Medidas comparativas do desempenho e da desigualdade nos exames de proficiência escolar em Minas Gerais por índices unidimensionais[1]

Victor Maia Senna Delgado
Centro de Estudos de Políticas Públicas – Fundação João Pinheiro-MG
e-mail: victor.maia@fjp.mg.gov.br
Adriana de Miranda-Ribeiro
Centro de Estatística e Informações – Fundação João Pinheiro-MG

RESUMO

Este artigo trata da medição conjunta do *desempenho* e da *desigualdade* nos exames de proficiência do PROEB (Programa de Avaliação da Educação Básica) de Minas Gerais. Atualmente, a educação passa por um momento importante no Brasil, com uma elevada, e cada vez mais permanente, atenção social para o problema. Desde 2006, com o movimento "Todos Pela Educação", metas de desempenho para a média dos exames de proficiência foram propostas para o país inteiro, ocorrendo também adesão de estados e municípios. Muitos estados adotam hoje seus próprios exames de proficiência em escalas e padrões similares àqueles presentes no SAEB (Sistema Nacional Educação Básica). Entretanto, a adoção de metas de desempenho por indicadores de tendência central podem mascarar efeitos não desejados na desigualdade entre os alunos, sendo necessária a proposição de metas que levem em conta a desigualdade do sistema. O artigo analisa três propostas de indicadores unidimensionais para os exames de proficiência, duas delas estão presentes na literatura e a terceira é derivada das duas primeiras e visa a tratar melhor os componentes de *desempenho* e *desigualdade* presente nos dados, fornecendo novas interpretações sobre índices ponderados de desempenho.

JEL: C43, I24, I21
Palavras-chave: Educação, Índices, Densidades Relativas, Proficiência, Desigualdade, PROEB-MG.
Tema 2: Economia Mineira

ABSTRACT

This article deals with the joint measurement of *performance* and *inequality* in the PROEB proficiency tests of Minas Gerais (PROEB stands for *Programa de Avaliação da Educação Básica*). Currently, education is going through an important chapter in Brazil, with a high continuous and increasingly attention to the problem. Since 2006, the movement "*Todos Pela Educação*" (*All For Education*) proposes, for the entire country, performance targets to average proficiency, this also brings states and municipalities to this targets. Many states now have adopted their own proficiency tests in SAEB scales (*Sistema Nacional Educação Básica*). However, the adoption of targets for indicators of central tendency may mask unwanted effects on inequality among students, being necessary to propose goals that take into account the inequality of the system. The article examines three proposals for one-dimensional proficiency indicators, two of them are found in the literature about this topic in Brazil and the third one is derived from the first two and seeks to better treat the components of *performance* and *inequality* present in data, giving new interpretations to performance weighted indices.

---

[1] Os autores agradecem a José Francisco Soares que colaborou com várias das ideias apresentadas neste trabalho e por comentários realizados sobre as versões finais, ao professor Eduardo Luiz G. Rios-Neto pelos comentários sobre o tema e à Clarissa Rodrigues, pesquisadora do IIASA *International Institute for Applied System Analysis* que auxiliou com dados do PISA internacional. Agradecemos também à secretaria de educação de Minas Gerais (SEE-MG) pelos trabalhos desenvolvidos no âmbito do Atlas da Educação de Minas Gerais. No entanto, nos responsabilizamos integralmente por erros e omissões que por ventura possam ter sido cometidos.

INTRODUÇÃO

No campo da educação, um problema importante que ainda resta ser resolvido trata da desigualdade de desempenho entre alunos. Uma educação desigual não oferece oportunidade a todos e pode ainda dar a falsa impressão de que as metas estão sendo cumpridas. Além disso, é conhecido o fato de que a educação brasileira é bastante desigual, tanto no que tange ao acesso dos alunos ao sistema público e privado, quanto no grau de aquisição cognitiva, medida pelos exames. Mesmo dentro da esfera pública, há desempenhos muito díspares nos testes de aprendizado. Esse artigo trata do tema da desigualdade educacional nessas medidas de desempenho dos alunos, em particular, resultados do PROEB-MG (Programa de Avaliação da Educação Básica de Minas Gerais), mostrando que medidas centrais (índices que não levam em conta medidas de dispersão e distribuição de notas) não necessariamente atendem os critérios de maior igualdade de oportunidades educacionais.

O Brasil vive um novo momento de atenção social para o tema da educação, a população está atenta para o problema e levanta questões que revelam, cada vez mais, um aprofundamento do conhecimento sobre o quadro educacional do país. Sob uma perspectiva normativa, o documento "*Todos pela Educação*" estabeleceu cinco metas para as quais o sistema educacional brasileiro precisa mirar para alcançar uma maior qualidade do ensino até o ano de 2022.[2] Existe um esforço para que todas essas metas possam ser bem medidas e cumpridas.

Neste artigo nos concentraremos sobre uma questão relacionada à terceira meta (toda criança com nível adequado de aprendizado) e os indicadores de desempenho a ela relacionados. Abordaremos também o problema de se desconsiderar a desigualdade ao lidar com medidas centrais. Existem diversas maneiras distintas de se alcançar a terceira meta, uma delas é ter todos os alunos com a nota de proficiência considerada adequada (todos os alunos com proficiência = 200 para leitura na 5ª série, por exemplo), outra maneira de se alcançar a mesma meta é obter todos os alunos com o nível avançado (todos com proficiência em exatos 250 em leitura 5ª série). Apesar de essas duas maneiras atenderem igualmente o critério da igualdade absoluta, todos com o mesmo desempenho e sem nenhuma desigualdade, a segunda forma é socialmente mais desejável, dado que a mesma igualdade é conseguida com um maior nível de desempenho.

O desafio é lidar com os dois objetivos ao mesmo tempo: *desempenho* e *desigualdade*. Ao contrário do postulado econômico, que estabelece um *trade-off* entre *equidade* e *eficiência* na **fronteira de produção**, Cunha, Heckman, Lochner e Masterov (2006) afirmam que para a educação fundamental, e particularmente para o ensino infantil, não há nada que diga não ser possível alcançar os dois objetivos simultaneamente. Para a teoria da produção econômica, só é possível alcançar maior eficiência e equidade em conjunto no caso em que a economia se encontra em um ponto **ineficiente de Pareto**, interno à fronteira de possibilidades de produção (figura 1). Nos casos de eficiência, ocorreria o *trade-off*, no qual uma redistribuição de recursos em prol da igualdade limitaria a fronteira máxima da possibilidades de produção.

Cunha et. al. (2006) estabelecem que no caso da educação esse *trade-off* não existe nos anos iniciais. A lógica por trás disso é de que para os anos iniciais da educação infantil, e também nos primeiros anos do nível fundamental, vale à pena investir em **todas** as crianças (mesmo as que estão atrasadas), pois é possível lhes garantir um bom desempenho educacional e com pouco esforço é possível equipara-las no desenvolvimento das habilidades. Para idades avançadas, esse ganho do investimento se torna cada vez menor e é mais eficiente investir nos alunos que já possuem uma formação prévia.

[2] As cinco metas são: 1) Toda Criança na Escola; 2) Toda criança plenamente alfabetizada até os 8 anos; 3) Todo aluno com nível de aprendizado considerado adequado para a série que cursa; 4) Todo jovem com ensino médio concluído até os 19 anos; e 5) Investimentos em Educação ampliados e bem geridos. Todos pela Educação: Nota técnica (2007).

Quando se mira o aumento de *desempenho* somente, pode existir um indesejável efeito onde o nível médio de *desempenho* mais elevado surge à custa de uma maior *desigualdade*. É possível melhorar o *desempenho* geral aumentando-se o conhecimento dos mais capazes, deixando os alunos que não adquiriram conhecimento prévio de lado e gerando *desigualdade*. Esse comportamento gerador de desigualdades pode ser contornado por um processo de aprendizado e por meio de um preciso acompanhamento das medidas centrais. Tem-se ainda que ao se voltar exclusivamente para a *igualdade* de desempenho, pode não se permitir o desenvolvimento livre das múltiplas habilidades individuais. No caso dos avaliadores de políticas educacionais, que precisam ter esses dois componentes em mente, o desafio é propor um indicador unidimensional (sobre uma variável y, a escala de proficiência) que una e capte tanto melhoras no *desempenho* quanto reduções na *desigualdade*.

O artigo apresenta três indicadores unidimensionais para lidar com o problema de *desempenho* e *igualdade*. O primeiro deles é o *IQE* (*Índice de Qualidade da Educação*), já proposto por Soares e Marotta (2009), usado como base para o Índice de desenvolvimento da Educação de São Paulo (IDESP) e como componente do IMRS (Índice Mineiro de Responsabilidade Social). O segundo índice é o *IPD* (*Índice Ponderado de Desempenho*), proposto inicialmente por Soares, Marotta e Delgado (2010) e derivado das técnicas de densidade relativa (Handcock e Morris, 1998; Rodrigues, 2009). O terceiro índice generaliza os dois primeiros e foi batizado de *IDR* (*Índice de Desempenho Relativo*), baseia-se na literatura de densidade relativa e possui relação com os índices de entropia de Theil (1967) e a divergência de Kullback e Leibler (1951). Os três índices tratam em conjunto os problemas de *desempenho* e *desigualdade* e podem ser encarados como índices derivados de indicadores com médias ponderadas ou em termos da densidade relativa, sendo importante a escolha de uma ponderação padrão que servirá de referência.

Acreditamos que a proposição dos três índices acrescenta também possibilidade de novas interpretações sobre o problema do *desempenho* e *desigualdade* na educação, por isso, maiores detalhes da motivação e do modelo conceitual serão tratados nas seções 2 e 3. A metodologia de construção dos indicadores é apresentada na 4ª seção. Os resultados dos indicadores com uso de dados do PROEB-MG são apresentados na seção 5, há uma crescente elevação do *desempenho* dos alunos da 5ª série, mas com um uma consequente elevação da *desigualdade*. A última e sexta seção também apresenta as possíveis implicações do trabalho.

2. MOTIVAÇÃO

Educação e Desigualdade são temas geralmente relacionados por pesquisadores da educação, mas apesar de serem muito discutidos, pouco se faz para medir a desigualdade educacional por meio dos indicadores de desempenho e pouco se fala da desigualdade intrínseca e interna ao sistema educacional, que deve ser amenizada por um sistema de ensino eficaz. O que se sabe é que há uma desigualdade na aquisição de anos de estudo (Thomas et al, 2001), porém para saber porque ela ocorre, é interessante estudar a desigualdade de desempenho e aquisição de educação pelo aluno, sendo importante perguntar porque as escolas brasileiras fracassam em oferecer um aprendizado mínimo a certa de 20 a 30% dos estudantes (Barbosa, 2009). Reconhece-se também que o fraco desempenho, o atraso e a repetência são as principais causas para evasão e abandono e, dessa forma, a desigualdade nos anos de estudo (Golgher, 2003; e Riani et. al., 2010).

Assim como na distribuição de renda, o Brasil possui uma apropriação dos conhecimentos escolares que é desigual. Dados do exame internacional do PISA, 2009, (*Programm for International Student Assessment*), presentes nas tabelas 1 e 2 abaixo, mostram que o desempenho do Brasil, além de estar entre um dos mais baixos obtidos no exame, é o mais desigual em comparação com EUA, México, Portugal e Finlândia, tomando como base o índice de Gini. Além disso, o país possui os porcentuais mais elevados de alunos abaixo da escala de nível 1 de proficiência do PISA (nota 359 para matemática e 335 para leitura). Isso quer dizer que 39% dos alunos brasileiros de 15 anos submetidos ao exame de matemática se situam no primeiro nível de competências na matéria. No caso dos exames de leitura o porcentual é menor: 22%, mas ainda o maior entre os países comparados.

Tabela 1 - Resultados do PISA-Matemática para países comparados

Pais	*Proficiência*	*Desvio-padrão*	*Gini*	***Porcentagem de alunos no Nível 1***
Finlândia	540.50	76.53	0.080	0.01
EUA	487.40	86.36	0.101	0.07
Portugal	486.89	86.78	0.101	0.08
Média OCDE	***488.40***	***93.18***	***0.109***	***0.09***
México	418.51	75.03	0.101	0.21
Brasil	**385.81**	**77.75**	**0.113**	**0.39**

Fonte: *PISA-OCDE*, 2009.

Tabela 2 - Resultados do PISA-Leitura para países comparados

Pais	*Proficiência*	*Desvio-padrão*	*Gini*	***Porcentagem de alunos no Nível 1***
Finlândia	535.88	83.01	0.087	0.01
EUA	499.83	93.54	0.107	0.05
Portugal	489.54	83.72	0.097	0.05
Média OCDE	***491.55***	***94.66***	***0.109***	***0.05***
México	425.27	81.44	0.108	0.15
Brasil	**411.76**	**90.75**	**0.125**	**0.22**

Fonte: *PISA-OCDE*, 2009.

No exame do PROEB em Minas Gerais, para a série de anos de 2006 a 2009, podemos observar uma elevação da média de desempenho para todas as séries, tanto para matemática quanto para leitura Gráfico 1. A evolução no ensino médio ainda é muito tímida enquanto que para 5ª série (antiga 4ª série do fundamental) essa evolução é mais nítida.

Como veremos nas seções 2 e 3 a seguir, há diversas maneiras de se mensurar a desigualdade sendo importante trabalhar as propriedades desejáveis desses indicadores. Um dos indicadores de desigualdade mais amplamente empregado é o índice de Gini, proposto pelo estatístico italiano Corrado Gini (1926).[3] Na tabela 3, vemos a evolução da proficiência do PROEB-MG.

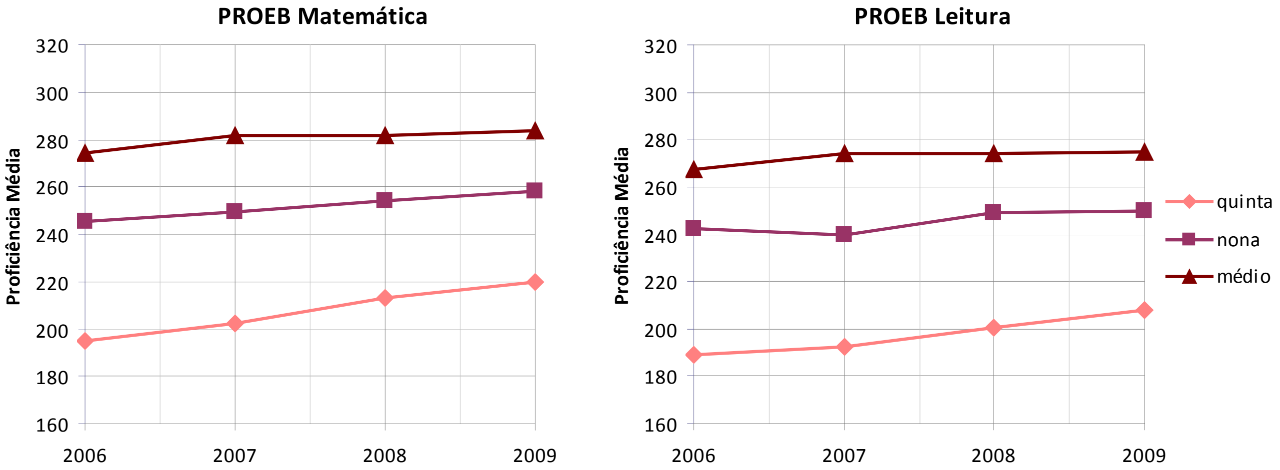
Sobre os dados de desigualdade das notas dos alunos em Minas Gerais, vemos que, pelo índice de Gini (tabela 3), o desempenho tem se tornado ligeiramente mais igualitário para quase todas as séries ao longo desses anos, variando pouco ano a ano. Nota-se também que a desigualdade é maior nos primeiros anos.O índice de Gini possui a propriedade de ser invariante a alterações proporcionais nas notas dos alunos, dessa forma, se de um ano para outro todos os alunos têm um acréscimo de 4% na nota, o índice não varia. Uma desvantagem é que o índice não pode ser decomposto, não é possível detectar de que parte da distribuição ocorre uma melhora ou piora.

Na seção 4, exploramos comparativamente propriedades de índices de desigualdade mais comuns e os índices propostos no artigo. A motivação do artigo se expressa no momento em que se vislumbra que a desigualdade educacional é muitas vezes medida como variáveis externas (resultantes

[3] Uma formalização mais matemática do índice aparece em Dalton (1920). Ceriani e Verme, 2011, publicaram recente artigo de comemorativo de 100 anos do índice de Gini no *Journal of Economic Inequality*.

indiretas) do quadro escolar (salários, anos de estudo, inserção no mercado) e poucas vezes medida a desigualdade no resultado pedagógico. Essa questão nos permite colocar uma consideração que de nada adianta expandir a educação se o sistema educacional (a aquisição do conhecimento) é *per si* desigual.

Gráfico 1 - Evolução do PROEB-MG, Matemática e Leitura (2006 a 2009)



Fonte: *PROEB-MG*, 2006 a 2009.

Tabela 3 - índices de Gini por Série e ano do PROEB-MG (2006 a 2009)

	2006	*2007*	*2008*	*2009*
Matemática				
5ª série	0.140	0.138	0.132	0.124
9ª série	0.118	0.114	0.112	0.106
3º ano EM	0.110	0.104	0.108	0.105
Leitura				
5ª série	0.137	0.136	0.116	0.116
9ª série	0.112	0.110	0.095	0.100
3º ano EM	0.105	0.099	0.095	0.093

Fonte: *PROEB-MG*, 2006 a 2009.

A seção a seguir elabora um modelo conceitual para o *desempenho* e a *igualdade* escolar, o que se pretende medir como *desigualdade* e qual é o *desempenho* que pode ser alcançado. Existe um grau de igualdade socialmente desejado? É possível partir de uma igualdade absoluta como referência para o campo educacional?

3. MODELO CONCEITUAL

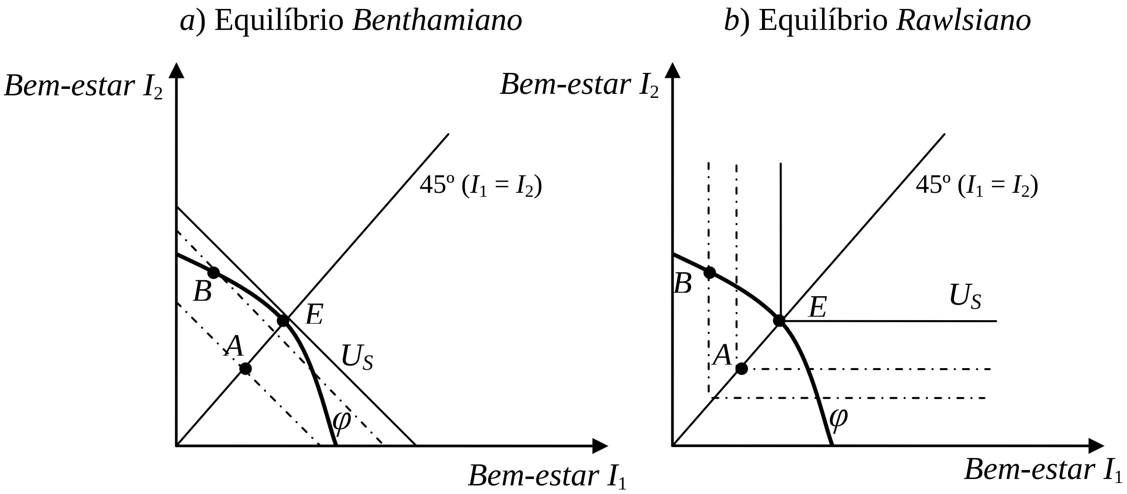
3.1. Modelo conceitual para o *desempenho* e a *igualdade*.

Umas das maneiras mais justas para se conceber um sistema de ensino é aquela em que todos recebem igual oportunidade para desenvolver plenamente suas habilidades. As habilidades possíveis de serem desenvolvidas pelo ciclo de vida escolar são múltiplas e idiossincráticas. Alunos submetidos a um mesmo regime de ensino podem desenvolver habilidades totalmente distintas, uns se desempenharão bem em leitura, geografia, biologia, outros terão desenvolvidos o lado da matemática, história, ciências. Essas habilidades não se restringem somente ao cognitivo, habilidades motoras, artísticas e de formação de caráter, também entram no rol de competências que a escola ajuda a formar.

Dessa maneira, nunca é possível estabelecer um parâmetro único de comparação para aquisição do conhecimento, toda medição do aprendizado será parcial e restrita e possuirá uma distribuição genérica com a maioria de alunos que se desempenham de acordo com a média esperada, uns poucos que se desempenham bem e outros poucos que se desempenham mal. O desempenho será aqui medido por meio de testes padronizados de acordo com a Teoria de Resposta ao Item (TRI). Esses exames permitem a comparabilidade por séries e pelos anos para cada matéria analisada. Os *escores* apresentam também uma graduação da compreensão do aluno sobre a matéria, sendo possível distinguir níveis (ver 4.2).

John Rawls (1971) estabelece que os parâmetros para julgar se uma instituição é justa são os princípios da *liberdade* e da *diferença*. O *princípio da liberdade* garante direitos de liberdade iguais entre os indivíduos. O *princípio da diferença* estabelece que o bem-estar do indivíduo I_1 não é igual ao bem-estar do individuo I_2, ou seja, o princípio da diferença reconhece que os indivíduos não são substitutos perfeitos e que a posição relativa é importante para a ponderação.[4] A Figura 1 apresenta o diagrama do Bem-estar de dois indivíduos e a curva com a **fronteira de possibilidades de produção** (φ). O segmento de reta de 45º representa o lócus em que o bem-estar dos dois indivíduos é equivalente. O ponto '*E*' representa o equilíbrio máximo da utilidade social (U_S) na perspectiva utilitarista *Benthamiana* (*a*), na qual a função de utilidade social é a reta perpendicular à reta de 45º, e o máximo da utilidade *Rawlsiana* (*b*), funções de utilidade-social em formato "L" com quinas.

Figura 1 – Pontos de equilíbrio para as funções de Utilidade Social nas perspectivas *Benthamiana* e *Rawlsiana*.



Fonte: *Elaboração dos autores*.

Além do ponto de equilíbrio '*E*' da figura 1, destacam-se os pontos '*A*' e '*B*'. O Ponto '*A*' é um ponto ineficiente e igualitário, pois é interior à fronteira de possibilidades e está sobre o segmento de reta de 45º. O Ponto '*B*' é eficiente e desigual, pois está sobre a fronteira de possibilidades de produção, mas apresenta bem-estar relativamente maior para o individuo I_2. Apesar das perspectivas *Benthamiana* e *Rawlsiana* terem preferências por um equilíbrio eficiente e igualitário

4 A representação analítica da curva de preferência social *Rawlsiana* é dada pela otimização *max-min*, escolhe-se maximizar o menor valor. O máximo social é alcançando maximizando-se o bem-estar do indivíduo mais desfavorecido. Nesse ponto, a ponderação pode ser vista como pertencente ao conjunto-binário {0,1}, com valor zero para o indivíduo (ou indivíduos) com posição relativa mais vantajosa e um para o indivíduo (ou indivíduos) com pior bem estar.

(solução de máximo no ponto 'E'), a classificação em cada um dos sistemas é diferente. De acordo com a utilidade social (U_S), o sistema *Benthamiano* classifica assim as preferências: c_B = $\{E \succ B \succ A\}$, lê-se 'E' é *estritamente preferível* ('$\succ$') à 'B' que é *estritamente preferível* à 'A'. No sistema *Rawlsiano* o ordenamento se torna: c_R = $\{E \succ A \succ B\}$. O sistema *Rawlsiano* prefere o ponto 'A', que é ineficiente porém equitativo, a um ponto 'B', que é eficiente mas desigual.

A exposição dos modelos da perspectiva utilitarista clássica e *Rawlsiana* é importante posto que acreditamos que a filosofia de Rawls é a mais adequada para propor a forma como os indicadores de educação devem ser interpretados. É neste ponto que conciliamos o que escreveu Rawls com a moderna teoria do desenvolvimento escolar da primeira infância de Cunha et al. (2006). Ao se referir aos princípios da *reparação* e da *diferença* no tocante à educação e no que diz respeito a um *trade-off eficiência* e *equidade* (aqui para nós transladado para *desempenho* e *igualdade*) Rawls (1971, pp. 120-121) ponderou:

> "*Na aplicação desse princípio* [o da reparação], *talvez se viessem a despender mais recursos com a educação dos menos inteligentes, e não dos mais inteligentes, pelo menos durante certo período da vida, digamos, os primeiros anos de escola"*. [...] *"**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 proporcionar a cada indivíduo um sentido seguro de seu próprio valor"*.
>
> Citação: J. *Rawls, "Uma Teoria da Justiça Social"*, pp. 120-121, 2008 [1971]. Grifos dos autores.

Nessa passagem Rawls afirma que as perspectivas de eficiência econômica e do bem-estar social não são as únicas a serem aplicadas para a justiça e equidade na educação. Posto isso, reconhecemos que a dimensão dos exames de desempenho para aferir conhecimento, a qual nos fiamos nesse artigo, se encaixa nessa abordagem mais ampla dos objetivos finais de uma educação de qualidade, que possua desempenho favorável e seja equitativa. O ideal seria ter aferição de conhecimentos para mais áreas além da matemática e leitura, isso ajudaria a compor um quadro mais completo.

Dessa maneira, não é possível se exigir igualdade absoluta em critérios de educação, e nesse ponto, apresentar os índices de Gini, tal como apresentamos nas tabelas 1, 2 e 3, não é uma medida tão apurada. Isso envolve reconhecer que não é possível se exigir que todos os alunos se desempenhem igualmente nas provas de matemática e leitura, até mesmo porque são medidas parciais do que deve ser o produto de um sistema educacional. O mais correto é reconhecer que os indivíduos são diferentes e desenvolvem, ao longo do ciclo de aprendizado, habilidades distintas. Como dito anteriormente, os indivíduos se dispersam em torno da média, alguns alunos serão bons em matemática, outros em línguas, e dentre aqueles que não se desempenham bem em nenhuma das aferições cognitivas é sempre possível instigar-lhes o desenvolvimento emocional e afetivo. Um índice de medição do desempenho educacional tem de tomar essa dispersão como característica do desenvolvimento humano natural.

Entretanto, a aceitação razoável de que os dados dispersam em torno de uma tendência central, não desencarrega a escola de garantir competências mínimas a todas as crianças. Assim, além de mirar a igualdade, o indicador de desenvolvimento educacional deve também tentar garantir os patamares de qualidade acima de um mínimo acordado socialmente das competências que podem ser desenvolvidas dentro da escola. Opomo-nos a uma referência que seja a **igualdade absoluta** (que é o referencial do índice de Gini) e propomos uma medição de **igualdade relativa** e acreditamos em um **patamar mínimo de desempenho** porque sabemos do potencial de cada criança se desenvolver cognitivamente, seja ela vinda de família pobre ou rica. Deve-se deixar claro, então, que quando medimos o desempenho da proficiência não estamos "acusando" ou "condecorando" o aluno, mas sim avaliando o sistema de ensino.

O motivo para termos uma educação desigual não está nos alunos nem em suas famílias, mas sim em todo o processo de ensino (ver figuras 2, 3 e 4), trata-se de um "*sistema gerador de desigualdade*" (*SGD*) do ensino brasileiro. Trataremos desse *SGD* na próxima subseção e na subseção seguinte os problemas advindos dele que os atuais indicadores não consideram. Na quarta seção propomos maneiras de se medir *desempenho* e a *igualdade* lidando com as considerações de toda a presente seção 3.

3.2. O sistema gerador de Desigualdades Educacionais: o que gera a desigualdade?

Dados os argumentos expostos na seção 3.1, podemos aceitar que os dados educacionais naturalmente apresentam uma distribuição e se dispersam em torno de uma tendência central, essa medida nunca poderá ser mantida ao mesmo tempo por todos (seria como imaginar todos os alunos com a mesma nota). É importante, portanto verificar o *quanto uma dispersão pode ser considerada uma desigualdade natural*.

Sugerimos um modelo conceitual no qual os dados de proficiência "revelados" ao pesquisador passam por um *Sistema Gerador de Desigualdade* (*SGD*). O *SGD* é um processo gerador de dados desconhecido revelado por uma função transportadora (t) que revela os dados "*visíveis*" (*observáveis*) para o pesquisador. O pesquisador só possui conhecimento da realização dos dados e tem de inferir o processo gerador que está por trás do que lhe é apresentado. A função geradora de desigualdade do *SGD* pode ser uma das funções de densidade probabilidade conhecidas da literatura, ou uma função desconhecida, ver equação (1) a seguir. Não importa qual seja a função, o pesquisador-observador terá de inferir esse processo para abstrair informações plausíveis dos dados.

Considere-se que estamos interessados na variável y, dada em função de uma série de características x, ($y = f(x)$), chamaremos essa função de "*função ideal*". O conceito envolve descobrir qual é a função-ideal (ou as funções) por trás do *SGD*, essa função $f(x)$ é "transportada" no ambiente por uma função t e, nesse processo, as desigualdades se revelam nos dados observados (y_p). O pesquisador tem acesso apenas aos dados observados em que $y_p = f_p(x)$. E a f_p é uma função inferida pelo pesquisador para a qual se quer o máximo de aderência à função ideal $f(x)$, ver figura 3.

Visto dessa forma, a função *ideal* "gera" os dados, essa dada função possui a **desigualdade natural** (ou a **igualdade relativa**, um conceito para o qual a sociedade pode considerar uma distribuição "justa"). Os "erros sistemáticos" e "não sistemáticos" são gerados (ou se revelam ao pesquisador) por meio de t e são capturados em processo de inferência que oferece os $\hat{y}_p$'s, obsevados e estimados. Dessa forma, o pesquisador poderá medir se a proficiência observada (y_p) revela *desigualdades naturais* ou, em outras palavras, *desigualdades sistemáticas* do *SGD*.

O processo da figura 2 abaixo pode ser descrito por funções e por meio das relações que se estabelecem entre elas, sabemos que a função *ideal* é dada por $y = f(x)$ para $x,y \in \Re^2$, tal que: $\int f(y)dy = 1$. Dessa forma, existe uma função de densidade para $f(y)$.[5] A variável y, ao se revelar no

[5] De fato, poderíamos nos concentrar apenas no conjunto da variável y, não sendo necessário mencionar que esta é uma função de x. No entanto, isso é importante para destacar que a variável y é dada em função de uma série de características x que serão abordadas mais à frente.

mundo observado, se torna y_p e o pesquisador terá de estimar a $f_p(x)$ que mais se ajusta a uma representação fiel e verdadeira de $f(x)$. Temos assim que $y_p = f_p(x)$ pode ser obtida da seguinte maneira:

$y_p = f_p(x) = f(x) + \varepsilon - u$ (1)

Onde $(\varepsilon - u) = t$ é a função de erros, ou função geradora de desigualdade. Temos que: $\varepsilon \sim N(0,\sigma_\varepsilon)$ e $u \sim N^+(0,\sigma_u)$. Ou seja, a função aferida pelo pesquisador $f_p(x)$ é igual à função *ideal*, porém, apresenta dois processos estocásticos embutidos: 'ε', que representa erros aleatórios normalmente distribuídos com média zero e variância determinada $N(0,\sigma_\varepsilon)$. São correspondentes aos erros não sistemáticos ou à *desigualdade natural*. O u, que representa erros aleatórios não negativos $N^+(0,\sigma_u)$ dados por uma *half-normal* ou alguma outra distribuição específica. Por meio da proposição em (1), podemos identificar a diferença entre a função real e a observada e propor a seguinte hipótese nula (H_0):

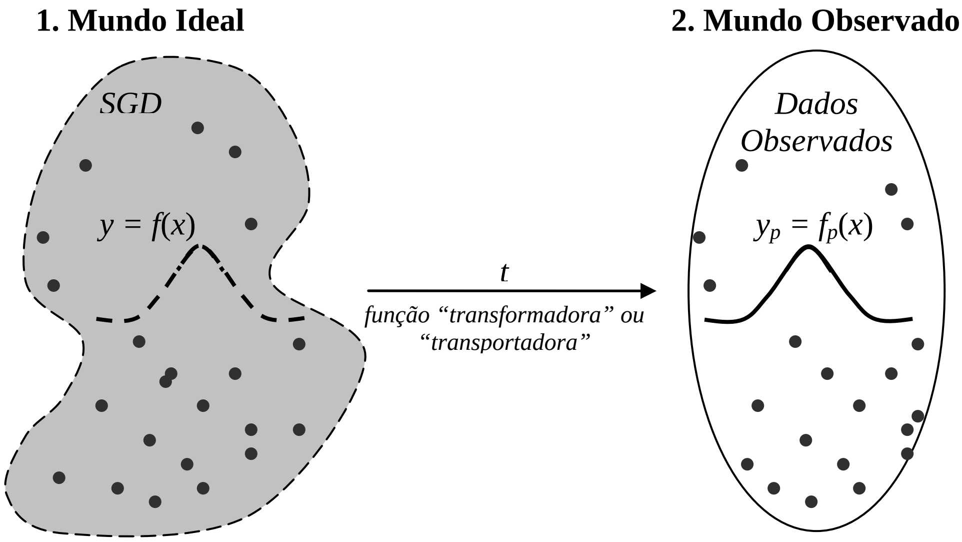
H_0: $y = y_p$ e a alternativa: H_1: $y \neq y_p$

Para que não se rejeite a hipótese nula, as duas funções dos erros propostos devem ser iguais em distribuição ($\varepsilon \sim N(0,\sigma) \overset{D}{=} u \sim N(0,\sigma)$). Nessa situação, rejeita-se a proposição de que o erro sistemático u seja de uma natureza diversa do erro não sistemático. É o mesmo que dizer que existe apenas o erro não sistemático, que não resulta em **desigualdade absoluta**.

A proposta desse trabalho é medir o *desempenho* e a *desigualdade* da proficiência dos alunos. Uma das maneiras de se medir a *desigualdade* é por meio da distância entre as funções de densidade e probabilidade da variável y, $f(y)$ e $f_p(y_p)$.[6] O índice de entropia obtido pelo indicador de divergência de Kullback-Leibler (1951) é uma das formas empregadas para compor o índice de desempenho e desigualdade:

$$D(F_p;F) = \int_{-\infty}^{\infty} f(y_P|x) \cdot \log\left(\frac{f(y_P|x)}{f(y|x)}\right) dy \quad (2)$$

Figura 2 – *Sistema Gerador de Desigualdade* (*SGD*) e função de transformação dos dados.

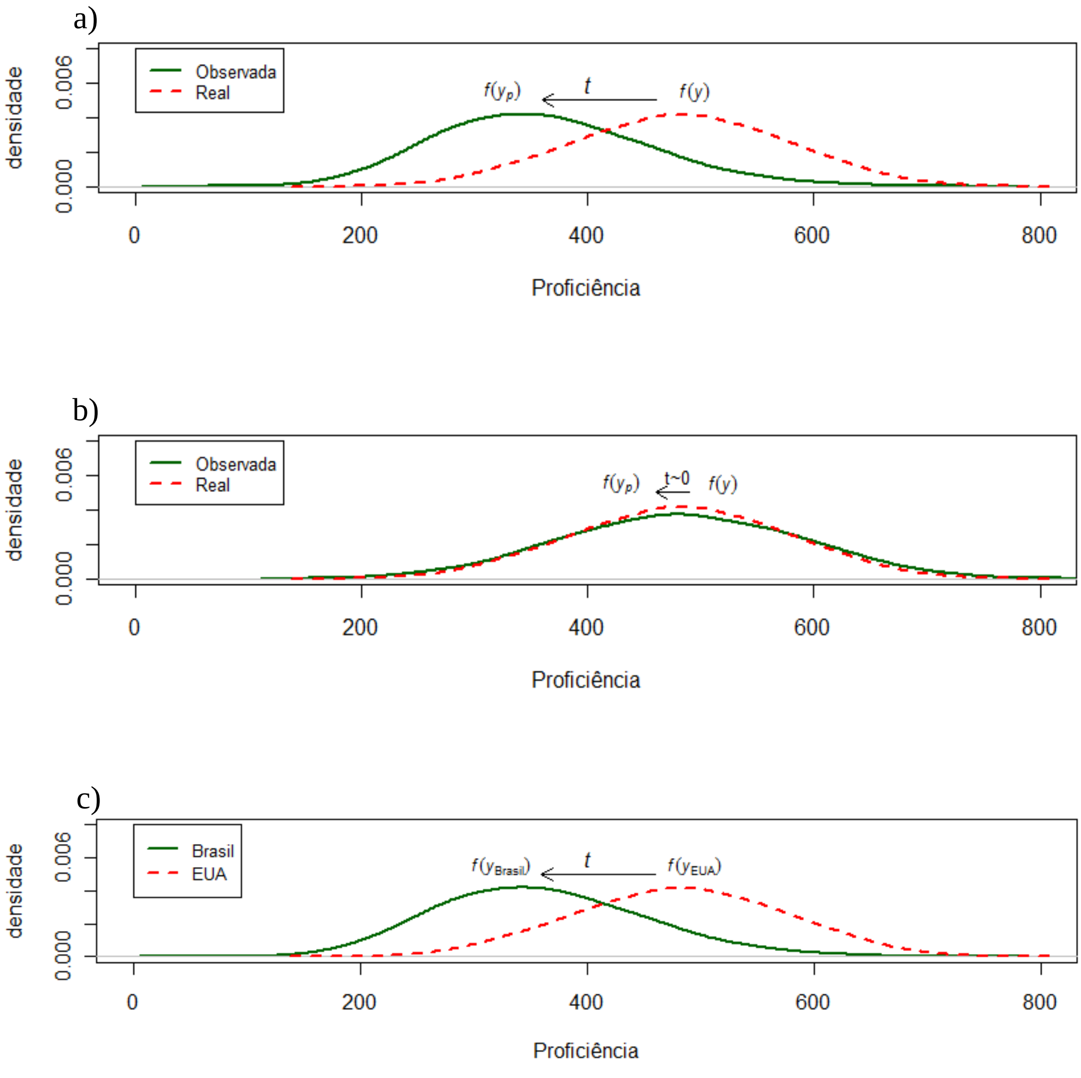


Fonte: *Elaboração dos autores.*

[6] As funções $f(y)$ e $f_p(y_p)$ podem ser condicionadas pelas características x: $f(y|x)$ e $f_p(y_p|x)$.

Voltaremos a dar mais detalhes sobre a estatística $D(F_p;F)$ na seção 4, por ora, é importante nos atermos à definição conceitual de proficiência e de sua distribuição padrão, a distribuição *ideal* $f(y)$. A distribuição "*ideal*" é uma distribuição que se deseja factível para um sistema educacional (função pontilhada da figura 3), essa função é a origem do *SGD* (a primitiva) e se torna observada após passar pela função transformadora $t(y)$. Se o sistema gera desigualdade, com $t = \varepsilon - u < 0$, a função observada será deslocada para a esquerda da *ideal*, tal como a função contínua da figura 5*a* e $D(F_p;F) > 0$. No caso em que $t = 0$, ou valores muito próximos de zero, a função não se distanciará muito da função real e $D(F_p;F) = 0$, figura 3*b*.[7]

Figura 3 - Comparação entre as funções de distribuição real e observada.



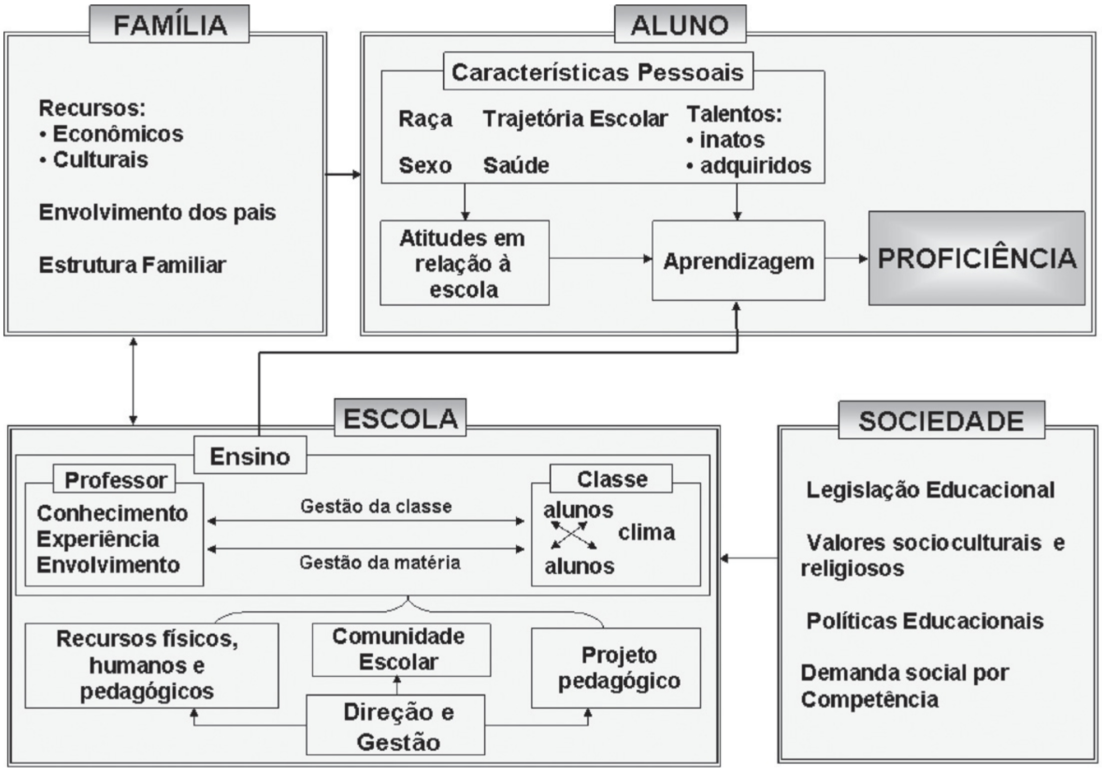
Fonte: *a*) e *b*) *Dados Aleatórios com elaboração dos autores* e *c*) *PISA, 2003.*

Estão contidas no *SGD* todas as variáveis *x* responsáveis pelo desempenho educacional (*y*). O sistema educacional, com todas as características *x* e seus reflexos nas variáveis de desempenho e

[7] No campo prático, se a função ideal é uma função factível, que pode ser alcançada pela gestão do sistema escolar, é possível descobrir porque um sistema em particular gera desigualdade e pior desempenho. Digamos que a referência para o ensino brasileiro seja a distribuição de notas dos Estados Unidos, por exemplo, como vimos nas tabelas 1 e 2, o desempenho do Brasil é mais baixo e mais desigual que o desempenho norte americano. Para alcançar o desempenho da função $f(y_{EUA})$ é necessário que o sistema educacional seja capaz de eliminar a função transformadora do *SGD* e possua $t(y) = 0$. Podemos interpretar que a função de desempenho brasileira $f(y_{Brasil})$ é mais desigual do que a norte americana, por conta do sistema não conseguir eliminar suas desigualdades.

outros produtos escolares é extremamente complexo. Soares (2007) sistematizou quatro dimensões importantes para descrever os reflexos de variáveis de contexto no aprendizado: *Família*, *Escola*, *Sociedade* e *Aluno*, essas dimensões e suas interações contribuem para alcançar o resultado final da aprendizagem e seus resultados captados pela proficiência (figura 4).

Figura 4 - Dimensões do Sistema Educacional e seu reflexo na Proficiência



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

3.3. O problema dos indicadores de medidas-centrais

De acordo com as perspectivas de *desempenho* e *igualdade* abordadas na subseção 3.1, e do modelo conceitual de 3.2, a interpretação dos indicadores de medição de desempenho da educação pode ganhar uma nova abordagem. Excetuando-se as soluções de canto e mantida a condição de Pareto, do ponto de vista utilitarista (revisite a figura 1), uma média maior é sempre preferida, mesmo que isso corresponda a maior desigualdade de desempenho. No caso das preferências Rawlsianas, aumentos de média não são suficientes para melhorias da utilidade social, apenas quando conseguem melhorar a situação do pior.

Para se medir adequadamente e interpretar indicadores educacionais de maneira a resolver os problemas da desigualdade revelados pelo ensino brasileiro, gerando uma massa crítica para a mudança, é preciso novos indicadores que superem o que estamos chamando de problema de medidas-centrais. Esses indicadores não levam em conta a dispersão. Olhar os indicadores de média implica aceitar uma substituição linear (uma preferência *Benthamiana*). Em outros termos, de acordo com essa visão, seria possível melhorar o desempenho médio sem que isso ocorra a um maior número de alunos e lhes confira um maior aprendizado.

Como exemplo simples, podemos pensar em um conjunto S de três alunos medidos por um indicador de aprendizado-proficiência (y_i) que varia de zero a dez: $S_1 = \{3, 3, 6\}$; média = 4. Dois alunos possuem desempenho igual a três e o terceiro igual a seis. Caso supuséssemos uma meta para média de desempenho igual a 5, uma maneira de se alcançar o resultado seria: S_2 = {2, 3, 10}; média = 5.

Esse conjunto atenderia à solução de máximo bem estar social do utilitarista clássico, mas visivelmente não atente à solução *Rawlsiana* dado que é uma solução que piora as condições de proficiência do aluno com pior desempenho. Uma meta tal como esta corre o risco de privilegiar alguns deixando a maioria para trás. De S_1 para S_2 ocorreu uma substituição na valoração de cada um dos alunos, essa substituição equivale a dizer que em troca de perder apenas um ponto para um dos alunos e manter o outro estável, é bom ganhar quatro pontos para o aluno mais capaz. O saldo assim é positivo (4-1 = 3), esses três pontos a mais são os que garantem um ponto extra na média.

Os pontos ganhos pelo terceiro aluno são vistos como substitutos para os dois primeiros, o que o mais apto ganha em aprendizado compensa o que os demais não ganham, e, caso se adote essa visão, tem-se o ótimo social em detrimento da igualdade. O aumento da desigualdade pode ser um efeito não desejado. Entretanto, a sociedade precisa ter em vista que ao colocar metas apenas para a média do desempenho, está implicitamente aceitando que a solução S_2 acima é uma das soluções possíveis do problema, e que o maior ganho de educação de um dos alunos (desde que suficientemente grande para elevar a média) pode ser transferido aos demais que deixam de aprender, não assume portanto nenhuma preferência por um sistema mais igualitário.

A proposição *Rawlsiana* envolve admitir que enquanto não se melhore a condição do mais atrasado, não há solução possível. Isso coloca um contrapeso e impede a obtenção de uma situação socialmente desigual. Para essa visão, o aprendizado não pode ser substituído entre os alunos e, nesse sistema, o requisito mínimo para a solução com média cinco é atendido tanto por S_3 (S_3 = {4, 5, 6}; média = 5) quanto por S_4 (S_4 = {5, 5, 5}; média = 5).

Na solução S_3, em relação à situação inicial S_1, nenhum dos alunos piorou, sendo que em S_4, pelo fato do primeiro aluno estar melhor, a solução é mais igualitária e preferencial. Um sistema de aproveitamento pleno é aquele que permite desenvolver capacidade máxima de cada um dos alunos, sendo que os melhores possuem liberdade para alcançar o seu potencial, outras soluções para os conjuntos *S* são possíveis.

Não é apenas a média simples que sofre deste problema, outras estatísticas de medidas centrais são menos sensíveis a valores extremos, mas não captam a dispersão dos dados. Caso se adotasse como meta proposta a mediana igual a cinco, as soluções S_3 e S_4 escritas acima seriam adotadas. No entanto, outra solução possível seria S_5 = {0, 5, 5}, a mediana estaria satisfeita, mas a média aritmética seria pior do que S_1. Indicadores de porcentuais de alunos abaixo de um padrão básico, ou outros tipos de média (ponderada, harmônica e geométrica) também sofrem do mesmo problema[8], entretanto, veremos que há como encontrar soluções melhores para o propósito de se medir resultados na educação.

Os indicadores propostos neste artigo se apoiam na constatação de que a substituição do aprendizado entre os alunos não é perfeita e de que, no quesito educacional, a desigualdade é um parâmetro importante a ser dimensionado e objetivado. A proposta para lidar com estas questões é adotar referências de desempenho que levem conta o valor social intrínseco de se ajudar a quem

[8] Suponha um conjunto *S* formado por subconjuntos S_i de tamanhos definidos s_i, em que s_i define o número de elementos em cada subconjunto e *N* é o total de elementos de todos os subconjuntos somados. Temos que $S = \{S_1, S_2, ..., S_K\}$. Se supusermos que esses subconjuntos são definidos pelo número de alunos abaixo de uma nota de corte (y_i), ao propor a meta de não termos nenhum aluno abaixo do básico teríamos $s_1 = 0$, ou seja, $S_1 = \emptyset$. A média é dada por:$\mu = s_1y_1/N + s_2y_2/N + ... + s_Ky_K/N$. Mantidas as notas de corte, ao se propor $s_1 = 0$, naturalmente se terá uma média mais alta para *S*, no entanto, nada se pode dizer da dispersão, dado que os outros subconjuntos S_i podem estar desigualmente distribuídos.

mais precisa e que isso pode ser mensurado por índices de desempenho ponderados e de desempenho relativo.

4. MÉTODO PARA TRÊS INDICADORES

Para o uso e interpretação do grande número de informações estatísticas é importante a construção de medidas sintéticas e indicadores que consigam captar grande parte da informação relevante em bancos de dados com milhares de alunos. Apresenta-se também a necessidade de resumir as informações de *desempenho* e *igualdade* de maneira que se permita propor metas para os dois objetivos, minimizando os problemas apresentados em 3.3.

4.1. Índice de Qualidade da Educação (IQE)

O *Índice de Qualidade da Educação* (*IQE*) é um índice discreto que pode ser obtido para grupos de alunos em turmas, escolas, municípios, etc. O *IQE* é construído por dois componentes, o componente de *defasagem*, que leva em conta o desempenho médio ponderado pela participação dos alunos nas diferentes faixas de proficiência e o componente da distância do indicador à referência. Quanto mais distante da referência, menor será o *IQE*:

$$IQE = \left(1 - \frac{defasagem}{3}\right) \qquad (3)$$

O indicador de defasagem pode ser representado por uma multiplicação de vetores:

$$defasagem = \mathbf{w}' \cdot \mathbf{p}$$

Em que **w**' é um vetor transposto, sua dimensão é (1 x *n*) e possui valores que atribuem importância (pesos) a cada um dos '*n*' níveis específicos. Na abordagem escolhida, em que n = 4, $\mathbf{w}_n$ é um vetor de valores monotônicos decrescentes. O nível n = 1 (chamado de 'abaixo do básico') possui valor três ($\mathbf{w}_1$ = 3), no nível 2, o valor é dois ($\mathbf{w}_2$ = 2), e assim por diante até $\mathbf{w}_4$ = 0. Visto na forma transposta temos o seguinte vetor:

$$\mathbf{w}' = \{3, 2, 1, 0\}$$

O **p**, por sua vez, é um vetor de porcentuais de alunos presentes em determinado nível. Esse vetor de porcentuais assume valores de acordo com o desempenho das turmas escolas ou municípios, e sua soma deve ser sempre igual a um ($\sum_{n=1}^{4} p_n = 1$). Supondo, como exemplo, uma turma em que a proporção dos alunos em cada nível é equivalente, teríamos o seguinte vetor **p**' = {0.25, 0.25, 0.25, 0.25}. A multiplicação (**w**'·**p**) desses dois vetores de dimensões (1 x *n*) e (*n* x 1) retornaria um escalar (1 x 1) com $\mathbf{w}'\mathbf{p} = 1.5$.

Para o caso acima, de porcentuais igualmente distribuídos entre os quatro níveis, o índice do *IQE* se torna igual a 0.5 (1 – 1.5/3). É fácil ver que o *IQE* varia entre **zero**, situação em que todos os alunos estão abaixo do básico (**p**' = {1, 0, 0, 0}), e **um**, valor para o qual todos os alunos estão situados no nível avançado (**p**' = {0, 0, 0, 1}). A interpretação do índice diz que o valor do *IQE* = 0 corresponde ao ponto em que todos os alunos estão defasados e se situam no nível abaixo do básico do desenvolvimento escolar. Quando o *IQE* = 1, a interpretação é de que todos os alunos estão acima no nível avançado.

A fórmula do *IQE* apresentada em (4) pode dispensar a *defasagem* e ser apresentada de maneira mais direta se invertermos o vetor $\mathbf{w}_n$ para um vetor crescente:

$$\mathbf{w}' = \{0, 1, 2, 3\}$$

$$IQE = \frac{\mathbf{w}'\mathbf{p}}{3} \qquad (4)$$

Outra forma de apresentação, útil para os fins desse artigo, encontramos em (5), em que $\sum_{n=1}^{4} \mathbf{w}_n = 1$, a soma do vetor de pesos é igual a 1, e o valor 3 é substituído por uma fórmula mais geral **w'p*** fornecendo a fórmula abaixo:

$$IQE = \frac{\mathbf{w}'\mathbf{p}}{\mathbf{w}'\mathbf{p}^*} \qquad (5)$$

$$onde\ \mathbf{w}' = \left\{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}\right\} \text{ e } p^{*}{}' = \{0,\ 0,\ 0,\ 1\} \quad (6)$$

4.2. Número de Níveis (*n*) e notas de corte (y_C).

Um ponto importante a ser acrescido trata da determinação do número de níveis de corte a serem utilizados para a construção do *IQE*. Os níveis são propostos por notas de corte, representadas por $y_{d,c}$, em que o subscrito '*d*' representa a disciplina a ser medida (*d* = {"*matemática*", "*leitura*"} ou, simplesmente, *d* = {"*m*","*L*"}) e 'c' pertence ao subconjunto *c* = {1, 2, ... , *n*-1}. Os cortes são adotados em uma escala progressiva para cada disciplina e cada série (tabela 4).

Tabela 4 - Níveis e notas de corte do PROEB-MG por série e matéria

	Abaixo do Básico (n=1)	Básico (n=2)	Intermediário (n=3)	Avançado (n=4)
Matemática				
5ª série	$y_{m,1} < 175$	$175 \le y_{m,2} < 225$	$225 \le y_{m,3} < 275$	$275 \le y_{m,4}$
9ª série	$y_{m,1} < 225$	$225 \le y_{m,2} < 300$	$300 \le y_{m,3} < 350$	$350 \le y_{m,4}$
3º ano EM	$y_{m,1} < 275$	$275 \le y_{m,2} < 350$	$350 \le y_{m,3} < 400$	$400 \le y_{m,4}$
Leitura				
5ª série	$Y_{L,1} < 150$	$150 \le y_{L,2} < 200$	$200 \le y_{L,3} < 250$	$250 \le y_{L,4}$
9ª série	$y_{L,1} < 200$	$200 \le y_{L,2} < 275$	$275 \le y_{L,3} < 325$	$325 \le y_{L,4}$
3º ano EM	$y_{L,1} < 250$	$250 \le y_{L,2} < 300$	$300 \le y_{L,3} < 375$	$375 \le y_{L,4}$

Fonte: *Elaboração dos autores com base em matriz proposta por Todos pela Educação, 2007.*

Os valores de corte são mantidos fixos para a comparabilidade entre os anos. O número de níveis pode ser maior do que 4 (*n* > 4), para esses casos, o *IQE* é mais refinado e representativo da distribuição das notas dos alunos. Essa é a proposta do *IPD* (*Índice Ponderado de Desempenho*) apresentado a seguir e serve de base para interpretação do *IDR* (*Índice de Desempenho Relativo*).

Empregando-se a ideia do *IQE,* dos níveis e das notas de corte, é possível construir uma "*família*" de indicadores relacionados. Considerar essa *família* de indicadores é de interesse na proposição de metas de *desempenho* e *igualdade,* pois se pode considerar que mesmo o *IQE* sofre dos problemas de medidas centrais apresentados em 3.3. Isso porque dentro dos níveis, a distribuição dos alunos pode diferir ligeiramente, e o *IQE* é insensível a essa mudança.

4.3. Índice Ponderado de Desempenho (IPD)

Acrescenta-se que o vetor **w**, base para a construção do *IQE*, apresenta uma ponderação crescente que pode ser não factível (ou no mínimo implausível) para se ponderar os alunos. Dificilmente se verá o comportamento do vetor de probabilidade **p** seguindo a mesma tendência do vetor **w** do *IQE*, ou na situação em que todos os alunos estão no nível avançado $\mathbf{p}_4 = 1$. O *Índice Ponderado de Desempenho* (*IPD*) se apoia noutro princípio: a estrutura do vetor de pesos **w** segue uma densidade de referência. Esta referência é a função *ideal* $f(y)$, referida na seção 3.2. O *IPD* também aumenta o número de níveis para um nível ótimo (n^*) em que $n^* > n = 4$. Imagine-se que ao invés de quatro grupos tivéssemos cem, à medida que aumentássemos o número de grupos, teríamos intervalos cada vez menores para as notas. Com o número de intervalos tendendo ao infinito, $n \to \infty$, a representação da distribuição fica mais pormenorizada (intervalos infinitesimais).[9]

Similar ao *IQE* dado pela fórmula apresentada em (1), o *IPD* é apresentado pela fórmula:

$$IPD = \frac{\mathbf{w}'\mathbf{p}}{\mathbf{w}'\mathbf{p}^*} \qquad (7)$$

$$onde\ \mathbf{w}' = \{p(y_1),\ p(y_2), \ldots ,\ p(y_{n^*})\},\ \text{em que} \sum_{i=1}^{n^*} p(y_i) = 1 \quad (8)$$

Em que temos que $p(y_1)$ é a probabilidade de y estar compreendido no intervalo do nível 1. Tomando-se um nível '*n*' qualquer, em que o limite inferior é $y = \alpha$ e o limite superior é $y = \beta$, temos que $p(y_n) = \int_\alpha^\beta f(y)dy$. Em todo o domínio de y temos $\int f(y)dy = 1$, definido conforme 3.2. O vetor **p**'* fornece o valor p = 1 para o nível no qual a probabilidade $p(y_n)$ é máxima: **p**'* = {0, 0,..., 1, ..., 0}.

A diferença do indicador *IPD* para o índice anterior é de que o vetor de pesos **w** tem extensão maior do que o vetor de quatro níveis de (6). Esse vetor constrói um número de níveis n^* ideal para a representação da função $f(y)$. Para esse artigo o número de *níveis* utilizados no *IPD* foi igual a 21 em matemática e 22 em português. Importante notar que de acordo com a função $f(y)$ os valores de **w** não são necessariamente uma sequência monotonicamente crescente. Em verdade, pode se generalizar os vetores **w** e **p** para funções contínuas e reescrever o *IPD*:

$$IPD = \frac{\int f(y) \cdot f_p(y_p) dy}{f(y^*)} \qquad (9)$$

Supondo uma integral definida para as funções acima multiplicadas, o *IPD* se torna um valor definido e limitado entre zero e um ($0 \leq IPD \leq 1$). Em palavras, o *IPD* é um somatório das funções ideal, $f(y)$, e observada, $f_p(y_p)$ multiplicadas uma sobre a outra e divididas pelo valor da função referencial quando $y = y^*$. Quanto mais próximo de **um**, mais igualitário é o desempenho dos alunos e mais similar é a função observada $f_p(y_p)$ da função referencial $f(y)$.

4.4. Índice de Desempenho Relativo (IDR)

O *IDR* é o mais sofisticado entre os três índices propostos, pois se baseia na distância da função observada em relação à função de referência (a *ideal*). Essa distância é similar às distâncias quantílicas, mas são tomadas em relação às probabilidades e não nos valores dos quantis. Para isso,

[9] Aumentar indeterminadamente o número de níveis acrescenta outras dificuldades: é impossível (ou pelo menos muito pouco prático) representar amostras finitas por uma densidade *kernell* com um *bandwidth* muito pequeno. A ideia por trás, quando apresentamos $n \to n^*$, é descobrir qual a possível distribuição referência para os dados de proficiência. Essa distribuição é a base de $\mathbf{w}_n$.

é necessário tomar uma base única de comparação em *y*, chamaremos essa base de *r*, que representa uma transformação de *y* e pode ser entendido como o percentil de *y*. O *r* está muito relacionado com o número de níveis *n*. Para *n*= 100, por exemplo, o *IDR* pode ser apresentado da forma:

$$IDR = \sum_{n=1}^{n=100} \mathbf{w_n} \cdot \log\left(\frac{\mathbf{p}_n}{\mathbf{p}_{p,n}}\right) \qquad (10)$$

Temos que $\mathbf{w}_n$ é um vetor de pesos com valores para cada nível *n*, e os valores $\mathbf{p}_n$ são probabilidades de referência para cada nível *n* e $\mathbf{p}_{p,n}$ é o vetor das probabilidades dos dados observados, obtidos pelo pesquisador. Para valores indefinidos, quando $\mathbf{p}_{p,n} = 0$, $IDR_n = \ln(100)$. Podemos adotar pesos iguais aos valores de referência: $\mathbf{w}_n = \mathbf{p}_n$, e sabendo-se que $n = r$, temos:

$$IDR = \sum_{r=1}^{r=100} \mathbf{p}_r \cdot \log\left(\frac{\mathbf{p}_r}{\mathbf{p}_{p,r}}\right) \qquad (11)$$

A essa altura, é mais conveniente e simples expressar o *IDR* por meio das funções de probabilidade $f(y)$, não sendo necessário restringir os valores de *r* aos percentis.

$$IDR = \int f(y_r)\cdot \log\left(\frac{f(y_r)}{f_p(y_{p,r})}\right)dy \qquad (12)$$

A função (12) é muito próxima da estatística de Kullback-Leibler apresentada em (2), $D(F_p, F)$, lembrando que *y* é função de uma série de características *x*: $y = f(x)$. Para valores condicionados em x, $f(y|x)$, podemos novamente obter (2).

Assim como a estatística *D*, os valores do *IDR* serão sempre maiores ou iguais a zero, $IDR \geq 0$. Adotando-se a notação proposta por Handcock e Morris (1998) é possível ver que o *IDR* é igual à média dos logaritmos de $g(r)$ dados na função *g*. De outra maneira, a área abaixo da função relativa $\log(g(r))\cdot g(r)$, integral no intervalo de 0 a 1:

$$g(r) = \frac{f(y)}{f_P(y_P)}$$

$$IDR = \int_0^1 \log(g(r))\cdot g(r)dr \qquad (13)$$

Um *IDR* = 0 significa uma situação em que as funções comparadas são idênticas, ou seja, $t = (\varepsilon - u) = 0$, não existem diferenças e não existe o *SGD*, não há geração de desigualdades. Esse caso está representado na figura 5*b* e não se rejeita a hipótese H_0. Quanto maior o *IDR*, maior a distância da função observada $f_p(y_p)$ em relação a função de referência $f(y)$. Isso significa que há um *SGD* e rejeita-se a hipótese nula H_0. Diferentemente dos dois primeiros índices, o *IDR* possui intervalo ($0 \leq IDR \leq \infty$) para o conjunto dos Reais não negativos, $\Re^+$.[10]

[10] Importante observar que o IDR pode ser maior do que zero por uma função à direita da função *ideal*, esse pode ser um problema importante para outros conjuntos de dados, mas não para o método aqui desenvolvido. Em síntese, podemos pensar que se os dados estão à direita da função de referência, esta não será mais uma referência *ideal*. Nesse trabalho, um *IDR* > 0 significará sempre uma função observada aquém do *ideal*. Em se pensando em funções contínuas, o *IDR* não tem um limite superior, no entanto, para um conjunto domínio fechado para a função $f(y)$ é possível o cálculo do valor **máximo** e **mínimo** do *IDR*.

4.5. Resumo dos Índices

O esforço desenvolvido nessa seção foi no sentido de construir índices que usem uma função de comparação *ideal*. Os três índices propostos podem ser decompostos em *desempenho* e *desigualdade*. Essa decomposição é feita utilizando-se uma transformação com uma função de referência $f_h(y)$ que mantém a média de $f(y)$, mas possui a dispersão e formato (*shape*) da função observada $f_p(y_p)$. Tal como proposto por Handcock e Morris:

$$\frac{f(y)}{f_P(y_P)} = \frac{f_h(y)}{f_p(y_p)} \times \frac{f(y)}{f_h(y)} \qquad (14)$$

Densidade Relativa Total	=	*Densidade Relativa para o efeito de 'location difference'*	x	*Densidade Relativa para o efeito de 'shape difference'*	(15)

Em termos da nomenclatura do artigo o *IDR* pode ser dividido em:

IDR = efeito '*desempenho*' x efeito '*desigualdade*' (16)

Podemos escrever o *IQE* e o *IPD* também em termos de funções, a função referencial para cada índice é o que os distingue: $f^{IQE}(y)$ para o *IQE* e $f^{IPD}(y)$ para o *IPD*:

$$\mathbf{w} \approx f(y) \text{ e } \mathbf{p} \approx f_p(y_p)$$
$$IQE = 2\int f^{IQE}(y) \cdot f_p(y_p)dy \qquad (17)$$

Lembrando que $\frac{1}{\sqrt{2\pi}} = Max\{\text{Prob}(y_i)\}$ é o ponto de máximo de uma função de referência $f(y)$ que seja normal, temos:

$$IPD = \frac{1}{\sqrt{2\pi}}\int f^{IPD}(y) \cdot f_p(y_p)dy \qquad (18)$$

Os efeitos de decomposição em *desempenho* e *desigualdade* são um dos principais interesses do trabalho. Pelos indicadores apresentados, se uma escola, município ou região apresenta uma evolução na média de desempenho em detrimento da equidade, os indicadores terão desempenho baixo. Caso exista equidade, mas com um desempenho muito aquém do desejado, o indicador tampouco será satisfatório. Portanto, os índices propostos sugerem que o melhor comportamento é o de melhorias *Rawlsianas* ou *quase-Rawlsianas*, tal como apresentadas em 3.1.

A seção a seguir apresenta uma seleção dos resultados dos três índices para Minas Gerais por ano e disciplina do PROEB. Ao final são apresentados possíveis desenvolvimentos futuros e mais decomposições propostas, assim como a relação desse índice com as demandas por metas pedagógicas e para melhoria da qualidade do ensino.

5. RESULTADOS

5.1. Resultados Gerais

Para cálculo dos três índices, utilizaremos banco de dados do PROEB-MG – Programa de Avaliação da Rede Pública da Educação Básica pertencente ao Sistema Mineiro de Avaliação da Educação Pública (SIMAVE). O *PROEB* é um exame de proficiência em leitura e matemática que começou a ser aplicado em 2002, e similar ao exame do SAEB nacional (Sistema de Avaliação da Educação Básica), com a diferença de que sua aplicação é censitária para escolas de rede estadual de Minas Gerais. Desde 2006, o exame do PROEB conta com a participação de escolas dos sistemas municipais e estaduais. A partir de 2008, todos os municípios de Minas Gerais estão presentes na amostra.

Os índices foram construídos com os resultados da 5ª série do ensino básico, que atende, em geral, crianças de 10 anos, com um intervalo adequado de 9 a 11 anos. A distorção idade-série (numero de alunos em idade fora da adequada sobre o total de matrículas) para a 5ª série se situa em torno de 19% em Minas Gerais, dados da Sinopse estatística da educação básica do INEP/MEC (2010). Cabe lembrar o leitor que nos referimos à 5ª série após a mudança na grade curricular, com a inserção da série inicial para crianças de 6 a 7 anos.

Os índices refletem resultados das provas de matemática e leitura em língua portuguesa no período de 2006 a 2009. Pelos dados das provas, podemos perceber a evolução das médias e medianas em cada uma das competências medidas ao longo dos anos, é possível ver que o desvio padrão aumenta em relação ao ano inicial, com exceção de 2008 em leitura (tabela 5).

Tabela 5 – Evolução do PROEB-MG, Matemática e Leitura.

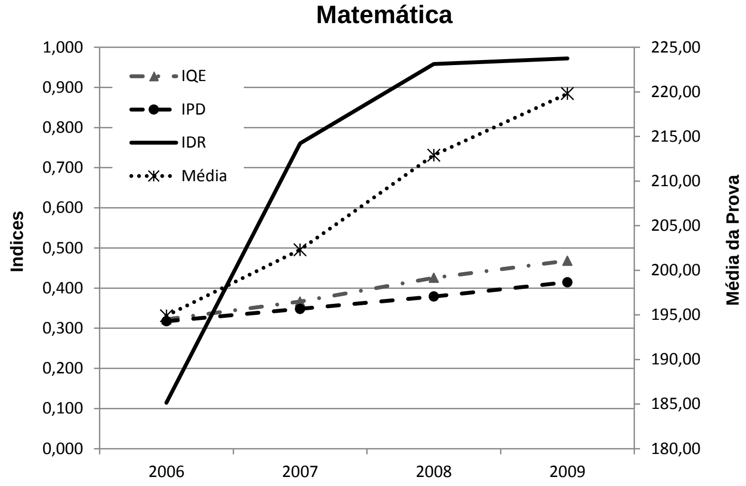
	Média	*Mediana*	*Desvio Padrão*	*Min*	*Max*
Matemática					
2006	194.90	193.60	48.14	51.66	352.60
2007	202.30	201.80	49.04	51.13	351.60
2008	212.90	211.30	49.28	53.27	360.20
2009	219.80	219.80	48.15	65.65	361.20
Leitura					
2006	189.00	188.50	45.67	52.02	333.60
2007	192.40	191.80	46.09	53.59	347.50
2008	200.30	198.90	41.16	75.88	358.90
2009	238.80	239.10	51.61	70.82	415.90

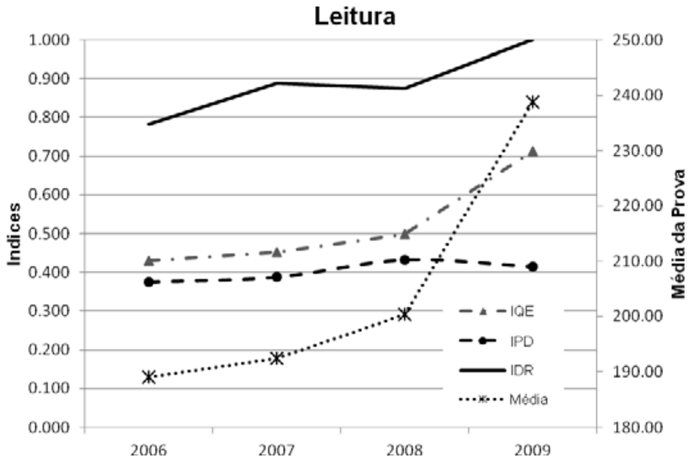
Fonte: *PROEB-MG, 2006 a 2009.*

Por conta dessa evolução das médias, os índices apresentam também a evolução do *desempenho*, o *IQE* de leitura mostra uma elevação de 66% em relação a 2006, e o *IDR* de matemática exibe uma evolução de 751% entre os quatro anos, explicada pelo enorme salto do índice entre os anos 2006 e 2007. Ao que se pode notar dos gráficos 3, há uma diferença de evolução entre as médias e os índices, principalmente no caso do *IDR*. Em matemática, a média, o *IQE* e o *IPD* apresentam evolução constante, ao passo que o *IDR*, apresenta uma evolução rápida nos primeiros anos e marginal nos anos de 2008 a 2009. Isso pode nos sugerir que o comportamento de matemática nos anos de 2008 e 2009, quando já próxima de um limite, se dá lentamente.

Sobre a prova de leitura, todos os índices partem de um patamar mais elevado. Em termos da função de referência, pode se dizer que a prova de leitura parte de um patamar mais próximo do referencial escolhido. A evolução da média em português é mais acentuada no biênio 2008 e 2009, isso se reflete no *IQE* e *IDR*, mas, em especial, não no *IPD*.

Gráfico 3 - Evolução dos três índices *IQE*, *IPD* e *IDR* e da média de 2006 a 2009.





Fonte: *Construção com os dados do PROEB-MG, 2006 a 2009. Obs.: IDR padronizado entre 0 e 1.*

A função de referência constitui-se de técnica importante na construção e decomposição dos índices de *desempenho* e *desigualdade*. Essa função de referência foi construída por meio das técnicas desenvolvidas em Soares, Marotta e Delgado (2010). Tal função é tida como *ideal* em termos *desempenho* e menor *desigualdade*. Para sua construção, adotou-se um conjunto de países de referência no *PISA*[11], todos os países referenciados possuem características positivas que servem de referência para o sistema brasileiro e se situam em um ponto mais avançado da escala de proficiência e com mais igualdade. A distribuição dos países é comparada com os resultados do Brasil em termos de desvio padrão, e são calculadas as distâncias quantílicas em termos de desvios: $\gamma = (Q_{i\%}^{países} - Q_{i\%}^{Brasil}) / \sigma(Q^{Brasil})$. Para transportar os valores para escala SAEB[12], o resultado de γ é multiplicado pela nota do SAEB em cada série. Ao final, os valores são ajustados para se adequarem aos níveis propostos conforme valores sa tabela 4 apresentada anteriormente.

[11] Essa distribuição de referência é formada pelos 24 países listados: Alemanha, Austrália, Áustria, Bélgica, Canadá, Coréia, Dinamarca, Espanha, Estados Unidos, Finlândia, França, Holanda, Inglaterra, Irlanda, Islândia, Itália, Japão, Luxemburgo, Noruega, Nova Zelândia, Polônia, Portugal, Suécia e Suíça.

[12] Haja vista que os métodos de Teoria da Resposta ao Item elaborado pelo PISA e SAEB são de metodologias distintas. O PISA emprega o método de "valores plausíveis". O TRI do SAEB emprega técnica aperfeiçoada por Klein (2003). A escala do PROEB (para qual utilizamos as funções de referência) é muito similar a do SAEB.

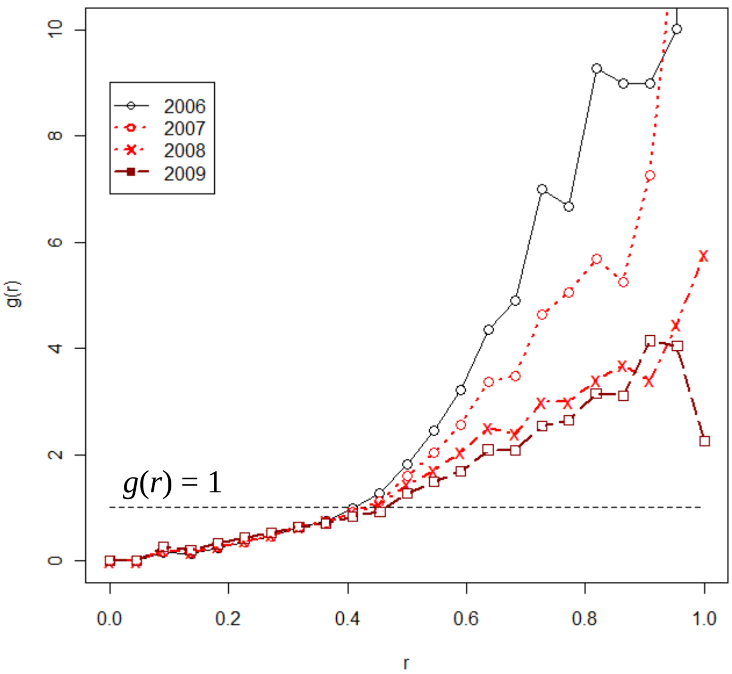
Essa função de referência possui a chave para interpretação e decomposição dos índices. Em um sistema ideal de aperfeiçoamento, o *Sistema Gerador de Desigualdade* (*SGD*) deve ser menor, e a função *observada* estará próxima da *ideal*.

5.2. Resultados de Decomposição

Nessa subseção mostraremos que a melhoria dos índices de *desempenho* não é acompanhada por diminuição dos resultados em *desigualdade*. Em outras palavras, o problema gerado pelos indicadores de medida central e relatado na subseção 3.3 ocorre nos dados do PROEB para Minas Gerais. Para as duas matérias (matemática e português) a desigualdade total do sistema é maior após os anos iniciais. Essa desigualdade pode ser decomposta conforme a equação (14) descrita em 4.5.

Para a construção do *IDR* (funções 12 e 13) e sua decomposição é interessante observamos a função relativa $g(r)$, quanto mais próxima de 1 essa função, menor será o valor total do *IDR*. Valores $g(r)$ muito maiores do que 1 indicam uma distância grande entre o referencial *ideal* e o *observado*. Podemos ver a evolução da função $g(r)$ ao longo dos anos para as notas de matemática, apresentadas no gráfico 4. O *IDR* é a área abaixo da função $\ln(g(r))\cdot g(r)$. Nota-se pela área abaixo de $g(r)$ que o *IDR*, de 2006 a 2009, será cada vez menor e por isso seu índice padronizado cada vez maior. A referência se encontra na linha contínua de $g(r) = 1$.

Gráfico 4 – Comportamento da função g(r) ao longo do tempo.

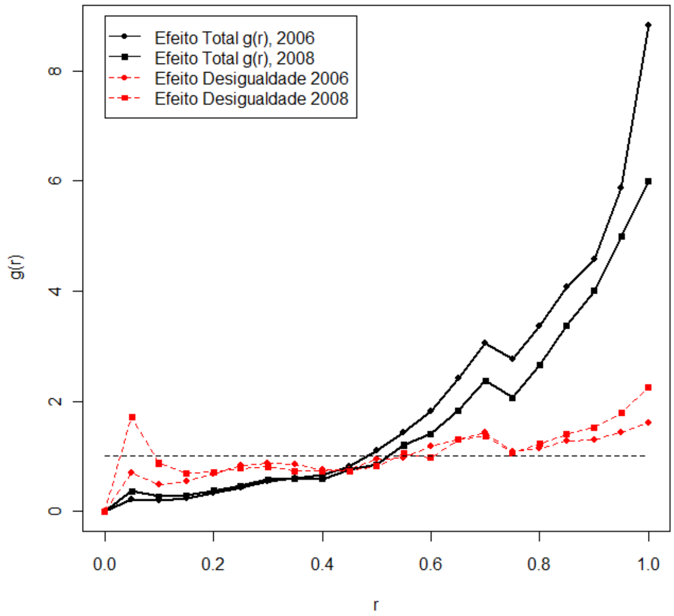


Fonte: *Construção com os dados do PROEB-MG, matemática, 2006 a 2009.*

A mesma função acima pode ser decomposta em seus efeitos de *desempenho* (nível) e *desigualdade* (dispersão), na verdade, a maneira mais fácil de obter a decomposição é encontrando o efeito total das densidades relativas $g(r)$-*total* e o efeito da decomposição (*shape*), ver gráfico 5. O efeito *desempenho* é dado pela diferença entre as duas. Para os dados agregados de Minas Gerais, o efeito desigualdade, em geral, não é elevado, ou seja, nos dados agregados do estado, o efeito de *desempenho* é preponderante. No entanto, como pode ser observado na tabela 6, o efeito *desigualdade* é muito importante para escolas e municípios.

A Tabela 6 abaixo apresenta a decomposição do efeito *desigualdade* do *IDR*. São os valores do *IDR-desigualdade* sobre o *IDR-total*. A *desigualdade* é maior para nível de agregação menor (escolas) e, em geral, a *desigualdade* se torna maior ao longo dos anos. A exceção é o ano de 2009 em leitura. Na tabela abaixo, as SRE's são as subregionais de ensino, ao todo 46, os pólos são agrupamentos destas regionais. Minas Gerais possui seis pólos: Central, Norte, Zona da Mata, Sul, Triângulo e Vale do Aço.

Gráfico 5 – Decomposição para a função g(r). Efeito Total e Efeito Desigualdade.



Fonte: *Construção com os dados do PROEB-MG, leitura, 2006 a 2009.*

Tabela 6 – Efeitos da decomposição IDR-Desigualdade de Matemática e Leitura.

	2006	2007	2008	2009
Matemática				
Escolas	61.57%	65.07%	75.58%	80.38%
Municípios	14.67%	13.37%	23.88%	32.24%
SRE's	0.45%	0.73%	4.42%	7.43%
Polos	0.14%	0.35%	1.33%	3.00%
MG	-14.63%	-13.12%	-5.50%	8.02%
Leitura				
Escolas	79.63%	82.39%	94.02%	76.39%
Municípios	29.65%	28.49%	55.95%	26.80%
SRE's	2.77%	5.17%	21.43%	14.39%
Polos	2.03%	3.26%	13.16%	11.36%
MG	7.28%	12.96%	31.32%	100.00%

Fonte: *Construção com os dados do PROEB-MG, 2006 a 2009.*

Para o ano de 2009 em leitura os resultados indicam que as observações praticamente alcançaram a referência ideal imposta, vide o porcentual de 100% na decomposição de MG. Lembrando que a média da prova de português se elevou muito de 2008 para 2009 (de 200.30 para 238.80), e que a média do referencial é de 221.53. Nesse caso em que as duas funções estão muito próximas, o *IDR* se torna muito próximo de zero e dessa maneira se torna mais difícil fazer a decomposição, pois as discrepâncias entre o *IDR-desigualdade* e o *IDR-total* se tornam magníficas.

No seu sentido prático é muito importante ressaltar qual seria a função de referência para essa disciplina, e se os valores com que os educadores trabalham até o momento são de fato exigentes ou podem ser alcançados sem maiores dificuldades pelos alunos. Pela conceituação adotada, temos que em leitura-2009 não houve geração de mais desigualdades, mas observou-se ainda grande discrepância entre os níveis, mais desigualdade em escolas e municípios e menos entre SRE's e pólos, e o porcentual ainda elevado.

Importante acrescentar que no cálculo dos três índices, em especial do *IDR*, há uma nova medição da *desigualdade* escolar. Lembrando os resultados do índice de Gini da Tabela 3, poder-se-ia acreditar que a desigualdade estaria reduzindo ano após ano para ambas as disciplinas. Isso pode ser visto no gráfico 6. Ao se comparar com a **igualdade absoluta** (seção 3.1), a desigualdade, de fato, se reduz. Mas essa perspectiva não é adequada para a comparação do ensino porque, conforme já exposto, acredita-se que a dispersão de competências é coisa natural e procuramos obter, portanto, uma **igualdade relativa**. Ao tratarmos desigualdade com Gini, obtemos uma difícil comparação anual, já que a média das notas evolui. O Gini compara patamares de igualdade absoluta em médias diferentes.

Ao propor o *IQE, IPD* e *IDR*[13], o objetivo é decompor os efeitos, o que não seria possível com o índice de Gini e não tão adequado aos objetivos aqui propostos com o índice de Theil.[14] Ao efetuarmos a decomposição enfrentamos uma série de questões de interpretação, significado, proposição e uso das referências. O gráfico 6 abaixo obtém as curvas de Lorenz tomando como exemplo os resultados de matemática. Ao contrário das construções usuais em que a reta de 45º representa a **igualdade absoluta**, a reta de 45º da figura abaixo é uma transformação da acumulada da função de referência, $F(y)$, obtêm-se então as outras curvas de Lorenz em relação a esta. Nesse caso, como o referencial não é a igualdade perfeita, as curvas de Lorenz podem estar abaixo ou acima do referencial (caso que seria impossível na construção usual). Temos que curvas de Lorenz abaixo do referencial são distribuições à esquerda de $f(y)$. Em outras palavras, $f(y)$ possui dominância estocástica de primeira ordem ($F(y) \leq F_p(y_p)$) sobre as observadas $f_p(y_p)$.

A desigualdade de dados desagregados como escolas e municípios, ser maior do que a desigualdade das Subregionais de ensino e regiões pólos é uma informação importante para as políticas públicas. Isso mostra que há escolas e municípios que são mais desiguais entre si, onde a educação e o Sistema Gerador de Desigualdade podem ser trabalhados como casos especiais que merecem atenção e reforço.

A conclusão a seguir desenvolve esse e outros apontamentos, em particular, o ponto acentuado na seção 3, de que os resultados da proficiência são advindos de uma série de características *x*

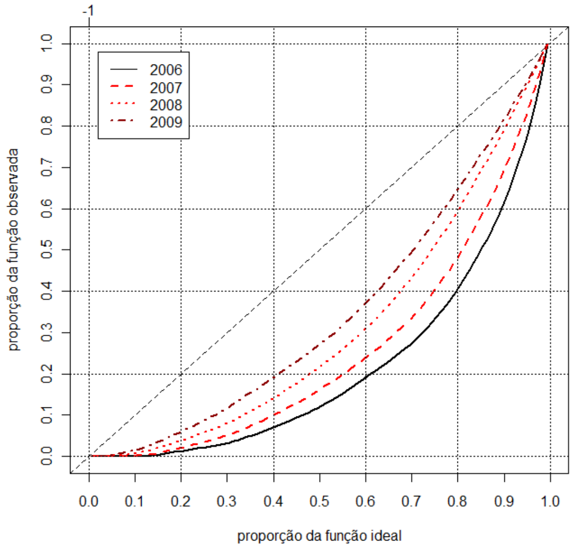
---

13 É possível também decompor os índices *IQE* e *IPD* para os efeitos de *desempenho* e *desigualdade*, embora seja uma decomposição mais grosseira do que a feita pelo *IDR*. Essa decomposição envolve colocar os dados observados com a mesma média do vetor **w** e formato (*shape*) do vetor **p,** obtendo-se assim o $IQE_{\text{desempenho}}$ (ou $IPD_{\text{desempenho}}$) e o efeito desigualdade é obtido pela diferença: $\text{IQE}_{\text{desigualdade}} = \text{IQE}_{\text{Total}} - IQE_{\text{desempenho}}$.

14 Observe que o índice de Theil (τ) guarda fortes relações com o *IDR*: $\tau = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\bar{y}}\cdot\ln\left(\frac{y_i}{\bar{y}}\right)$, a diferença está no referencial que, assim como no Gini, é a média da variável observada ($\bar{y}$). Não são de se admirar as semelhanças com a estatística $D(F_p;F)$ presente em (2), dado que Theil (1967) tirou sua inspiração da teoria da informação muito popular nos anos 50 e 60, com várias referências comuns encontradas em Kullback-Leibler. Uma generalização possível que se pode ter do índice de Theil e o aproximaria da distância de Kullback-Leibler (mas o afastaria da desigualdade absoluta) seria adotar outro referencial que não a média: $\tau_{KL} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_{1,i}}{y_{2,i}}\cdot\ln\left(\frac{y_{1,i}}{y_{2,i}}\right)$, em que y_1 é a observada e y_2 o seu referencial.

(características do aluno, da família, das escolas, etc.) que podem ser mensuráveis com $y = f(x)$. Melhor ainda, cada y é dado por um x sendo que temos na verdade uma distribuição condicional para y: $f(y|x)$. Como enfatizamos anteriormente, é possível também que exista uma diferença entre o *x-ideal* e o *x-observado*, isso pode ser tratado com apuro nas medições de x e com melhores especificações dos modelos que relacionam x à y. Com mais dados das características x, é possível uma série de decomposições por fatores, além de um tratamento mais geral para o desenvolvimento da desigualdade no tempo.

Gráfico 6 – Obtenção das curvas de Lorenz para anos em proporção da função ideal.



Fonte: *Construção com os dados do PROEB-MG, 2006 a 2009.Observação:ao contrário da forma usual a reta de 45º que ora se apresenta indica a função de referência transformada e não a igualitária absoluta.*

6. CONCLUSÕES

Este trabalho dá sequência a uma linha de pesquisa que trata das *desigualdades* educacionais e do *desempenho*, dando continuidade a uma literatura que aborda esses dois conceitos e constrói indicadores usando uma função de referência de desempenho internacional e comparativo. Introduz-se, ademais, uma percepção sobre o *Sistema Gerador de Desigualdade* (*SGD*), em um esforço para se mensurar o quanto o sistema de ensino está distante do ideal de aprendizado para os alunos. Sendo o Brasil um país com fortes desigualdades, propôs-se mensurar o quão distante o sistema está de um ideal internacional de comparação. Para isso, foram construídos três índices: *IQE*, *IPD* e *IDR*, *Índice de Qualidade da Educação*, *Índice Ponderado de Desempenho* e *Índice de Desempenho Relativo*. Os três índices captam o quanto as unidades de análise (escolas, municípios e regiões) estão distantes do referencial pretendido, sendo o terceiro índice o mais apropriado para captar a distância em relação a toda função de maneira completa e decompor o *desempenho* e *desigualdade*.

Os resultados foram obtidos por meio de dados de proficiência de Minas Gerais para 5ª série do ensino básico envolvendo os anos de 2006 a 2009 para matemática e leitura. Os três indicadores apontam para uma melhoria do *desempenho* dos alunos, ou seja, a nota média está melhorando. Por meio da decomposição do *IDR* foi possível analisar quanto da melhoria geral do índice era devida ao aumento na dispersão (alunos melhores tirando notas cada vez maiores). Observou-se que a desigualdade aumentou em todos os quatro níveis de análise: nas escolas, municípios, sub-regionais de ensino e macrorregiões pólos, tanto em matemática quanto em português. Importante ressaltar que ao realizar o índice de Gini tenderíamos a acreditar que a desigualdade estaria diminuindo, mas isso seria ilusório, pois o índice de Gini compara a desigualdade sempre à média atual, sendo impossível decompô-lo para interpretar o que é diferença de média e dispersão ano a ano.

A maior parte do efeito *desigualdade* ocorre quando se compara escolas e municípios, estes níveis de análise possuem observações com bastante dispersão. À medida que o nível de agregação aumenta, torna-se mais difícil notar a *desigualdade relativa*, tornando-se preponderante o efeito de média. No nível agregado do estado, ela se torna praticamente nula em matemática e em torno de 7% em português. Para a decomposição dos índices adotou-se uma função referencial fixa em 2009, essa função referencial foi alcançada e, em certos pontos, ultrapassada pelos dados observados. Isso acrescenta um tema de pesquisa para a construção dos referenciais ideais.

As escolas e municípios podem aprender muito uns com os outros, a dispersão desses níveis de agregação apontam tanto problemas quanto soluções. Escolas com baixa desigualdade e alto desempenho podem servir de comparação para as demais que se encontram em pior situação. Além disso, uma política baseada nos índices propostos evita que a melhoria dos índices se dê em detrimento da desigualdade. Ao adotar metas que levam em conta medidas centrais como média do desempenho ou proporção de alunos que não estão abaixo do básico, corre-se o risco de se ter essas metas cumpridas sem que o quesito da igualdade seja considerado.

No que tange à qualidade do ensino, as definições são variadas, mas as melhores conceituações que se pode fazer disso surgem quando se considera que um sistema de qualidade é o que oferece as melhores oportunidades a todos os alunos, independente do seu *background* familiar, social ou individual. Um ensino de qualidade é aquele que ensina mais e melhor a um maior número de alunos. Os exames de proficiência, agora bastante disseminados no Brasil todo, são uma ferramenta essencial para se alcançar isso.

Demais desenvolvimentos possíveis estão na decomposição dos índices em relação a aspectos importantes na formação da proficiência dos alunos, assim como a proposição de testes para os valores e para as hipóteses formuladas da função transformadora $t(y)$. Entender como a desigualdade avança de acordo com as características “pessoais”, da “família”, “escola” e “sociedade” é extremamente importante para obter mais conclusões dos resultados aqui apresentados. O artigo se concentrou nos dados obtidos para 5ª série no PROEB-MG. Acreditamos que o estudo para outras séries e dos exames nacionais pode acrescentar grandes avanços.

Por fim, se acrescenta que este trabalho procurou uma perspectiva nova de conceituação do problema da *desigualdade* quando aplicada aos alunos. Quando se trata de notas e coeficientes que medem *desempenho*, a igualdade teórica em que todos os alunos tiram a mesma nota (a **igualdade absoluta**) se torna não só impossível, mas um fato para o qual sabemos que a construção da proficiência nunca será suficiente. Acrescenta-se que os alunos possuem desenvolvimentos, gostos e aprendizados diferenciados e colocá-los todos em cima de um referencial pontual seria reducionista e pouco plausível. A ideia básica do artigo se concentrou em adotar uma distribuição ideal, tanto no desempenho, quanto na dispersão. E só a sociedade poderá julgar qual é essa distribuição *ideal* é esta, para o que convocamos os especialistas a nos ajudarem nesta discussão.

7. REFERÊNCIAS

BARBOSA, M. L. de O.; “Desigualdade e Desempenho: uma introdução à sociologia da escola brasileira”. Editora Argvmentvm. 1ª Edição, Belo Horizonte (2009).

CERIANI, L.; VERME, P. “The origins of the Gini índex: extracts from Variabilità e Mutabilità (1912) by Corrado Gini”. *Journal of Economic Inequality*. Vol 9. Jun. (2011).

CUNHA, F.; HECKMAN, J. J.; LOCHNER, L. J.; MASTEROV, D. V. “Interpreting the Evidence on Life Cycle Skill Formation”. In *Handbook of the Economics of Education*, orgs.: Hanushek, E. A. e Welch, F. pp. 697-818. North-Holland, Amsterdam (2006).

DALTON, H. “The Measurement of the Inequality of Incomes”. *The Economic Journal*, Vol. 30, No. 119. Sep. (1920). pp. 348-361.

FUNDAÇÃO JOÃO PINHEIRO. “Índice Mineiro de Responsabilidade Social (IMRS)”. Disponível em: http://www.fjp.gov.br/index.php/servicos/82-servicos-cepp/956-indice-mineiro-de-responsabilidade-social-imrs (último acesso, 23/10/2011).

GINI, C. “The Contributions of Italy to Modern Statistical Methods”. *Journal of The Royal Statistical Society*, Vol. 89, No. 4. Jul (1926). pp 703-724.

GOLGHER, A. B. “Indicadores Alternativos de Educação” em: RIOS-NETO, E. G.; RIANI, J. L. R. (Org.) “*Introdução à Demografia da Educação*” Campinas: ABEP, 2003.

HANDCOCK, M.S.; MORRIS, M. “Relative distribution methods”. Sociological Methodology, Vol. 28, (1998). pp. 53-97.

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

KULLBACK, S.; LEIBLER, R. A. “On Information and Sufficiency” *The Annals of Mathematical Statistics*, Vol. 22, No.1, Mar. (1951). pp. 79-86.

RAWLS, J. “Uma teoria da justiça” Editora: Martins Fontes, 3ª Edição, São Paulo, (2008).

RIANI, J. L. R.; SILVA, V. C.; SOARES, T. M. Repetir ou progredir? Uma análise da eficiência de repetência nas escolas públicas de Minas Gerais. In: *XIV Seminário sobre a Economia Mineira*, Diamantina (2010).

RODRIGUES, C.G. “A Relação entre a expansão do acesso ao ensino e o desempenho escolar no Brasil evidências com base no SAEB período de 1997-2005”. Tese. Cedeplar/UFMG (2009).

SOARES, J. F. “Melhoria do desempenho cognitivo dos alunos do ensino fundamental”. Cadernos de Pesquisa (Fundação Carlos Chagas), v. 37, (2007). pp. 135-160.

SOARES, J. F.; MAROTTA, L. “Desigualdades no Sistema de Ensino Fundamental Brasileiro”. In: VELOSO, F.; PESSÔA, S.; HENRIQUES, R.; GIAMBIAGI F. (Org.). Educação Básica no Brasil. 1 ed. Rio de Janeiro: Campus (2009) v.1. pp. 73-91.

SOARES, J.F.; MAROTTA, L.; DELGADO, V.M.S. "Measuring the Quality and Equity of Basic Education Systems", *XVII International Sociological Association World Congress* - Gothenburg, Sweden, (2010).

THEIL, H. “Economic and Information Theory”. Chicago, Rand McNally and Company, 1967.

THOMAS, V.; WANG, Y.; FAN, X. “Measuring Education Inequality: Gini Coefficients of Education”. *Policy Research Working Paper*, WPS 2525. World Bank (2001).

TODOS PELA EDUCAÇÃO. Nota Técnica Preliminar metodologia para a obtenção das metas finais e parciais. Comissão técnica do compromisso Todos pela Educação. São Paulo, 2007.Disponível em:http://www.todospelaeducacao.org.br//arquivos/arquivo/notatecnica_22112007.pdf

(ultimo acesso em 23/10/2011).