,31	43,33
Outra	4.389	41,42	244,36	44,89

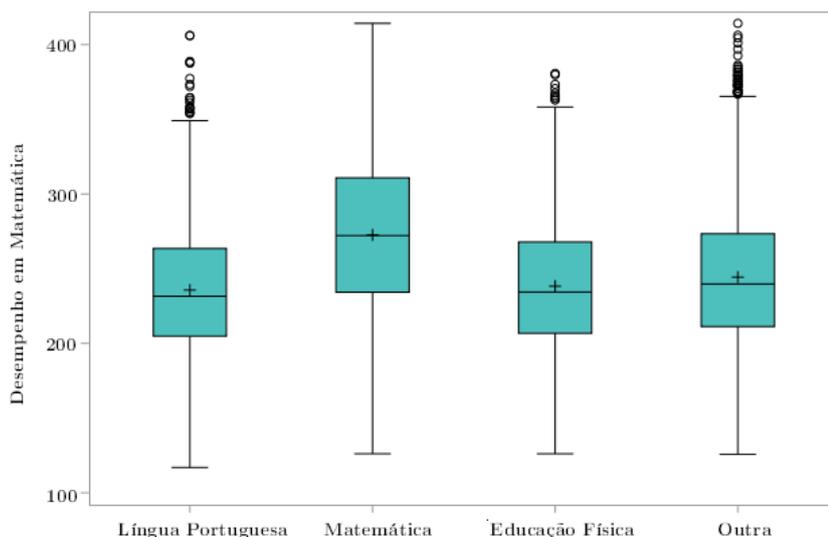


Figura 8 – Desempenho em matemática dos alunos segundo a disciplina que mais gosta.

Computador em casa

A partir da Tabela 10 e da Figura 9 nota-se que os alunos que afirmam ter acesso a um computador em casa têm melhor desempenho, sendo que o uso da internet não promove um melhor resultado na prova de matemática, apesar da internet possibilitar ao aluno acesso a inúmeras informações que podem trazer ganho para o seu aprendizado. É importante ressaltar também a grande quantidade de alunos sem computador em casa (72,69%).

Tabela 10 – Medidas resumo da variável desempenho em matemática segundo computador em casa

Computador em casa	n	%	Média	Desvio Padrão
Com acesso à internet	1.948	18,39	252,35	51,48
Sem acesso à internet	942	8,91	253,79	51,73
Não possui computador	7.700	72,69	244,78	46,20

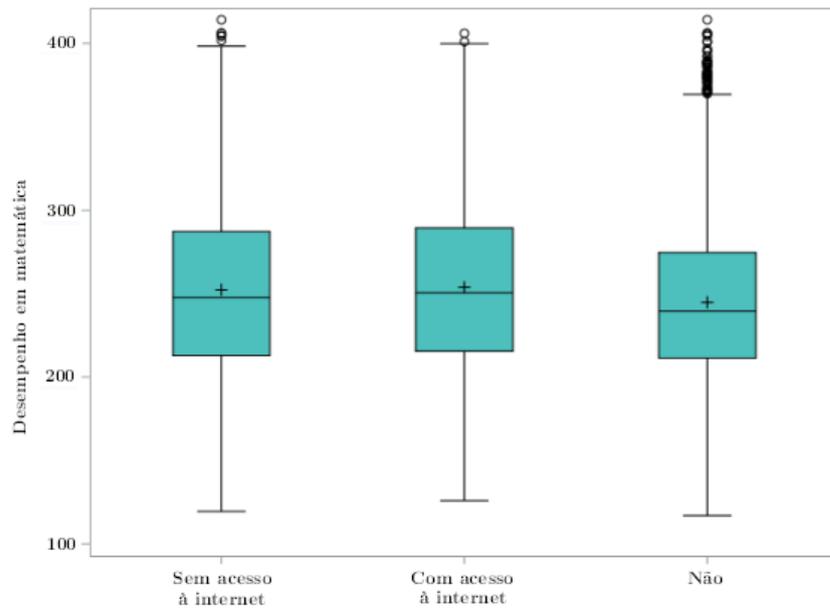


Figura 9 – Desempenho em matemática dos alunos segundo computador em casa.

De forma geral, apesar das diferenças nos desempenhos entre os grupos, pode-se dizer que essas diferenças são consideravelmente baixas, principalmente entre os alunos com acesso à internet e sem acesso à internet.

3.2 Análise Fatorial

Esta análise tem como finalidade a construção de fatores que representam as condições socioeconômicas dos estudantes e clima em sala de aula. Além de verificar, de forma exploratória, a relação destes fatores com o desempenho obtido pelos alunos.

Como base para construção dos fatores, foram utilizadas as variáveis socioeconômicas e variáveis que avaliam aspectos relacionados a dedicação e empenho do professor e comportamento dos estudantes.

Nível socioeconômico

Para geração do fator relativo a condição socioeconômica utilizou-se o método de extração por componentes principais. Somente um fator foi identificado (sugerido pelo *scree plot*), e explica cerca de 35% da variância total dos dados.

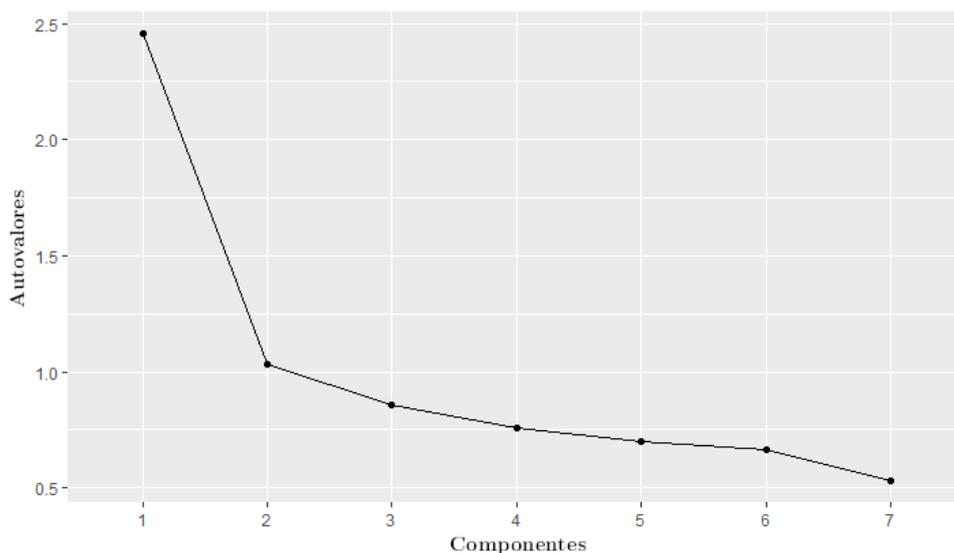
Figura 10 – *Scree plot*

Tabela 11 – Cargas fatoriais

Variáveis	Fator 1
Banheiro	0,60831
Rádio	0,59205
Geladeira	0,50491
Televisão a cores	0,70584
Máquina de lavar roupas	0,51085
Aparelho de DVD	0,66382
Automóvel	0,53069

Observa-se, na Tabela 11, que o único fator encontrado é composto por sete variáveis, sendo que a variável que mais contribui para este fator é a televisão à cores (item com maior carga).

Afim de avaliar, então, se o fator encontrado é relacionado com o desempenho dos alunos, foi calculado o coeficiente de correlação de Pearson entre os escores gerados pelo fator socioeconômico e a variável desempenho, obtendo uma correlação fraca (0,052). Graficamente, a associação entre ambos é dado pelo diagrama de dispersão da Figura 11.

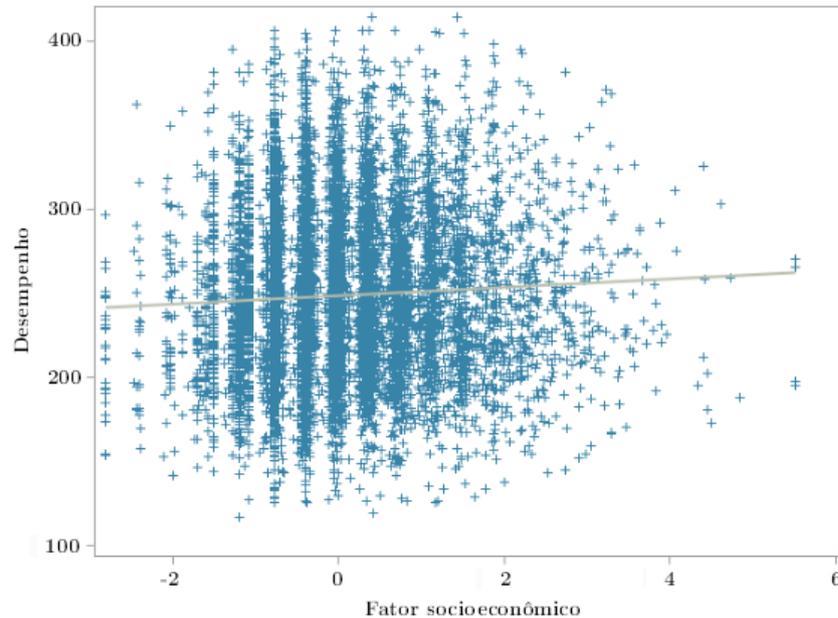


Figura 11 – Diagrama de dispersão do fator socioeconômico *versus* desempenho em matemática

Ao observar a Figura 11, nota-se que os pontos no diagrama de dispersão não mostram uma tendência fortemente crescente, indicando que a associação entre as variáveis analisadas é bastante fraca. Este efeito observado pode ser explicado devido as características da amostra considerada no estudo, essa amostra é composta apenas por estudantes da rede pública, o que pode não evidenciar o efeito do nível socioeconômico entre os alunos.

Sala de aula

Para a extração dos fatores relacionados ao clima em sala de aula utilizou-se o método de componentes principais. A solução utilizada foi obtida através do método de rotação VARIMAX (rotação ortogonal). Dois fatores foram identificados (sugeridos pelo *scree plot*), e juntos explicam cerca de 45% da variância total dos dados.

Através da matriz fatorial rotacionada, apresentada na Tabela 12, pode-se observar que o Fator 1 é composto por cinco itens que se referem ao relacionamento entre o professor e os alunos. Assim, ele pode ser denominado de relação professor-aluno. Já o Fator 2 é composto por cinco itens que se referem ao comportamento dos alunos em sala de aula. Logo, pode ser denominado de indisciplina.

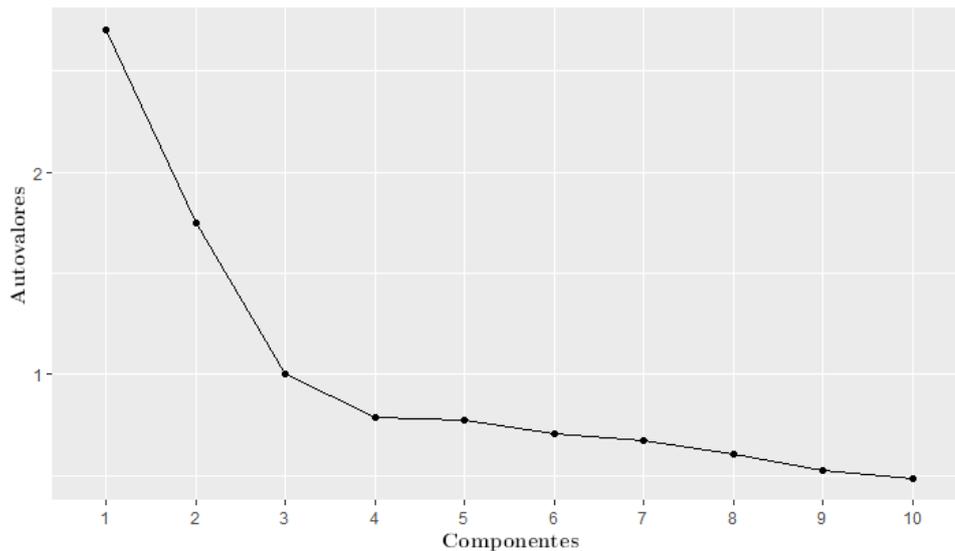


Figura 12 – Scree plot

Tabela 12 – Cargas fatoriais rotacionadas

Variáveis	Fator 1	Fator 2
Com que frequência o(a) seu(sua) professor(a) exige que os alunos estudem e prestem atenção nas aulas?	0,64446	0,00624
Com que frequência o(a) seu(sua) professor(a) mostra interesse no aprendizado de todos os alunos?	0,75278	-0,10849
Com que frequência o(a) seu(sua) professor(a) está disponível para esclarecer as dúvidas dos outros alunos?	0,72961	-0,02649
Com que frequência o(a) seu(sua) professor(a) dá as notas de maneira justa?	0,63741	-0,05749
Com que frequência o(a) seu(sua) professor(a) precisa esperar muito tempo até que os alunos façam silêncio?	0,23416	0,70471
Com que frequência há barulho e desordem na aula ?	0,02705	0,80675
Com que frequência os alunos saem da aula antes do término?	-0,16082	0,63376
Com que frequência o(a) seu(sua) professor(a) falta às aulas?	-0,24906	0,27082
Com que frequência o(a) seu professor(a) dá oportunidade de os alunos expressarem?	0,62438	-0,07022
Com que frequência os alunos não prestam atenção ao que o professor fala?	-0,34624	0,48267

A associação entre o Fator 1 e 2 e o desempenho em matemática dos alunos, dado pelo coeficiente de correlação de Pearson, resultou em 0,11 e -0,08, respectivamente. Através do diagrama de dispersão apresentado na Figura 13 é possível observar uma tendência crescente dos pontos, o que indica que o relacionamento entre os professores e alunos está positivamente relacionado com o desempenho destes. Já na Figura 14, onde é apresentada a relação entre a indisciplina e o desempenho dos estudantes, observa-se um comportamento levemente decrescente dos pontos, indicando que a indisciplina está negativamente relacionada ao desempenho dos alunos.

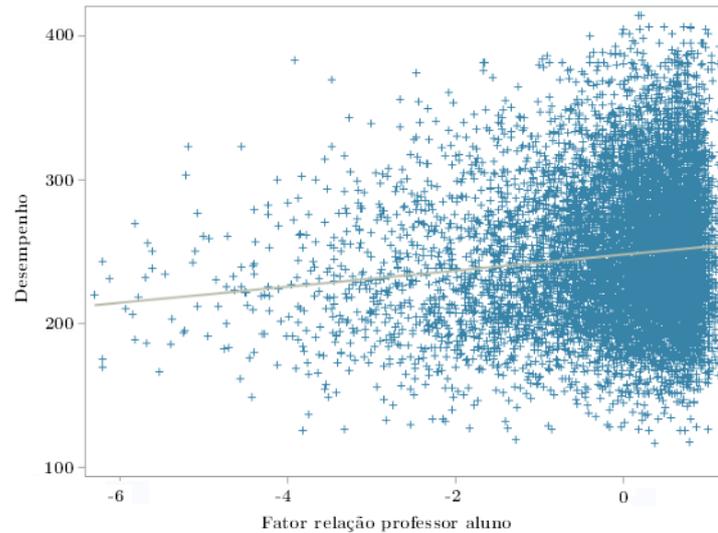


Figura 13 – Diagrama de dispersão do fator relação professor-aluno *versus* desempenho em matemática

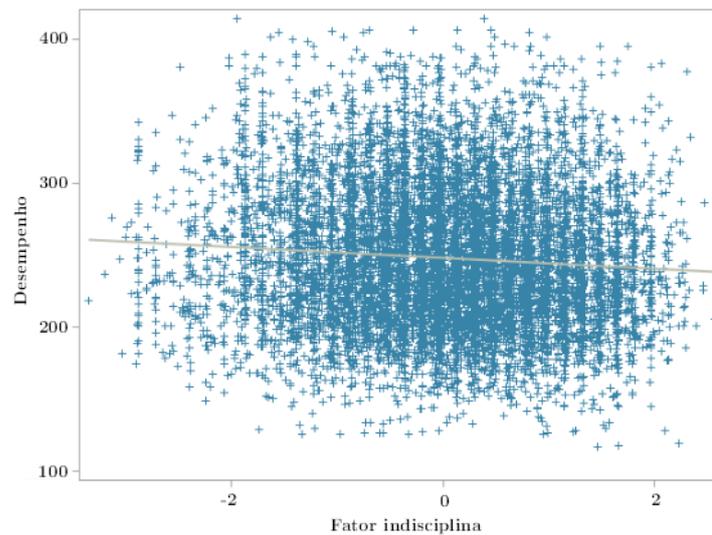


Figura 14 – Diagrama de dispersão do fator indisciplina *versus* desempenho em matemática

3.3 Análise Multinível

Para a realização da análise de regressão multinível, as variáveis qualitativas foram recodificadas assumindo somente valores 0 ou 1. A saber, *sexo* (0= feminino; 1= masculino), *etnia* (0= negro, pardo e indígena; 1= branco e amarelo), *reprovação* (0= nenhuma; 1= uma vez ou mais), *bolsa família* (0= não; 1= sim), *computador em casa* (0= não; 1= sim) e *disciplina que mais gosta* (0= outras; 1= matemática). Esta recodificação, que resultou em um agrupamento de classes de algumas variáveis, foi feita baseada nas análises anteriores e da literatura já existente (Vinha, Karino e Laros, 2016). Além disso, para a construção dos modelos, a variável *idade* foi desconsiderada por apresentar muitos valores

ausentes (11,39%). Por fim, todas as variáveis explicativas (qualitativas e quantitativas) foram padronizadas com média no valor zero e desvio padrão igual a 1, porque isso facilita a interpretação quando se quer comparar os efeitos de variáveis que possuem diferentes escalas (Hox, 2010).

Assim, seguindo os passos de modelagem proposto por Hox (2010), primeiramente foi ajustado o modelo sem variáveis explicativas, denominado de modelo nulo. De acordo com este modelo, a proficiência média em matemática dos alunos do 1º ano do Ensino Médio representado pelo intercepto é de 244,03 pontos.

Tabela 13 – Modelo nulo

Efeito Fixo	Efeito	EP ¹	Estatística t	p-valor
Intercepto	244,03	1,50	162,66	<0,001
Efeito randômico				
Variância do Nível 1 - aluno	2.014,46	27,55	73,11	<0,001
Variância do Nível 2 - escola	287,65	37,94	7,58	<0,001
ICC			0,12	
Deviance			113.483,9	

¹ Nota: Erro Padrão

A partir do modelo nulo (Tabela 13) pode-se calcular o coeficiente de correlação intraclasse dado por

$$ICC = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2} = \frac{287,65}{287,65 + 2014,46} = 0,12. \quad (12)$$

O resultado obtido mostra que 12% da variância do desempenho dos alunos está relacionada ao nível da escola, o que justifica o uso do método multinível. No entanto, se comparado a outros trabalhos (Andrade e Laros, 2007; Laros, Marciano e Andrade, 2012) o ICC encontrado é relativamente baixo, esse fenômeno pode ser atribuído ao fato de as escolas neste estudo serem mais homogêneas, todas públicas e de uma mesma cidade (Fortaleza).

Entre o primeiro e o segundo passo proposto por Hox (2010), foram inseridas as variáveis de controle: *nível socioeconômico*, *nível socioeconômico agregado*, *etnia* e *escola-*

ridade da mãe (Tabela 14). O objetivo é verificar quais são as outras fontes de variação entre as escolas ao controlar essas variáveis (Fletcher, 1998). No entanto, ao adicionar *escolaridade da mãe*, a variável *nível socioeconômico* teve efeito negativo. Este efeito pode ser consequência de multicolinearidade, já que quando analisado separadamente, o efeito do *nível socioeconômico* está positivamente relacionado à proficiência em matemática. Assim, optou-se por retirar a variável *nível socioeconômico* do nível do aluno dos modelos posteriores.

Tabela 14 – Modelo com as variáveis de controle (Modelo 1)

Efeito Fixo	Efeito	EP	Estatística t	p-valor
Intercepto	245,67	1,26	195,75	<0,001
Variáveis controle				
NSE agregado	34,48	5,19	6,65	<0,001
Etnia	0,78	0,44	1,78	0,075
Escolaridade da mãe	4,83	0,45	10,81	<0,001
Efeito randômico				
Variância do Nível 1 - aluno	1.987,73	27,48	72,32	<0,001
Variância do Nível 2 - escola	186,86	26,26	7,12	<0,001
ICC			0,09	
Deviance			110.905,8	
Nº de parâmetros			6	
Diferença do Deviance			2.578,1	
Diferença de parâmetros (g.l.)			3	
Teste χ^2 (p-valor)			<0,001	
Variância explicada				
Nível 1 - aluno			1,3%	
Nível 2 - escola			35,0%	

Na Tabela 14 pode-se observar pelo p-valor obtido no teste χ^2 (<0,001) que a diferença entre os deviances do Modelo 1 em relação ao Modelo nulo é significativa, indicando uma melhora no ajuste do modelo. Também pode-se observar os efeitos das variáveis de controle no desempenho dos alunos, sendo que o efeito com maior destaque é o *nível socioeconômico agregado*. O aumento de um desvio padrão do *nível socioeconômico da*

escola, aumenta, em média, 34,48 pontos o desempenho do aluno.

É interessante verificar no Modelo 1 que as variáveis controle explicaram 35% da variância do nível da escola. Em outras palavras, 35% da variância do desempenho médio dos alunos entre as escolas deve-se ao *nível socioeconômico da escola, etnia e escolaridade da mãe*. Este resultado indica o quão grande é o efeito da composição socioeconômica no desempenho dos alunos. Já em relação ao nível dos alunos, apenas 1% da variância foi explicada.

O segundo passo, sugerido por Hox (2010), consiste na inserção das variáveis explicativas do nível 1 (nível do aluno). Assim, as variáveis utilizadas para este passo foram *sexo, etnia, reprovação, bolsa família, tem computador em casa, disciplina que mais gosta, fator relação professor-aluno e fator indisciplina*.

A Tabela 15 apresenta os resultados do modelo proposto, sendo que a variável *tem computador em casa* foi retirada do modelo porque não teve uma contribuição significativa.

Assim, os resultados apresentados na Tabela 15 (Modelo 2) revelam que a variável com maior influência no desempenho dos alunos é a *disciplina que mais gosta*. O fato do aluno gostar de matemática aumenta, em média, 10,80 pontos no desempenho obtido no teste de proficiência em matemática.

Outras variáveis que também apresentam uma influência grande no desempenho são *reprovação* e *sexo*. Alunos que já reprovaram uma vez ou mais têm seu desempenho médio reduzido em 9,75 pontos em relação aos alunos que nunca reprovaram. Já em relação ao *sexo*, nota-se pelo efeito dessa variável que alunos do *sexo masculino* possuem um desempenho médio superior aos alunos do *sexo feminino* (7,90 pontos).

Quanto a variável *bolsa família*, chama a atenção o efeito negativo (-1,04). Assim, os alunos que recebem bolsa família possuem desempenho médio reduzido em 1,04 pontos em relação aqueles que não recebem. É importante ressaltar que este resultado não deve ser interpretado erroneamente, no sentido de que a *bolsa família* causa um desempenho fraco em matemática, pois este estudo é apenas observacional com corte transversal em que não se pode estabelecer relações de causa e efeito.

Ainda, de acordo com o Modelo 2, o *fator relação professor-aluno* contribui de forma positiva para o desempenho dos alunos. Os alunos com professores interessados no aprendizado da turma, que exigem atenção, que estão disponíveis para esclarecer dúvidas e dão as notas de maneira justa, possuem um desempenho médio maior do que aqueles

que possuem professores descompromissados.

Nota-se também que a indisciplina do aluno produz impacto negativo sobre o seu desempenho. Alunos que percebem barulho e desordem na sala, saídas antes dos términos das aulas e falta de atenção ao que o professor fala têm, em média, uma queda de 1,09 pontos no teste de proficiência em relação aos alunos que percebem maior comprometimento e disciplina na sala de aula.

Tabela 15 – Modelos com as variáveis explicativas (Modelo 2 e Modelo 3)

Efeito Fixo	Modelo 2			Modelo 3		
	Efeito	EP	p-valor	Efeito	EP	p-valor
Intercepto	246,99	1,07	<0,001	247,47	1,04	<0,001
Variáveis controle						
NSE agregado	27,77	4,45	<0,001	24,35	4,45	<0,001
Etnia	1,16	0,42	0,006	1,16	0,42	0,006
Escolaridade da mãe	3,83	0,43	<0,001	3,81	0,43	<0,001
Variáveis do nível aluno						
Sexo	7,90	0,42	<0,001	7,90	0,42	<0,001
Reprovação	-9,75	0,43	<0,001	-9,75	0,43	<0,001
Bolsa família	-1,04	0,43	0,015	-1,05	0,43	0,014
Disciplina que mais gosta	10,80	0,42	<0,001	10,80	0,42	<0,001
Fator relação professor-aluno	3,89	0,44	<0,001	3,90	0,44	<0,001
Fator indisciplina	-1,09	0,44	0,013	-1,04	0,44	0,018
Variáveis do nível da escola						
Salas				5,81	1,55	<0,001
Número de funcionários				-1,99	1,68	0,236
Efeito randômico						
Variância do Nível 1 - aluno	1.687,25	24,27	<0,001	1.687,21	24,27	<0,001
Variância do Nível 2 - escola	128,74	19,03	<0,001	111,45	16,91	<0,001
Deviance		100.959,6			100.943,1	
Nº de parâmetros		12			14	
Diferença do Deviance		9.946,2			16,5	
Diferença de parâmetros (g.l.)		6			2	
Teste χ^2 (p-valor)		<0,001			<0,001	
Variância explicada						
Nível 1 - aluno		16,2% (15,1%)			16,2% (15,1%)	
Nível 2 - escola		55,2% (31,1%)			61,3% (40,4%)	

Por fim, verifica-se que o Modelo 2 explica 16,2% da variância do nível do aluno e 55,2% do total de variância em relação ao nível da escola. É importante ressaltar que o

percentual da variância explicada do nível do aluno (15,1%) e da variância explicada pelo nível da escola (31,1%) em relação ao modelo com as variáveis controle (Modelo 1), estão apresentadas entre parênteses.

Após a inserção das variáveis do nível do aluno, o próximo passo é incluir as variáveis explicativas do nível da escola (Modelo 3). Foram consideradas nesta etapa as variáveis *infraestrutura*, *número de salas de aula* e *número de funcionários*. Somente a variável *número de salas de aula* foi significativa ao nível de 5%, no entanto o *número de funcionários* foi mantido no modelo por apresentar, no ajuste do Modelo 5, interação significativa com o *fator relação professor-aluno*.

O único efeito significativo apresentado no nível da escola indica que o *número de salas de aula* contribui de forma positiva no desempenho do aluno. Isto é, escolas maiores tendem a ter desempenho médio maior dos alunos em 5,81 pontos.

Ainda no Modelo 3, apresentado na Tabela 15, verifica-se que 61,3% da variância total do nível 2 foi explicada com a adição das variáveis *número de salas* e *número de funcionários*. Já em relação ao modelo controle, 40,4% da variância do nível da escola foi explicada.

O quarto passo consiste em verificar se existe efeito randômico dos coeficientes de inclinação das variáveis do nível do aluno. Isto é, se as variáveis do nível do aluno se comportam de maneira diferente entre as escolas. Para testar o efeito randômico foram consideradas as variáveis de controle *etnia* e *escolaridade da mãe* e as demais variáveis do nível do aluno. No Modelo 4, apresentado na Tabela 16, apenas a variável *disciplina que mais gosta* apresentou efeito randômico significativo. Isso significa que o aluno gostar de matemática tem efeito diferenciado sobre o seu desempenho dependendo da escola em que estuda.

No último passo verifica-se os efeitos de interação entre os níveis. Nesse passo, é importante ressaltar que as variáveis do nível da escola testadas no Modelo 3 que deram não significativas (*infraestrutura* e *número de funcionários*) também foram consideradas, o mesmo ocorreu com a variável do nível 1 (*computador em casa*).

De todas as possíveis interações entre as variáveis do nível do aluno e do nível da escola apenas uma apresentou efeito significativo (Tabela 16, Modelo 5), a interação entre o *fator relação professor-aluno* e *número de funcionários*. Este efeito de interação indica que a dedicação e disponibilidade do professor para com os alunos tem efeito reduzido em

escolas com maior número de funcionários.

Tabela 16 – Modelo com efeito randômico e interação (Modelo 4 e Modelo 5)

Efeito Fixo	Modelo 4			Modelo 5		
	Efeito	EP	p-valor	Efeito	EP	p-valor
Intercepto	247,48	1,04	<0,001	247,44	1,04	<0,001
Variáveis controle						
NSE agregado	24,26	4,44	<0,001	24,16	4,42	<0,001
Etnia	1,15	0,42	0,006	1,15	0,42	0,006
Escolaridade da mãe	3,80	0,43	<0,001	3,80	0,43	<0,001
Variáveis do nível aluno						
Sexo	7,91	0,42	<0,001	7,92	0,42	<0,001
Reprovação	-9,73	0,43	<0,001	-9,71	0,43	<0,001
Bolsa família	-1,04	0,43	0,015	-1,05	0,43	0,014
Disciplina que mais gosta	10,82	0,48	<0,001	10,82	0,48	<0,001
Fator relação professor-aluno	3,93	0,44	<0,001	4,04	0,44	<0,001
Fator indisciplina	-1,05	0,44	0,016	-1,07	0,44	0,014
Variáveis do nível da escola						
Salas	5,78	1,55	<0,001	5,65	1,54	<0,001
Número de funcionários	-1,94	1,67	0,247	-1,77	1,67	0,290
Efeito de interação						
Fator relação professor aluno x N° de funcionários				-1,07	0,45	0,017
Efeito randômico - Nível 1						
Variância do Nível 1 - aluno	1.681,89	24,33	<0,001	1.681,18	24,32	<0,001
Efeito Randômico - Nível 2						
Variância do Nível 2 - escola	110,84	16,86	<0,001	109,52	16,71	<0,001
Variância Disciplina que mais gosta	5,52	3,47	0,055	5,46	3,46	0,057
Deviance	100.939,2			100.933,5		
N° de parâmetros	15			16		
Diferença do Deviance	3,9			5,7		
Diferença de parâmetros (g.l.)	1			1		
Teste χ^2 (p-valor)	0,0483			0,017		

Para verificar se o modelo final (Modelo 5) atende as suposições de normalidade, homocedasticidade e linearidade, foi feita uma análise de resíduos dos níveis 1 e 2. Através dos gráficos apresentados nas Figuras 15 e 16, constata-se que para os resíduos do nível 1 não houve violações das suposições. Na Figura 15 observa-se que não existe desvios dos pontos em relação a reta indicando adequação a normalidade e na Figura 16 os pontos estão distribuídos de forma aleatória em torno no zero indicando linearidade e variância constante.

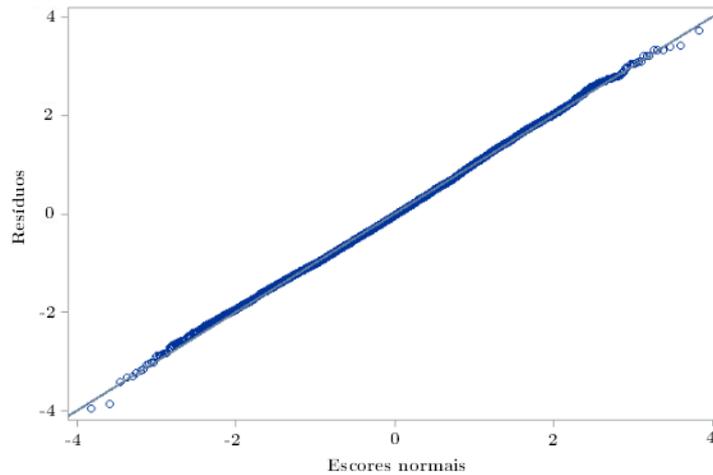


Figura 15 – Gráfico dos resíduos padronizados *versus* escores normais

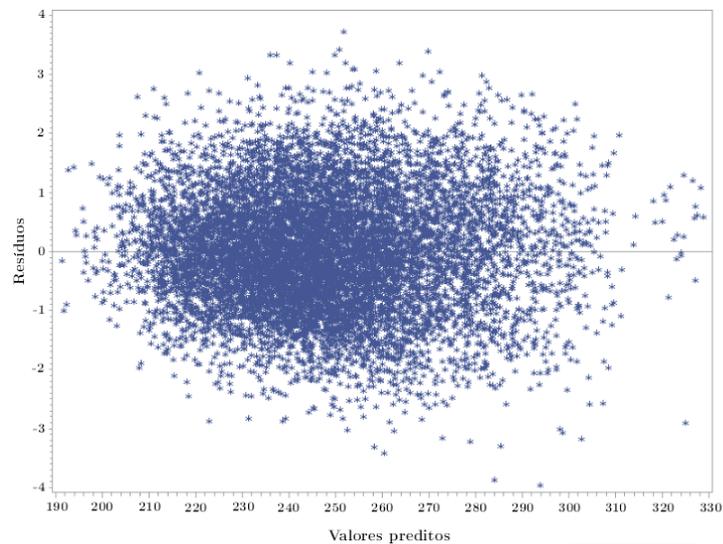


Figura 16 – Gráfico dos resíduos *versus* valores preditos

Com relação aos resíduos do nível 2, os gráficos apresentados nas Figuras 17 e 18, os quais dizem respeito as componentes aleatórias do nível 2, revelam que não houve fortes violações da normalidade. Os pontos se ajustam de forma razoável a reta exceto pela cauda superior onde se distanciam consideravelmente, indicando caudas pesadas. Quanto a Figura 19 pode-se observar que os pontos se comportam de forma aleatória atendendo a suposição de homocedasticidade.

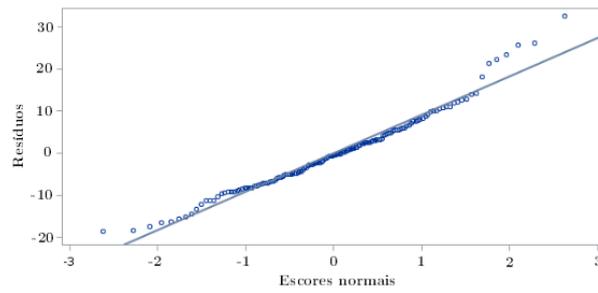


Figura 17 – Gráfico dos resíduos *versus* escores normais para o Intercepto

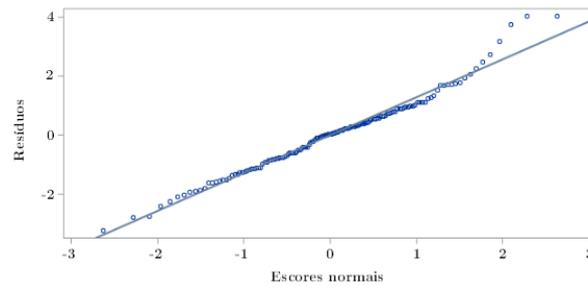
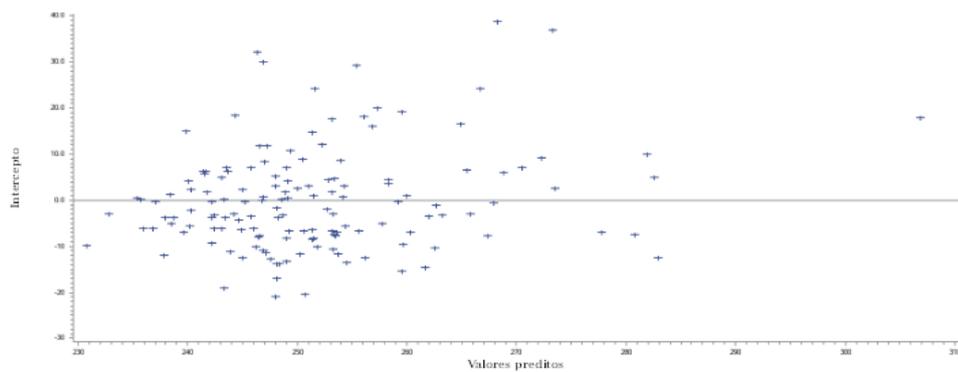
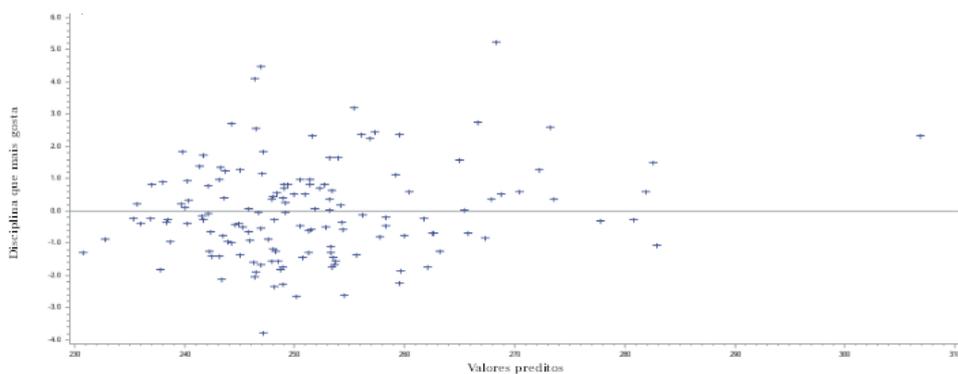


Figura 18 – Gráfico dos resíduos *versus* escores normais para a *disciplina que mais gosta*



(a)



(b)

Figura 19 – Gráfico dos resíduos *versus* valores preditos para o intercepto (a) e *disciplina que mais gosta* (b)

4 CONCLUSÃO

O presente estudo propõe avaliar, através da análise multinível, o efeito dos fatores associados ao clima em sala de aula, a indisciplina e a relação entre alunos e professores, sobre o desempenho dos estudantes considerando os fatores contextuais que influenciam na aprendizagem.

Com relação às variáveis de controle observou-se que estas explicaram 35% da variância do desempenho dos alunos no nível da escola e 1,3% da variância do desempenho no nível dos alunos, sendo que a variável com maior efeito foi o *nível socioeconômico do aluno agregado*. Este efeito evidencia o quanto o nível socioeconômico pode influenciar no desempenho do alunado.

No que diz respeito aos fatores associados ao clima em sala de aula, observou-se que estes afetam significativamente o desempenho dos estudantes. Os resultados do presente estudo mostraram que a indisciplina do aluno gera impacto negativo no seu desempenho, enquanto a relação entre professor e aluno afeta positivamente a nota obtida no teste de proficiência em matemática.

No nível 1, verificou-se que a variável com maior efeito no desempenho em matemática do aluno foi *disciplina que mais gosta*. Já em relação as variáveis do nível 2, apenas o *número de salas de aula* apresentou efeito significativo. Além disso, o estudo apresentou efeito randômico para a variável *disciplina que mais gosta* e efeito de interação entre *fator relação professor-aluno* e *número de funcionários*.

Ademais, o resíduo do nível do aluno atendeu plenamente as suposições de normalidade, linearidade e homocedasticidade. Entretanto, no nível da escola, os gráficos dos resíduos padronizados *versus* os escores normais apresentaram uma oscilação em relação a reta indicando um desvio da normalidade.

Por fim, vale ressaltar que ainda existe um número restrito de pesquisas no Brasil que buscam avaliar a influência do clima em sala de aula sobre o desempenho obtido pelos alunos. Sugere-se que estudos futuros realizem análises similares considerando também o desempenho em português e séries do Ensino Fundamental.

REFERÊNCIAS

- ANDRADE, J. M.; LAROS, J. A. Fatores associados ao desempenho escolar: um estudo multinível com dados do SAEB/2001. *Psicologia: Teoria e Pesquisa*, v. 23, n.1, p. 33-42, 2007.
- FRANCO, C. et al. Qualidade e equidade em educação: reconsiderando o significado de “fatores intra-escolares”. *Ensaio: Avaliação e Políticas Públicas em Educação*, v. 15, n 55, p. 277-298, 2007.
- FLETCHER, P. R. À procura do ensino eficaz. Rio de Janeiro: Ministério da Educação e Cultura, Departamento da Avaliação da Educação Básica, 1998.
- HOX, J. J. *Multilevel analysis: Techniques and Applications*. 2 ed. Grã - Bretanha: Routledge, 2010. 392 p.
- LAROS, J. A.; MARCIANO, J. L.; ANDRADE, J. M. Fatores associados ao desempenho escolar em Português: um estudo multinível por regiões. *Ensaio: Avaliação e Políticas Públicas em Educação*, v.20, p. 623-646, 2012.
- MINGOTI, S. A. *Análise de dados através de métodos de estatística multivariada – uma abordagem aplicada*. Belo Horizonte: UFMG, 2005. 295p.
- SOARES, J. F.; ALVES, M. T. G. Desigualdades raciais no sistema brasileiro de educação básica. *Educação e Pesquisa*, São Paulo, v. 29, p. 147-165, 2003.
- VINHA, L. G. A.; J. A.; KARINO, C. A.; J. L. LAROS, J. A. Fatores associados ao desempenho em matemática no ensino fundamental no Brasil. *Revista Psico - USF*, v. 21, n.1, p. 87-100, 2016.