



Universidade de Brasília - UnB  
Instituto de Ciências Exatas - IE  
Departamento de Estatística - EST

**Estudo da relação entre a percepção de violência e conflitos nas  
escolas e o desempenho acadêmico**

Marina Macedo Gonçalves da Paixão

Brasília  
2018



Marina Macedo Gonçalves da Paixão

**Estudo da relação entre a percepção de violência e conflitos nas escolas e o desempenho acadêmico**

Orientador:

Prof. Dr. Luís Gustavo do Amaral Vinha

Relatório Final do Trabalho de Conclusão do Curso 2 apresentado para o Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito necessário para obtenção do grau de Bacharel em Estatística.

Brasília

2018



*"O sucesso é ir de fracasso em fracasso  
sem perder o entusiasmo."*

*Winston Churchill*



---

## Agradecimentos

Concluir uma monografia não é algo fácil e acredito que estar cercada de pessoas boas me trouxe (e continuará trazendo) bons frutos, afinal, "nenhum de nós é tão bom quanto todos nós juntos".

Primeiramente, gostaria de agradecer a Deus, pois sei que, sem o apoio d'Ele, nada disso seria possível. Sua bondade, seu amor e sua confiança me deram forças para concluir essa etapa.

Agradeço aos meus pais, Fátima e José, que sempre me incentivaram a dar o meu melhor em tudo que faço. Foram inúmeras noites chorando em razão de provas, de estresse e das mais diversas dificuldades, porém eles sempre estiveram ali me apoiando. O meu amor por vocês é infinito e eu agradeço todos os dias por tê-los ao meu lado.

Agradeço, ainda, ao meu orientador, Prof. Dr. Luís Gustavo do Amaral Vinha, que sempre foi muito interessado, paciente e atencioso ao longo de todo esse período de orientação. Obrigada por sempre estar disponível para tirar dúvidas e por me incentivar nessa reta final do curso.

Agradeço à Universidade de Brasília e, principalmente, aos professores e servidores do Departamento de Estatística por me proporcionarem a realização desse curso de bacharelado.

Agradeço a todos os meus amigos pela amizade, pois quando precisava desabafar, me distrair, estudar ou apenas conversar, vocês estavam sempre presentes. Enfatizo, aqui, um agradecimento especial a Bárbara, Taísa, Regina, Juliana, Enya, Caio.

Tive a oportunidade de realizar um intercâmbio pelo Ciência sem Fronteiras (CsF) na Trinity College Dublin, o que mudou minha perspectiva de vida pessoal e profissional, logo, não poderia deixar de agradecer Ana Luíza, Hercules, Larissa, George, Paulo e Jorge, que foram meus companheiros nessa época.

Agradeço também à todos que fizeram parte da minha jornada na ESTAT Consultoria, onde passei alguns dos melhores anos da minha vida. O crescimento e o aprendizado que obtive nessa empresa júnior é incomparável.

Obrigada aos meus colegas de curso, em especial, Bruno, Eduarda, Ludimila, Luíza, Thayanne e Matheus. Por fim, gostaria de agradecer à Laura Teixeira, que esteve presente em todos os momentos mais importantes da UnB, e à Leylanne e ao Lucas por estarem comigo nessa reta final de curso. Todos os fins de semana estudando finalmente estão chegando a um fim, amigos.



---

## Resumo

Dentro de um contexto no qual o direito à educação é uma garantia constitucional e as instituições escolares tem tido problemas frequentes com violência e conflitos envolvendo seus estudantes, o presente estudo buscou analisar a relação entre a percepção de violência e conflitos nas escolas na visão dos diretores com o desempenho acadêmico dos estudantes do 5º ano do ensino fundamental para o estado de São Paulo. Ajustou-se um modelo de regressão multinível com efeitos fixos e randômicos a partir de microdados provenientes do SAEB de 2015. As variáveis respostas foram as proficiências em língua portuguesa e em matemática. As variáveis sexo, etnia, idade, escolaridade da mãe, reprovação, dever de língua portuguesa, dever de matemática, nível econômico, incentivo social e nível cultural entraram no modelo final como variáveis explicativas do nível do aluno e a violência como variável explicativa do nível da escola. Observou-se que a reprovação, o dever de matemática, a idade e a escolaridade da mãe são os fatores com maior influência sob o desempenho acadêmico dos estudantes nas disciplinas de língua portuguesa e de matemática. Constatou-se que a violência exerce um impacto negativo no desempenho dos alunos em língua portuguesa e em matemática. Os resultados deste relatório podem contribuir para a descoberta dos principais fatores que influenciam o aprendizado escolar, buscando seu constante aprimoramento.

**Palavras-chave:** Regressão Multinível; Avaliação Educacional; Violência e conflitos; Desempenho acadêmico; SAEB.



## Sumário

<b>1 Introdução</b>	15
1.1 Direito à educação	15
1.2 Sistema de avaliação da educação básica (SAEB)	17
1.3 Violência e conflitos nas escolas	17
1.4 Revisão de literatura	19
1.5 Problematização e objetivos do estudo	20
<b>2 Referencial teórico</b>	21
2.1 Modelos de regressão multinível	21
2.2 Passos para construção	23
2.3 Métodos de estimação	25
2.4 Comparação de modelos	25
2.5 Análise de resíduos	26
<b>3 Metodologia</b>	28
3.1 Dados	28
<b>4 Resultados e discussões</b>	33
4.1 Análise descritiva unidimensional	33
4.1.1 Escolas	33
4.1.2 Alunos	35
4.1.3 Reprovação	38
4.1.4 Desempenho escolar	38
4.2 Análise descritiva bidimensional	39
4.2.1 Escala de violência vs. desempenho escolar	39
4.3 Modelagem	42
<b>5 Conclusão</b>	53
<b>Referências</b>	55



## Lista de Figuras

1	Hierarquia presente em Modelos de Regressão Multinível, no qual os indivíduos do nível 1 encontram-se agrupados em um nível 2. . . . .	21
2	Distribuição da variável dependência administrativa para as escolas do 5º ano do ensino fundamental. . . . .	33
3	Distribuição da variável localização para as escolas do 5º ano do ensino fundamental. . . . .	34
4	Variabilidade da escala de violência nas escolas no estado de São Paulo.	35
5	Distribuição da variável sexo para os alunos do 5º ano do ensino fundamental. . . . .	36
6	Distribuição da variável faixa etária para os alunos do 5º ano do ensino fundamental. . . . .	37
7	Distribuição da variável etnia autodeclarada para os alunos do 5º ano do ensino fundamental. . . . .	37
8	Distribuição da variável do número de reprovações dos alunos do 5º ano do ensino fundamental. . . . .	38
9	Proficiência dos alunos em Língua Portuguesa e Matemática. . . . .	39
10	Escala de violência relacionada com o desempenho escolar em língua portuguesa. . . . .	40
11	Escala de violência relacionada com o desempenho escola em matemática. . . . .	41
12	Análise de homocedasticidade para os valores ajustados e os resíduos do nível do aluno no modelo final. . . . .	49
13	Análise de normalidade para os resíduos do nível do aluno no modelo final. . . . .	49
14	Análise de normalidade para os resíduos do nível da escola no modelo final. . . . .	50
15	Análise de independência para os resíduos do nível do aluno no modelo final. . . . .	50



---

## Lista de Tabelas

1	Proficiência de Língua Portuguesa e Matemática. . . . .	29
2	Informações do banco de dados de São Paulo. . . . .	29
3	Variáveis do Questionário do Aluno. . . . .	30
4	Fatores relacionados ao nível econômico, nível cultural e incentivo social. . . . .	31
5	Fatores relacionados à violência segundo o diretor. . . . .	32
6	Coeficiente de Correlação Intraclasse para 3 níveis . . . . .	42
7	Modelo Nulo . . . . .	43
8	Modelo 1 . . . . .	44
9	Modelo 2 . . . . .	45
10	Modelo 3 . . . . .	46
11	Modelo 4 . . . . .	48
12	Modelo para São Paulo completo . . . . .	51



# 1 Introdução

## 1.1 Direito à educação

Ao longo das últimas décadas, diversos avanços tecnológicos e sociais marcaram a sociedade. Essas mudanças demonstraram uma certa inquietude quanto à garantia dos direitos fundamentais. Nesse cenário, a educação é constantemente vista como o principal veículo de transformação para uma sociedade mais igualitária e mais digna.

Os princípios da igualdade e da dignidade humana, de acordo com a Constituição Federal Brasileira de 1988<sup>1</sup>, são extremamente importantes para todos que fazem parte da sociedade. A garantia de um tratamento igualitário e de uma proteção à pessoa humana findada na redução de quaisquer desigualdades ou discriminações estão asseguradas por meio deles.

Dentro de um contexto que busca a igualdade de oportunidades, encontra-se inserida a educação. Para que sejam atingidos a igualdade de chance para o trabalho e o exercício pleno da cidadania, a educação passa a ser elencada como um meio obrigatório para o alcance desses objetivos. A cidadania está fortemente relacionada à compreensão do contexto no qual a pessoa está inserida, seja ele político, econômico, social, ambiental ou cultural, pressupondo, em todos os casos, uma educação básica acerca de assuntos comuns a todos (Bittar, 2014).

A educação está presente na Carta Magna como um direito social público e subjetivo, além de também estar assegurada por meio de diversas outras leis que buscam a garantia e a efetividade do sistema educacional. Diante disso, a educação pode ser vista como um instrumento essencial para a formação do caráter cidadão e da identidade individual. Observa-se, então, o artigo 6º da Constituição Federal:

Art. 6º. São direitos sociais a educação, a saúde, a alimentação, o trabalho, a moradia, o transporte, o lazer, a segurança, a previdência social, a proteção à maternidade e à infância, a assistência aos desamparados, na forma desta Constituição.

Com isso, o Estado possui uma obrigação formal de garantir a educação de qualidade a todos os brasileiros. Porém, não é o único responsável, sendo um dever tanto do Estado quanto da família de forma a proporcionar um desenvolvimento tanto para o país quanto para o indivíduo. Cabe, além disso, à sociedade incentivar e colaborar para o alcance dessa finalidade. Conforme artigo 227, verifica-se:

---

<sup>1</sup>Constituição Federal do Brasil de 1988

Art. 227. É dever da família, da sociedade e do Estado assegurar à criança, ao adolescente e ao jovem, com absoluta prioridade, o direito à vida, à saúde, à alimentação, à educação, ao lazer, à profissionalização, à cultura, à dignidade, ao respeito, à liberdade e à convivência familiar e comunitária, além de colocá-los a salvo de toda forma de negligência, discriminação, exploração, violência, crueldade e opressão.

Portanto, o direito à educação é um direito constitucional fundamental e universal. Nesse sentido, a educação brasileira pode ser classificada em dois níveis: básico e superior. De acordo com a Lei de Diretrizes e Bases da Educação (LDB)<sup>2</sup>, esse primeiro é formado pelos ensinamentos infantil, fundamental e médio.

O ensino básico tem sua importância pautada na aprendizagem de competências que permitam com que o aluno desenvolva atividades essenciais para as próximas etapas de seus estudos e também para o mercado de trabalho. Sendo assim, ao final do ensino básico, o objetivo é que o aluno tenha uma formação ética e moral bem definidas e uma plena capacidade de compreender o ambiente natural, social, político, cultural e tecnológico no qual está inserido. Em suma, as disciplinas que compõem a base nacional dizem respeito às áreas de linguagem, matemática, ciências naturais e ciências humanas e permitem a composição de um vasto conhecimento ao aluno.

Durante as fases iniciais, isto é, entre o ensino infantil e o 5º ano do ensino fundamental, o estudante começa a desenvolver uma rotina de leitura e escrita, além de, por exemplo, aprender a realizar operações matemáticas simples. Nos últimos anos desse ciclo básico, época na qual o aluno já tem as funções básicas de leitura, escrita e cálculo bem consolidadas, esse estudante adquire, aos poucos, uma visão crítica do mundo e ganha conhecimentos de contexto histórico e de ciências mais avançadas, podendo, eventualmente, persistir no caminho acadêmico por meio do ingresso em nível superior (Oliveira, 2014).

É importante ressaltar que a garantia constitucional para com a educação somente diz respeito ao ensino básico, não constando o ensino superior no rol de dever do Estado Brasileiro. Dessa forma, a educação escolar pública tem provimento obrigatório e gratuito aos jovens entre 4 e 17 anos de acordo com as etapas de pré-escola, ensino fundamental e ensino médio.

Com isso, a educação tem a constante necessidade de investigar e conhecer quais métodos teriam maior influência na qualidade de ensino a ser ofertado, traçando, assim, uma análise do que seria considerado uma boa prática dentro do ambiente escolar e de quais fatores estariam mais fortemente relacionados a um alto padrão educacional.

---

<sup>2</sup>Lei 9.394/1996

## 1.2 Sistema de avaliação da educação básica (SAEB)

É interessante notar que todas os níveis de ensino, com exceção da educação infantil, atualmente, passam por algum tipo de avaliação padronizada e organizada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP)<sup>3</sup>, que está vinculado ao Ministério da Educação e Cultura (MEC)<sup>4</sup>.

Para avaliação da educação básica, por exemplo, existe o SAEB, que é composto pela Avaliação Nacional da Educação Básica (Aneb), pela Prova Brasil e pela Avaliação Nacional da Alfabetização (ANA), enquanto para a avaliação do ensino superior, existe o Sistema Nacional de Avaliação da Educação Superior (Sinaes). No presente trabalho, a proposta principal será em torno dos dados obtidos via SAEB.

Frente a necessidade do Brasil em obter informações a respeito da educação brasileira, o Sistema de Avaliação da Educação Básica (SAEB) foi implantado no ano de 1990 para suprir essa carência de dados. Dessa forma, os objetivos do SAEB são melhorar a educação brasileira, descobrir quais fatores mais influenciam o desempenho estudantil, averiguar a qualidade do ensino das escolas e permitir a adoção de programas sociais voltados para uma melhora da educação.

O SAEB corresponde a uma série de avaliações que buscam analisar e melhorar a educação brasileira, além de descobrir quais fatores mais influenciam o desempenho do estudantil e de averiguar a qualidade do ensino das escolas. A partir das informações coletadas pelo SAEB, torna-se possível a adoção de programas sociais dirigidos a melhora da educação pelo governo federal. Enfatiza-se, aqui, que o SAEB não busca avaliar os alunos individualmente, e sim avaliar os sistemas de ensino brasileiros.

Atualmente, o sistema de avaliação em questão é aplicado a cada dois anos, abrangendo estudantes das redes públicas e privadas do país matriculados no 5º ano e 9º ano do ensino fundamental e, também, aqueles que cursam o 3º ano do ensino médio.

## 1.3 Violência e conflitos nas escolas

A violência sempre esteve, de alguma forma, associada a humanidade, pois os conflitos surgem a partir de uma convivência diária com outros seres humanos, podendo resultar ou não em um problema de violência. Nesse sentido, existem os defensores de que os atos violentos são características inerentes e naturais dos seres humanos e aqueles que defendem a violência como um produto e resultado do ambi-

---

<sup>3</sup>Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

<sup>4</sup>Ministério da Educação e Cultura

ente histórico no qual os seres humanos estiveram inseridos, trazendo essa bagagem negativa consigo (Marriell et al., 2006).

Com base nisso, atribuiu-se a essa palavra o conceito de emprego de força física, uso de intimidação e ameaça, atentado à liberdade individual ou uma relação de poder contra alguém. Basicamente, a definição de violência corresponde a uma transgressão da ordem e regras da vida em sociedade, implicando na ausência e desrespeito do direito de um terceiro.

No Brasil, a violência é um problema diário, sendo possível detectar sua ocorrência nos mais diversos cenários e formas, tais como delinquência, criminalidade, abusos ou mesmo prática de justiça extralegal (Marriell et al., 2006). A violência não pode e não deve ser confundida com agressividade ou com ato criminoso. A agressividade decorre de um comportamento advindo de alguém que vê a agressão como uma forma de defesa, o crime é a prática de um ato ilegal conforme previsão em lei e a violência é uma conduta reativa ao ambiente no qual a pessoa está inserida e que causa dano a um terceiro.

A mídia, em geral, vem noticiando, de forma cada vez mais frequente, os casos constantes de violência relacionada às escolas, o que tem gerado força e notoriedade para o tema. As instituições escolares tem lidado com violência e conflitos envolvendo crianças e adolescentes com certa regularidade, o que demonstrou contribuir para o afastamento desses jovens da escola (UNICEF, 2012). A violência na escola pode resultar em inúmeras consequências ao desenvolvimento saudável das crianças e dos adolescentes, tanto para as vítimas quanto para os agressores.

A violência é um problema social que está presente nas escolas e se manifesta de variadas formas e envolve todos que estão inseridos no sistema escolar. No âmbito escolar, a violência pode ocorrer entre os próprios alunos, entre os estudantes e o patrimônio da escola e entre alunos e professores ou diretores. É, de certa forma, contraditório abordar a violência na escola, pois ela é tida como um ambiente de formação da ética e da moral das crianças e adolescentes que ali se encontram, mas ela está presente e é um motivo atual de preocupação.

A escola pode ser apontada como o primeiro ambiente social a ser experimentado pela criança e pelo adolescente, uma vez que, antes disso, o ambiente social, em tese, estaria restrito a famílias e aos locais frequentados em conjunto com os pais, tais como igrejas e vizinhos. Logo, a escola constitui o primeiro círculo social onde a criança e o adolescente aprendem a conviver com pessoas diferentes.

Diante disso, a violência na escola pode ser elencada como o primeiro espaço a ser trabalhado frente a um combate à violência e conflitos, buscando incentivar o aspecto crítico do ser humano, despertar a característica inovadora dentro de cada

um, renovar a capacidade de diálogo e propor um novo projeto de sociedade no qual o bem de todos esteja realmente em vista.

É importante salientar que, no Sistema de Avaliação de Educação Básica (SAEB), as principais informações coletadas pelo INEP acerca da violência nas escolas estão, em geral, relacionadas a possíveis agressões verbais ou físicas, a supostas ameaças, a tentativas de crime contra a vida, ao porte de arma de fogo e ao uso de drogas lícitas ou ilícitas e são coletadas a partir de um ponto de vista tanto do diretor quanto do professor, e não do aluno.

## 1.4 Revisão de literatura

Um dos grandes desafios na educação está relacionado ao fortalecimento da proficiência em matérias tradicionais, como Língua Portuguesa e Matemática, que são áreas constantemente avaliadas por pesquisas feitas dentro do sistema educacional, conforme mostra a literatura. Laros et al. (2010), por exemplo, analisaram as relações entre o desempenho na área da matemática com as características das escolas e dos alunos, concluindo que a previsão do desempenho em matemática dos estudantes está fortemente ligada a fatores como nível socioeconômico, cobrança e incentivo dos pais, etnia autodeclarada, etc.

Tendo isso em vista, a literatura vem demonstrando já ter explorado outros fatores relacionados ao desempenho escolar. Nesse sentido, Vinha and Laros (2016), por exemplo, apresentaram uma revisão de literatura rica em artigos que reuniram fatores como nível socioeconômico, abandono escolar ou uso de eletrônicos e os relacionaram ao rendimento acadêmico, mostrando, ainda, que a principal técnica utilizada para tal foi, em sua maioria, o modelo de regressão multinível.

A associação entre fatores de violência e conflito com o desempenho escolar do estudante também já foi explorado. Por exemplo, o estudo de Vinha et al. (2014) analisou escalas utilizadas em estudos referentes ao SAEB, chegando a mencionar possíveis correlações entre os fatores de violência e aspectos como recursos da escola, experiência do diretor, carência de recursos humanos e outros.

O estudo de Candian (2009) buscou associar o desempenho acadêmico com a violência por meio de modelos hierárquicos e obteve, como resultado, que a presença de violência ao ambiente escolar resulta em menores desempenhos mesmo que haja um controle de características tais como nível socioeconômico e qualidade de ensino. Esse artigo, no entanto, foi restrito à violência segundo os professores, não fazendo menção à visão do diretor da escola.

## 1.5 Problematização e objetivos do estudo

Diante do exposto acima, é possível observar que a educação é uma característica essencial para o progresso da sociedade. Todavia, o fenômeno da violência está cada vez mais presente nas escolas e pode estar afetando esses alunos.

Este estudo tem sua importância pautada no desempenho acadêmico do estudante do ensino básico e nas variáveis que podem influenciá-lo, em especial, a violência e os conflitos escolares. Ou seja, o presente trabalho busca analisar a relação entre a percepção de violência e conflitos nas escolas na visão dos diretores com o desempenho acadêmico dos estudantes do 5º ano do Ensino Fundamental para o estado de São Paulo. Para a realização dessa análise, foram traçados como objetivos específicos a análise descritiva dos dados provenientes dos questionários do SAEB 2015 dos alunos do 5º ano do Ensino Fundamental e diretores, o estudo das relações entre percepção de violência e conflitos com outras características das escolas e dos estudantes, além de sua relação com o desempenho acadêmico por meio de modelos de regressão multinível.

Nesse sentido, o presente tema se justifica em razão da importância do contexto brasileiro frente à melhoria da educação e ao combate à violência, buscando uma possível relação dos dois em ambiente escolar. Além disso, os dados utilizados para este estudo são provenientes de um banco de dados *online*, o que também justifica a escolha do tema em questão. Portanto, a transparência no site do INEP facilita a obtenção dos dados, que serão analisados ao longo desta pesquisa de modo a relacionar a percepção de violência e conflitos escolares com o desempenho acadêmico.

Destarte, questiona-se se existe, de fato, uma tendência a quanto maior a percepção de violência e conflitos na escola, menor o desempenho acadêmico demonstrado pelos estudantes ali envolvidos. Por fim, a presente proposta inicia o estudo da relação entre a percepção de violência e conflitos nas escolas e o desempenho acadêmico.

## 2 Referencial teórico

### 2.1 Modelos de regressão multinível

Os modelos de regressão multinível consistem em uma evolução do modelo de regressão linear e também são conhecidos por denominações como modelo hierárquico linear, modelo de coeficiente randômico ou modelo linear misto. Esse modelo, basicamente, consiste em agrupamentos de variáveis de acordo com níveis hierárquicos, permitindo, assim, uma análise dentro de cada unidade (Hox, 2010).

Nos modelos de regressão clássica, assume-se a independência entre os indivíduos. No entanto, os estudos com regressão multinível envolvem dados de natureza hierárquica, ou seja, o agrupamento das observações de um nível mais baixo forma um nível mais alto, e assim por diante.

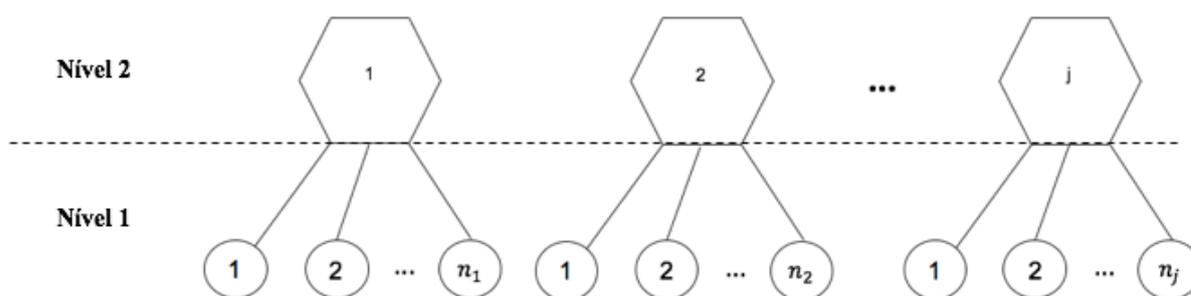


Figura 1: Hierarquia presente em Modelos de Regressão Multinível, no qual os indivíduos do nível 1 encontram-se agrupados em um nível 2.

As observações relativas a indivíduos dentro de um agrupamento (nível 2), muitas vezes, não atendem o pressuposto da independência, resultando, assim, na necessidade da aplicação de modelos multinível. Essa técnica é utilizada por diversas razões. A primeira razão seria o uso correto de inferências, pois é preciso considerar a existência de níveis hierárquicos nos dados e isso não seria possível por meio da regressão clássica, uma vez que ela pressupõe observações independentes e isso afetaria o erro-padrão dos coeficientes da regressão. Uma segunda razão seria o interesse em efeitos de grupos, pois um ponto-chave de uma pesquisa pode envolver uma variável que diz respeito a uma junção de grupos, cujos efeitos somente poderiam ser propriamente calculados via regressão multinível. Por fim, uma terceira razão seria a precisão na estimação de efeitos randômicos e fixos dos grupos (Rasbash, 2007).

O modelo de regressão multinível linear necessita de uma variável dependente quantitativa, uma vez que é preciso calcular seus efeitos, discriminá-la e analisá-la. Caso haja uma variável dependente que seja, por exemplo, dicotômica ou ordinal, técnicas como regressão logística ou multinomial seriam mais apropriadas. Segundo

Hox (2010), as variáveis independentes podem pertencer a qualquer nível, isto é, em um modelo multinível com dois níveis, as variáveis explicativas podem assumir tanto o nível 1 quanto o nível 2.

Um das vantagens principais dessa técnica é justamente o fato de ela facilitar o manuseio de dados hierárquicos, pois permite analisar as correlações em diferentes níveis e escolher um modelo que faça as associações mais adequadas. Além disso, a análise multinível necessita de um agrupamento claro e de uma classificação de nível para cada variável observada, permitindo, inclusive, uma distinção entre níveis individuais e agregados como classificação.

Por exemplo, em um modelo de regressão com dois níveis e uma variável explicativa pertencente ao nível 1, a equação de regressão será dada por

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

onde

- $\beta_{0j}$  é o intercepto da  $j$ -ésima unidade do nível 2;
- $\beta_{1j}$  é o coeficiente de regressão para a variável  $X_1$  na  $j$ -ésima unidade do nível 2;
- $e_{ij}$  é o erro aleatório.

É importante ressaltar que, para esses modelos, supõe-se que  $e_{ij} \sim N(0, \sigma^2)$ , isto é, os erros seguem uma distribuição normal com média zero e variância constante. Observa-se, ainda, que os erros de um nível são independentes dos erros dos demais níveis.

Considere agora a introdução de uma variável explicativa  $Z_j$  no segundo nível. Ao inseri-la, os coeficientes da regressão, apresentados na equação (1), podem ser dados por

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

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

O coeficiente  $\beta_{0j}$  pode ser interpretado como o valor médio de  $Y$  para a  $j$ -ésima unidade do nível 2. O coeficiente  $\beta_{1j}$  demonstra que a relação entre  $Y$  e a variável explicativa pertencente ao nível 1 depende da variável explicativa do nível 2. Sendo assim, o modelo de 2 níveis com uma variável explicativa do nível 1 e uma variável explicativa do nível 2 pode ser reescrito da seguinte forma

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

Nota-se que os quatro primeiros termos da equação (4) correspondem à parte fixa do modelo, enquanto os três últimos termos configuram a parte aleatória do modelo.

Observa-se, ainda, que a variância explicada do nível mais baixo é dada por

$$R_1^2 = \frac{\sigma_{e|b}^2 - \sigma_{e|m}^2}{\sigma_{e|b}^2}, \quad (5)$$

onde  $\sigma_{e|b}^2$  é a variância residual do nível mais baixo do modelo de base, que é, em geral, o modelo nulo sem quaisquer variáveis explicativas, e  $\sigma_{e|m}^2$  é a variância residual do nível mais alto do modelo de comparação. A aplicação e a interpretação para a variância explicada do nível mais alto é semelhante.

É interessante observar que a extensão dessa técnica e interpretações para 3 ou mais níveis ocorre de forma análoga, porém mais complexa em razão da existência de um número maior de parâmetros.

## 2.2 Passos para construção

De acordo com a descrição de Hox (2010), os modelos de regressão multinível são construídos do mais simples para o mais complexo e podem ser construídos ao longo de 5 passos. Na apresentação desses passos, assume-se, como níveis 1 e 2, respectivamente, os níveis do aluno e da escola.

### • Modelo Nulo

O primeiro passo ao proceder com a construção do modelo de regressão multinível é a o ajuste de um modelo nulo, que corresponde ao modelo sem variáveis explicativas. Esse modelo, também conhecido como modelo vazio, é obtido com

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}, \quad (6)$$

onde,

- $\gamma_{00}$  é o intercepto da regressão;
- $u_{0j}$  é o erro aleatório no nível da escola;
- $e_{ij}$  é o erro aleatório no nível do aluno.

Os modelos de regressão multinível são necessários para dados que envolvem observações agrupadas com aspectos mais semelhantes entre alguns grupos do

que em outros, o que implica na na violação do pressuposto da independência entre as observações. O grau de dependência entre as observações pode ser observado com base em um coeficiente de correlação, também conhecido por coeficiente de correlação intraclasse  $\rho$ , que é dado por

$$\rho = \frac{\sigma_{u_0}^2}{\sigma_{u_0}^2 + \sigma_e^2}. \quad (7)$$

Essa etapa é importante para o cálculo do coeficiente de correlação intraclasse (ICC). Esse coeficiente, dado pela equação (7), deve ser interpretado como a proporção da variação de  $Y$  que é explicada pelo agrupamento dos alunos, ou seja, as escolas. O valor de  $\rho$  varia de 0 a 1, sendo que, quanto mais próximo ao valor 1, maior será a proporção da variância que é explicada pelo nível mais alto, indicando, portanto, a necessidade da aplicação da modelagem multinível.

#### • Modelo 1

O próximo passo corresponde ao modelo com a inclusão das variáveis do nível 1. Nesse caso, é inserida uma variável explicativa com efeito fixo e sua contribuição para o modelo é calculada por

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

onde,  $X_{1ij}$  é a variável explicativa do nível do aluno.

#### • Modelo 2

Em seguida, é adicionada ao modelo uma variável explicativa do nível 2 (nível mais alto) com efeito fixo. Esse modelo é dado por

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

onde  $Z_j$  é a variável explicativa do nível da escola.

#### • Modelo 3

O próximo passo corresponde ao modelo com atribuição de um efeito randômico aos coeficientes relativos às variáveis do nível 1 e é dado por

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

onde,  $u_{1j}X_{1ij}$  é o efeito aleatório da variável  $X_{1ij}$ .

- **Modelo 4**

O último passo permite testar interações entre as variáveis do nível aluno e do nível escola, pois são adicionadas as interações ao modelo. Dessa forma, a equação final para o modelo completo é dada por

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

## 2.3 Métodos de estimação

Com base nos dados, será possível estimar os parâmetros do modelo de regressão multinível e seus coeficientes. A estimação pode ocorrer por máxima verossimilhança, método dos mínimos quadrados generalizados, equações de estimação generalizadas, métodos bayesianos e *bootstrapping*, sendo a primeira a mais comum dentre esses métodos. Esse tipo de estimação via máxima verossimilhança é, geralmente, mais robusto e produz estimadores mais eficientes assintoticamente e mais consistentes. Existem dois tipos de funções de máxima verossimilhança empregadas na regressão multinível: função de máxima verossimilhança completa e função de máxima verossimilhança restrita. Ambas funcionam de forma semelhante ao estimar parâmetros, associá-los ao erro-padrão e calcular a *deviance*, o que, na prática, faz com que existam poucas diferenças entre os dois métodos (Hox, 2010).

## 2.4 Comparação de modelos

Conforme abordado anteriormente, a construção de modelos de regressão multinível engloba desde um modelo mais simples até um modelo mais complexo, que possui interações e efeitos randômicos. Dessa forma, faz-se necessário comparar os modelos observados, de modo a avaliar e escolher o melhor modelo para o estudo em questão.

Existem três medidas essenciais para a comparação entre modelos de regressão multinível:

- **AIC (Critério de Informação de Akaike)**

Esse critério foi proposto por Akaike no ano de 1974 ao utilizar a informação proveniente de Kullback-Leiber, que é uma medida de distância entre um modelo verdadeiro e seu candidato, para sua definição. Ele desenvolveu uma estimativa da medida de Kullback-Leiber com base na função de máxima verossimilhança por meio de uma penalidade que buscava corrigir um viés proveniente da com-

paração de modelos. Dessa forma, o AIC é calculado por:

$$AIC = -2 \ln(L(\hat{\theta})) + 2 * (p - 1) \quad (12)$$

A melhor escolha seria aquela que apresenta o menor valor de AIC dentre todos os modelos analisados.

- **BIC (Critério de Informação Bayesiano )**

Esse critério foi proposto por Schwarz em 1978. Sua diferença para o AIC consiste em um maior penalidade com o maior número de parâmetros presentes no modelo. O BIC pode ser calculado assim:

$$BIC = -2 \ln(L(\hat{\theta})) + (p - 1) * \ln(n) \quad (13)$$

De forma semelhante, o modelo é escolhido com base no menor valor de BIC.

- **Deviance**

A deviance é uma estatística produzida pelo procedimento da máxima verossimilhança e também serve para indicar o ajuste (ou desajuste) do modelo para o banco de dados observado.

$$Deviance = -2 \ln(L(\hat{\theta})) \quad (14)$$

De forma geral, um valor menor de deviance é mais adequado para a escolha do modelo. A diferença entre deviance de modelos multinível é frequentemente utilizada para comparar modelos, uma vez que segue uma distribuição  $\chi^2$ , o que permite explorar a importância dos efeitos randômicos no momento da comparação dos modelos.

## 2.5 Análise de resíduos

Alguns procedimentos devem ser realizados para que os resultados gerados pelo modelos de regressão multinível sejam considerados confiáveis. Para isso, avalia-se a veracidade dos pressupostos de homocedasticidade, normalidade, a linearidade e a independência dos erros. Além disso, é necessário avaliar a ausência de multicolinearidade.

Em relação ao pressuposto de homocedasticidade, sua validade será testada por meio de análise gráfica. O gráfico de resíduos comparado com os valores ajustados para análise do modelo multinível serve para verificar o comportamento dos pontos e da homogeneidade na dispersão dos dados.

No caso do pressuposto de normalidade dos resíduos, a análise gráfica também se torna a melhor solução, por meio da qual espera-se como resultado uma

distribuição dos pontos bem próximos a uma reta com inclinação 1 e intercepto 0.

Em relação à linearidade e à independência, espera-se obter uma nuvem homogênea e dispersa da distribuição dos pontos em torno do 0 para cada nível via gráfico. Dessa forma, é possível constatar a existência ou não de alguma tendência que venha a comprometer a qualidade dos resultados.

## 3 Metodologia

### 3.1 Dados

O dados utilizados para esta pesquisa foram obtido a partir do SAEB (Sistema de Avaliação de Educação Básica) referente ao ano de 2015, que é o ano mais recente disponibilizado pelo INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira)<sup>5</sup>. A partir dele, foram extraídas as informações referentes a proficiência acadêmica em língua portuguesa e em matemática, a violência nas escolas, ao perfil socioeconômico, ao incentivo cultural dos pais e responsáveis para com os alunos, a escolas e seus docentes, dentre outras informações analisadas neste relatório.

Tendo isso em vista, um dos instrumentos de avaliação usados para este estudo foram os questionários contextuais disponíveis no portal eletrônico do INEP. Foi dado um foco maior às 10 perguntas relacionadas à escala de violência nos questionários aplicados aos professores e aos diretores. Para este trabalho, foram utilizadas as respostas dos diretores em razão deste ser mais adequado à análise proposta, uma vez que o diretor possui uma visão geral sobre a escola e os alunos, enquanto o professor possui uma visão restrita às turmas para as quais ministra aulas.

Para este estudo, foram avaliados somente os alunos do 5º ano do Ensino Fundamental, logo, apenas os dados obtidos com os questionários aplicados a essa série foram utilizados. Além disso, foram analisados somente aqueles alunos que responderam a 70% ou mais do questionário aplicado.

Outro instrumento importante para este estudo são as provas de proficiência contendo questões de múltipla escolha. Essas provas são aplicadas dentro do SAEB e consistem em avaliações com questões acerca das disciplinas de língua portuguesa e matemática, buscando, assim, avaliar o desempenho acadêmico, no caso, dos alunos do 5º ano do Ensino Fundamental. Aqui, foram utilizadas proficiências em língua portuguesa e em matemática transformadas em escala única do SAEB com média igual a 250 e desvio-padrão de 50.

---

<sup>5</sup>Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP)

Tabela 1: Proficiência de Língua Portuguesa e Matemática.

Variável Dependente	Questionário do SAEB	Descrição
<b>Proficiência em Língua Portuguesa</b>	PROFICIENCIA_LP_SAEB	Desempenho de Língua Portuguesa estimado pelo INEP com média 250 e desvio-padrão 50, tendo como referência o 9º ano do Ensino Fundamental de 1997.
<b>Proficiência em Matemática</b>	PROFICIENCIA_MT_SAEB	Desempenho de Matemática estimado pelo INEP com média 250 e desvio-padrão 50, tendo como referência o 9º ano do Ensino Fundamental de 1997.

Outra restrição adotada é o estado de São Paulo como unidade de federação a ser estudada. Para o presente relatório, foi analisada uma amostra de 10% retirada do banco de dados do estado de São Paulo completo. Essa amostra foi utilizada para a análise descritiva e para a modelagem multinível. Ao final, o modelo multinível escolhido foi ajustado ao estado de São Paulo completo.

Tabela 2: Informações do banco de dados de São Paulo.

Dados	São Paulo Completo	Amostra de São Paulo
Alunos	411.943	41.194
Escola	5.684	5.582
Diretores	5.684	5.582

Para a análise de regressão multinível, as variáveis utilizadas do questionário aplicado aos alunos foram sexo, etnia, idade, número de pessoas em casa, escolaridade da mãe, reprovação, dever de língua portuguesa, dever de matemática, nível econômico, incentivo social, nível cultural e violência.

A literatura já demonstrou ter associado o desempenho acadêmico com variáveis que possam influenciá-lo (Karino, 2016). Para o presente relatório, foram consideradas as variáveis do aluno (sexo, etnia, idade, número de pessoas em casa, escolaridade da mãe, reprovação, dever de língua portuguesa, dever de matemática,

nível econômico, incentivo social e nível cultural) e as variáveis da escola (violência).

Tabela 3: Variáveis do Questionário do Aluno.

<b>Variável do nível aluno</b>	<b>Pergunta</b>	<b>Código de origem</b>	<b>Descrição e Codificação</b>
<b>Sexo</b>	Qual é o seu sexo?	TX_RESP_Q001	Masculino = 0 Feminino = 1
<b>Etnia</b>	Como você se considera?	TX_RESP_Q002	Branco/Amarelo = 1 Caso contrário = 0
<b>Idade</b>	Qual a sua idade?	TX_RESP_Q004	10 ou 11 anos = 1 Caso contrário = 0
<b>Número de Pessoas em Casa</b>	Quantas pessoas vivem atualmente em sua casa?	TX_RESP_Q016	3 ou 4 pessoas = 1 Caso contrário = 0
<b>Escolaridade da Mãe</b>	Até que série sua mãe/mulher responsável por você estudou?	TX_RESP_Q019	Ensino Médio = 1 Caso contrário = 0
<b>Reprovação</b>	Você já foi reprovado?	TX_RESP_Q045	Sim = 1 Não = 0
<b>Dever de Português</b>	Você faz o dever de casa de Língua Portuguesa?	TX_RESP_Q047	Sim = 1 Não = 0
<b>Dever de Matemática</b>	Você faz o dever de casa de Matemática?	TX_RESP_Q049	Sim = 1 Não = 0

Além disso, o perfil de desempenho acadêmico foi traçado com base nos questionários direcionados aos alunos do ensino básico. Os questionários relacionados aos estudantes buscam relacionar suas características à vida escolar, ao nível socioeconômico e à cultura. Em relação às variáveis que dizem respeito ao nível econômico, incentivo social e nível cultural, foram analisados os itens provenientes do questionário preenchido pelo aluno do 5º do ensino fundamental via Teoria da Resposta ao Item (TRI). Os escores obtidos foram utilizados para análise do modelo de regressão multinível. De forma similar, foram atribuídos scores à violência via TRI (Vinha, 2018).

Tabela 4: Fatores relacionados ao nível econômico, nível cultural e incentivo social.

<b>Fator</b>	<b>Item</b>	<b>Código de Origem</b>	<b>Resposta de Origem</b>
<b>Nível Econômico</b>	Televisão em cores.	TX_RESP_Q005	Não tem/ Sim, uma/ Sim, duas/ Sim, três/ Sim, quatro ou mais.
	Aparelho de rádio.	TX_RESP_Q006	
	Videocassete e/ou DVD.	TX_RESP_Q007	
	Geladeira.	TX_RESP_Q008	
	Freezer como parte da geladeira.	TX_RESP_Q009	
	Freezer separado da geladeira.	TX_RESP_Q010	
	Máquina de lavar roupa.	TX_RESP_Q011	
	Carro.	TX_RESP_Q012	
	Computador.	TX_RESP_Q013	
	Banheiro.	TX_RESP_Q014	
	Número de quartos para dormir.	TX_RESP_Q015	
<b>Incentivo Social</b>	Incentivo a estudar.	TX_RESP_Q027	Não/Sim
	Incentivo a fazer o dever de casa e trabalhos da escola.	TX_RESP_Q028	
	Incentivo a leitura.	TX_RESP_Q029	
	Incentivo a ir a escola e/ou não faltar às aulas.	TX_RESP_Q030	
	Conversa sobre o que acontece na escola.	TX_RESP_Q031	
<b>Nível Cultural</b>	Frequência de leitura de jornais.	TX_RESP_Q032	Sempre/ De vez em quando/ Nunca.
	Frequência de leitura de livros.	TX_RESP_Q033	
	Frequência de leitura de revistas.	TX_RESP_Q034	
	Frequência de leitura de revistas em quadrinhos.	TX_RESP_Q035	
	Frequência de leitura de notícias na internet.	TX_RESP_Q036	
	Frequência de ida à biblioteca.	TX_RESP_Q037	
	Frequência de ida ao cinema.	TX_RESP_Q038	
	Frequência de ida a espetáculo ou exposição.	TX_RESP_Q039	

Tabela 5: Fatores relacionados à violência segundo o diretor.

<b>Fator</b>	<b>Pergunta</b>	<b>Código de Origem</b>	<b>Resposta de Origem</b>
<b>Violência</b>	Agressão verbal ou física de alunos a professores/funcionários da escola.	TX_RESP_Q090	Sim/Não
	Agressão verbal ou física de alunos a outros alunos da escola.	TX_RESP_Q091	Sim/Não
	Você foi vítima de atentado à vida.	TX_RESP_Q092	Sim/Não
	Você foi ameaçado por algum aluno.	TX_RESP_Q093	Sim/Não
	Você foi vítima de furto.	TX_RESP_Q094	Sim/Não
	Você foi vítima de roubo.	TX_RESP_Q095	Sim/Não
	Alunos frequentaram a escola sob efeito de bebida alcoólica.	TX_RESP_Q096	Sim/Não
	Alunos frequentaram a escola sob efeito de drogas ilícitas.	TX_RESP_Q097	Sim/Não
	Alunos frequentaram a escola portando arma branca.	TX_RESP_Q098	Sim/Não
	Alunos frequentaram a escola portando arma de fogo.	TX_RESP_Q099	Sim/Não

O *software* utilizado para a pesquisa é denominado R<sup>6</sup>. Esse sistema computacional permitiu a aplicação das técnicas estatísticas mencionadas acima ao banco de dados proveniente dos questionários aplicados no SAEB 2015. Os pacotes principais desse *software* que possibilitou a aplicação multinível foram *lme4* e *nlme*.

<sup>6</sup>The R Project for Statistical Computing

## 4 Resultados e discussões

### 4.1 Análise descritiva unidimensional

#### 4.1.1 Escolas

O número total observado dentro da amostra de São Paulo é de 5.582 escolas, para as quais foram avaliadas as características de dependência administrativa e localização.

##### 4.1.1.1 Dependência administrativa

Em relação à dependência administrativa das escolas observadas, 73,6% são municipais, seguidas por 25,5% de escolas estaduais.

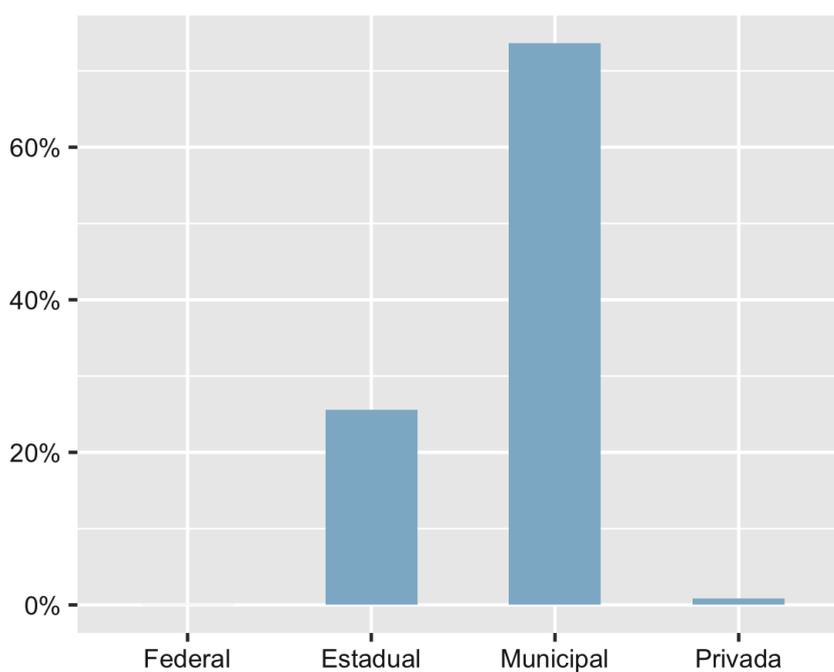


Figura 2: Distribuição da variável dependência administrativa para as escolas do 5º ano do ensino fundamental.

#### 4.1.1.2 Localização

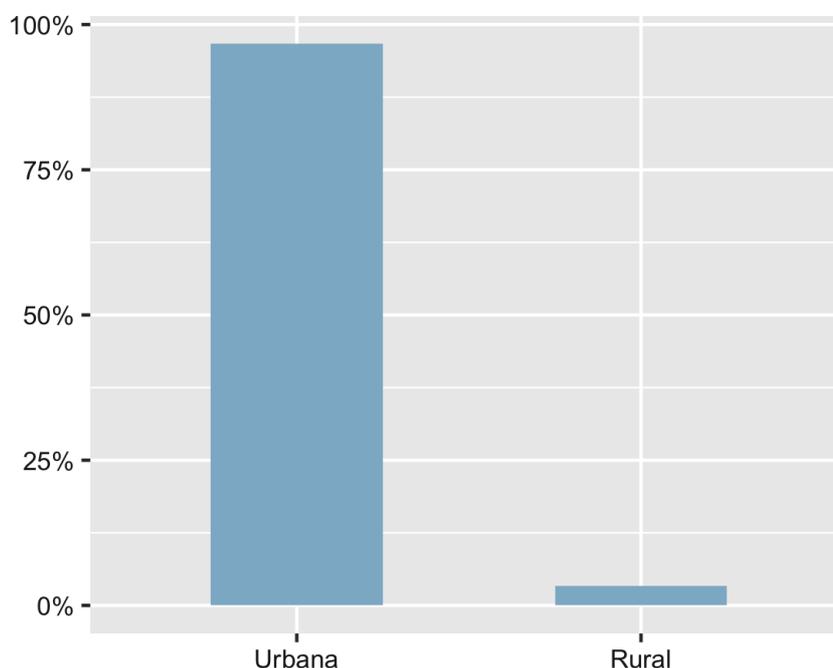


Figura 3: Distribuição da variável localização para as escolas do 5º ano do ensino fundamental.

Conforme demonstrado pelo Figura 3, as escolas são, em sua maioria, urbanas (96,4%). As escolas em ambiente rural correspondem à apenas 3,6% do total.

#### 4.1.1.3 Escala de violência

Em relação à escala de violência proveniente do questionário do diretor, as ocorrências mais frequentes são de agressões verbais ou físicas de alunos a professores ou funcionários da escola e agressões verbais ou físicas de alunos para com outros alunos. No primeiro caso, cerca de 53,0% dos diretores responderam já ter ocorrido uma violência do tipo na escola. Em relação à segunda pergunta, cerca de 74,0% dos diretores também responderam positivamente para a ocorrência do fato.

Nenhum caso de atentado à vida, ameaça por aluno, vítima de roubo (com uso de violência), alunos sob efeito de bebida alcoólica, alunos sob efeito de drogas ilícitas, alunos portando arma branca (facas, canivetes etc) ou alunos portando arma de fogo chegou a superar 20,0% de resposta afirmativa pelos diretores.

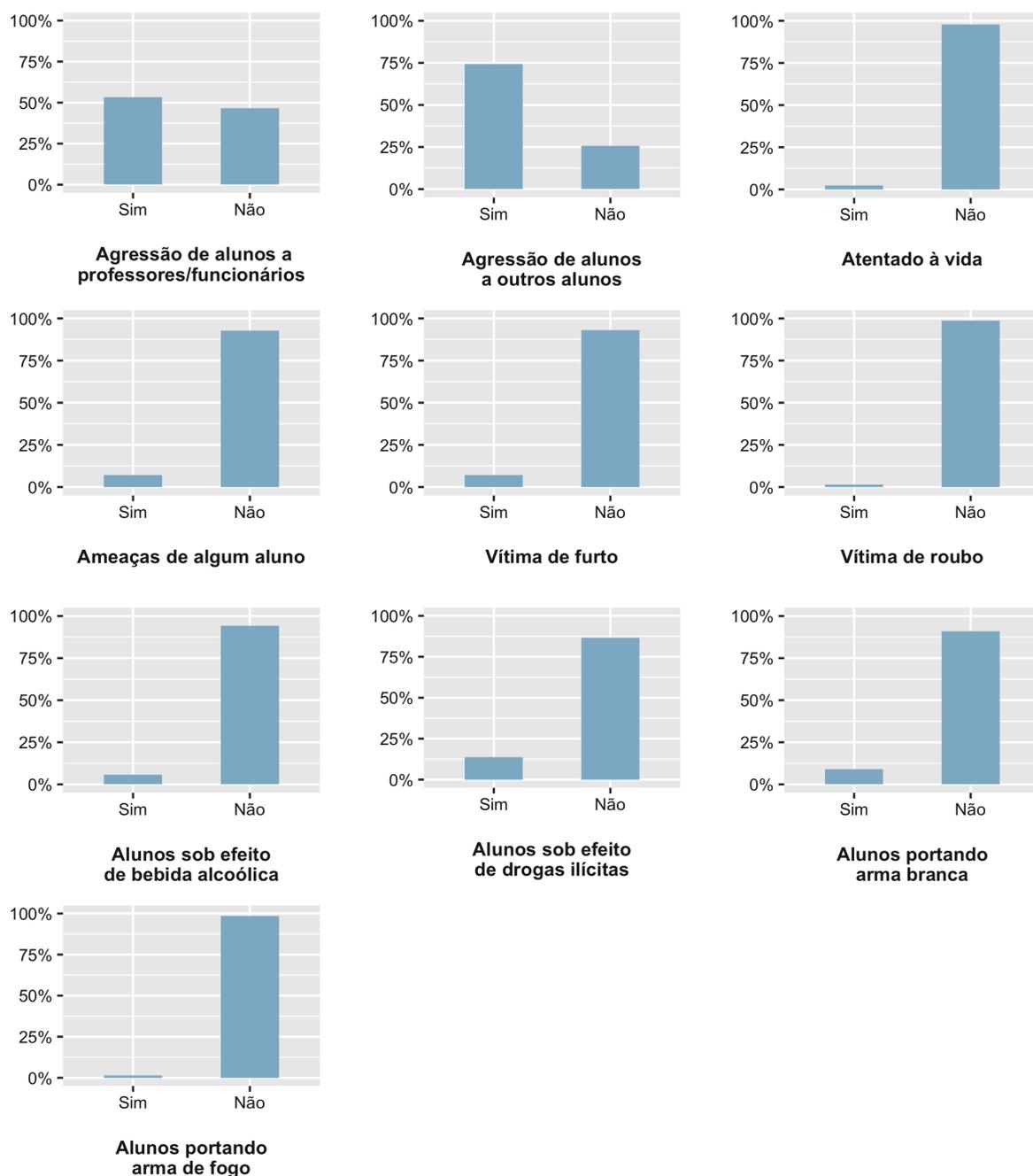


Figura 4: Variabilidade da escala de violência nas escolas no estado de São Paulo.

#### 4.1.2 Alunos

O tamanho da amostra considerado para esta análise é de 41.194 alunos para os quais foram analisadas as variáveis gênero, idade e etnia autodeclarada.

#### 4.1.2.1 Gênero

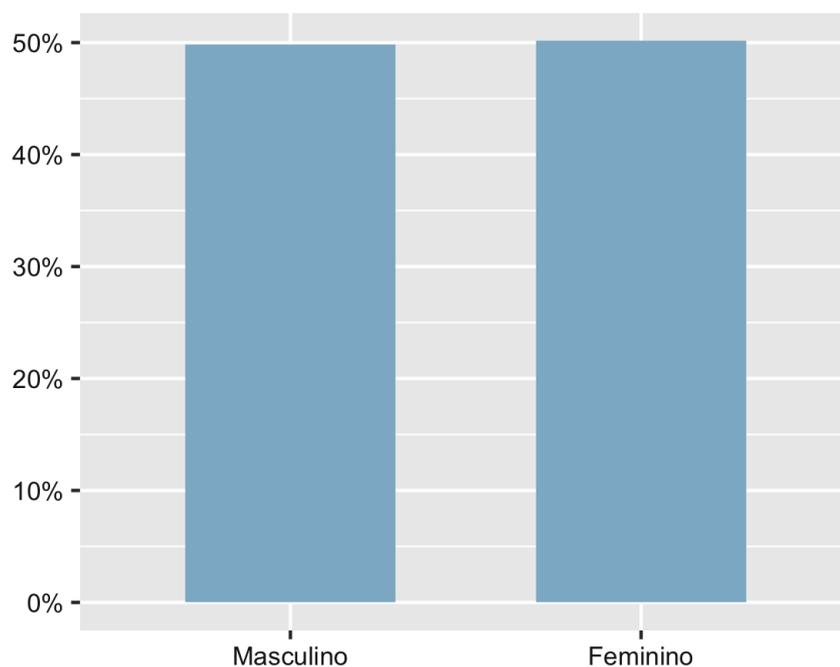


Figura 5: Distribuição da variável sexo para os alunos do 5º ano do ensino fundamental.

Para a amostra obtida, 49,9% alunos do 5º ano do ensino fundamental são do sexo masculino e 50,2% alunos do 5º ano do ensino fundamental são do sexo feminino.

#### 4.1.2.2 Idade

No que diz respeito à faixa etária dos alunos do 5º ano do ensino fundamental, a idade considerada adequado para um aluno que está em dia com sua educação é entre 10 e 11 anos. Observa-se que os alunos analisados possuem uma maioria enquadrada justamente entre essas idades.

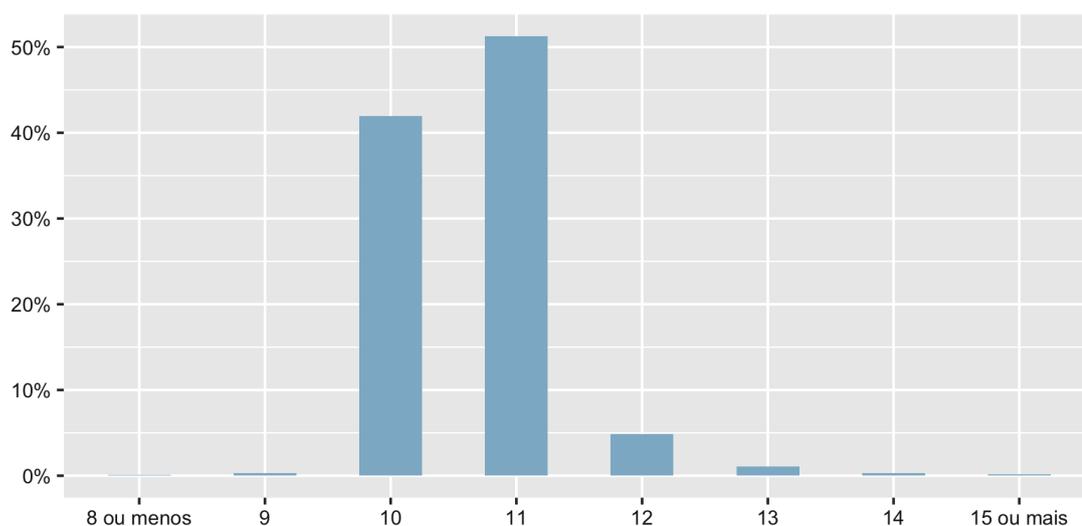


Figura 6: Distribuição da variável faixa etária para os alunos do 5º ano do ensino fundamental.

#### 4.1.2.3 Etnia autodeclarada

Conforme demonstrado pela Figura 7, é possível notar que 42,8% dos alunos se consideram pardos e 33,6% dos alunos se autodeclararam brancos. O terceiro maior percentual (11,6%) observado demonstra que uma parcela dos alunos não sabe declarar sua etnia.

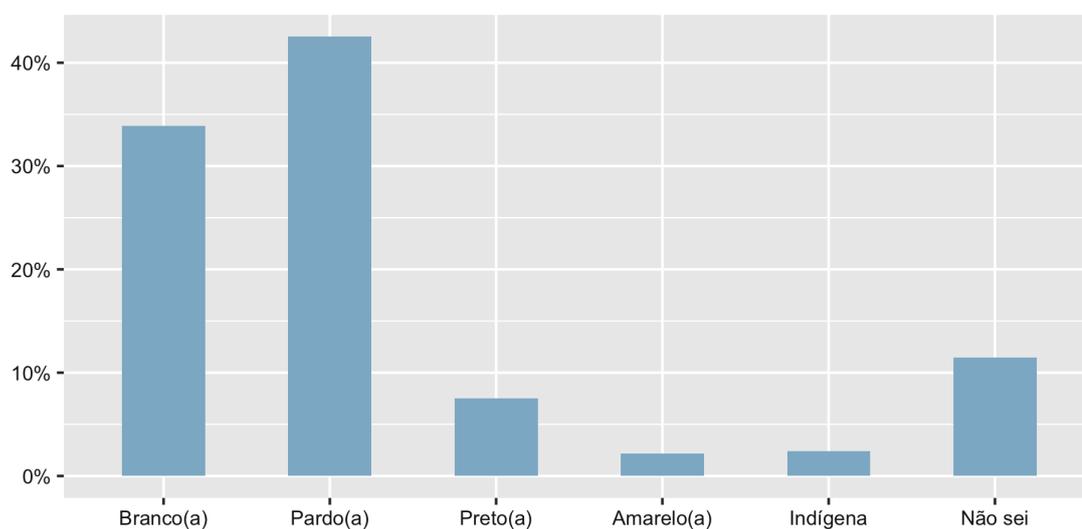


Figura 7: Distribuição da variável etnia autodeclarada para os alunos do 5º ano do ensino fundamental.

### 4.1.3 Reprovação

Conforme demonstrado pelo gráfico à esquerda na Figura 8, aproximadamente 88,1% dos alunos não possuem reprovações, enquanto 9,5% já reprovaram apenas uma vez e 2,5% já reprovaram mais de uma vez. Foi calculado o percentual de alunos que afirmam ter reprovado por escola e, conforme demonstrado pelo gráfico à direita abaixo, pode-se verificar que grande parte das escolas tem até 20% de alunos que afirmam ter reprovado.

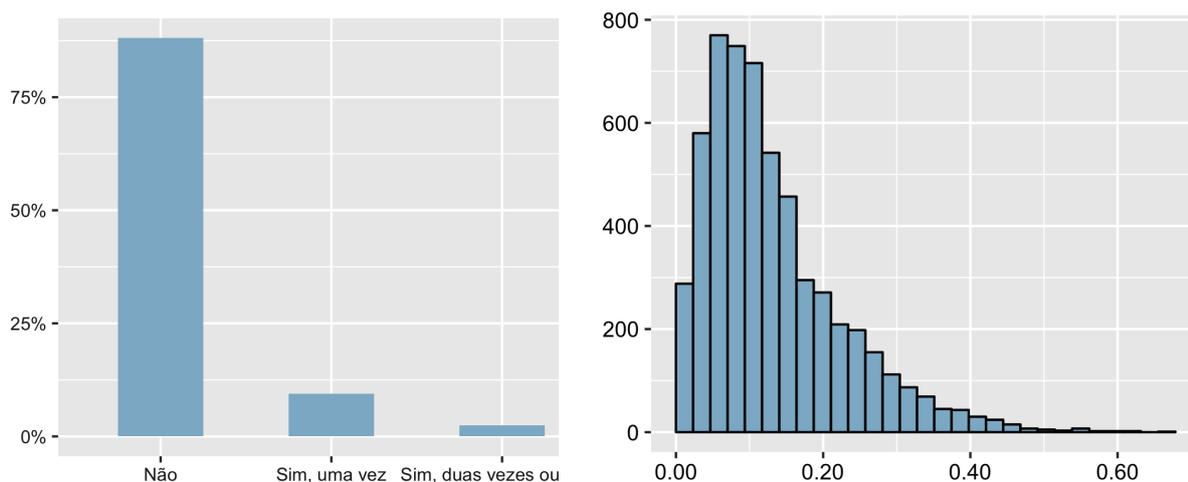


Figura 8: Distribuição da variável do número de reprovações dos alunos do 5º ano do ensino fundamental.

### 4.1.4 Desempenho escolar

É possível notar que a proficiência em matemática foi maior que a proficiência em língua portuguesa. Por exemplo, 50,0% dos alunos atingiram uma pontuação de 234 para matemática enquanto 50,0% dos alunos atingiram 219 pontos para língua portuguesa.

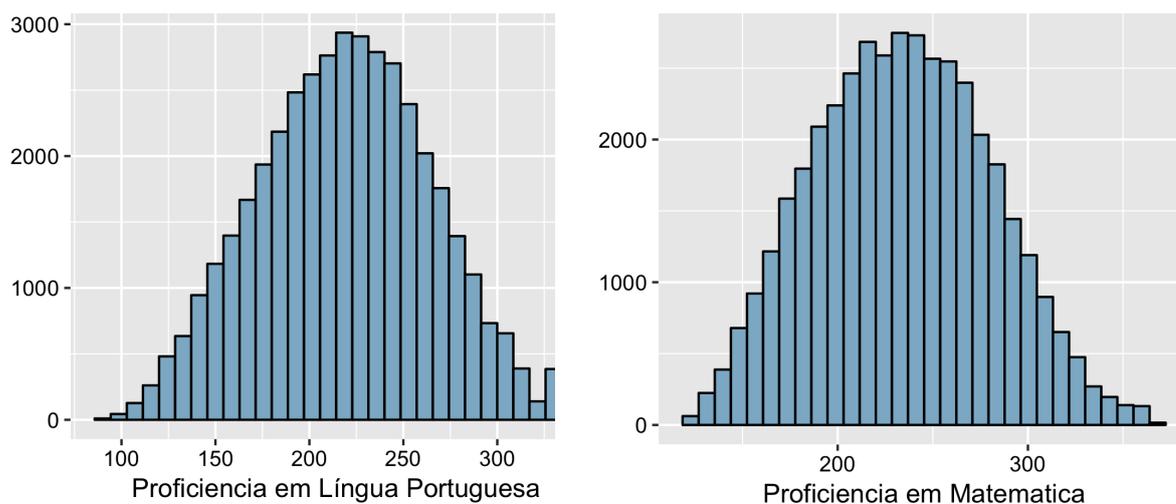


Figura 9: Proficiência dos alunos em Língua Portuguesa e Matemática.

## 4.2 Análise descritiva bidimensional

### 4.2.1 Escala de violência vs. desempenho escolar

#### 4.2.1.1 Escala de violência vs. notas em língua portuguesa

Como o objetivo principal deste trabalho é o estudo da relação entre a percepção de violência e conflitos nas escolas com o desempenho acadêmico, a análise bidimensional entre essas duas variáveis foi necessariamente estudada. Primeiramente, foi analisado o desempenho estudantil dos alunos do 5º ano do ensino fundamental de acordo com a nota média da disciplina de língua portuguesa. Por meio da Figura 10, é possível observar que a nota obtida para a disciplina de língua portuguesa não apresenta forte associação com violência e conflitos nas escolas, uma vez que o desempenho nessa disciplina não apresentou mudanças significativas para a resposta referente à escala de violência.

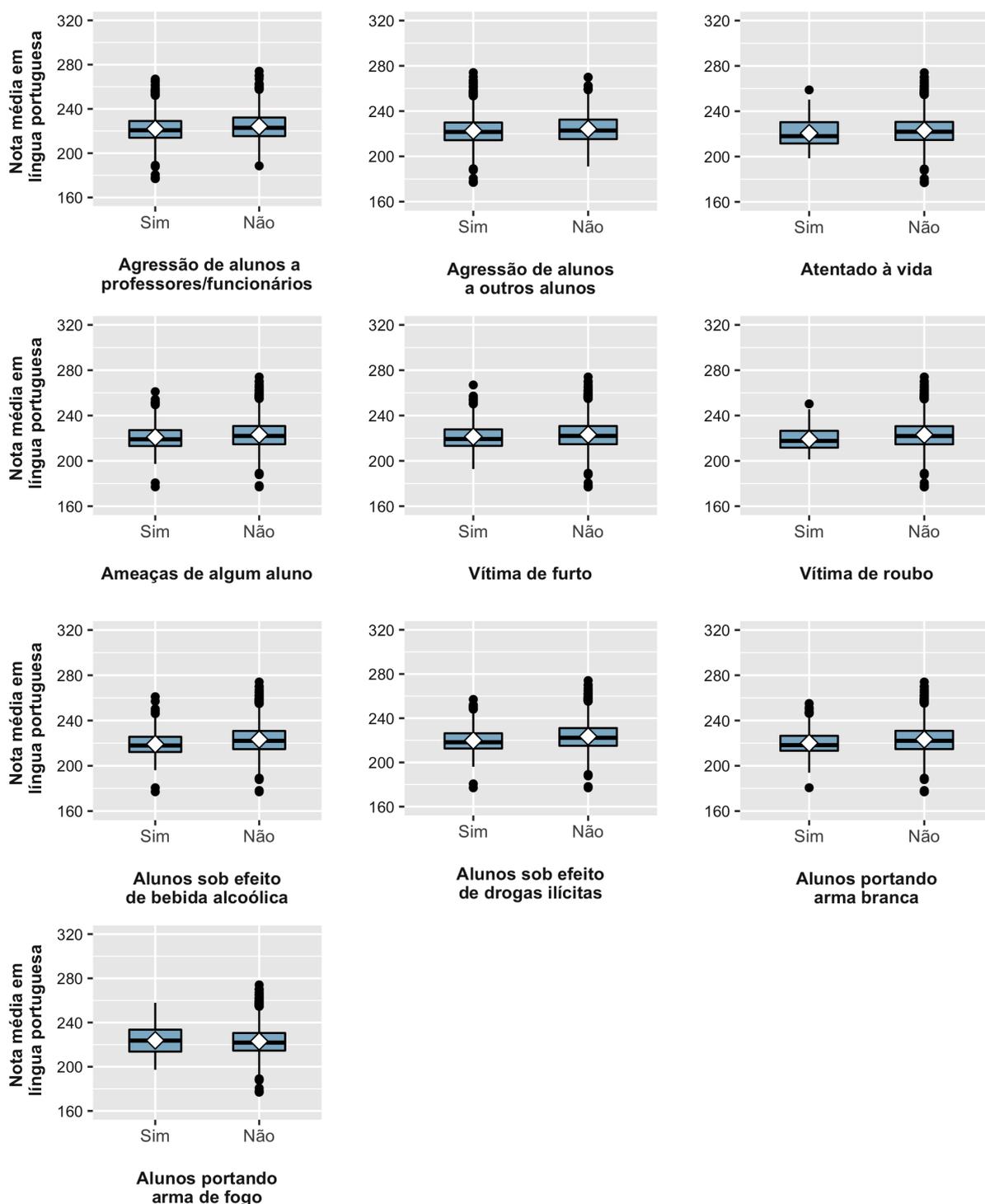


Figura 10: Escala de violência relacionada com o desempenho escolar em língua portuguesa.

#### 4.2.1.2 Escala de violência vs. notas em matemática

Em seguida, foram elaborados *boxplots* do desempenho estudantil representado pela nota média na disciplina de matemática dos alunos do 5º ano do ensino fundamental com a percepção de violência e conflitos. Com análise semelhante à dis-

ciplina de língua portuguesa, a escala de violência não apresentou forte associação, pois à medida em que a nota do aluno se modificava, os casos de violência e conflitos não sofriam grandes alterações, sendo observado um comportamento semelhante para as dez escalas de violência.

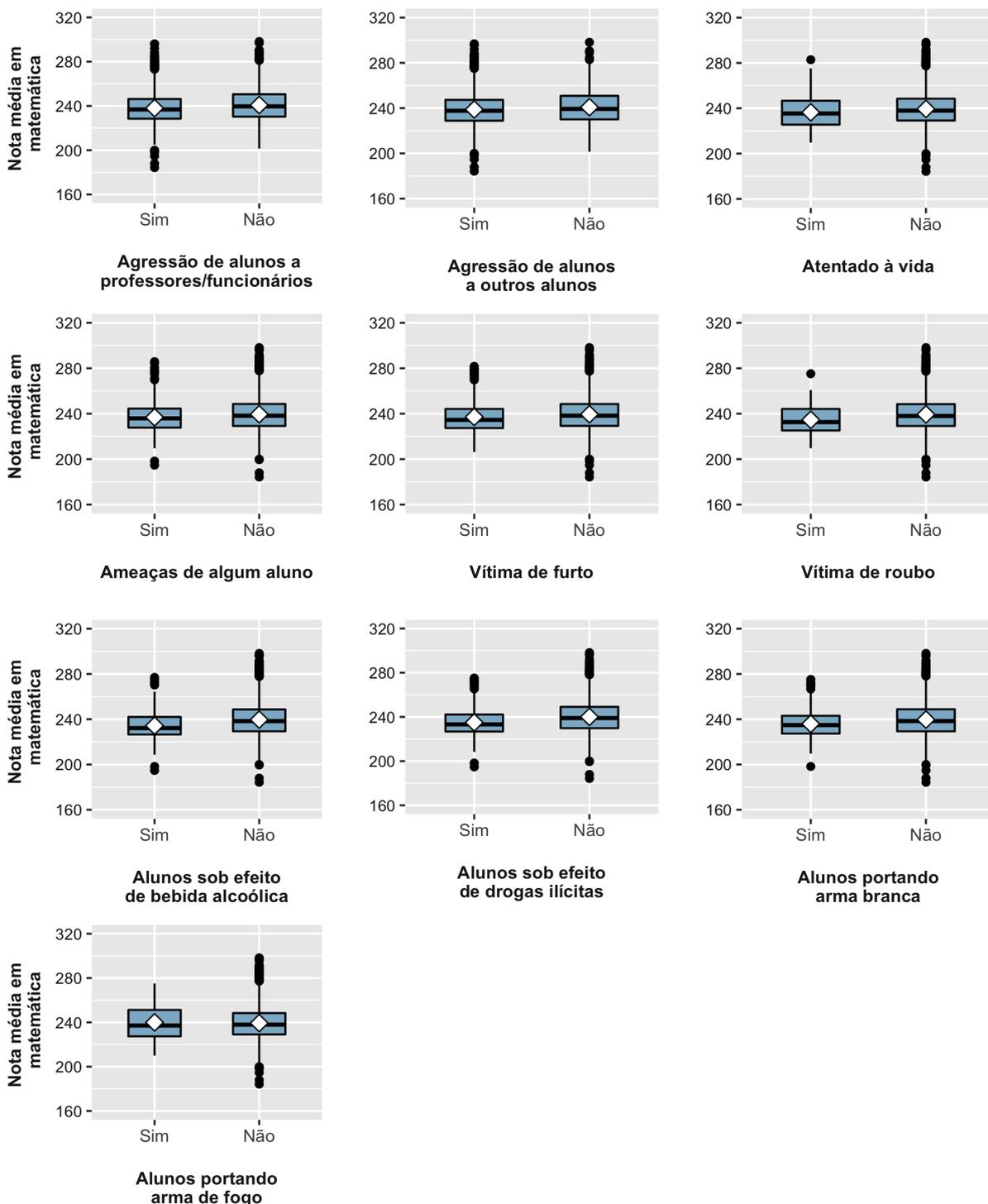


Figura 11: Escala de violência relacionada com o desempenho escola em matemática.

### 4.3 Modelagem

Neste estudo, inicialmente, foi avaliada a utilização de um modelo com três níveis. A inserção de um terceiro nível se daria com a inclusão de distritos da cidade de São Paulo. Nesse caso, o banco de dados corresponderia à cidade de São Paulo, e não ao estado de São Paulo. Para tal, seriam elaborados três níveis onde os distritos de São Paulo seriam o terceiro nível, as escolas localizadas nesses distritos seriam o segundo nível e os alunos que ali estudavam seriam o primeiro nível.

A escolha dos distritos de São Paulo poderia ocorrer de três formas: todos os distritos, distritos mais violentos ou distritos aleatórios da cidade de São Paulo. Ao estudar essas possibilidades, a opção que traria mais coerência à análise da relação entre violência e conflitos com o desempenho acadêmico dos alunos seria o uso de todos os distritos, uma vez que eles diferem em indicadores de criminalidade e violência.

No entanto, considerando todos os distritos, por meio do ajuste do modelo nulo, verificou-se que o coeficiente de correlação intraclasse (ICC) não assumiu um valor significativo, indicando que menos de 1% da variância do desempenho acadêmico dos alunos em matemática e língua portuguesa poderia ser explicado no nível dos distritos.

Tabela 6: Coeficiente de Correlação Intraclasse para 3 níveis

Nível	Língua Portuguesa		Matemática	
	$\sigma^2$	ICC	$\sigma^2$	ICC
Distrito	19,51	0,0087	18,67	0,0097
Escola	107,19	0,0481	98,27	0,0513
Aluno	2.101,25	0,9431	1.796,76	0,9388

O coeficiente de correlação intraclasse (ICC) encontrado para ambas as disciplinas é baixo e não implica no uso de um terceiro nível. Portanto, aqui se encerra a possibilidade de um terceiro nível, excluindo sua hipótese. A partir desse momento, a regressão multinível foi aplicada a apenas dois níveis, aluno e escola, para o estado de São Paulo.

No ajuste do modelo nulo, no qual existe somente o intercepto sem a presença de quaisquer variáveis explicativas, os resultados obtidos confirmaram a implementação de dois níveis para este estudo. O coeficiente de correlação intraclasse (ICC) assumiu o valor de 0,0872 no caso da proficiência em língua portuguesa, o que indica que cerca de 8,72% da variância do desempenho escolar dos alunos nessa matéria pode ser explicado pela escola. No caso da proficiência em matemática, o coeficiente de correlação intraclasse (ICC) assumiu o valor de 0,0876, o que também sugere que

Tabela 7: Modelo Nulo

	Língua portuguesa		Matemática	
	Efeito	EP	Efeito	EP
<b>Efeito fixo</b>				
Intercepto	219,22	0,2990	234,30	0,3217
<b>Efeito randômico</b>				
Variância do nível aluno	1.960,10		1.884,70	
Variância do nível escola	187,40		266,50	
<b>ICC</b>	0,0872		0,0876	
<b>Deviance</b>	431.968,70		431.266,40	
<b>AIC</b>	431.974,70		431.272,40	
<b>BIC</b>	432.000,60		431.298,30	

cerca de 8,76% da variância do desempenho escolar nessa matéria dos alunos está relacionado ao nível das escolas. Portanto, encontra-se justificado o uso dos modelos de regressão multinível. O valor do coeficiente pode ser apontado como baixo por Vinha et al. (2016), porém, para outros autores como Hox (2010), esse valor é suficiente para o uso de um modelo multinível.

A Tabela 7 apresenta um valor do intercepto de 219,22 e 234,30, respectivamente, para língua portuguesa e matemática. Isso pode ser interpretado como o valor médio geral da proficiência nessas disciplinas para a amostra de alunos do 5º do ano do ensino fundamental.

Em seguida, o próximo passo adotado foi a inserção das variáveis explicativas do nível 1. No modelo 1, nota-se que as alunas têm um desempenho superior em língua portuguesa e inferior em matemática. Ressalta-se que a variável reprovação produz um grande impacto sob o desempenho acadêmico, assim como as variáveis dever de matemática e idade do aluno.

O modelo 1 mostra que alunos do sexo feminino, de etnia branca ou amarela, com idade ideal para o 5º ensino fundamental, com mães que possuem ensino médio, que fazem o dever de língua portuguesa e que fazem o dever de matemática apresentam um melhor desempenho nas proficiências de língua portuguesa e matemática, assim como o nível econômico e cultural desses alunos também os influenciam positivamente. Os fatores referentes à reprovação do aluno e ao incentivo social influenciam negativamente as proficiências em língua portuguesa e em matemática. Todos esses efeitos são significativos a 5%.

A estatística *deviance* e os critérios AIC e BIC diminuíram com o modelo 1, o que significa que o modelo 1 se ajustou muito melhor aos dados do que o modelo nulo.

Tabela 8: Modelo 1

	Língua portuguesa		Matemática	
	Efeito	EP	Efeito	EP
<b>Efeito Fixo</b>				
Intercepto	182,39	1,3330	199,90	1,3323
<b>Variáveis do nível aluno</b>				
Sexo	7,94	0,4426	-5,47	0,4400
Etnia	5,98	0,4559	4,45	0,4537
Idade	10,07	1,0280	8,47	1,0226
Escolaridade da mãe	9,44	0,4938	9,62	0,4910
Reprovação	-24,70	0,7953	-24,11	0,7919
Dever de língua portuguesa	3,91	0,3612	2,51	0,3635
Dever de matemática	10,25	0,4835	14,87	0,4818
Nível econômico	3,56	0,2567	4,87	0,2559
Incentivo social	-5,62	0,4178	-3,30	0,4154
Nível Cultural	3,55	0,2966	4,87	0,2559
<b>Efeito randômico</b>				
Variância do nível aluno	1.706,40		1.662,20	
Variância do nível escola	126,50		191,20	
<b>Deviance</b>	391.876,10		391.739,30	
<b>AIC</b>	391.902,10		391.765,30	
<b>BIC</b>	392.013,20		391.876,40	

O terceiro passo consiste no ajuste do modelo 2, no qual foi acrescentada uma variável explicativa do nível da escola, a violência. Com a introdução dessa variável, a variância do nível da escola sofreu uma diminuição. Para o modelo 2, em comparação com o modelo nulo, 12% é o percentual da variância explicada no nível do aluno e 34% é o percentual da variância explicada no nível da escola.

Observa-se, no modelo 2, que os alunos que já reprovaram uma ou mais vezes tem seu desempenho tanto em língua portuguesa quanto em matemática reduzidos em, respectivamente, 24,75 e 24,16 pontos (Tabela 9).

Tabela 9: Modelo 2

	Língua portuguesa		Matemática	
	Efeito	EP	Efeito	EP
<b>Efeito fixo</b>				
Intercepto	182,30	1,3319	199,75	1,3311
<b>Variáveis do nível aluno</b>				
Sexo	7,94	0,4424	-5,48	0,4397
Etnia	5,93	0,4557	4,40	0,4535
Idade	10,03	1,0274	8,44	1,0221
Escolaridade da mãe	9,45	0,4935	9,61	0,4907
Reprovação	-24,75	0,7948	-24,16	0,7915
Dever de língua portuguesa	4,01	0,3610	2,62	0,3631
Dever de matemática	10,15	0,4833	14,77	0,4816
Nível econômico	3,51	0,2567	4,83	0,2558
Incentivo social	-5,59	0,4175	-3,27	0,4152
Nível Cultural	3,55	0,2965	4,11	0,2951
<b>Variável do nível escola</b>				
Violência	-2,84	0,3816	-3,39	0,4108
<b>Efeito randômico</b>				
Variância do nível aluno	1.706,00		1662,40	
Variância do nível escola	123,00		185,10	
<b>Deviance</b>	391.821,10		391.671,80	
<b>AIC</b>	391.849,10		391.699,80	
<b>BIC</b>	391.968,70		391.819,40	

Além disso, pela Tabela 9, o fato de o aluno fazer o dever de língua portuguesa aumenta aproximadamente 4,01 em seu desempenho em língua portuguesa e 2,62 em seu desempenho em matemática. Ao analisar o dever de matemática, este aumenta 10,15 o desempenho em língua portuguesa e 14,77 o desempenho em matemática.

Tabela 10: Modelo 3

	Língua portuguesa		Matemática	
	Efeito	EP	Efeito	EP
<b>Efeito fixo</b>				
Intercepto	182,25	1,3256	199,26	1,3136
<b>Variáveis do nível aluno</b>				
Sexo	7,90	0,4422	-5,57	0,4391
Etnia	5,91	0,4556	4,37	0,4530
Idade	9,97	1,0273	8,37	1,0203
Escolaridade da mãe	9,45	0,4932	9,67	0,4900
Reprovação	-24,70	0,7942	-24,04	0,7894
Dever de língua portuguesa	4,01	0,3617	2,75	0,3661
Dever de matemática	10,21	0,4840	14,91	0,4834
Nível econômico	3,52	0,2626	4,82	0,2629
Incentivo social	-5,58	0,4177	-3,25	0,4189
Nível cultural	3,55	0,3054	4,09	0,3039
<b>Variável do nível escola</b>				
Violência	-2,83	0,3809	-3,27	0,4074
<b>Efeito randômico</b>				
Dever de língua portuguesa	5,33		17,23	
Dever de matemática	9,05		22,54	
Nível econômico	10,48		12,99	
Incentivo social	0,94		9,76	
Nível cultural	19,23		19,11	
Variância do nível aluno	1.683,18		1.627,45	
Variância do nível escola	81,56		73,14	
<b>Deviance</b>	391.799,40		391.595,00	
<b>AIC</b>	391.837,40		391.633,00	
<b>BIC</b>	391.999,70		391.795,30	

Em seguida, é executado o quarto passo, por meio do qual é analisado se os coeficientes das variáveis explicativas possuem variância significativa entre os grupos. Durante essa etapa, o efeito randômico foi atribuído às variáveis escolaridade da mãe, dever de língua portuguesa, dever de matemática, nível econômico, incentivo social e nível cultural. No entanto, a escolaridade da mãe não apresentou um coeficiente randômico significativo, foi retirada do modelo e foi atribuído apenas o efeito fixo. As outras cinco variáveis do nível aluno apresentaram coeficiente randômico significativo, ou seja, o efeito do dever de língua portuguesa, dever de matemática, nível econômico, incentivo social e nível cultural não é igual para todas as escolas (Tabela 10).

Para o quinto e último passo, foram testadas possíveis interações entre as variáveis do nível aluno e do nível escola, sendo elas a violência interagindo com dever de língua portuguesa, dever de matemática, nível econômico, incentivo social e nível cultural, conforme demonstrado pela Tabela 11. Entre as possíveis interações observadas com as variáveis do nível aluno e do nível escola, apenas duas apresentaram efeito significativo para o desempenho acadêmico em língua portuguesa, que foram as interações entre a violência e o dever de língua portuguesa e a violência e o dever de matemática. Tal fato demonstra, por exemplo, que quanto maior a violência, maior o efeito da variável dever de língua portuguesa. Para o desempenho acadêmico em matemática, apenas a interação entre violência e o dever de matemática teve efeito significativo e, nesse caso, quanto menor a violência, maior o efeito da variável dever de matemática.

Tabela 11: Modelo 4

	Língua portuguesa		Matemática	
	Efeito	EP	Efeito	EP
<b>Efeito fixo</b>				
Intercepto	182,21	1,3263	199,25	1,3141
<b>Variáveis do nível aluno</b>				
Sexo	7,90	0,4422	-5,57	0,4391
Etnia	5,90	0,4556	4,35	0,4530
Idade	9,98	1,0273	8,41	1,0203
Escolaridade da mãe	9,44	0,4932	9,67	0,4899
Reprovação	-24,67	0,7942	-24,00	0,7894
Dever de língua portuguesa	4,11	0,3665	2,81	0,3706
Dever de matemática	10,12	0,4858	14,83	0,4851
Nível econômico	3,50	0,2634	4,80	0,2638
Incentivo social	-5,59	0,4180	-3,24	0,4194
Nível cultural	3,57	0,3060	4,11	0,3043
<b>Variável do nível escola</b>				
Violência	-1,95	1,1834	-1,27	1,1687
<b>Efeito randômico</b>				
Dever de língua portuguesa	5,15		17,18	
Dever de matemática	9,09		22,50	
Nível econômico	10,33		13,02	
Incentivo social	0,35		9,63	
Nível cultural	19,17		18,99	
Variância do nível aluno	1.683,26		1.627,26	
Variância do nível escola	81,76		73,10	
<b>Efeitos de interação</b>				
Violência : dever de matemática	-1,28	0,6770	-1,61	0,6763
Violência : dever de língua portuguesa	0,90	0,5134	-	-
<b>Deviance</b>	391.791,80		391.586,00	
<b>AIC</b>	391.839,80		391.634,00	
<b>BIC</b>	392.044,80		391.839,00	

Dessa forma, o modelo ideal foi escolhido com base nos valores de AIC e BIC. Além disso, ao observar o modelo 4, os efeitos interação entre a violência e os deveres de língua portuguesa e matemática têm valores contrários, o que pode gerar uma interpretação duvidosa. Logo, o modelo 3 foi escolhido como modelo final, pois apresenta o melhor conjunto geral para explicar a relação entre as variáveis explicativas mencionadas e o desempenho acadêmico em língua portuguesa e matemática.

Com o intuito de verificar os pressupostos do modelo, foi realizada a análise de resíduos para o modelo final.

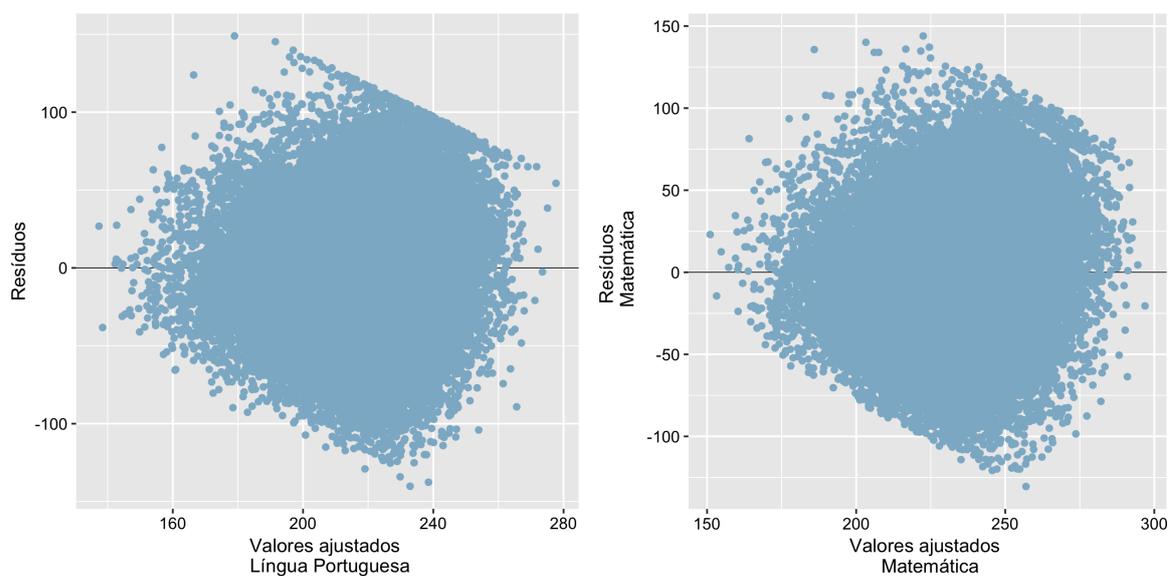


Figura 12: Análise de homocedasticidade para os valores ajustados e os resíduos do nível do aluno no modelo final.

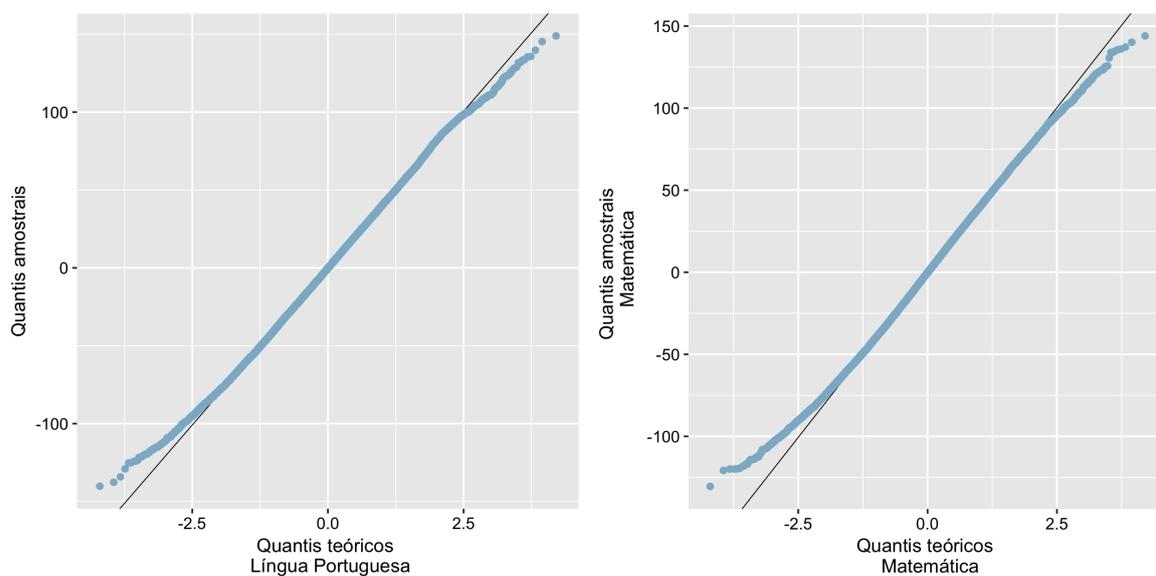


Figura 13: Análise de normalidade para os resíduos do nível do aluno no modelo final.

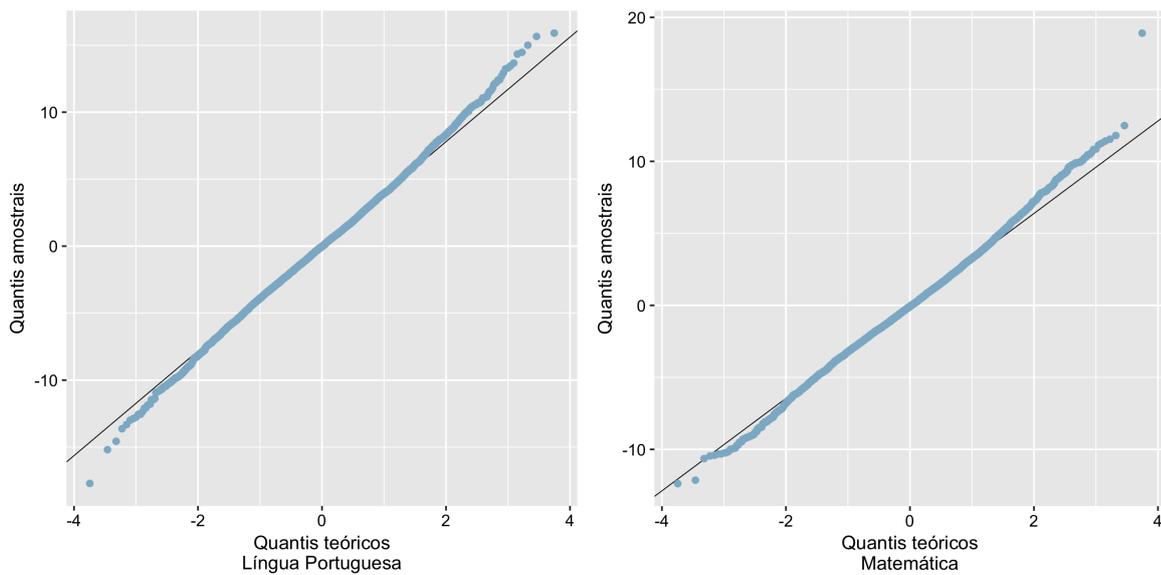


Figura 14: Análise de normalidade para os resíduos do nível da escola no modelo final.

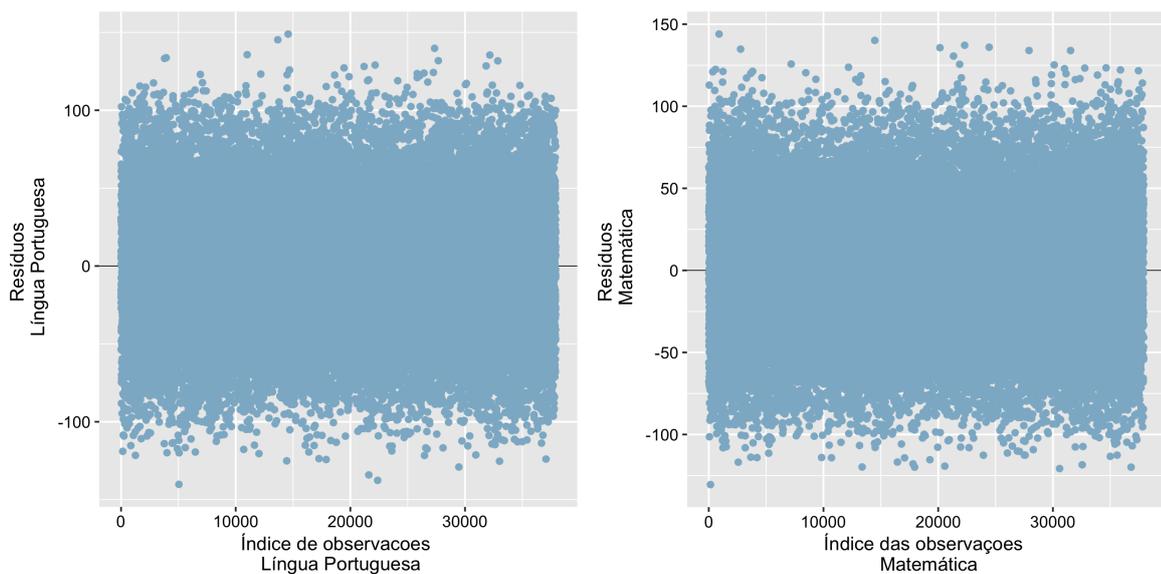


Figura 15: Análise de independência para os resíduos do nível do aluno no modelo final.

Com base nas figuras acima, não há evidências para rejeitar os pressupostos de linearidade e homocedasticidade, uma vez que os pontos se encontram dispersos de forma aleatória e bem distribuída. Para o pressuposto da normalidade, observa-se um efeito de causa pesada, pois os pontos se distanciam consideravelmente da reta em suas extremidades, tanto para a normalidade dos erros do nível do aluno quanto para os erros do nível da escola.

Dessa forma, procedeu-se com a aplicação do modelo final ao banco de dados

completo de São Paulo.

Tabela 12: Modelo para São Paulo completo

	Língua portuguesa		Matemática	
	Efeito	EP	Efeito	EP
<b>Efeito fixo</b>				
Intercepto	184,68	0,4308	200,66	0,4327
<b>Variáveis do nível aluno</b>				
Sexo	7,43	0,1300	-6,15	0,1352
Etnia	5,11	0,1445	3,76	0,1403
Idade	9,45	0,3161	8,99	0,3115
Escolaridade da mãe	8,94	0,1532	9,12	0,1509
Reprovação	-24,55	0,2472	-24,11	0,2437
Dever de língua portuguesa	5,05	0,1390	2,87	0,1408
Dever de matemática	8,48	0,1567	14,11	0,1625
Nível econômico	2,48	0,0855	3,92	0,0833
Incentivo social	-5,60	0,1358	-3,35	0,1336
Nível cultural	3,96	0,1019	4,64	0,1007
<b>Variável do nível escola</b>				
Violência	-2,90	0,2394	-3,33	0,2738
<b>Efeito randômico</b>				
Dever de língua portuguesa	20,77		22,04	
Dever de matemática	4,31		14,10	
Nível econômico	3,25		2,42	
Incentivo social	6,11		5,67	
Nível cultural	7,12		7,14	
Variância do nível aluno	1.683,69		1.630,99	
Variância do nível escola	82,03		110,46	
<b>Deviance</b>	3.905.537,00		3.895.644,00	
<b>AIC</b>	3.905.575,00		3.895.682,00	
<b>BIC</b>	3.910.782,00		3.895.889,00	

Ao comparar o modelo final da amostra de 10% de São Paulo com sua aplicação para o estado de São Paulo completo, é possível perceber que, apesar de os valores das estimativas de efeito fixo e de efeito randômico não serem iguais, o comportamento é semelhante. Por exemplo, os alunos do sexo feminino, brancos ou amarelos, que fazem o dever de casa e com mães que possuem ensino médio completo possuem uma influência positiva no desempenho médio em língua portuguesa. Os alunos do sexo feminino, que já reprovaram uma ou mais vezes e que não são in-

centivados socialmente tem seu desempenho acadêmico afetado negativamente em matemática.

Além disso, observa-se que os efeitos randômicos também foram atribuídos significativamente à escolaridade da mãe, ao dever de língua portuguesa, ao dever de matemática, ao nível econômico, ao incentivo social e ao nível cultural, os quais tem efeito diferenciado sobre o desempenho escolar a depender da escola em que estuda.

Por fim, enfatiza-se aqui que não foram encontrados quaisquer indícios de multicolinearidade.

## 5 Conclusão

A importância desse estudo está pautada na necessidade de uma educação mais igualitária, digna e justa para todos os cidadãos brasileiros, conforme previsto pela Constituição Federal de 1988. O direito à educação é uma garantia do Estado, da família e da sociedade.

Atualmente, uma forma de avaliação do ensino adotada pelo INEP é o Sistema de Avaliação da Educação Básica (SAEB), que busca estudar quais fatores influenciam o desempenho acadêmico, aprimorar a educação e sua qualidade e possibilitar a implantação de programas sociais educacionais.

Dentro do âmbito escolar, a percepção de violência e conflitos pode se manifestar das mais diversas formas, envolvendo alunos, professores, diretores e patrimônio escolar. Logo, o presente tema se mostrou recorrente e atual, uma vez que o SAEB acontece desde 1990 e busca provir insumos para a educação brasileira. Além disso, a relação entre violência e conflitos nas escolas com o desempenho escolar foi pouco abordado pela literatura, levando ao questionamento sobre o pensamento comum da sociedade de que quanto maior a percepção de violência e conflitos na escola, menor o desempenho acadêmico desses alunos.

Ao proceder com este estudo, foi utilizada a análise de regressão multinível devido à natureza hierárquica dos dados, constituídos por alunos (nível 1) agrupados em escolas (nível 2). A existência desses níveis exige uma técnica que saiba lidar com efeitos randômicos e fixos provenientes desse agrupamento.

O ajuste do modelo multinível ocorreu em conformidade com as 5 etapas propostas por Hox (2010), composto pela elaboração de um modelo nulo sem variáveis explicativas, seguido por um modelo apenas com variáveis explicativas do nível do aluno, um modelo com variáveis explicativas dos níveis do aluno e da escola, um modelo com efeito randômico atribuído às variáveis explicativas e um último modelo com possíveis interações entre as variáveis explicativas.

Durante a fase de ajuste dos modelos, as variáveis explicativas eram significativas em quase todas as etapas, o que levou a uma necessidade de olhar mais profundamente para a estatística *deviance* e para os critérios AIC e BIC.

O modelo final inclui variáveis explicativas do nível do aluno (sexo, etnia, idade, escolaridade da mãe, reprovação, dever de língua portuguesa, dever de matemática, nível econômico, incentivo social e nível cultural), variáveis explicativas do nível da escola (violência) e atribui efeito randômico significativo ao dever de língua portuguesa, dever de matemática, nível econômico, incentivo social e nível cultural, o que implica que sua influência sobre o desempenho acadêmico não é igual para todas as institui-

ções acadêmicas observadas.

É interessante observar que ambos os deveres exercem impacto positivo no desempenho das disciplinas de língua portuguesa e de matemática, pois, provavelmente, estão relacionados ao fato de os alunos que fazem a tarefa de casa das matérias ministradas pela escola costumam se sair melhor no geral.

As variáveis que se destacaram ao ter efeitos sob o desempenho acadêmico em língua portuguesa e em matemática são a idade, a escolaridade da mãe, o dever de matemática e a reprovação, sendo esta última um efeito negativo.

Em relação à violência, é possível notar que essa apresenta um impacto negativo sob o desempenho dos alunos tanto em língua portuguesa quanto em matemática. Logo, a relação entre elas é inversa, isto é, quanto maior o efeito da violência, menor se torna o desempenho acadêmico desse aluno.

O incentivo social apresentou um impacto negativo em relação ao desempenho escolar. Esse resultado, apesar de inesperado, já foi observado por Laros et al. (2010). Uma possível explicação para esse comportamento seria o fato de os pais incentivarem e cobrarem mais de seus filhos quando estes apresentam um desempenho acadêmico não satisfatório, sendo errado interpretar tal fator como uma relação na qual quanto mais os pais incentivam seus filhos, menor o desempenho escolar.

O uso dos modelos de regressão multinível foi ideal para a estrutura hierárquica escolar. Essa técnica se mostrou extremamente condizente com o tema proposto, porque permite que as variáveis dos alunos e das escolas sejam verificadas quanto a existência de um efeito fixo, de um efeito randômico e de um efeito de interação com o desempenho escolar.

Por fim, ressalta-se que esse tema vem ganhando força e, em razão disso, recomenda-se que sejam realizadas futuras pesquisas relacionando a percepção de violência e conflitos com o desempenho escolar para alunos de outras séries, tais como 9º ano do ensino fundamental e 3º ano do ensino médio.

## Referências

- Bittar, C. B. (2014). *Educação e direitos humanos no Brasil*. São Paulo: Saraiva.
- Candian, J. F. (2009). Violência escolar e desempenho: As evidências do saeb 2003. *IUPERJ- UCAM*.
- Hox, J. J. (2010). *Multilevel Analysis: Techniques and Applications. Quantitative Methodology Series*, volume 2.ed. Routledge.
- Karino, C. A. (2016). Avaliação da igualdade, equidade e eficácia no sistema educacional brasileiro. Tese de doutorado, Universidade de Brasília.
- Laros, J. A., Marciano, J. L. P., and de Andrade, J. M. (2010). Fatores que afetam o desempenho na prova de matemática do saeb: um estudo multinível. *Avaliação Psicológica*, 9(2):173–186.
- Marriell, L. C., AssisII, S. G., AvancillI, J. Q., and Oliveira, R. V. C. (2006). Violência escolar e auto-estima de adolescentes. *Caderno de Pesquisa*, 36(127).
- Oliveira, J. B. A. (2014). *Repensando a Educação Brasileira: O que Fazer para Transformar nossas Escolas*. São Paulo: Atlas.
- Rasbash, J. (2007). Why use multilevel modelling? University of Bristol, <http://www.bristol.ac.uk/cmm/learning/multilevel-models/>.
- UNICEF, F. N. U. I. (2012). Acesso, permanência, aprendizagem e conclusão da educação básica na idade certa. direito de todas e de cada uma das crianças e dos adolescentes. Technical report, UNICEF. Iniciativa Global Pelas Crianças Fora da Escola - Brasil.
- Vinha, L. G. A. (2018). Validação das escalas de percepção de violência e conflitos nas escolas. Manuscrito não publicado.
- Vinha, L. G. A., Karino, C. A., and Laros, J. A. (2014). Os questionários do saeb: o que eles realmente medem. *Estudos em Avaliação Educacional*, 5(59):270–297.
- Vinha, L. G. A., Karino, C. A., and Laros, J. A. (2016). Fatores associados ao desempenho em matemática no ensino fundamental no brasil. *Psico - USF*, 21(1):87–100.
- Vinha, L. G. A. and Laros, J. A. (2016). El uso de datos en evaluaciones educativas en brasil, chile y argetina. *Revista Iberoamericana de Evaluación Educativa*, 9(2):127–148.