



UFOP

Universidade Federal
de Ouro Preto

UNIVERSIDADE FEDERAL DE OURO PRETO
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS

DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS

Márcio Vinícius De Oliveira

**ANÁLISE DE DESIGUALDADE EDUCACIONAL: COMPARAÇÃO ENTRE
UMA DIVERGÊNCIA KULLBACK-LEIBLER CONTÍNUA E UMA
DEGENERADA DISCRETA**

Mariana

2017

Márcio Vinícius De Oliveira

**ANÁLISE DE DESIGUALDADE EDUCACIONAL: COMPARAÇÃO ENTRE
UMA DIVERGÊNCIA KULLBACK-LEIBLER CONTÍNUA E UMA
DEGENERADA DISCRETA**

Monografia apresentada ao curso de Ciências Econômicas da Universidade Federal de Ouro Preto como parte dos requisitos para a obtenção do grau de Bacharel em Ciências Econômicas.

Orientador: Prof. Dr. Victor Maia
Senna Delgado

Mariana

DEECO/ICSA/UFOP

2017

0482a Oliveira, Márcio Vinicius de
Análise de desigualdade educacional [recurso eletrônico]
: comparação entre uma divergência Kullback-Leibler
contínua e uma degenerada discreta / Márcio Vinicius
de Oliveira.-Mariana, MG, 2017.
1 CD-ROM; (4 3/4 pol.).

Trabalho de Conclusão de Curso (graduação) - Universidade
Federal de Ouro Preto. Instituto de Ciências Sociais
Aplicadas. Departamento de Ciências Econômicas - DEECO/ICSA/UFOP

1. Educação - Teses. 2. MEM. 3. Igualdade na educação
- Teses. 4. Monografia. 5. Desigualdade - Brasil -
Teses. 6. Qualidade (Educação) - Teses. 7. Estatística
matemática - Teses. I. Delgado, Victor Maia Senna.
II. Universidade Federal de Ouro Preto - Instituto
de Ciências Sociais Aplicadas - Departamento de Ciências
Econômicas. III. Título.

CDU: Ed. 2007 -- 519.2
: 15
: 1419074



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO – UFOP
DEPARTAMENTO DE CIÊNCIAS ECONÔMICAS E GERENCIAIS
COLEGIADO DE CIÊNCIAS ECONÔMICAS



MÁRCIO VINÍCIUS DE OLIVEIRA

Curso de Ciências Econômicas - UFOP

Análise de desigualdade educacional em 2007 e 2011: comparação entre uma divergência Kullback-Leibler contínua e uma degenerada discreta

Trabalho apresentado ao Curso de Ciências Econômicas do Instituto de Ciências Sociais e Aplicadas (ICSA) da Universidade Federal de Ouro Preto como requisito para a obtenção do grau de Bacharel em Ciências Econômicas, sob orientação do Prof. Dr. Victor Maia Sena Delgado.

Banca Examinadora:

Prof. Dr. Victor Maia Sena Delgado

Profa. Dra. Fernanda Faria Silva

Prof. Dr. Héder Carlos de Oliveira

Mariana, 06 de setembro de 2017.

Agradecimentos

Agradeço primeiro a Deus por ter me dado força, paciência e sabedoria para chegar neste momento e não ter me deixado desistir. Aos meus pais e familiares pelo apoio incondicional, sem vocês essa conquista não teria sido alcançada.

Agradeço aos meus amigos pelo companheirismo seja para comemorar as coisas boas, seja para apoiar nos momentos difíceis. Agradeço em especial aos que dividiram o lar comigo em Mariana, a República Mocambos pelos momentos vividos.

Agradeço a Universidade Federal de Ouro Preto, pelo ensino oferecido pelo seu corpo docente em especial ao meu orientador ao professor Victor pela convivência e aprendizado durante as aulas e no período de escritura da monografia.

RESUMO

A educação brasileira tem obtido importantes avanços nas últimas décadas, pois se no fim da A educação brasileira continua a ser um grande desafio a ser vencido, por dentre muitas questões a falta das condições equitativas de aprendizado. Este trabalho tem por objetivo apresentar essas questões através de revisão na literatura que busca contextualizar como os fatores externos, ou seja, fora do ambiente escolar influencie menos no desempenho escolar. A divergência Kullback-Leibler foi o método usado para medir o comportamento da desigualdade de aprendizado no 5º e no 9º ano do ensino fundamental das escolas brasileiras nos anos de 2007 e 2011 e para fazer esse cálculo, foi necessário extrair os dados do Sistema Nacional de Avaliação da Educação Básica (SAEB). A divergência KL é o cálculo da distância entre a referência que possui média zero de uma distribuição observada, neste trabalho dos municípios brasileiros. No trabalho foi, realizado uma comparação entre uma divergência contínua, ou seja, que apresenta o comportamento geral da desigualdade e uma divergência discreta que similar ao índice de Theil que adotou pesos maiores aos municípios com desempenho no SAEB aquém do esperado e foi-se reduzindo a zero quando o desempenho do município estava além do esperado. Os resultados mostraram que houve um leve aumento na desigualdade entre os municípios brasileiros de 2011 em relação a 2007 evidenciando que enquanto as condições familiares influenciar negativamente no desempenho escolar, o Brasil ficará estagnado quanto à educação.

Palavras chave: educação, desigualdade, qualidade

ABSTRACT

The Brazilian education continues to be a great challenge to be overcome, among many issues the lack of equitable conditions of learning. This paper aims to present these questions through a review in the literature that seeks to contextualize how external factors, that is, outside the school environment, influence less in school performance. The Kullback-Leibler divergence was the method used to measure the behavior of the learning inequality in the 5th and 9th years of elementary school in Brazil in the years 2007 and 2011, and to do this calculation, it was necessary to extract the data from the National System of Assessment of Basic Education (SAEB). The divergence KL is the calculation of the distance between the reference that has zero mean of an observed distribution, in this work of the Brazilian municipalities. In the work, a comparison was made between a continuous divergence, that is, that presents the general behavior of the inequality and a discrete divergence similar to the Theil index that adopted greater weights to the municipalities with SAEB performance lower than expected and was reduced to zero when the performance of the municipality was beyond expectations. The results showed that there was a slight increase in inequality among Brazilian municipalities in 2011 compared to 2007, showing that while family conditions negatively influence school performance, Brazil will be stagnant in terms of education.

Key words: education, inequality, quality

LISTA DE FIGURAS

Figura 1: Dimensão reduzida do sistema educacional e seu reflexo na proficiência	19
Figura 2: Sistema Nacional de Avaliação da Educação Básica (SAEB).....	22
Figura 3: Conceitos de extração de recursos com base na divergência do K-L	27

LISTA DE GRÁFICOS

Gráfico 1: Evolução dos resultados do Brasil no SAEB entre 1995 e 2011 de Leitura e Matemática do 5º e do 9º ano do Ensino Fundamental	1
Gráfico 2: Uma possibilidade de curva de Lorenz	7
Gráfico 3: Representação da densidade acumulada para o nível de leitura em 2003 e 2015	8
Gráfico 4: Escala de Proficiência em Infraestrutura baseada na escala da informação .	13
Gráfico 5: Dispersão de leitura do KL 2007 e Kl 2011 do 5º ano do ensino fundamental	32
Gráfico 6: Histograma de leitura, 5º ano do Ensino Fundamental, KL 2007 e KL 2011	32
Gráfico 7: Comparação das Divergências Kullback-Leibler para o 5º ano em leitura de 2007 e 2011	33
Gráfico 8: Dispersão de leitura do IDE 2007 e do IDE 2011 do 5º ano do ensino fundamental	33
Gráfico 9: Histograma de leitura, 5º ano do ensino fundamental, IDE 2007 e IDE 2011	34
Gráfico 10: Comparação da Divergência Kullback-Leibler Discreta ou Índice de Desigualdade Educacional (IDE) para o 5º ano de leitura de 2007 e 2011	34
Gráfico 11: Dispersão de matemática do KL 2007 e KL 2011 do 5º ano do ensino fundamental	35
Gráfico 12: Histograma de matemática, KL 2007 e KL 2011 do 5º ano do ensino fundamental	35
Gráfico 13: Comparação das Divergências KL de Matemática 5º ano para 2007 e 2011	36
Gráfico 14: Dispersão de matemática do IDE 2007 e IDE 2011 do 5º ano do ensino fundamental	36
Gráfico 15: Histograma de matemática, 5º ano do Ensino Fundamental, IDE 2007 e IDE 2011	37
Gráfico 16: Comparação do IDE para matemática 5º ano do ensino fundamental de 2007 e 2011	37

Gráfico 17: Dispersão de leitura do KL2007 e KL 2011 do 9º ano do ensino fundamental	39
Gráfico 18: Histograma de leitura para o 9º ano do ensino fundamental para o KL 2007 e KL 2011	39
Gráfico 19: Comparativo da KL para leitura 9º ano de 2007 e 2011	40
Gráfico 20: Dispersão de leitura para o 9º ano do ensino fundamental para o IDE 2007 e IDE 2011	40
Gráfico 21: Histograma de leitura para o 9º ano do Ensino Fundamental para o IDE 2007 e IDE 2011	41
Gráfico 22: Comparação IDE para leitura, 9º ano do Ensino Fundamental para 2007 e 2011	41
Gráfico 23: Dispersão de Matemática para o 9º ano do Ensino Fundamental para a KL 2007 e KL 2011	42
Gráfico 24: Histograma de Matemática para o 9º ano do Ensino Fundamental para a KL 2007 e KL 2011	42
Gráfico 25: Comparação da KL para matemática 9º ano de 2007 e 2011	43
Gráfico 26: Dispersão de Matemática para o 9º ano do Ensino Fundamental para o IDE 2007 e IDE 2011	43
Gráfico 27: Histograma de Matemática para o 9º ano do Ensino Fundamental para o IDE 2007 e IDE 2011	44
Gráfico 28: Comparativo do IDE de Matemática 9º ano do Ensino Fundamental de 2007 e 2011	44

LISTA DE TABELAS

Tabela 1: Número de participante das avaliações Saeb/Prova Brasil entre 2007 e 2011	10
Tabela 2: Descrição de como é classificada a infraestrutura escolar	11
Tabela 3: Estatísticas Descritivas dos questionários dos alunos do 5º e do 9º ano do Ensino Fundamental do Saeb de 2007 e 2011	17
Tabela 4: Proficiência para a Língua Portuguesa. Conforme a fonte, os intervalos dos níveis das escalas de proficiências incluem o primeiro ponto e excluem o último ponto	24
Tabela 5: Níveis de Proficiência para a Matemática. Conforme a fonte, os intervalos dos níveis das escalas de proficiências incluem o primeiro ponto e excluem o último ponto	25
Tabela 6: Estatística Descritiva dos dados da KL e do IDE de 2007 e 2011 para o 5º ano do ensino fundamental.....	31
Tabela 7: Estatística Descritiva dos dados da KL e do IDE de 2007 e 2011 para o 9º ano do ensino fundamental.....	38
Tabela 8: Valores da distância observada até a distribuição de referência dos grupos definidos por região da Prova Brasil do 5º ano do Ensino fundamental de 2007 e 2011	45

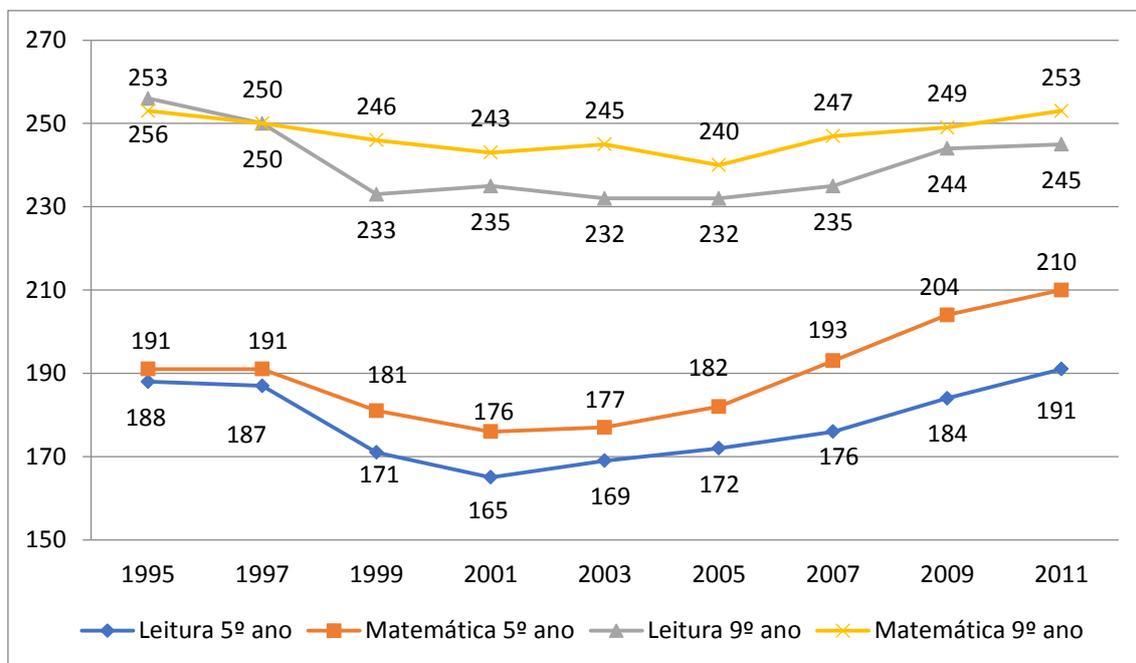
SUMÁRIO

1	INTRODUÇÃO	1
2	A MENSURAÇÃO DA DESIGUALDADE NO CONTEXTO EDUCACIONAL BRASILEIRO.....	5
2.1	Os avanços institucionais da educação brasileira.....	5
2.2	A construção dos indicadores de desigualdade	6
2.2.1	Nível Socioeconômico	15
2.2.2	Função de Produção da Educação e o Efeito Escola.....	19
3	MÉTODOS	21
4	RESULTADOS.....	31
5	CONCLUSÃO	46
6	REFERÊNCIAS BIBLIOGRÁFICAS	48

1 INTRODUÇÃO

A discussão sobre como oferecer uma educação universal e de qualidade para todos os alunos tem conseguido avançar durante as últimas décadas, mas ainda os desafios para chegar ao modelo ideal têm requerido vários estudos na área. O desafio nas décadas de 1980 e 1990 era o de como reduzir as taxas de analfabetismo, e para isso vieram um conjunto de medidas como a priorização da educação como um dos objetivos da nação na Constituição de 1988, a vinculação orçamentária e os programas de avaliação. Essas ações atreladas a outros programas governamentais fizeram levar a 92% da população com até 15 anos o acesso à escola primária em 2015, segundo o banco de dados Data World Bank em 2017. Um fato importante também foi o aumento da despesa em educação de 4,77% em 1998 para 6% em 2013 do PIB em educação, adotando medidas como piso nacional para o magistério e o aumento do gasto per capita nacional que passou de 10,68% para 20,04%. Percebe-se também uma evolução da proficiência média dos alunos, principalmente após 2005 quando o Sistema de Avaliação da Educação (SAEB) passou pela reformulação que é usada nos dias atuais.

Gráfico 1: Evolução dos resultados do Brasil no SAEB entre 1995 e 2011 de Leitura e Matemática do 5º e do 9º ano do Ensino Fundamental



Fonte: Diretoria de Avaliação da Educação Básica – DAEB/INEP, 2017.

Apesar de o Brasil estar apresentando avanços quanto aos investimentos, acesso e níveis de aprendizado apresentado nos indicadores, os mesmos estão aquém se comparados ao resto do mundo, pois segundo um artigo produzido pela equipe do Centro de Políticas Públicas do Insper em 2016, a escolaridade média brasileira cresceu menos em relação a dos países desenvolvidos, a ponto de ter sido ultrapassado pela Coreia do Sul, que na década de 1950 estava em posição inferior do Brasil. Fato que se repete também em relação a seus pares na América do Sul, pois a distância em relação ao Chile, país que detém a maior escolaridade média da região, começou a reduzir apenas a partir da década de 1990.

No Programa Internacional de Avaliação dos Alunos (PISA)¹, organizado pela Organização de Cooperação e Desenvolvimento Econômico (OCDE), que serve para acompanhar o nível de aprendizado para jovens de 15 em leitura, ciências e matemática, o Brasil, segundo o relatório apresentado da prova de 2015, apresenta uma estagnação na pontuação desde 2000, além da mesma estar abaixo da média dos países da OCDE.

Vários fatores explicam o atraso educacional brasileiro, dentre os quais, a desigualdade de aprendizado dos alunos brasileiros, tema desta monografia, pois conforme Ricardo Paes de Barros afirmou em uma entrevista a revista *Época* em 07 de agosto de 2015:

“as condições sociais influenciam muito além do que deveria no aprendizado do aluno gerando o que ele afirma ser uma desigualdade de oportunidade e funciona como um alimentador de desigualdade no futuro”. (BARROS, R.P (2015), REVISTA ÉPOCA.

Um fato que Paes de Barros afirma também na entrevista é a questão de as boas práticas educacionais como, por exemplo, em Sobral, no Ceará não são disseminadas pelo resto do país.

A desigualdade de aprendizado neste estudo é analisada pelos métodos da divergência Kullbak-Leibler (1951) em que é medida a distância entre a distribuição de

¹ Mais informações sobre o PISA disponível <http://portal.inep.gov.br/artigo/-/asset_publisher/B4AQV9zFY7Bv/content/o-que-e-o-pisa/21206>

referência e a distribuição observada dos municípios de 2007 a 2011. Sobre a igualdade educacional:

“... uma situação de igualdade em educação ocorre quando a distribuição dos aprendizados dos estudantes, pertencentes a quaisquer grupos, é igual à distribuição de referência. Diante disso, a medida de qualidade da educação deve ser a distância da distribuição real observada até a distribuição de referência e, conseqüentemente, a desigualdade é estudada comparando-se o valor dessa distância para diferentes grupos de estudantes. Ou seja, as especificidades de uma análise educacional impõem que o relevante educacionalmente é a equivalência entre distribuições e não a igualdade absoluta do aprendizado.” (SOARES e DELGADO, 2016. p. 2)

A questão enfatizada por Delgado; Ribeiro e Soares (2013) é que a partir do momento em que o desempenho se baseia na média ele pode se manter à custa de mais desigualdade, pelo fato de que se é possível aumentar o aprendizado daqueles que são mais capazes e deixar os que possuem dificuldades em situação inferior gerando desigualdades cada vez maiores. Delgado; Ribeiro e Soares (2013) afirmam que para buscar uma melhoria para esse problema, é necessário um programa de acompanhamento mais efetivo e preciso quanto aos resultados, além da possibilidade para desenvolver as diversas habilidades que um aluno pode desenvolver na escola.

Os estudos no campo da desigualdade de aprendizado buscam suas informações no Sistema Nacional de Avaliação Básica (SAEB) que conforme Alves e Xavier (2016), vem abordando questões tais como a aprovação, a situação do docente quanto a formação, as complexidades do diretor, e o nível socioeconômico dos alunos que se mostram um fator de geração de desigualdades.

Para explicar os fatores que geram a desigualdade, esse estudo escolheu usar os dados do SAEB dos anos de 2007 e 2011 para medir a distância de aprendizagem dos municípios brasileiros em relação ao ponto de referência que é zero (explicado no capítulo 3). O estudo está dividido em 4 capítulos além deste: referencial teórico, métodos, resultado e conclusão. O referencial teórico, explora-se a literatura do tema, passando desde a contextualização das origens dos indicadores de desigualdade de renda como o Gini e sua aplicação na área educacional, passando pela apresentação dos indicadores socioeconômicos, o impacto da infraestrutura escolar no aprendizado até a

função de produção educacional que é uma adaptação da microeconomia tradicional para a educação. No capítulo que trata dos métodos, o estudo vai buscar apresentar o processo de construção do SAEB usado atualmente e em como ele será usado no estudo pelo software R, além de explicar a divergência Kullback-Leibler, que é a usada no cálculo. No capítulo dos resultados, será feita uma descrição dos resultados, tanto na distância KL quanto no IDE. O capítulo da conclusão vai além de buscar dar um apanhado geral no texto, tem como objetivo propor soluções realistas ao estudo, capaz de ajudar na melhoria dos indicadores educacionais e que facilite no cumprimento das metas estabelecidas no Plano Nacional de Educação, aprovado em 2014, em que constam as diretrizes de buscar uma educação inclusiva e de qualidade para todos, mas que enfrenta atualmente dificuldades na sua execução.

2 A MENSURAÇÃO DA DESIGUALDADE NO CONTEXTO EDUCACIONAL BRASILEIRO

2.1 Os avanços institucionais da educação brasileira

Melhorar a educação brasileira tem sido um dos grandes desafios das últimas décadas, pois o avanço nessa área reflete em aumento na produtividade da economia, e leva a uma população melhor instruída que terá melhor capacidade em conviver os desafios que lhe cercam. Na Constituição Federal promulgada em 1988, o direito a educação está presente no seu artigo 3º como um dos objetivos fundamentais da nação para a construção e fortalecimento da cidadania e mais especificamente entre os artigos 205 e 214 em se que enfatiza:

“A educação, direito de todos e dever do Estado e da família, será promovida e incentivada com a colaboração da sociedade, visando ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho. ”

Artigo 205 da Constituição Federal de 1988.

Em 1995 foi promulgado através da Emenda Constitucional nº 14 e regulamentada pela lei nº 9.424/96 a instituição do Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e Valorização do Magistério (FUNDEF) no qual estava atribuído instituir a vinculação dos orçamentos da União, estados e municípios para a educação básica. Em 2006, o FUNDEF foi substituído, segundo o Ministério da Educação, pelo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (FUNDEB) através da Emenda Constitucional nº 53 e regulamentado pela [Lei nº 11.494/2007](#) e o [Decreto nº 6.253/2007](#) com duração até 2020. Em 2014 foi aprovado o Plano Nacional de Educação através da Lei nº 13.005 de 2014 norteando os objetivos e como chegar aos mesmos em um prazo de dez anos, 2024. Segundo o Ministério da Educação, o plano tem três eixos de ação, reafirmação do direito ao acesso ao ensino, redução das desigualdades sociais e regionais e valorização dos professores, além da ambiciosa meta da aplicação de 10% do PIB em educação.

Vale ressaltar também que a gestão do Ensino Fundamental Público é de responsabilidade dos estados e municípios, estes são responsáveis pela aplicação das diretrizes determinadas por todo esse processo enfatizado de alterações orçamentárias.

Além das mudanças na área orçamentária, o Brasil nos últimos anos tem buscado várias metodologias para diagnosticar os desafios e as conquistas na área da educação. O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) é responsável por diversas avaliações e diagnósticos da educação básica brasileira, entre outras pesquisas mais importantes podemos constatar o Censo Escolar, o Sistema de Avaliação da Educação Básica (SAEB), através de processos que mudaram ao longo do tempo como Prova Brasil, Avaliação Nacional da Educação Básica (ANEB) e a Avaliação Nacional de Alfabetização (ANA), que juntos levam a formação do Índice de Desenvolvimento da Educação Básica (IDEB), e a Prova Docente, que avalia os professores do magistério.

Além das iniciativas governamentais para melhorar o acesso e a qualidade, existem também ações por parte de pessoas da sociedade civil como o movimento “Todos pela Educação” e a Fundação Lemman, ambos buscando o que a sociedade brasileira também: desejam uma educação universal e de qualidade para todos.

Com um processo de mudanças institucionais na gestão educacional, o Brasil segundo o Data World Bank aumentou sua taxa de alfabetização para pessoas acima de 15 anos de 86% em 2000 para 92% em 2015. O IDEB para os anos iniciais do ensino segundo o INEP passou de 3,8 em 2005 para 5,5 em 2015 superando a meta estabelecida pelo MEC, mas ainda abaixo do índice para comparação com os países desenvolvidos. No Programa Internacional de Avaliação de Estudantes (PISA) que é usado para avaliar jovens no fim do ensino fundamental, o Brasil ainda aparece nas piores colocações nos últimos anos.

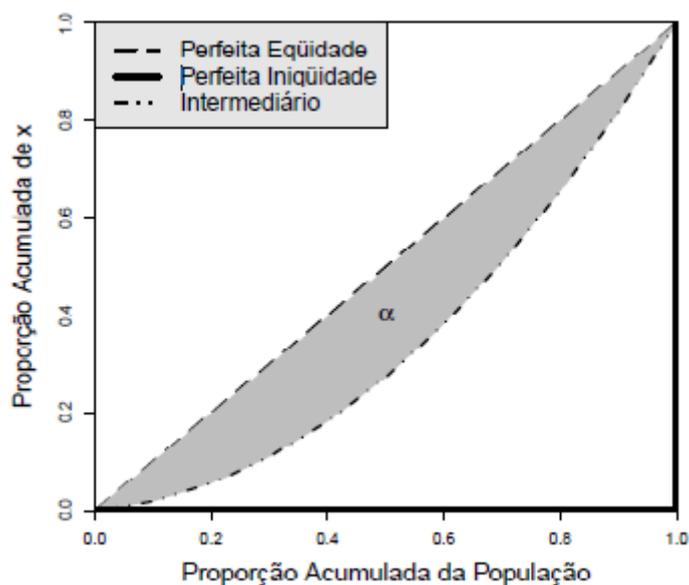
2.2 A construção dos indicadores de desigualdade

O Brasil conseguiu evoluir no aprendizado escolar nas últimas décadas, com o aumento das taxas de alfabetização até os 15 anos conforme dados da Data Bank World. Contribuíram para este avanço a criação de índices de avaliação do aprendizado em nível nacional e em algumas avaliações estaduais, porém, esses avanços não ocorreram

de maneira homogênea regionalmente e em todos os estratos sociais. Para essas conclusões existem estudos, como uma pesquisa feita em 2013 pela Fundação Victor Civita chamada de Análise das Desigualdades Interescolares no Brasil, que busca explicar as causas desse desafio da equidade através da análise de uma literatura baseada no problema da desigualdade por questões do interior das escolas, análise dos bancos de dados dos índices de avaliação educacional como o IDEB e análise dos questionários aplicados a alunos e docentes junto a essas avaliações de desempenho.

Os estudos derivados da desigualdade de aprendizado se baseiam em publicações já consolidados em desigualdade de renda como a curva de Lorenz e o Índice de Gini. Medir a curva de Lorenz sem desigualdade é quando o ponto x é igual a média, ou seja, a reta se encontra em perfeita equidade e as diagonais (0,0) e (1,1) são as mesmas da curva. Caso o ponto x seja desigual ao extremo, pode se concluir que a curva estará no eixo horizontal até o encontro com a proporção acumulada para x . A área denominada α é aquela que se encontra entre a linha de perfeita equidade e a curva, conforme apresentada no Gráfico 2.

Gráfico 2: Uma possibilidade de curva de Lorenz



Fonte: Centro de Políticas Sociais da Fundação Getúlio Vargas, 2012.

O Índice de Gini pode ser realizado sem recorrer ao uso da curva de Lorenz baseando-se em uma distribuição de x aleatória e conforme Lerman e Yitzhaki (1984) a curva tem em seu desenho geométrico pelo menos duas vezes a área entre a Lorenz e a reta da de perfeita equidade:

$$\text{Gini} = 2x(L_{0,5}) - 1 \quad (1)$$

Temos que $x(L_{0,5})$ é o valor da proporção acumulada da população quando a proporção acumulada da renda é 0,5. Essa metodologia, já consolidada, foi usada para estudos mais recentes para avaliar o desempenho dos estudantes e comparar o quão ainda está desigual o nível de aprendizagem. Outra forma de medição das desigualdades são as distribuições de probabilidade apresentadas por Soares e Delgado (2016). Os autores utilizaram-se de uma medida chamada de Kullback-Leibler, devido ao trabalho seminal de Kullback e Leibler (1951) que será chamada de agora em diante de KL neste trabalho.

A KL é usada para medir a ‘distância’ entre o ponto de referência e os desempenhos obtidos, ela é similar à curva de Lorenz, mas que na verdade, leva a uma medida mais geral que o Índice de Gini. Soares e Delgado (2016) afirmaram que o uso desse tipo de distribuição é importante para referenciar medir a distância entre os municípios e com isso, ganhar mais focalização quanto ao planejamento das políticas públicas. O gráfico 3 apresenta bem o que representa esse tipo de distribuição, evidenciando a redução da desigualdade de aprendizado dos alunos entre 2005 e 2013.

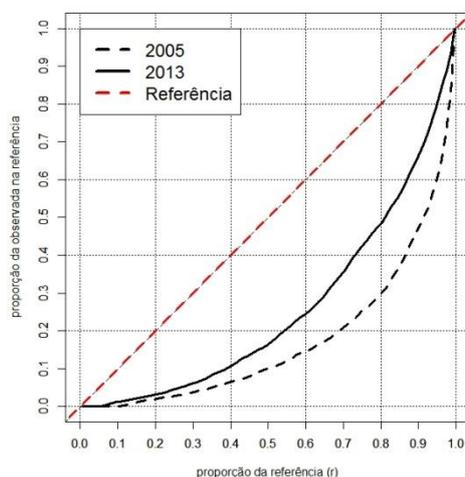


Gráfico 3: Representação da densidade acumulada para o nível de leitura em 2003 e 2015

Fonte: Prova Brasil, 2005 e 2013.

Soares e Delgado (2016) também usaram o conceito de densidade relativa aplicado por Handcock e Morris (1998 e 2006) por ter uma utilidade mais elevada se a

densidade acumulada for bem compreendida. Além disso, Delgado et. al. (2013) usaram uma medição de desigualdade similar à usada na medição de aprendizagem pelo teste do PISA, usada também em índice de desempenho escolar no estado de São Paulo.²

Para encontrar essa desigualdade de aprendizado os estudiosos educacionais têm usado os dados do SAEB a partir da Prova Brasil e da Avaliação Nacional da Educação Básica (ANEB), além de alguns estudos terem abrangência estadual, como por exemplo, a Avaliação da Rede Pública de Educação Básica (PROEB) em Minas Gerais. Além de dados referentes aos níveis de proficiência, essas avaliações adotam questionários a alunos, professores e diretores tornando-se possível um mapeamento da realidade das escolas.

Segundo Alves e Xavier (2016), o desempenho dos alunos na Prova Brasil, principal fonte de dados para este estudo, vai de 0 a 500 e esse resultado é medido pelo modelo de Teoria da Resposta ao Item, o que segundo Klein (2009), com a equalização entre as séries, evitam-se erros disseminados quando se analisa cada série escolar presente no teste. Novamente, segundo Alves e Xavier (2016) quando se estima os indicadores pela Teoria da Resposta ao Item (TRI)³, são levadas em conta apenas as respostas que foram respondidas pelos alunos que preencheram o questionário.

Para conhecer o que impacta o desempenho dos alunos, César e Soares (2001) afirmam que fatores individuais como renda, situação familiar e localização, por exemplo, são mais determinantes do que fatores escolares. Apesar disso, eles afirmam que os estudos para se analisar o impacto da escola no nível de aprendizado são fundamentais para diagnosticar a desigualdade de proficiência.

O nível socioeconômico (NSE), segundo Alves; Soares e Xavier (2014), representa, um estudo amplo da realidade educacional e podem explicar bem os impactos externos no desempenho dos alunos e têm sido usado com muita frequência

² é a medição da distância através dos desvios padrão de um grupo de países que tiveram agrupados seus respectivos desempenhos e a obtenção da média aritmética deles. A média desse grupo de países, chamado país típico, é o ponto de referência para comparação entre os demais pertencentes.

³ A TRI baseia-se em modelos matemáticos que permitem a elaboração de provas diferentes com o mesmo grau de dificuldade. O modelo é baseado em três parâmetros gerais — grau de dificuldade da questão, nível de discriminação (o quanto o item consegue diferenciar o quanto o candidato sabe ou não) e a probabilidade de acerto ao acaso (chute). Mais informações no site do MEC <<http://portal.mec.gov.br/ultimas-noticias/222-537011943/17189-teoria-de-resposta-ao-item-garante-isonomia-das-provas>>

por estudiosos. O NSE divide os alunos em estratos sociais. O dado para a construção das estatísticas descritivas para o nível socioeconômico vem dos questionários dos métodos de avaliação de desempenho aplicado aos alunos como SAEB, Prova Brasil usando-se a TRI como já mencionado. O Brasil tem dados recentes sobre a área, esses dados começaram a ser coletados com maior focalização e eficiência a partir da segunda metade dos anos 1990. Abaixo a tabela 1 com os participantes das avaliações entre 2007 e 2011 do 5º ano do ensino fundamental:

Tabela 1: Número de participante das avaliações Saeb/Prova Brasil entre 2007 e 2011

AVALIAÇÃO	ALUNOS	ESCOLAS	PROFESSORES	DIRETORES
Saeb 2007	2.983.647	48.149	292.828	47.596
Saeb 2011	3.238.506	58.960	316.668	58.960

Fonte: Elaborado pelo autor com dados do Saeb/Prova Brasil de 2007 a 2011.

Os dados do Saeb começaram a ser coletados a partir de 1995, é um sistema de dados que reúne também a ANEB e a Prova Brasil disponível em forma de microdados na página do INEP. A Prova Brasil, segundo Alves e Xavier (2016), aplica suas avaliações para os alunos do 5º e do 9º ano do ensino fundamental para medir a proficiência em Português e Matemática nas escolas públicas que possuem pelo menos 20 alunos na sala de aula, já a ANEB usa mesmos critérios só que para as escolas particulares e para o 3º ano do ensino médio de maneira amostral.

O mecanismo de avaliação da infraestrutura escolar, segundo Neto et al (2013), vem ganhando ênfase nos últimos anos no Brasil pelo fato de envolver a eficiência do gasto público destinado à manutenção e melhoria da infraestrutura das escolas públicas na qualidade do nível de aprendizagem dos alunos. Em seu artigo, os autores publicaram uma escala para medir a infraestrutura escolar, concluíram que apesar de uma melhoria parcial na estrutura física das escolas, o impacto no aprendizado foi baixo. Além disso, segundo eles, em um estudo feito por Soares; Razo e Farinãs (2006), os baixos índices de aprendizados na zona rural tem importante causa na falta de estrutura.

Neto et al (2013) afirmaram também que no Brasil a oferta do ensino básico é predominantemente por estados e municípios e esses são muito desiguais dependendo da

região do país, o que impacta diretamente na gestão educacional. O estudo feito buscou ver qual é a situação da atividade das escolas se elas estão em atividades, paralisadas ou extintas e se as mesmas possuem infraestrutura elementar, básica, adequada ou avançada e abrangeu todo país se baseando nos dados do Censo Escolar de 2011.

Tabela 2: Descrição de como é classificada a infraestrutura escolar

NÍVEL	INTERVALO	PORCENTAGEM	DESCRIÇÃO DOS NÍVEIS DE PROFICIÊNCIA
1 – Elementar	0 < 20	0,0	Estão neste nível escolas que possuem somente aspectos de infraestrutura elementares para o funcionamento de uma escola, tais como água, sanitário, energia, esgoto e cozinha.
	20 < 30	2,8	
	30 < 40	17,8	
	40 < 50	23,9	
2 - Básica	56 < 60	40,0	Além dos itens presentes no nível anterior, neste nível as escolas já possuem uma infraestrutura básica, típica de unidades escolares. Em geral, elas possuem: sala de diretoria e equipamentos como TV, DVD, computadores.
3 – Adequada	60 < 70	14,9	Além dos itens presentes nos níveis anteriores, às escolas deste nível, em geral, possuem uma infraestrutura mais completa, o que permite um ambiente mais propício para o ensino e aprendizagem. Essas escolas possuem, por exemplo, espaços como sala de professores, biblioteca, laboratório de informática e sanitário para educação infantil. Há também espaços que permitem o convívio

			social e o desenvolvimento motor, tais como quadra esportiva e parque infantil. Além disso, são escolas que possuem equipamentos complementares como copiadora e acesso à internet.
4 – Avançada	70 < 80	0,6	As escolas neste nível, além dos itens presentes nos níveis anteriores, possuem uma infraestrutura escolar mais robusta e mais próxima do ideal, com a presença de laboratório de ciências e dependências adequadas para atender estudantes com necessidades especiais.

Fonte: (NETO et al, 2013). Uma escala para medir a infraestrutura escolar.

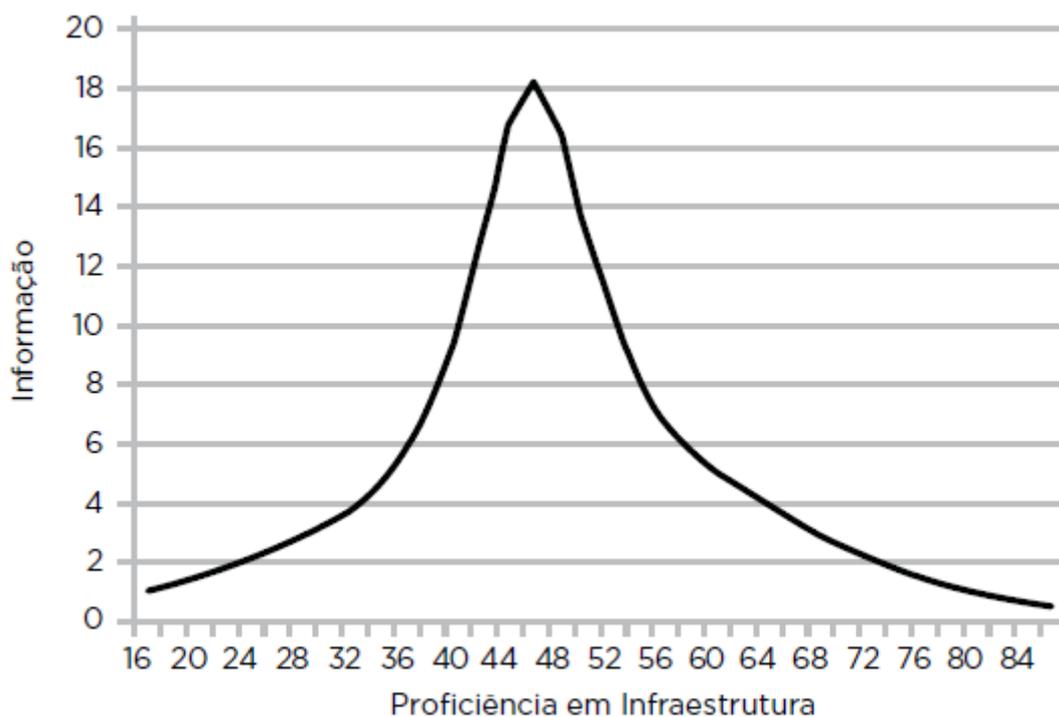
As conclusões deste artigo são que 44,5% das escolas estão classificadas como elementares, 40% das escolas estão com o nível de infraestrutura classificado com básico e apenas 14,9% com o nível de infraestrutura classificado como elementar e apenas 0,6% é classificada com infraestrutura avançada. As conclusões dos autores afirmam que 51,3% das escolas estaduais estão no nível básico, 61,8% das escolas municipais no nível elementar e 72,3% das escolas particulares estão nos níveis elementares e básicos. Essas diferenças são maiores à medida que as escolas se distanciam dos centros urbanos, evidenciando assim um processo de perpetuação das desigualdades na formação escolar, pois as condições ainda estão longe das ideais e necessárias para o desenvolvimento equitativo do aprendizado.

Um estudo realizado pelo Departamento de Análise em Políticas Públicas da Fundação Getúlio Vargas (DAPP – FGV) denominado Universo Escolar, avaliou a infraestrutura das escolas em que seus alunos participaram do Exame Nacional do

Ensino Médio (ENEM), e foi notada a percepção de que alguns instrumentos básicos são necessários para que o aluno obtenha bons níveis de aprendizagem.

No gráfico 4, foi calculado os escores da proficiência da infraestrutura escolar com base nos dados do Censo Escolar de 2011. Conforme NETO et. al. (2013), esta medida baseia-se na referência dos parâmetros dos itens que compõem a curva de informação. Para cada ponto percentual da escala, o erro de informação representa o inverso do valor de informação. NETO et. al., 2013, afirmam também que para definir os itens âncora, se leva em consideração o percentual de resposta positiva empírica em torno de cada ponto. Essa metodologia se baseia em um estudo de Beaton e Allen (1992) e no estudo foi usada a proporção empírica dos itens.

Gráfico 4: Escala de Proficiência em Infraestrutura baseada na escala da informação



Fonte: (NETO et al, 2013).

O estudo “Remuneração dos docentes em exercício na educação básica” apresentado pelo INEP neste ano de 2017 foi baseado no cruzamento de dados do Censo Escolar e da Relação Anual de Informações Sociais (RAIS), alcançou 93,3% dos docentes da educação básica e concluiu que a rede municipal que é 45 vezes maior do

que a rede federal e paga o equivalente a menos da metade do salário das federais, e que a rede privada tem os piores vencimentos. O estudo conclui também que há disparidade regional entre o salário dos docentes, o que evidencia que o problema da desigualdade vai para além dos alunos, chega aos professores, aumentando os desafios para o debate educacional.

Na microeconomia tradicional, o grande dilema abordado pelos estudiosos da área envolve o *trade-off* entre eficiência e equidade Delgado, Ribeiro e Soares (2013) usaram uma citação de Rawls para com a relação desempenho e igualdade

“ Na aplicação desse princípio [o da reparação], talvez se viessem a despendar mais recursos com a educação dos menos inteligentes, e não dos mais inteligentes, pelo menos durante certo período de vida, digamos, os primeiros anos de escolar. [...] o princípio da diferença alocaria recursos para a educação, digamos, para elevar as expectativas de longo prazo dos menos desfavorecidos. Se tal fim for alcançado dando-se mais atenção aos mais talentosos, é permissível; caso contrário, não. E ao tomar essa decisão, não se deve aferir o valor da educação apenas no tocante à eficiência econômica e ao bem-estar social. Tão ou mais importante é o papel da educação de capacitar uma pessoa a desfrutar de uma cultura de sua sociedade e participar de suas atividades, e desse modo de propiciar a cada indivíduo um sentido seguro de seu próprio valor. (RAWLS, 2008, p. 120-121.)

Vale lembrar que justiça e equidade da educação não serão as únicas coisas observadas, pois não há como se chegar a uma igualdade absoluta que é a que se busca o índice de Gini. Delgado, Ribeiro e Soares (2013) propõem é uma medição da igualdade relativa em busca do que chamam de patamar mínimo de desempenho. Os autores afirmam que os dados apresentam uma tendência de média central e se dispersam em torno de uma média e que seria impossível manter a ideia de que todos os alunos tenham a mesma nota. Eles comparam este conceito com os de **erro sistemático**, que é aquele que erro que persiste na relação entre o valor estimado e valor real, esse seria uma deficiência real do sistema educacional e o **erro não sistemático**, que é algo provido de uma dispersão aleatória de dados e com viés igual a zero disperso em torno da média.

O Sistema Gerador de Desigualdade (SGD), segundo Delgado, Ribeiro e Soares (2013) representa um processo desconhecido que gera dados ao pesquisador. O pesquisador que tem o objetivo de construir o processo gerador que dá origem aos dados. Esse tipo de função é oriundo da função densidade de probabilidade (f.d.p.), e observa-se que:

$$1) f(x) \geq 0 \quad (2)$$

$$2) \int_{-\infty}^{\infty} f(x)dx = 1, \quad \text{sendo } x \text{ uma variável aleatória contínua em } \mathbb{R}.$$

Essas funções, segundo Delgado, Ribeiro e Soares (2013), possuem a igualdade relativa que busca uma distribuição a mais justa possível. Esta questão será especificada com mais detalhes no capítulo 3:

2.2.1 Nível Socioeconômico

Barros (2012) afirma que apesar de o Brasil ter passado por alguma evolução na infraestrutura, a mesma ainda está longe do ideal, e que as pesquisas não conseguem mensurar o impacto dessas melhorias de maneira efetiva na qualidade do aprendizado dos alunos. No entanto para o autor claro que é necessária uma infraestrutura adequada para o aprendizado.

Para (BARROS, 2012) o ambiente familiar é fator preponderante no desempenho dos alunos, pois segundo um relatório do Centro de Estudos e Pesquisas em Políticas Públicas da Faculdade de Educação da Universidade de São Paulo (USP), em parceria com a Fundação Victor Cívita esse impacto é mais forte em países pobres do que nos mais ricos, pois as mesmas podem proporcionar os filhos algo como aulas de reforço, cursos de línguas que geram uma formação complementar a esses alunos. A questão é que esta formação complementar é restrita e difusora das desigualdades, pois os alunos não capazes de ter essa complementação fazem a escola ser o total protagonista do seu aprendizado:

“Este estudo apresenta a dimensão ainda ignorada da desigualdade no nível nacional, que parece ser de importância crucial para entender onde as escolas podem ser mais importantes. De fato, relativamente à família, a escola parece

ser mais importante em contextos desiguais, mas não em contextos de pobreza. Tal descoberta implica que, mesmo em um país rico, mas relativamente desigual (como os Estados Unidos ou Singapura), as escolas podem desempenhar um papel mais importante do que tem sido reconhecido. ” (CHUDGAR; LUSCHEI, 2009, pg. 652).

Uma reportagem feita pelo jornal Valor Econômico, em abril de 2017, mostra que um estudo feito pelo movimento Todos Pela Educação, baseado em dados da Avaliação Nacional de Alfabetização (ANA) confirma esse impacto da situação das famílias na educação, pelo fato de o estudo concluir que os alunos cujas as famílias estão em um nível socioeconômico elevado, ou seja, cuja renda familiar é acima de sete salários, mínimos possuem 98,3% do nível adequado de leitura, 95,4% na escrita e 85,9% na matemática, enquanto as famílias com nível socioeconômico baixo, ou seja, cuja renda familiar é de até um salário mínimo, 45,4% apresenta o nível de leitura adequado, 24,9% na escrita e 14,3% na matemática. Essa avaliação, segundo o INEP, é aplicada para alunos do 3º ano do ensino fundamental e mostra que o problema da desigualdade de aprendizado começa quando o aluno entra na escola, e aí tem de se analisar além da renda familiar, a localização, cor, sexo, dependência administrativa e a turma.

Albernaz, Ferreira e Franco (2002), em um estudo sobre Qualidade e Equidade no Ensino Fundamental, concluíram que além da situação socioeconômica de família impactar no desempenho dos alunos, o nível socioeconômico médio dos alunos da escola é determinante por envolver questões de agrupamento social que mensuram de maneira mais eficaz a realidade escolar, pois as escolas tendem a ser homogênea quanto às características socioeconômicas dos seus alunos, apesar de se ter uma ou outra exceção à regra. Os autores afirmam também que quanto maior o nível socioeconômico médio da escola, menor será o efeito da realidade familiar no desempenho dos alunos e por consequência da escola.

A construção desses dados é derivada dos questionários aplicados junto as avaliações do SAEB e incluem, além dos alunos, professores e diretores. Por meio delas é possível construir um mapeamento municipal, estadual e nacional da realidade educacional.

Alves, Soares e Xavier, (2014), enfatizam a correlação da renda familiar com o desempenho dos alunos e, por consequência, a desigualdade nos níveis de aprendizados justificando as várias evidências empíricas ao redor do mundo, e cita o famoso relatório de Coleman⁴, publicado em 1966, como exemplo dessa associação. Os autores afirmam da necessidade dos questionários aplicados terem clareza e precisão adequada a idade dos alunos para que se consiga a real compreensão da realidade educacional brasileira.

Segundo Andrade e Laros (2007), o SAEB iniciado em 1995, foi um grande avanço na mensuração da realidade educacional, apesar de os mesmos sugerirem em forma de hipótese que a relação família-aluno pode não estar de maneira adequada no questionário respondido pelos alunos, sugerindo inclusive uma readequação no questionário para abordar questões psicológicas para se atuar de maneira mais ampla nessa relação.

Tabela 3: Estatísticas Descritivas dos questionários dos alunos do 5º e do 9º ano do Ensino Fundamental do Saeb de 2007 e 2011

Variáveis	5º ano		9º ano	
	2007	2011	2007	2011
SEXO DO ALUNO (%)				
Masculino	46,2%	49,3%	45,7%	46,3%
Feminino	45,7%	47,1%	52,9%	52,4%
Sem Informação	8,1%	3,6%	1,4%	1,3%
COR (%)				
Branco	32,0%	29,2%	35,4%	32,5%
Pardo	42,5%	43,7%	45%	44,7%
Negro	10,7%	9,3%	10,7%	10,2%
Amarelo	2,9%	1,8%	3,8%	3,1%
Indígena	3,6%	2,3%	3,4%	2,1%

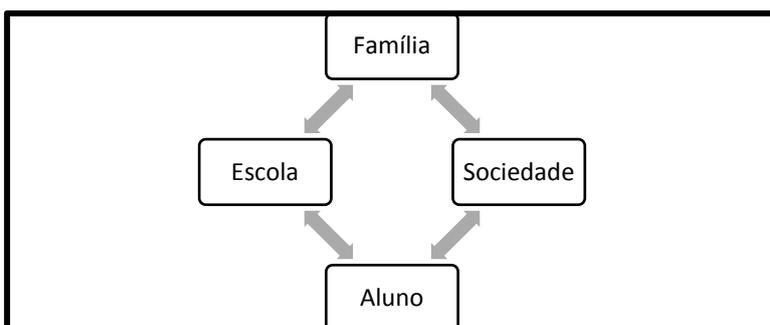
⁴ O relatório de Coleman foi publicado em 1956 por James Samuel Coleman baseando-se em dados de alunos, professores e escolas e buscou relacionar, o se o desempenho escolar está relacionado aos investimentos em infraestrutura, composição socioeconômica das escolas.

Sem Informação	8,4	13,7	1,7%	7,3%
ATRASO ESCOLAR (%)	2007	2011	2007	2011
Regular	63,3%	74,8%	66,2%	44,3%
Atrasado	34,3%	24,2%	33,2%	25,5%
Sem Informação	2,4	0,9	0,6%	0,2%

Fonte: Santos, J., 2014. Baseado nos microdados do Saeb/ Prova Brasil.

De maneira resumida, a figura 1 busca sintetizar o que foi publicado por Delgado, Ribeiro e Soares (2013). Ela apresenta as dimensões do reflexo educacional e seus impactos na proficiência. O sistema é integrado pela família, aluno, escola e a sociedade. Todos estão de maneira interligada uns aos outros, sobre os quais ocorrem impactos de todas as ações dos agentes inseridos. No sistema, a família é representada pela estrutura como ela é constituída socialmente, o envolvimento dos pais na atividade escolar do filho e como é a questão da renda, sendo, portanto, uma relação aparentemente simples, mas que impacta diretamente no desempenho do aluno.

Figura 1: Dimensão reduzida do sistema educacional e seu reflexo na proficiência



Fonte: Consolidado de Scheerens e Bosker (1997), Lee, Bryk e Smith (1993) e Gautthier (1997) apud Soares (2007)

Já a escola possui uma forte integração de funcionamento do sistema porque nela é onde são aplicadas as diretrizes pedagógicas e sua aplicação nas classes pelos professores e de maneira mais abrangente pelos diretores, e para tal aplicação vale

lembrar que depende também do envolvimento da comunidade escolar como toda a estrutura física e a relação interna da turma como o professor e entre os alunos.

O aluno é a parte mais interessada na focalização nos projetos educacionais possui suas características próprias oriundas de fatores importantes usados neste estudo como cor, sexo, raça e o desempenho escolar ao longo dos anos e isso impacta no nível de aprendizagem refletido nas escalas de proficiência. Em relação à sociedade, dela se demanda as questões legais, como as mudanças propostas nas questões pedagógicas, gestão orçamentária e a interação entre a escola e a comunidade.

2.2.2 Função de Produção da Educação e o Efeito Escola

Segundo Albernaz, Ferreira e Franco (2002), para o Brasil a função de produção para a educação começou a ter mais sentido após a criação do SAEB em 1995, pois antes havia uma limitação de informações para análise do desempenho escolar, já que existia apenas a variável taxa de escolaridade, gerando um problema na formulação desse modelo que lembra o da microeconomia tradicional. O modelo, segundo os autores, busca entender a relação entre os insumos e o produto da formação educacional e conforme o modelo temos o seguinte:

$$y = F(c, m, g, p, s)$$

Em que y representa o desempenho dos alunos, c representa o aluno e suas características próprias, m representa a características das famílias, g representa o nível socioeconômico, p representa as características médias dos alunos da escola, s representa o salário e a experiência dos professores.

Se fosse para se basear na teoria microeconômica tradicional, Araújo (2013) afirma que determinados custos dos insumos usados pelas instituições de ensino e os resultados obtidos para determinar o nível de tecnologia das escolas, o conhecimento disto seria suficiente para conseguir elevar o seu produto final, que no caso é o desempenho da escola e dos alunos. Apesar disso, Barr (2004) afirma que os outputs não têm uma precisão concreta do produto, pelo fato de a educação não ter um conceito único, e por isso contar apenas com os testes padronizados, que podem gerar uma avaliação imperfeita da realidade educacional, além de gerar fortes críticas por parte de setores

como a sociologia, pedagogia, por esse tipo de ação ser limitada e quantificada ao ponto de não gerar os efeitos esperados.

Araújo (2013) afirma então que o produto mensurado por essa função é basicamente a capacidade da escola em desenvolver a geração de conhecimento técnico, ou seja, leitura e matemática e o conhecimento cognitivo, que é a capacidade de aprimorar a convivência em comunidade dos alunos.

Com os insumos fornecidos de maneira monetária através da infraestrutura física através da aquisição de computadores, livros, carteiras, por exemplo ou dos insumos fornecidos de maneira não monetária através do efeito dos pares que segundo César e Soares (2011) representa o padrão médio do grupo de alunos a qual a escola está inserida por exemplo através do nível socioeconômico médio dos alunos, a motivação em frequentar a escola por exemplo.

Já o efeito-escola entra como insumo não monetário, ou seja, o aluno em si. Segundo César e Soares (2011), buscou-se nesse estudo apresentar a relação entre o desempenho de um aluno e a uma escola específica independente da formação dos professores ou do desempenho médio dos alunos ou comparar com uma relação ao desempenho dos alunos com as práticas pedagógicas e administrativas de uma determinada escola. Usando variáveis socioeconômicas como sexo, cor, dependência administrativa ao qual a escola está inserida, trabalha-se além das estatísticas de média e desvio padrão evidenciando-se que a desigualdade entre os alunos de uma mesma escola é prejudicial para todos, assim como um problema de concentração escolar, pouco observado em países desenvolvidos.

3 MÉTODOS

Os dados utilizados no estudo foram extraídos do Sistema Nacional de Avaliação da Educação Básica (SAEB), criado em 1995, com o objetivo de apresentar dados referentes ao nível de aprendizagem escolar dos alunos brasileiros. Esses dados são usados por gestores públicos para a formulação de políticas públicas orientadas de acordo com cada região do país, segundo afirma no site do INEP onde ficam concentrados esses dados.

Segundo o manual do usuário do SAEB de 2007, o sistema busca colher de maneira censitária os dados de acesso, qualidade, eficiência e equidade dos ensinos fundamental e médio do Brasil e com isso obter acesso à evolução da rede educacional brasileira ao longo dos anos. O órgão responsável por administrar o SAEB é o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), que segundo o site da instituição é uma autarquia do Ministério da Educação (MEC) orientada para formular políticas públicas para todos os níveis de ensino e governo na busca de melhorar a qualidade e o acesso da rede educacional.

Nos manuais do usuário⁵ do SAEB a avaliação educacional é feita sobre dois pressupostos: se o desempenho dos alunos vai impactar na qualidade do ensino oferecido e se não há algo isolado que impacte na qualidade, gerando o que as notas técnicas afirmam serem informações mais precisas para orientar em melhores políticas educacionais. Em 2005, o SAEB passou por uma reestruturação⁶ com o objetivo segundo o manual do usuário de 2007, de oferecer informações mais consistentes para avaliar a rede educacional buscando ver a questão do rendimento escolar e do sistema de ensino.

Essa reestruturação visava de maneira amostral avaliar o sistema público e privado, as redes federal, estaduais e municipais e os entes da federação. Nessa reestruturação foi criado a Avaliação Nacional da Educação Básica (ANEB) e Avaliação Nacional do Rendimento Escolar (ANRESC) conhecida também como Prova Brasil. A ANEB segundo o manual do usuário de 2007 representa a manutenção do que era feito no

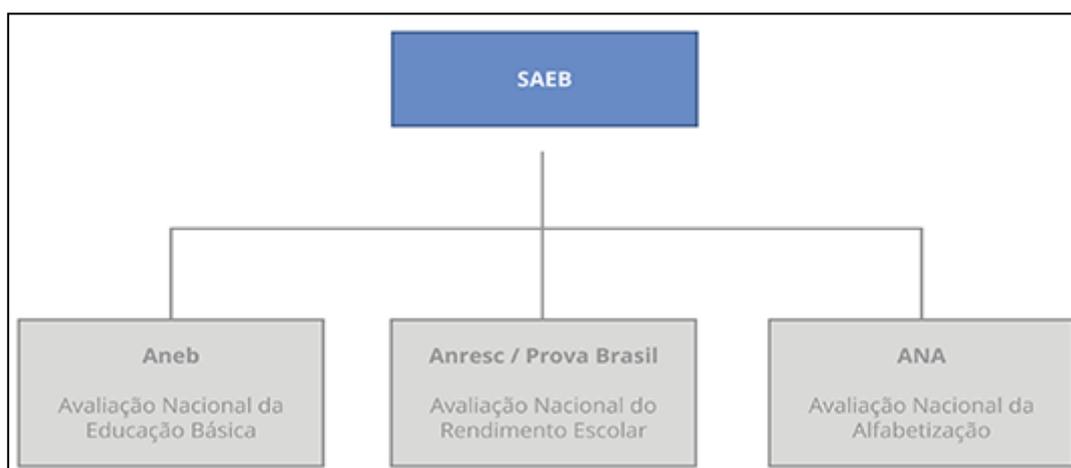
⁵ Os manuais do usuário citados são dos anos de 2009, 2011, 2013 e 2015.

⁶ A reestruturação ocorreu por meio da Portaria Ministerial nº 931 de 21 de março de 2005.

SAEB anterior a reestruturação de 2005 e serviu, conforme o manual do usuário de 2009, para auxiliar os executores de políticas públicas em educação na questão orçamentária, ações pedagógicas e a estipulação de metas para melhoria dos níveis de aprendizagem dos alunos e por consequência das escolas. Já a (ANRESC) - Prova Brasil, segundo o mesmo manual de 2009, veio com um caráter de aprofundamento do sistema por avaliar a qualidade do ensino ofertado na rede pública de educação baseado nas habilidades e competência obtidas pelos alunos, indo então além da aprendizagem. Em 2013 foi criada a Avaliação Nacional de Alfabetização, e que segundo o site do INEP tem o objetivo de avaliar os níveis de leitura e letramento dos alunos do 3º ano do ensino fundamental.

O SAEB, depois de ter passado por várias reestruturações que formularam o sistema de avaliação usado atualmente, possui 3 tipos de avaliação de aprendizagem dos alunos brasileiros conforme a figura 2, e essa é a estrutura geral:

Figura 2:Sistema Nacional de Avaliação da Educação Básica (SAEB)



Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2013.

Essas avaliações das provas usadas no estudo são realizadas a cada 2 anos a partir de 2005. Como dito anteriormente, os dados são obtidos de maneira censitária e amostral aos alunos do 5º e 9º ano do ensino fundamental e 3º ano do ensino médio das escolas brasileiras. Os dados para a realização do SAEB são providos do Censo Escolar realizado anualmente⁷, através do Educacenso que é o sistema que guarda os dados

⁷ Conforme a Portaria Ministerial nº 264/2007.

coletados para utilização nas avaliações, cruzamento de dados e orientações de políticas educacionais. Segundo o site do INEP, o Censo Escolar abrange a educação regular que representa a educação infantil, ensino fundamental e ensino médio, a educação especial, a educação de jovens e adultos (EJA) e a educação profissional que abrange os cursos técnicos, cursos de educação profissional e cursos de formação continuada com a colaboração⁸ entre a União, os entes federados e os municípios com suas atribuições específicas⁹.

A junção do SAEB e do Censo Escolar geram o Índice de Desenvolvimento da Educação Básica (IDEB) que conforme informa o site do INEP serve para avaliar em um índice a presença do aluno na escola e seu desempenho nas avaliações de aprendizagem, que por consequência é um agente orientador de acompanhamento das metas pré-estabelecidas pelo Plano Nacional de Educação (PNE)¹⁰ e no caso do IDEB serve para o acompanhamento da meta de que o Brasil em 2022 tenha a nota 6,0, comparável a de países desenvolvidos. Conforme a nota explicativa do IDEB publicada em 2007, o indicador é calculado da seguinte maneira:

$$IDEB_{ji} = N_{ji}P_{ji} \quad 0 \leq N_j \leq 10; 0 \leq P_j \leq 1 \text{ e } 0 \leq IDEB_j \leq 10$$

No qual,

i = ano do exame (Saeb e Prova Brasil) e do Censo Escolar.

N_{ji} = média da proficiência em Língua Portuguesa e Matemática, padronizada para um indicador entre 0 e 10 dos alunos da unidade j , obtida em determinada edição do exame realizado ao final da etapa de ensino.

P_{ji} = indicador de rendimento baseado na taxa de aprovação da etapa de ensino dos alunos da unidade j , varia entre 0 e 1.

As avaliações do SAEB, segundo o manual do usuário de 2013, buscam avaliar o desempenho escolar em Língua Portuguesa, no qual o foco é a leitura e o de Matemática é na resolução de problemas. Em ambas as provas há uma divisão por cada

⁸ Conforme a Portaria Ministerial nº 316/2007.

⁹ As atribuições dos envolvidos no processo estão disponíveis no site do INEP na parte do censo escolar <<http://inep.gov.br/web/guest/censo-escolar>>.

¹⁰ Plano de Desenvolvimento da Educação é um conjunto de metas e ações orientadas para a melhoria da qualidade educacional <<http://portal.mec.gov.br/arquivos/livro/livro.pdf>>.

etapa curricular do aluno através das Matrizes de Referência¹¹, as provas são montadas baseada no sistema Blocos Incompletos Balanceados (BIB) que significa haver uma combinação de blocos por área a ser avaliada.

Conforme já mencionado no referencial teórico deste estudo, o SAEB se baseia na Teoria de Resposta ao Item (TRI) que conforme o manual do usuário de 2009 é usado desde 1995 podendo ser feitas comparações de resultados de até dez edições anteriores. Os manuais do usuário do SAEB afirmam que o sistema de avaliação não tem como objetivo ranquear os alunos ou realizar uma avaliação individual de desempenho, mas apresentar informações conjuntas de uma turma, escola, município ou estado dos níveis de aprendizagem. A avaliação individual não é possível pelo fato de a metodologia não permitir o mesmo, pois a prova possui itens diferentes e no caso da avaliação da série, tem de ter pelo menos 20 alunos. Abaixo as classificações dos níveis de proficiência:

Tabela 4: Proficiência para a Língua Portuguesa. Conforme a fonte, os intervalos dos níveis das escalas de proficiências incluem o primeiro ponto e excluem o último ponto

Leitura		
Níveis	5º Ano	9º ano
Abaixo do Básico	Até 150	Até 200
Básico	De 150 a 200	De 200 a 275
Intermediário	De 200 a 250	De 275 a 325
Avançado	Maior que 250	Maior que 325

Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais, 2015.

¹¹ Conforme o site do INEP, as matrizes de referência são elaboradas para organizar na aplicação dos testes do SAEB, ocorrendo uma subdivisão que é chamada de descritores que é uma associação entre o conteúdo oferecido e o desenvolvimento mental do aluno. Vale ressaltar que essas matrizes servem para fazer um recorte curricular e dividir a prova em níveis realistas de aprendizagem intelectual do aluno. Mais informações na base de dados do SAEB, disponível nos microdados do site do INEP <<http://provabrasil.inep.gov.br/web/guest/microdados>>.

Tabela 5: Níveis de Proficiência para a Matemática. Conforme a fonte, os intervalos dos níveis das escalas de proficiências incluem o primeiro ponto e excluem o último ponto

Matemática		
Níveis	5º ano	9º ano
Abaixo do Básico	Até 175	Até 225
Básico	De 175 a 225	De 225 a 300
Intermediário	De 225 a 275	De 300 a 350
Avançado	Maior que 275	Maior que 350

Fonte: Instituto Nacional de Estudos e Pesquisas Educacionais.

Os resultados dos testes do SAEB são incorporados aos questionários aplicados a alunos, professores e diretores sobre diversas questões, e assim como os testes, os questionários são aplicados baseados no Censo Escolar do ano anterior e é feito um trabalho amostral adequado com o objetivo de obter uma boa precisão dos resultados.

Conforme a citação:

“Além dos instrumentos de medida de desempenho, são aplicados quatro tipos de questionários: de alunos, de professores, de diretores e de escolas. Os alunos respondem a perguntas sobre o ambiente e nível socioeconômico familiar, hábitos de estudo e de leitura, motivação, trajetória escolar, entre outros aspectos. O diretor da escola e os professores de cada uma das disciplinas testadas são convidados a fornecer informações sobre sua formação profissional, nível socioeconômico e cultural, estilo de liderança, formas de gestão, práticas pedagógicas, clima acadêmico, clima disciplinar, recursos humanos e pedagógicos. Procurou-se não realizar alterações nos questionários em relação à aplicação do ano de 2013, embora possam ser encontradas algumas atualizações pontuais, como as relacionadas a faixas salariais referenciadas no salário mínimo presentes nos questionários de professor e diretor. O aplicador preenche ainda questionários com informações sobre a escola, que estão relacionadas ao ambiente físico, infraestrutura, recursos, entre outros aspectos.” (MANUAL DO SAEB DE 2015, p 8, 2016)

Os resultados dos testes feitos pelos alunos e os questionários respondidos são usados na geração da base de dados através dos microdados¹², que agregam as bases dos alunos, professores, diretores, escolas, municípios, estados e Brasil. Esses microdados vêm passando por alterações ao longo dos anos, o que exigiu diferentes rotinas para leitura no R. Foram necessárias adaptações para se realizar este trabalho como por exemplo, em 2007 os microdados foram gerados no formato TXT e desde 2011 é usado o formato CSV, além dos microdados terem suas bases de dados constituídas de maneira diferente.

Nos microdados de 2007 e 2011 agregou-se o desempenho de todos os estudantes, ou seja, do 5º ano e 9º ano do Ensino Fundamental e do 3º ano do ensino médio em duas bases: uma de proficiência e questionário. Os dados passam por um processo de mascaramento, que segundo os manuais do usuário são uma codificação fictícia usada para facilitar a leitura dos dados, essa codificação é explicada em um dicionário existente nos microdados do SAEB.

A leitura dos dados foi feita no *software R*, em que as rotinas de 2007 e 2011, foram rodadas e recodificadas de acordo com os seus formatos e suas particularidades, e após a leitura passaram por um processo denominado agregação por município, em que foi gerada a média de desempenho dos alunos e que gerou uma média para os municípios das variáveis analisadas. Foi encontrada a proficiência e as demais variáveis estatísticas médias e o percentual de uma das opções das perguntas dos questionários por meio de um processo de manipulação dos dados que era similar a uma *dummy*.

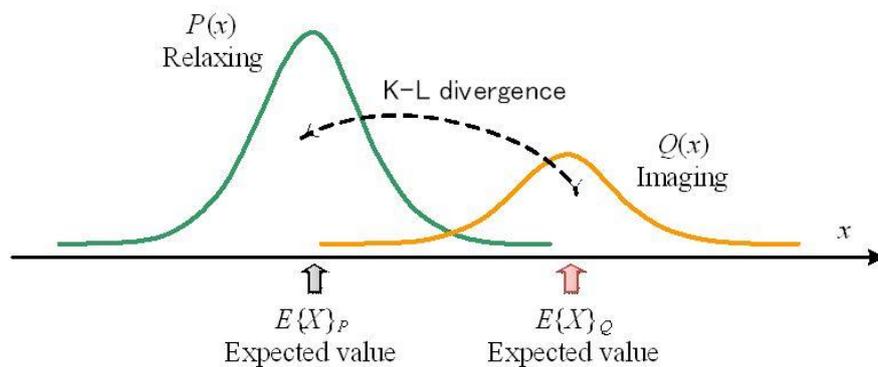
Vale ressaltar que inicialmente, a ideia seria contar com os microdados do SAEB de 2005, mas na discussão para construção desse estudo foi excluído, por entre outros motivos ter sido o primeiro SAEB dentro da nova metodologia, o que pode ter trago consigo alguns desajustes, sendo isto uma hipótese, nada que se comprovou ou se buscou para comprovar tal feito. Outra dificuldade é que a base de 2005 não está disponível de maneira direta no SAEB, sendo necessária obtê-la de maneira indireta para a comparação com a Prova Brasil.

¹² Segundo o site do INEP, os microdados são dados constituídos com o menor nível de desagregação e podem ser usados para leitura de softwares estatísticos. Os microdados são aplicados em todas as pesquisas, exames e avaliações do INEP.

O objetivo de um estudo de desigualdade é observar a distância em que se encontra uma pessoa em melhores condições, daquele que se encontra em piores condições. Na educação, segundo Soares e Delgado (2016) é feito o estudo da distância dos diferentes grupos de estudantes em relação a uma distribuição de referência e para tal e neste estudo será usado a distribuição de distância Kullback-Leibler que também é chamada de entropia relativa.

Segundo Pavão (2011), porque significa medir a distância entre duas distribuições de probabilidades. De acordo com o autor, a entropia é um medidor de incerteza de uma variável aleatória em que a distribuição de probabilidade de uma variável é discreta e os valores da entropia são sempre maiores ou iguais a zero e menores ou iguais a 1. Pavão (2011) afirma também que os eventos com a mesma probabilidade tanto na distribuição esperada como na distribuição verdadeira significa que a ineficiência¹³ é igual à zero.

Figura 3: Conceitos de extração de recursos com base na divergência do K-L



Fonte: INTECH, 2017.

Para a questão educacional, Soares e Delgado (2016) afirmam que é importante considerar a distância entre o desempenho alto com a distribuição de referência alto e com o desempenho baixo dos estudantes e nas duas situações o que deve ser enfatizado é a medida de qualidade. A medida KL para a educação busca aprofundar o desempenho dos estudantes para além do IDEB, que é gerado pela média das proficiências dos alunos em relação a uma escola que atende os quesitos básicos.

¹³ Ineficiência aqui entendida no sentido de que nenhuma informação é perdida, não há erro de informação.

Neste estudo, a ênfase da distância será para os municípios brasileiros, ou seja, medir o processo de evolução dos municípios dos níveis de aprendizagem e com isso ver também a evolução da distância dos municípios em relação à média brasileira e de municípios referência de aprendizagem. O cálculo é feito da seguinte maneira:

$$KL_m = \int_{-\infty}^{\infty} f_{Ref}(y) \cdot \log\left(\frac{f_{Ref}(y)}{f_{Obs}(y_m)}\right) dx \quad (3)$$

Em que:

KL_m , representa a distância de cada município m em relação à referência, $f_{Ref}(y)$, temos então $f_{Ref}(y)$ que representa a função de densidade distribuição da referência, $f_{Obs}(y_m)$ representa a função de densidade de distribuição observada dos municípios. Soares e Delgado (2016) afirmam, baseando-se em Shlens (2014) que essa expressão tem origem pouco intuitiva, sendo na realidade uma aproximação da função de máxima verossimilhança, ou seja, as distribuições $f_{Ref}(y)$ e $f_{Obs}(y_m)$ caso sejam discretas, terão uma KL que é uma aproximação da máxima verossimilhança em que se observa $f_{Obs}(y_m)$, sendo que o modelo gerador de dados é $f_{Ref}(y)$.

Vale ressaltar algumas observações a respeito da distância KL, que são assumidas algumas convenções conforme Esquef, Albuquerque e Albuquerque (2003) que $0 \cdot \log \frac{0}{f_{Obs}(y_m)} = 0$, e $f_{Ref}(y) \cdot \log \frac{f_{Ref}(y)}{0} = \infty$, além de afirmar que a divergência KL sempre será positiva, satisfazendo a desigualdade de Gibbs que:

$$\int_{-\infty}^{\infty} f_{Ref}(y) \cdot \log\left(\frac{1}{f_{Ref}(y)}\right) dx \leq \int_{-\infty}^{\infty} f_{Ref}(y) \cdot \log\left(\frac{1}{f_{Obs}(y_m)}\right) dx \quad (4)$$

Considerando que $f_{Ref}(y)$ é uma distribuição de probabilidade qualquer, e $f_{Obs}(y_m)$ outra distribuição que satisfaz a seguinte condição:

$$\int_{-\infty}^{\infty} f_{Obs}(y_m) dx \leq 1 \quad (5)$$

Vale ressaltar que essa distância apesar do nome não apresenta simetria, e essa questão segundo Jeffreys (1961) e foi aprofundada por Esquef, Albuquerque e Albuquerque (2003). Soares e Delgado (2016) afirmam que essa distância se tornou usual nos trabalhos nas áreas de ciências sociais, principalmente após Handcock e

Morris (1998, 2006) que apresentaram o conceito de densidade relativa para esse tipo de estudo com o objetivo de melhor entendimento.

A representação gráfica da distância KL se torna algo similar à curva de Lorenz, enfatizada no referencial teórico. Os autores afirmam perceber que a divergência KL é igual à média dos logaritmos $g(r)$, ou sua integral, que está no intervalo entre 0 e 1:

$$g(r) = \frac{f_{Ref}(y)}{f_{Obs}(y_m)} \quad (6)$$

E a notação fica da seguinte maneira:

$$KL_m = \int_0^1 \log(g(r)) \cdot g(r) dr \quad (7)$$

Se a divergência KL_m for igual à zero, representa que não há um gerador de desigualdades (SGD). Conforme Delgado, Ribeiro e Soares (2013), quanto maior for a distância KL_m , maior será a distância da $f_{Obs}(y_m)$ em relação à $f_{Ref}(y)$, que representa a função de densidade distribuição da referência.

$$KL(f_{Ref}(y), f_{Obs}(y_m)) = \int_0^1 \log(g(r)) \cdot g(r) dr = 0$$

Delgado, Ribeiro e Soares (2013) chamam a divergência Kullback-Leiber (1951) de Índice de Desempenho Relativo (IDR) que junto com o Índice de Qualidade da Educação (IQE) e o Índice Ponderado de Desempenho (IPD) interpretam a relação qualidade e igualdade da educação em Minas Gerais. O IDR, objeto metodológico do estudo pode se dividir em:

$$IDR = \textit{efeito 'desempenho'} \times \textit{efeito 'desigualdade'}$$

Delgado, Ribeiro e Soares (2013) afirmam também que conforme são analisados os indicadores, se um município obtiver uma evolução média de desempenho sem equidade, mas abaixo do esperado o indicador não atingirá seus objetivos.

Completando o cálculo da análise da divergência KL, ou o IDR, foi calculado o Índice de Desigualdade Escolar (IDE), desenvolvido a partir de uma distribuição degenerada discreta, que conforme Handcock e Morris (1998) significa ser uma distância entre a distribuição referencial e as distribuições observadas dos municípios possuem a proporção idêntica aos níveis de proficiência.

Neste estudo, a divergência KL possui quatro categorias: abaixo do básico que são aqueles que possuem o desempenho no SAEB aquém do que se espera e que seja necessário para um mínimo de conhecimento, básico, que são aqueles que possuem um desempenho mínimo de conhecimento, mas aquém do que se considera ideal, intermediário, que possuem um nível de conhecimento de acordo com o esperado para a série e avançado, que desenvolvem conhecimentos um pouco além do esperado para a série que cursam. Podemos escrever a KL de acordo com essas quatro categorias:

$$KL = \sum_{i=1}^4 p_i \ln\left(\frac{p_i}{1/n}\right) = p_1 \ln\left(\frac{p_1}{p_1^0}\right) + p_2 \ln\left(\frac{p_2}{p_2^0}\right) + p_3 \ln\left(\frac{p_3}{p_3^0}\right) + p_4 \ln\left(\frac{p_4}{p_4^0}\right)$$

As proporções passam por uma linearização que assumem os valores abaixo:

$$\ln\left(\frac{p_i}{1/n}\right) = 4, 2, 1 \text{ e } 0$$

Em uma metodologia criada para este trabalho, os valores representam o quão distante da referência após a linearização eles se encontram, e percebe-se que os municípios abaixo do básico assumem proporções maiores que os demais porque eles estão mais distantes da distribuição de referência, e que os municípios que apresentam um desempenho avançado assumem peso zero, pelo fato da divergência dos mesmos está próxima à distribuição de referência, diferentemente do que é apresentado nos demais grupos de municípios, então o IDE é calculado da seguinte maneira.

$$IDE = p_1 \cdot 4 + p_2 \cdot 2 + p_3 \cdot 1 + p_4 \cdot 0$$

O IDE é um caso discreto e particular do KI, no qual os valores de $\ln(p_i/p_0)$ assumem valores específicos dados acima

4 RESULTADOS

Os resultados apresentados neste estudo, feito através da divergência Kullback-Leibler contínua e a degenerada discreta (IDE), apresentam uma relativa redução da distância média observada de 2011 em relação a 2007. Os resultados mostram também que a KL discreta, ou IDE, por haver o critério de proporcionalidade, ou seja, de criar pesos maiores para os grupos com maior geração de desigualdade e vai reduzindo os pesos até zerar o desempenho dos geram menos desigualdade, apresenta uma menor dispersão em relação à divergência KL.

Especificamente para o 5º ano do ensino fundamental, as distribuições apresentam uma redução na distância média observada para a distribuição de referência tanto para leitura, quanto para matemática. Observa-se também uma melhoria quanto à distância mínima observada em relação à distribuição de referência, mas também um aumento da dispersão, o que evidencia que a desigualdade entre os municípios brasileiros segue aumentando tanto para leitura, quanto para matemática. Apresentada na tabela 6:

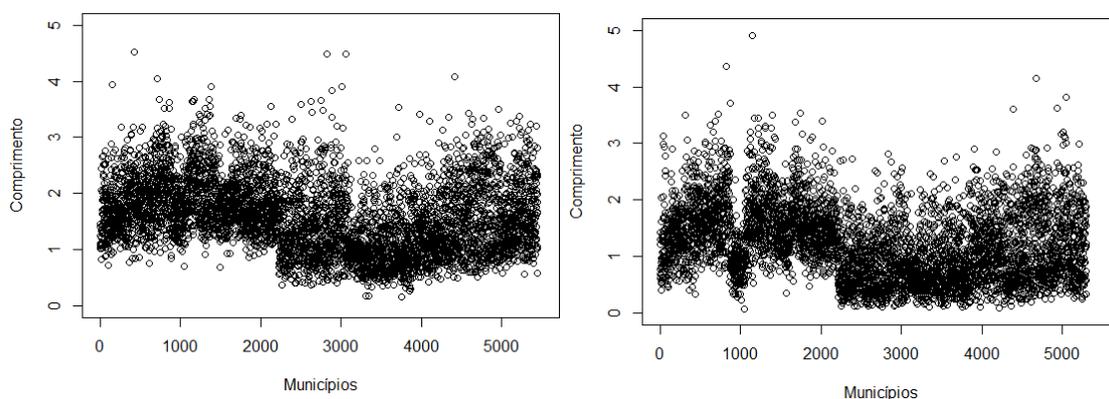
Tabela 6: Estatística Descritiva dos dados da KL e do IDE de 2007 e 2011 para o 5º ano do ensino fundamental

Leitura					
		Média	Desvio padrão	Mínimo	Máximo
KL	2007	1,58347	0,6297038	0,1529928	4,5116223
	2011	1,239409	0,6351811	0,07480164	4,91294549
IDE	2007	2,396679	0,4664598	0,556962	3,904762
	2011	2,045279	0,5294857	0,3150685	3,6
Matemática					
		Média	Desvio padrão	Mínimo	Máximo
KL	2007	1,749519	0,7249091	0,1253	4,7276
	2011	1,416353	0,7478593	0,08448424	4,82771060
IDE	2007	4,945878	0,2123093	4,178571	5,714286
	2011	4,988626	0,1822516	4,2	5,576923

Fonte: Elaborado pelo autor.

No gráfico 5, observa-se um leve aumento na dispersão dos municípios observados na KL de 2011 em relação a 2007, fato confirmado pelo desvio padrão que passou de 0,6297 para 0,6351. Vale afirmar também que a distância entre o melhor município para o pior também aumentou, pois o valor mínimo passou de 0,1529 para 0,0748 e o valor máximo passou de 4,5116 para 4,9113 evidenciando um aumento da desigualdade entre os municípios brasileiros, para leitura no 5º ano do ensino fundamental.

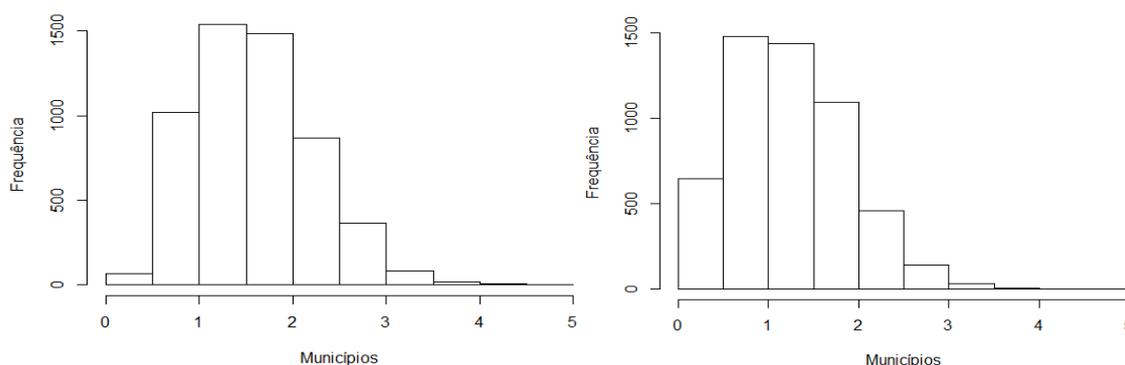
Gráfico 5: Dispersão de leitura do KL 2007 e KI 2011 do 5º ano do ensino fundamental



Fonte: Elaborado pelo autor.

No entanto, apesar do aumento da desigualdade, no gráfico 6, observa-se que mais municípios estão próximos da distribuição de referência, fato que a média da KL dos municípios brasileiros passou 1,5834 em 2007 para 1,2394 em 2011.

Gráfico 6: Histograma de leitura, 5º ano do Ensino Fundamental, KL 2007 e KL 2011

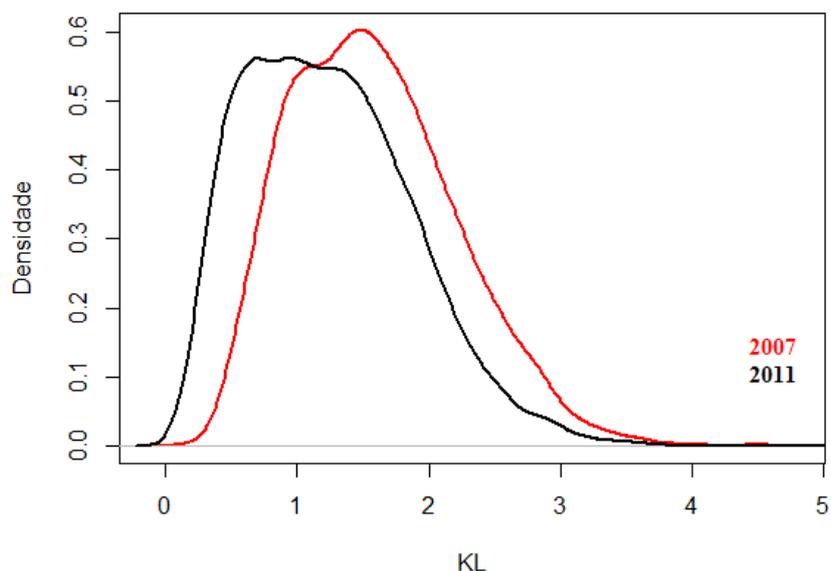


Fonte: Elaborado pelo autor.

O gráfico 7 representa a estrutura gráfica do cálculo da divergência KL para o 5º ano de leitura comparando o desempenho da distância observadas pelos municípios de

2007 e 2011, assim como dito na explicação do histograma e da dispersão que observou-se uma melhoria, mas que não acompanhada da redução da desigualdade.

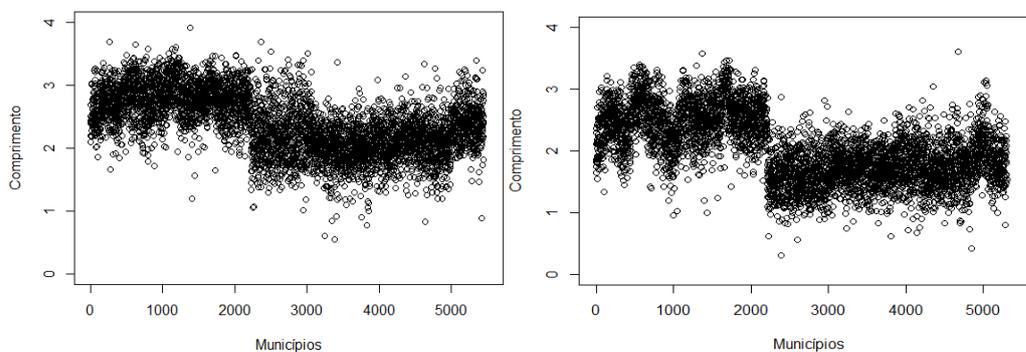
Gráfico 7: Comparação das Divergências Kullback-Leibler para o 5º ano em leitura de 2007 e 2011



Fonte: Elaborado pelo autor.

Para o IDE, observa-se no gráfico 8 um aumento maior na dispersão, ou seja, da desigualdade entre os municípios para o aprendizado em leitura, dado que o desvio padrão passou de 0,4664 para 0,5294 de 2007 em relação a 2011, apesar disso houve uma redução tanto no valor mínimo, que passou de 0,5569 para 0,3150, quanto no valor máximo passou de 3,9047 para 3,6, havendo uma melhora no desempenho nacional apesar de um aumento na desigualdade.

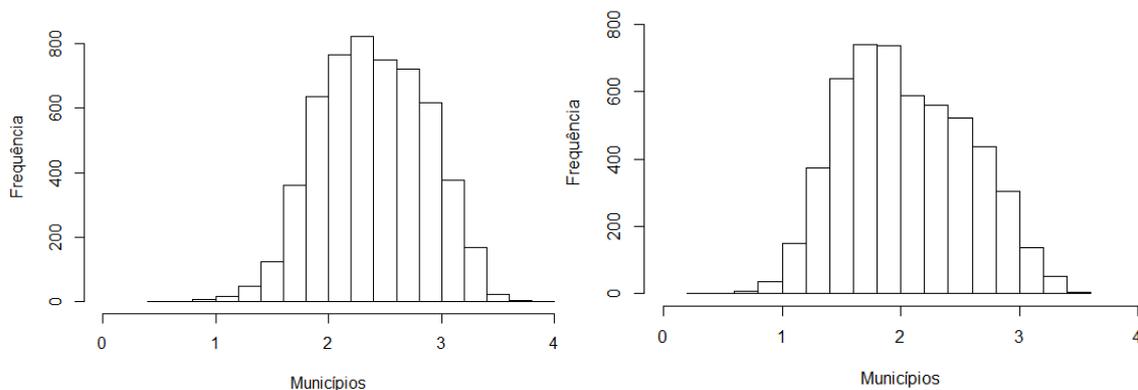
Gráfico 8: Dispersão de leitura do IDE 2007 e do IDE 2011 do 5º ano do ensino fundamental



Fonte: Elaborado pelo autor.

Houve um aumento dessa dispersão e uma menor amplitude de 2011 em relação a 2007, pelo fato deles estarem mais em uma posição intermediária da desigualdade segundo o cálculo do IDE presente no gráfico 9 abaixo:

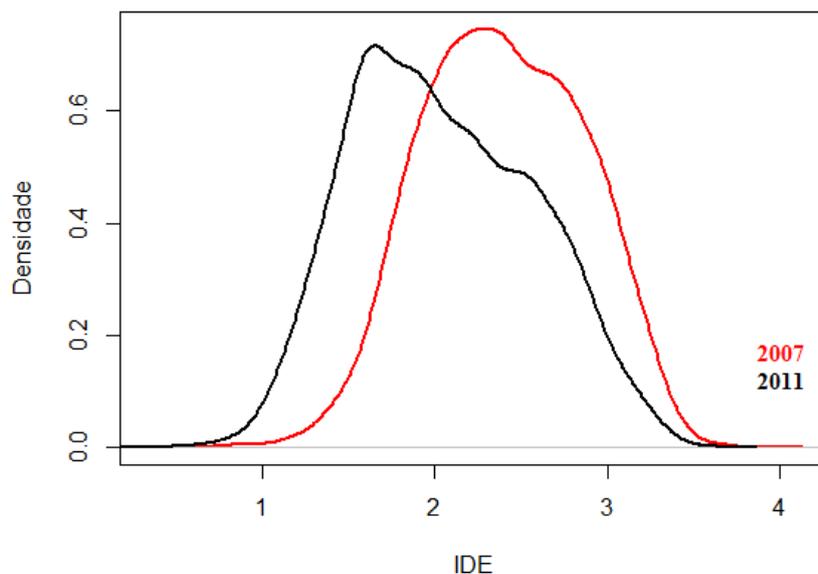
Gráfico 9: Histograma de leitura, 5º ano do ensino fundamental, IDE 2007 e IDE 2011



Fonte: Elaborado pelo autor.

No gráfico 10, a comparação entre o IDE de 2007 e 2011 para leitura no 5º ano do ensino fundamental, evidenciando uma melhoria educacional, apesar de a desigualdade também persistir, mesmo que com os pontos extremos menores.

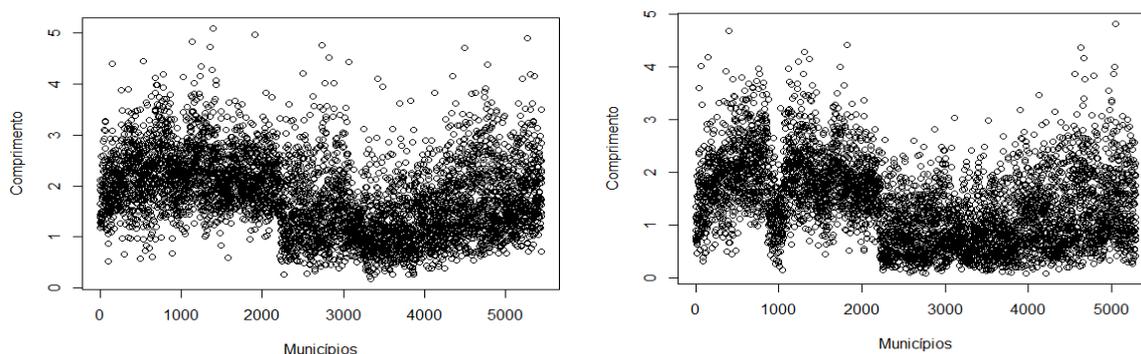
Gráfico 10: Comparação da Divergência Kullback-Leibler Discreta ou Índice de Desigualdade Educacional (IDE) para o 5º ano de leitura de 2007 e 2011



Fonte: Elaborado pelo autor.

Para matemática no 5º ano do ensino fundamental a distribuição KL apresenta um aumento na dispersão, ou seja, na diferença entre o melhor para o pior desempenho dos municípios brasileiros de 2007 em relação a 2011, o desvio padrão aumentou passando de 0,7249 para 0,7478, enquanto que o valor mínimo passou de 0,1253 para 0,08448 e o valor máximo passou de 4,7276 para 4,8277 ou seja, um aumento da diferença entre o valor mínimo e máximo, gerando por tanto um aumento na desigualdade de aprendizado de matemática neste ciclo.

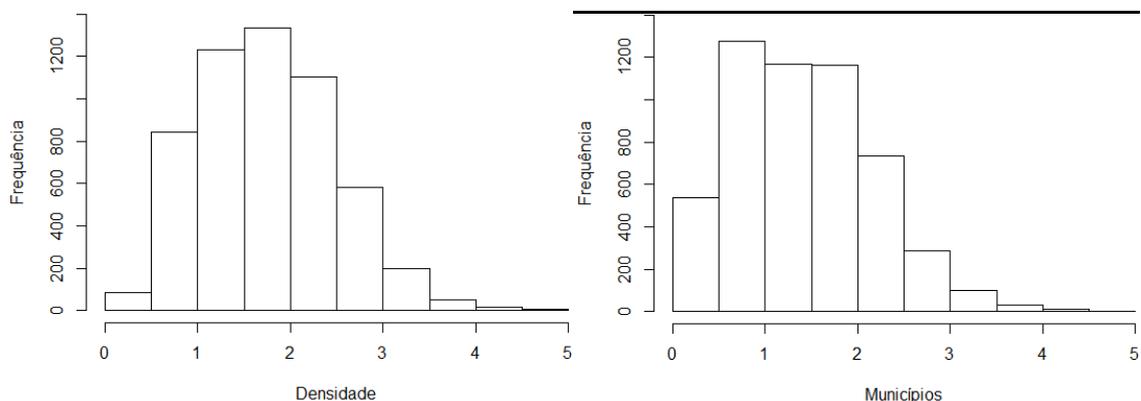
Gráfico 11: Dispersão de matemática do KL 2007 e KL 2011 do 5º ano do ensino fundamental



Fonte: Elaborado pelo autor.

Observa-se, apesar do aumento da desigualdade entre os municípios brasileiros para matemática no 5º ano, que os mesmos têm melhorado quanto ao desempenho na média, fato que a média passou de 1,7495 em 2007 para 1,4163 em 2011 e observa-se no gráfico 12, que os histogramas têm apresentado que mais municípios estão mais próximos da referência.

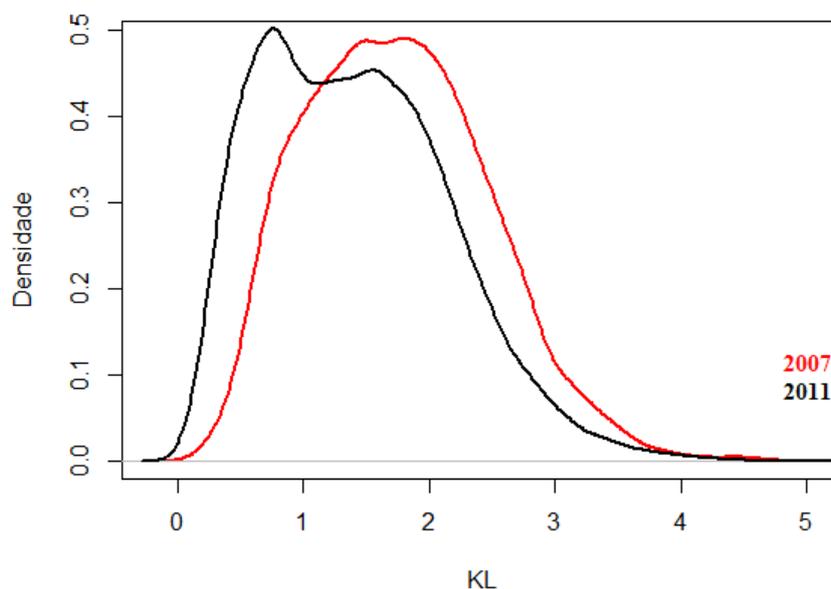
Gráfico 12: Histograma de matemática, KL 2007 e KL 2011 do 5º ano do ensino fundamental



Fonte: Elaborado pelo autor.

No gráfico 13, a comparação da divergência KL do 5º ano para matemática para 2011 evidencia o que já dito no histograma e na dispersão que a melhoria na KL média não se reflete na redução da desigualdade.

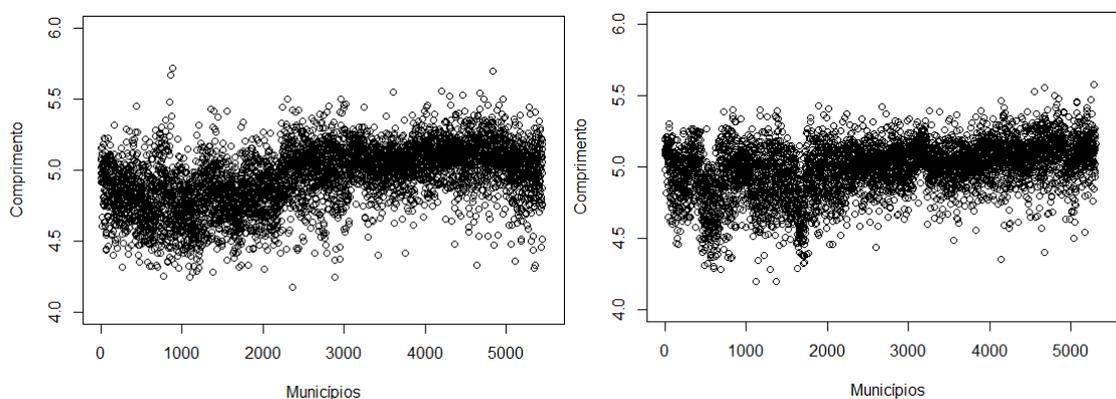
Gráfico 13: Comparação das Divergências KL de Matemática 5º ano para 2007 e 2011



Fonte: Elaborado pelo autor.

Quanto ao IDE, observa-se um comportamento contrário quanto à dispersão conforme o gráfico 14, pois o desvio padrão reduziu de 0,2123 para 0,1822 de 2007 para 2011 respectivamente. O valor mínimo passou de 4,1785 para 4,2 enquanto o valor máximo passou de 5,7142 para 5,5769, portanto uma redução da distância entre o pior desempenho com o melhor. Observa-se também uma piora quanto à situação do melhor município e uma relativa melhora do município com pior desempenho educacional.

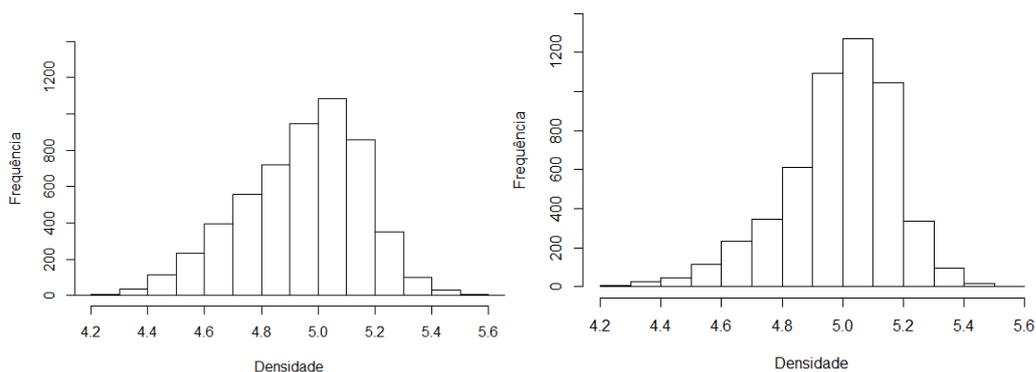
Gráfico 14: Dispersão de matemática do IDE 2007 e IDE 2011 do 5º ano do ensino fundamental



Fonte: Elaborado pelo autor.

Apesar da redução da distância entre o pior município e o melhor ocorreu uma piora segundo o IDE dos municípios com relação a matemática do 5º ano do ensino fundamental que pode ser visto pelo gráfico 15 nos histogramas e pelo fato de a média dos municípios ter passado de 4,9458 para 4,9886 de 2007 para 2011, um comportamento diferente da KL contínua.

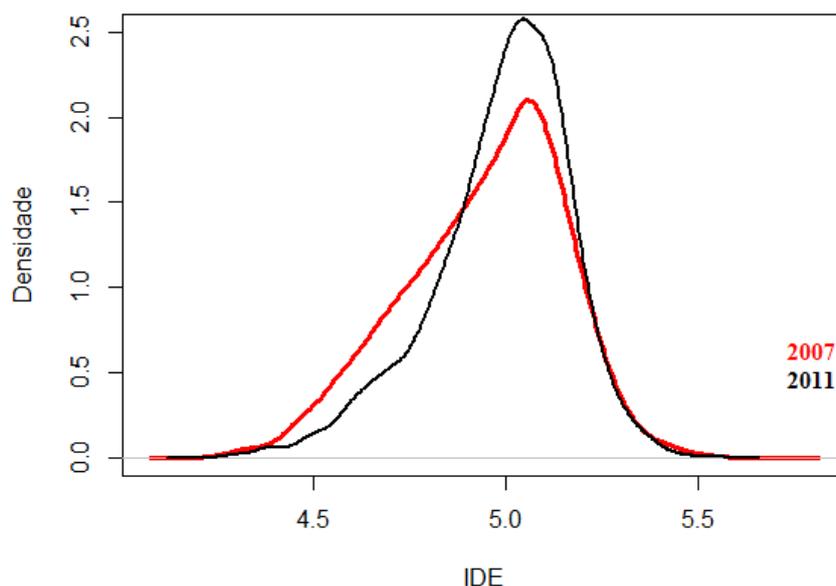
Gráfico 15: Histograma de matemática, 5º ano do Ensino Fundamental, IDE 2007 e IDE 2011



Fonte: Elaborado pelo autor.

O gráfico 16, possui o comparativo do IDE para 2007 e 2011 para matemática no 5º ano apresenta um comportamento de redução da desigualdade em detrimento da piora do desempenho médio do conjunto dos municípios brasileiros.

Gráfico 16: Comparação do IDE para matemática 5º ano do ensino fundamental de 2007 e 2011



Fonte: Elaborado pelo autor.

Já para o 9º ano do ensino fundamental, as distribuições apresentam uma redução na distância média observada para a distribuição de referência tanto para leitura, quanto para matemática de 2011 em relação a 2007. Observa-se também uma melhoria quanto à distância mínima observada em relação à distribuição de referência, mas um aumento da dispersão, o que evidencia que a desigualdade entre os municípios brasileiros segue aumentando tanto para leitura, quanto para matemática. Uma exceção ocorre para matemática na divergência KL de 2007 e 2011, em que se observou uma redução da dispersão, podendo afirmar baseado nessa metodologia que houve uma redução da desigualdade. Nesse caso, tabela 7:

Tabela 7: Estatística Descritiva dos dados da KL e do IDE de 2007 e 2011 para o 9º ano do ensino fundamental

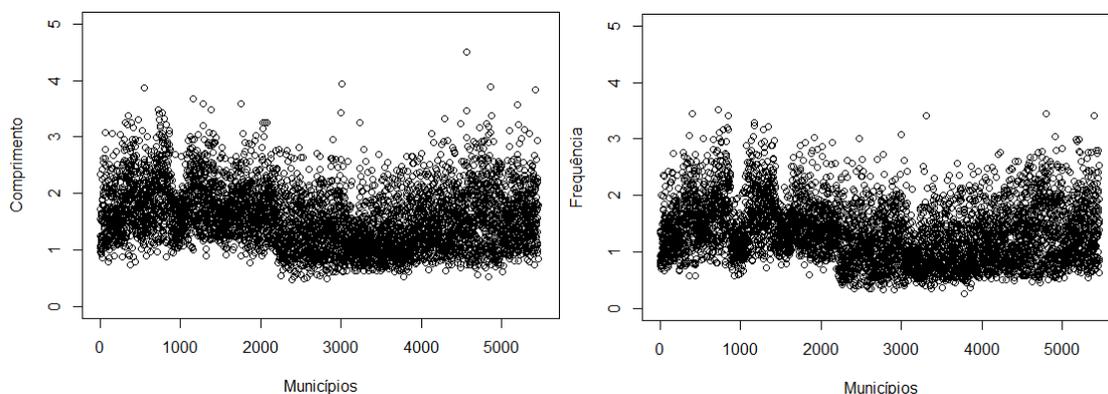
		Leitura			
		Média	Desvio padrão	Mínimo	Máximo
KL	2007	1,584627	0,5333506	0,474813	4,500448
	2011	1,351593	0,5245865	0,2750177	3,5148090
IDE	2007	2,439151	0,3436139	1,333333	3,684211
	2011	2,219658	0,363514	1,048	3,4642986
		Matemática			
		Média	Desvio padrão	Mínimo	Máximo
KL	2007	1,802011	0,6201057	0,3450177	4,8475251
	2011	1,575833	0,5933406	0,244055	4,405360
IDE	2007	2,67687	0,4090152	0,8461538	3,6923077
	2011	2,541169	0,4372935	0,7105263	3,8214286

Fonte: Elaborado pelo autor.

Para o 9º ano do ensino fundamental, a divergência KL apresenta uma redução da dispersão de 2007 para 2011 dos municípios brasileiros conforme o gráfico 17. O desvio padrão passou de 0,5333 para 0,5245 enquanto que o valor mínimo passou de

0,4748 para 0,275 e o valor máximo passou de 4,5 para 3,5148. Portanto ocorreu uma queda da desigualdade de aprendizado se baseado na KL.

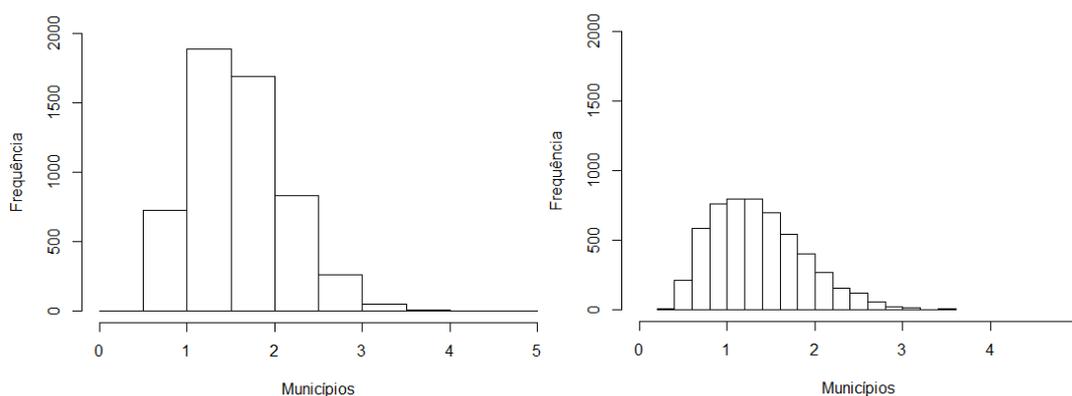
Gráfico 17: Dispersão de leitura do KL2007 e KL 2011 do 9º ano do ensino fundamental



Fonte: Elaborado pelo autor.

Essa queda é acompanhada da queda da média entre a KL dos anos de 2007 e 2011 passando de 1,802 para 1,5758. O comportamento do gráfico 18, mostra nos histogramas que os municípios estão mais próximos da distribuição de referência em 2011 em relação a 2007.

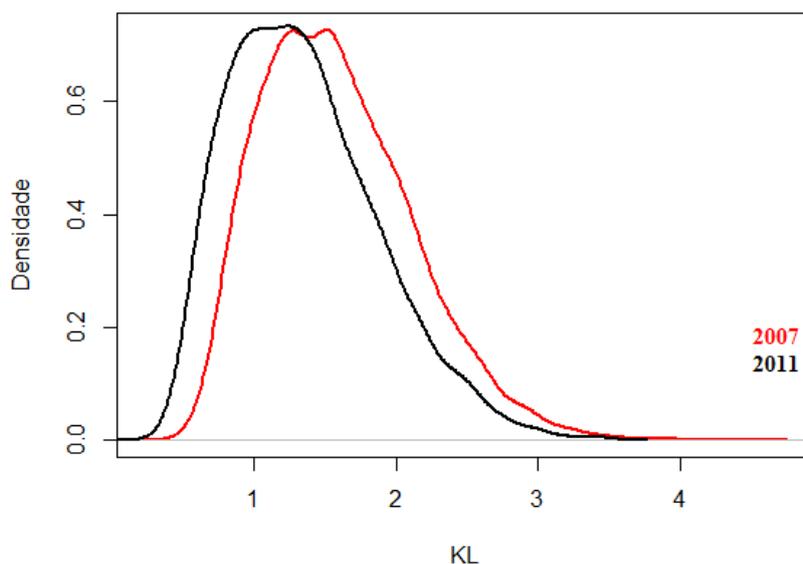
Gráfico 18: Histograma de leitura para o 9º ano do ensino fundamental para o KL 2007 e KL 2011



Fonte: Elaborado pelo autor.

Diferentemente da leitura para o 5º ano, a divergência KL apresentou uma redução de desigualdade, acompanhada de uma melhoria no desempenho médio da KL de 2007 para 2011, conforme o gráfico 19 (página a seguir):

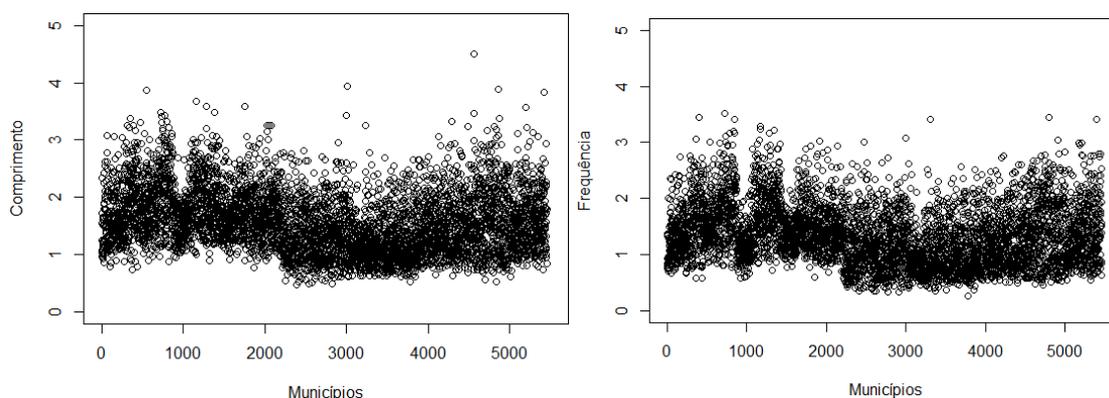
Gráfico 19: Comparativo da KL para leitura 9º ano de 2007 e 2011



Fonte: Elaborado pelos autores.

Para o IDE, observa-se um aumento na dispersão de 0,3436 em 2007 para 0,3635 em 2011, enquanto o valor mínimo passou de 1,333 para 1,048 e o valor máximo passou de 3,6842 para 3,4642. Apesar de o valor do melhor município está mais próximo da distribuição de referência, o município com pior desempenho obteve uma melhoria mais lenta, o que justifica o aumento da dispersão e por consequência da desigualdade.

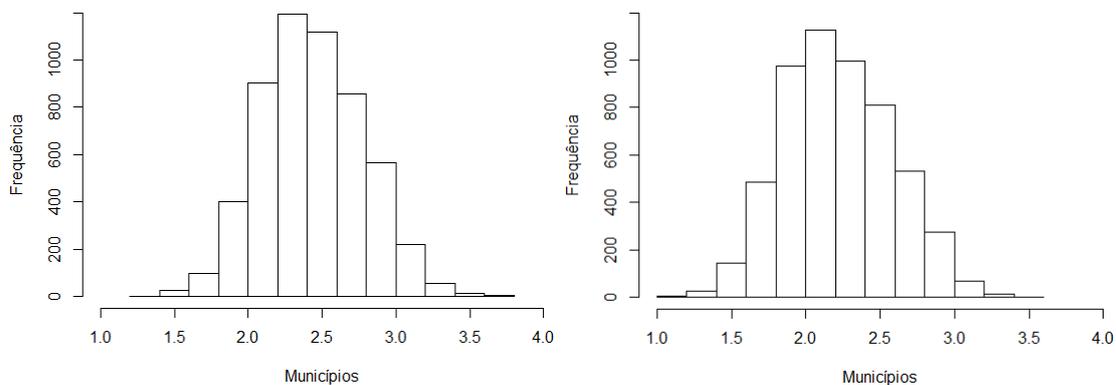
Gráfico 20: Dispersão de leitura para o 9º ano do ensino fundamental para o IDE 2007 e IDE 2011



Fonte: Elaborado pelo autor.

Para o 9º ano de matemática do ensino fundamental, observa-se uma melhoria no comportamento médio dos municípios brasileiros, pois o IDE passou de 2,4391 em 2007 para 2,2196 em 2011, o que mostra que apesar de ter ocorrido um aumento da desigualdade, no geral os municípios têm apresentado uma melhoria de desempenho.

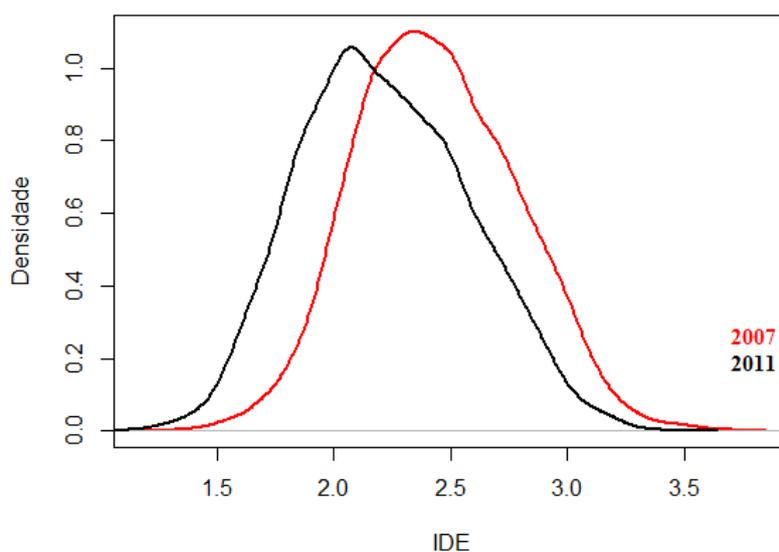
Gráfico 21: Histograma de leitura para o 9º ano do Ensino Fundamental para o IDE 2007 e IDE 2011



Fonte: Elaborado pelo autor.

O gráfico 22 representa IDE para o 9º ano de leitura comparando o desempenho da distância observadas pelos municípios de 2007 e 2011, assim como dito na explicação do histograma e da dispersão, observou-se uma melhoria, mas que não acompanhada da redução da desigualdade.

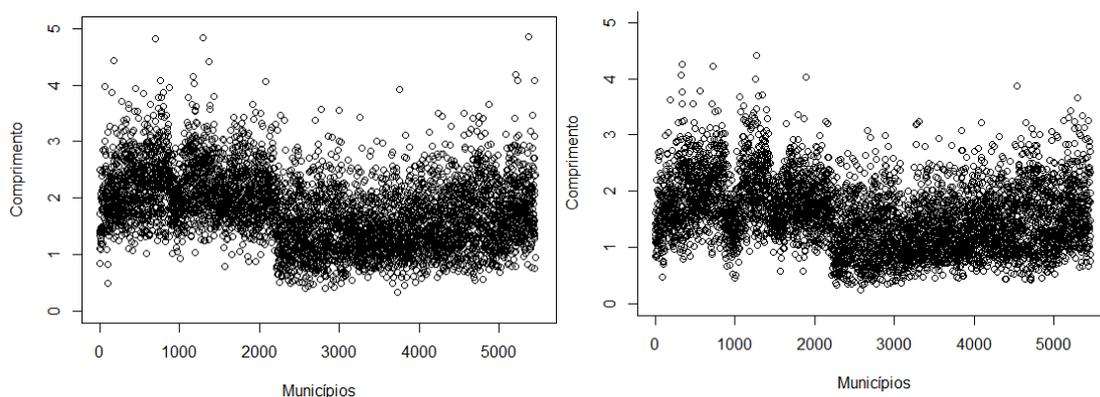
Gráfico 22: Comparação IDE para leitura, 9º ano do Ensino Fundamental para 2007 e 2011



Fonte: Elaborado pelo autor

Para a distribuição KL de matemática, observa-se uma diminuição da dispersão, pois o desvio padrão passou de 0,6201 em 2007 para 0,5933 em 2011, o valor mínimo passou de 0,3450 para 0,2440 e o valor máximo passou de 4,8475 para 4,4053, ou seja uma leve queda na distância entre o melhor município e o pior município. Isso leva-se a concluir que houve uma queda na desigualdade de aprendizado baseado na KL para matemática no 9º ano do ensino fundamental.

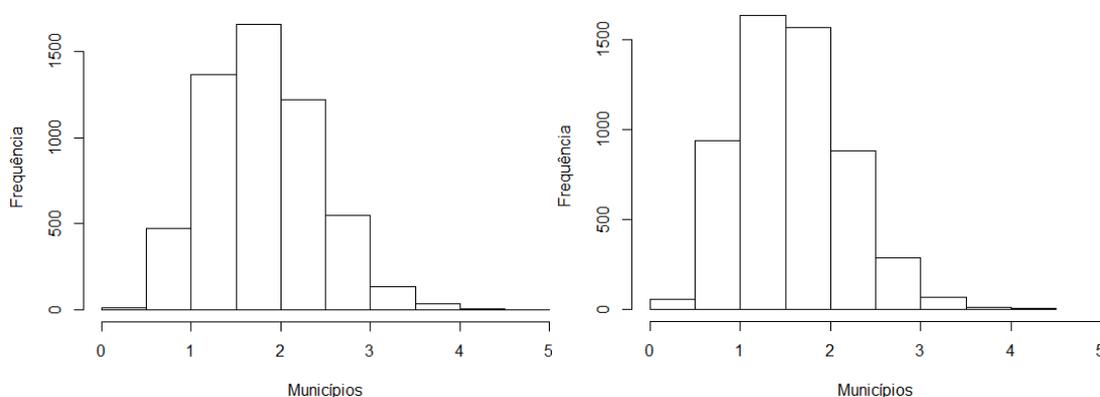
Gráfico 23: Dispersão de Matemática para o 9º ano do Ensino Fundamental para a KL 2007 e KL 2011



Fonte: Elaborado pelo autor.

O gráfico 24, apresenta, o histograma do KL para 2007 e 2011 para o 9º ano do ensino fundamental mostra que os municípios têm se aproximado da distribuição de referência, e a média da KL dos municípios brasileiros caiu de 2,6768 em 2007 para 2,5411 em 2011 evidenciando nessa situação uma redução da desigualdade de maneira geral, diferente das outras situações apresentadas acima.

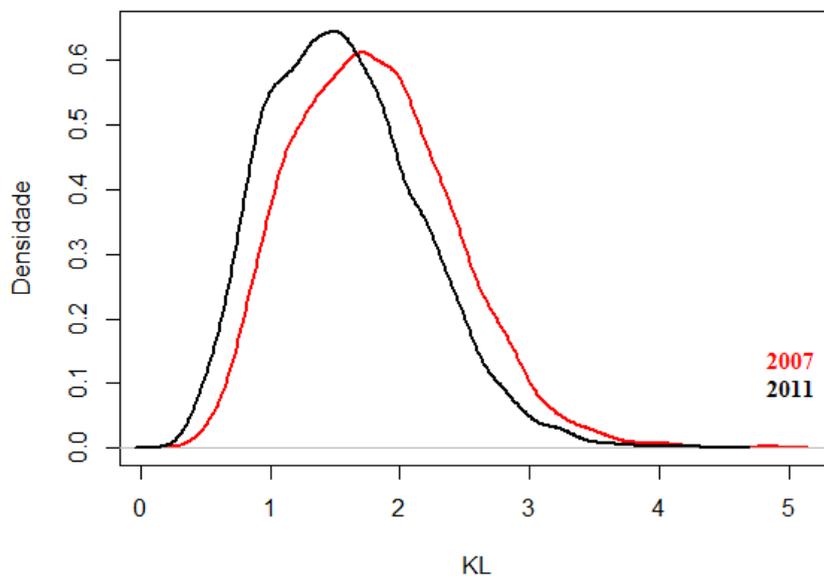
Gráfico 24: Histograma de Matemática para o 9º ano do Ensino Fundamental para a KL 2007 e KL 2011



Fonte: Elaborado pelo autor.

Diferentemente da leitura para o 5º ano, a divergência KL apresentou uma redução de desigualdade, acompanhada de uma melhoria no desempenho médio da KL de 2007 para 2011, conforme o gráfico 25:

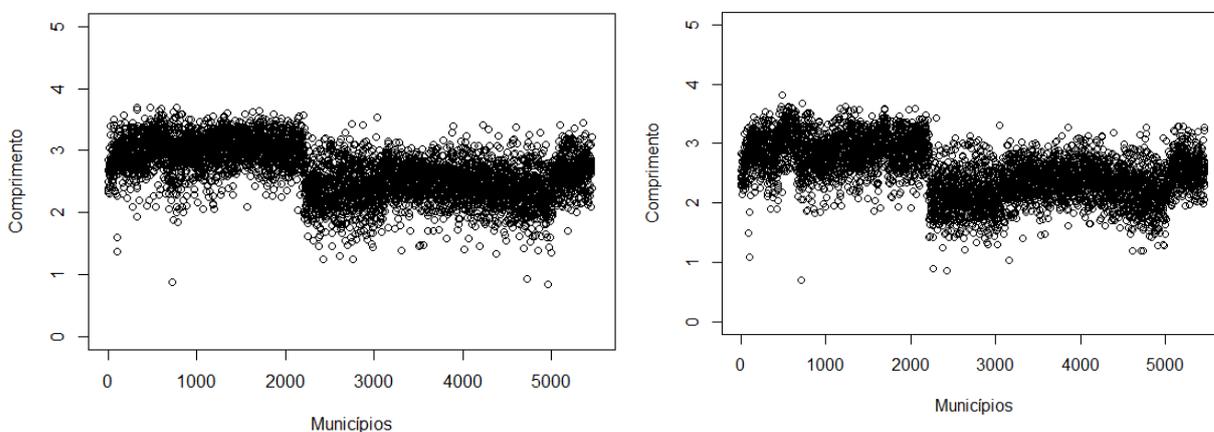
Gráfico 25: Comparação da KL para matemática 9º ano de 2007 e 2011



Fonte: Elaborado pelo autor.

Para o IDE do 9º ano do ensino fundamental de matemática, observa-se um efeito contrário, um aumento da dispersão com o desvio padrão passando de 0,4090 em 2007 para 0,4372 em 2011, uma queda no valor mínimo de 0,8461 para 0,7105 e um aumento no valor máximo, passando de 3,6923 para 3,8214 evidenciando um aumento da distância entre o melhor e o pior município brasileiro, ou seja, da desigualdade de aprendizado de matemática.

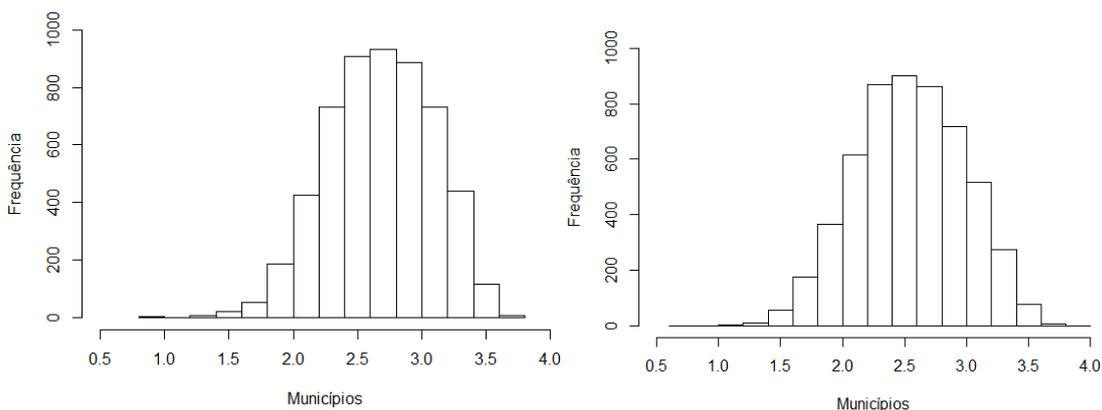
Gráfico 26: Dispersão de Matemática para o 9º ano do Ensino Fundamental para o IDE 2007 e IDE 2011



Fonte: Elaborado pelo autor.

Apesar de notada um aumento na desigualdade dos municípios, observou-se uma redução do IDE médio dos municípios brasileiros passando de 2,6768 em 2007 para 2,5411 em 2011, ou seja, apesar do aumento da dispersão houve uma redução média da desigualdade entre os municípios brasileiros conforme também está demonstrado no gráfico 27 do histograma do IDE.

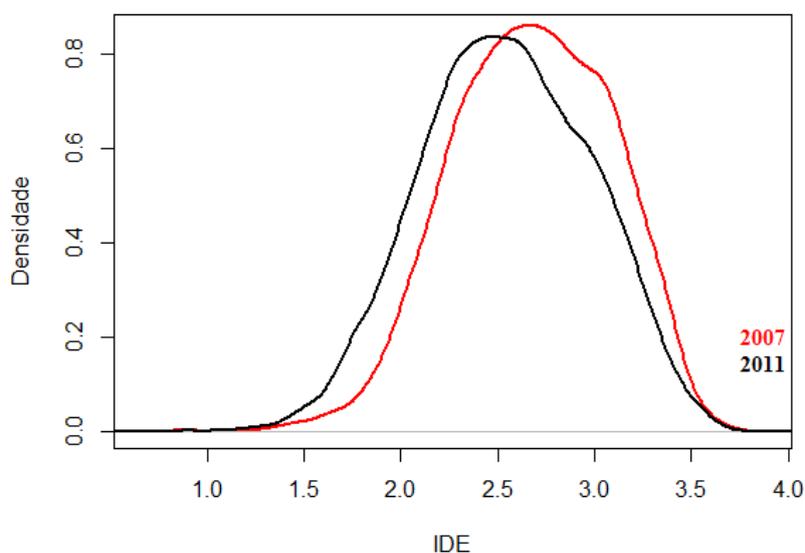
Gráfico 27: Histograma de Matemática para o 9º ano do Ensino Fundamental para o IDE 2007 e IDE 2011



Fonte: Elaborado pelo autor

O gráfico 28 representa IDE para o 9º ano de matemática comparando o desempenho da distância observadas pelos municípios de 2007 e 2011, assim como dito na explicação do histograma e da dispersão, onde se observou uma melhoria, que porém não acompanhada da redução da desigualdade.

Gráfico 28: Comparativo do IDE de Matemática 9º ano do Ensino Fundamental de 2007 e 2011



Fonte: Elaborado pelo autor.

As tabelas e gráficos acima evidenciaram que tanto para a KL, quanto para o IDE, os melhores municípios têm reduzido cada vez mais a distância observada para a distância de referência aumentando sua diferença para os piores municípios que evoluem a níveis menores. Isso se reflete quando se agrupa os resultados para as regiões brasileiras como as utilizadas por Soares e Delgado (2016), temos a tabela 8, que denuncia o do jeito que os resultados se apresentam a cada exame. A continuar esse passo, a diferença entre as regiões com os melhores indicadores em relação aos piores se manterá ao longo das próximas décadas.

Tabela 8: Valores da distância observada até a distribuição de referência dos grupos definidos por região da Prova Brasil do 5º ano do Ensino fundamental de 2007 e 2011

Região/disciplina	2007	2011
Leitura – Norte	1,28	0,92
Leitura – Nordeste	1,40	1,04
Leitura – Sudeste	0,78	0,38
Leitura – Sul	0,82	0,40
Leitura - Centro-Oeste	0,94	0,48
Matemática – Norte	1,67	1,20
Matemática – Nordeste	1,73	1,31
Matemática – Sudeste	0,88	0,42
Matemática – Sul	0,87	0,42
Matemática – Centro-Oeste	1,12	0,61

Fonte: (SOARES e DELGADO, 2016).

5 CONCLUSÃO

Este trabalho teve como objetivo foi diagnosticar o nível de desigualdade educacional dos municípios brasileiros no 5º e do 9º ano do ensino fundamental brasileiro. Para construção da argumentação, foi necessário referenciar os principais medidores de desigualdade de renda como o Índice de Gini, o Índice de Theil e a Curva de Lorenz. Além disso, explicar o quanto as questões internas, ou seja, infraestrutura da escola, questões pedagógicas impactam no desempenho dos estudantes. O resultado nas discussões teóricas foram que o efeito é limitado, mas não menos importante. A família possui influência muito maior do que deveria ter no rendimento escolar.

Os indicadores de desigualdade apresentados neste trabalho mostram o quão complexo é a situação educacional brasileira, demonstrando que apesar dos importantes avanços institucionais ocorridos, o desafio educacional requer dos gestores melhores práticas em políticas públicas, com focalização no ponto central que é a redução dos fatores externos a escola dentro do aprendizado.

A divergência KL, foi de importante aprendizado por se apresentar como uma das tantas ferramentas estatísticas que podem ajudar a desenvolver importantes trabalhos em todas as áreas de conhecimento. Sem essa teoria, dificilmente acompanharia o comportamento da desigualdade educacional de 2007 e 2011 de maneira tão robusta quanto foi neste trabalho, além de ter apresentado um campo de estudo pouco citado dentro da sala de aula, que é a questão educacional e seus dilemas como a desigualdade. Um debate que vem ganhando cada vez mais força no meio acadêmico e pedagógico pela importância do tema que impacta no dia a dia da sociedade.

Ou seja, os resultados apresentados e opinião de importantes estudiosos do tema mostram que o dilema educacional brasileiro vai além do que é usado hoje como política pública educacional e se demonstra necessário uma discussão de medidas que busquem equalizar o aprendizado para todas as crianças e jovens do Brasil, pois mantendo do jeito que está atualmente, dificilmente o país conseguirá o tão sonhado salto educacional.

Os resultados indicam um leve aumento na desigualdade dos municípios de 2011 em relação a 2007. Isso evidencia o desafio que os gestores públicos têm para se buscar a educação inclusiva em que os alunos obtenham o aprendizado condizente com suas capacidades

6 REFERÊNCIAS BIBLIOGRÁFICAS

ALBERNAZ, Ângela; FERREIRA, Francisco H. G.; FRANCO, Creso. **Qualidade e equidade no ensino fundamental brasileiro.** Disponível em <<http://ppe.ipea.gov.br/index.php/ppe/article/viewFile/139/74>>. Acesso em 17/06/2017.

ALVES, M.T.G.; SOARES, J.F.; XAVIER, F.P. **Índice socioeconômico das escolas de educação básica brasileiras.** Ensaio: aval. pol. públ. Educ., Rio de Janeiro, v.22, n. 84, p. 671-704, jul./set. 2014.

ALVES, M.T.G.; XAVIER, F.P. **Construção de Indicadores para descrever a desigualdade de aprendizado na Prova Brasil.** Est. Aval. Educ., São Paulo, v. 27, n. 66, p. 782-815, set./dez. 2016.

ANDRADE, Josemberg M. de; LAROS, Jacob A.. **Fatores associados ao Desempenho Escolar:** Estudo Multinível com Dados do SAEB/2001. Psic.: Teor. e Pesq., Brasília, Jan-Mar 2007, Vol. 23 n. 1, pp. 033-042. Disponível em <<http://www.scielo.br/pdf/ptp/v23n1/a05v23n1.pdf>>. Acesso em 17/06/2017.

ARAÚJO, Rayane Carmo. **Determinantes do Aprendizado no Brasil:** uma análise das funções de produção de educação. 62 f. Monografia de Bacharelado – Faculdade de Economia da Universidade Federal Fluminense, 2013. Disponível em <<http://www.proac.uff.br/cede/sites/default/files/Monografia-RayanneAraujo.pdf>>. Acesso em 18/06/2017.

BRASIL. Lei nº 13.005, de 25 de junho de 2014. **Aprova o Plano Nacional de Educação – PNE e dá as providências.** Disponível em <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/113005.htm>. Acesso em 25/05/2017.

CAMINHOS PARA MELHORAR O APRENDIZADO. **O que importa para o aprendizado escolar. Qualidade da infraestrutura.** Disponível em <<http://www.paramelhoraroaprendizado.org.br/Conteudo/verbete.aspx?canal=20100701145550501160&subtema=20110419134526287427&verbete=20110419154657163205>>. Acesso em 15/06/2017.

CENTRO DE POLÍTICAS PÚBLICAS DO INSPER. **O atraso educacional brasileiro numa perspectiva histórica.** Disponível em <<https://www.insper.edu.br/blogdocpp/artigo-o-atraso-educacional-brasileiro-numa-perspectiva-historica/>>. Acesso em 13/08/17.

CENTRO DE POLÍTICAS PÚBLICAS DO INSPER. Artigo: **Desenvolvimento infantil e desigualdade.** Disponível em <<https://www.insper.edu.br/blogdocpp/artigo-desenvolvimento-infantil-e-desigualdade/>>. Acesso em 17/08/17.

CÉSAR, C.C.; SOARES, J.F. **Desigualdades acadêmicas induzidas pelo contexto escolar.** Revista Brasileira de Estudos de População, v.18, n. ½, jan/dez. 2001.

DELGADO, V.M.S.; MIRANDA-RIBEIRO, A.; SOARES J.F. **Desigualdade Escolar e Desempenho.** In: FAHEL, M. et al. (Org.). **Desigualdades educacionais e pobreza.** Belo Horizonte: PUC Minas, 2013. p. 163-208.

DEPARTAMENTO DE FINANCIAMENTO DA EDUCAÇÃO BÁSICA /FUNDEF:
O QUE É FUNDEF. Disponível em: <
<http://mecsrv04.mec.gov.br/sef/fundef/funf.shtm>>. Acesso em 25/05/2017.

ESQUEF, Israel Andrade; ALBUQUERQUE, Márcio Portes de; ALBUQUERQUE, Marcelo Portes. **Fundamentos de Teoria da Informação.** Disponível em <<http://www.cbpf.br/cat/pdsi/pdf/cap2webfinal.pdf>>. Acesso em 02/08/2017.

FILHO, N.M. **Como melhorar o aprendizado nas escolas.** Centro de Políticas Públicas do Insper, 2017. Disponível em <<https://www.insper.edu.br/blogdocpp/artigo-como-melhorar-o-aprendizado-nas-escolas/>>. Acesso em 17/08/17.

FUNDAÇÃO VICTOR CIVITA. **Análise das desigualdades intraescolares no Brasil.** São Paulo: Estudos e Pesquisas Educacionais, 2013.

HANDCOCK, Mark S.; MORRIS, Martina. **Relative distribution methods.** *Sociological Methodology*, v. 28, n. 1, p. 53-97, 1998.

HANDCOCK, Mark S.; MORRIS, Martina. **Relative distribution methods in the social sciences.** Springer Science & Business Media, 2006.

HOFFMAN, RODOLFO (1988) **Distribuição de Renda, Medidas de Desigualdade e Pobreza,** São Paulo. Editora da Universidade de São Paulo. 1998, Capítulo 3.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo Escolar.** Disponível em <<http://inep.gov.br/web/guest/censo-escolar>>. Acesso em 26/07/2017.

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Remuneração dos docentes em exercício na educação básica: um pareamento das dbases de dados do Censo Escolar e da RAIS.** Brasília, Junho de 2017.

INEP – Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Saeb.** Disponível em <<http://inep.gov.br/web/guest/educacao-basica/saeb>>. Acesso em 25/07/2017.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Sobre o INEP.** Disponível em <<http://inep.gov.br/web/guest/sobre-o-inep>>. Acesso em 25/05/2017.

INTECH OPEN. **Perfomance assessment:** Evaluation of features by Kullback-Leibler Divergence. Disponível em <<https://www.intechopen.com/books/brain-computer-interface-systems-recent-progress-and-future-prospects/optimal-fractal-feature-and-neural-network-eeg-based-bci-applications>>. Acesso em 31/07/2017.

KULLBACK, S.; LEIBLER R.A. (1951). **On information and sufficiency.** *Annals of Mathematical Statistics*, 22(1) 79-86.

KLEIN, R. **Utilização da Teoria de Resposta ao Item no Sistema Nacional de Avaliação da Educação Básica (SAEB).** *Revista Ensaio*, n.40, v.11, p. 283-296, jul./set. 2003.

LERMAN, R.I.; YITZHAKI S. **A note on the calculation and interpretation of the Gini Index.** Elsevier Science Publishers B.V. (North-Holland) *Economics Letters* 15 (1984) 363-368.

NETO, J.J.S; JESUS, G.R. de; KARINO, C.A.; ANDRADE, D.F de. **Uma escala para medir a infraestrutura escolar**. Est. Aval. Educ., São Paulo, v. 24, n. 54, p. 78-99, jan./abr. 2013.

OSHIMA, F.Y.; EVELIN, G. Ricardo Paes de Barros: **“A crise da educação é mais grave do que a da pobreza”**. Revista Época, 07/08/2015. Disponível <<http://epoca.globo.com/ideias/noticia/2015/08/ricardo-paes-de-barros-crise-da-educacao-e-mais-grave-do-que-da-pobreza.html>>. Acesso em 13/08/17.

PAVÃO, Rodrigo. **Teoria da Informação**. Disponível em <http://www.ib.usp.br/~rpavao/Teoria_da_Informacao.pdf>. Acesso em 31/07/2017.

PORTAL FNDE. **Sobre o Fundeb. O que é?** Disponível em: <<http://www.fnde.gov.br/financiamento/fundeb/fundeb-apresentacao%20-%20Acesso%20em%2025/05/2017%20%C3%A1s%2010:02>>. Acesso em 25/05/2017.

PORTAL DO PNE. **Plano Nacional de Educação (2014/2024) em movimento**. Disponível em <<http://pne.mec.gov.br/>>. Acesso em 25/05/2017.

PROVA BRASIL, SAEB. **Escala de Proficiência**. Disponível em <http://download.inep.gov.br/educacao_basica/prova_brasil_saeb/escala/escala_proficiencia/2013/escalas_ensino_fundamental_2013.pdf>. Acesso em 29/07/2017.

PROVA BRASIL, INEP. **Matrizes e Escalas**. Disponível em <<http://provabrasil.inep.gov.br/web/guest/educacao-basica/saeb/matrizes-e-escalas>>. Acesso em 28/07/2017.

SANTOS, J.C. **Porque Parou? Investigações sobre a recente estagnação da proficiência de Minas Gerais durante o período 2010-2012**. 48f. Monografia de Bacharelado. Departamento de Ciências Econômicas da Universidade Federal de Ouro Preto, 2014.

SOARES, J.F.; DELGADO, V.M.S. **Medida das desigualdades de aprendizado entre estudantes do ensino fundamental**, 2016.

THE WORLD BANK. **Databank** Disponível em <<http://data.worldbank.org/topic/education>>. Acesso em 13/08/17.

VALOR ECONÔMICO. **Alfabetização entre crianças ricas é ate 6 vezes maior que pobres**, 2017