

<https://doi.org/10.18222/eae.v37.12621>

CONSIDERAÇÕES SOBRE A NOTA DO PROF. DELGADO

 RODOLFO HOFFMANN¹

¹ Universidade de São Paulo (USP), Piracicaba-SP, Brasil; hoffmannr@usp.br

RESUMO

Em artigo anterior usei indevidamente “notas” como sinônimo de “proficiências”. Decerto a metodologia de avaliação do aprendizado avançou muito com a Teoria da Resposta ao Item (TRI). O debate sobre a escolha de medidas apropriadas se refere à comparação de distribuições de proficiências. O T de Theil não é um caso especial da divergência de Kullback-Leibler. Sigo defendendo o uso de medidas básicas de tendência central, dispersão e desigualdade para comparar distribuições de proficiências. Poder-se-ia acrescentar uma medida de assimetria. Representações gráficas podem ser esclarecedoras. É possível experimentar outros métodos, mas não faz sentido defender o uso da divergência de Kullback-Leibler para substituir as médias e variâncias ou evitar o uso de medidas comuns de desigualdade. Para ilustrar, analisam-se microdados do Sistema de Avaliação da Educação Básica.

PALAVRAS-CHAVE INDICADORES EDUCACIONAIS • DESIGUALDADES EDUCACIONAIS • DIVERGÊNCIA DE KULLBACK-LEIBLER.

COMO CITAR:

Hoffmann, R. (2026). Considerações sobre a nota do Prof. Delgado. *Estudos em Avaliação Educacional*, 37, Artigo e12621. <https://doi.org/10.18222/eae.v37.12621>

CONSIDERACIONES SOBRE LA NOTA DEL PROF. DELGADO

RESUMEN

En un artículo anterior, utilicé indebidamente "notas" como sinónimo de "competencias". Ciertamente, la metodología de evaluación del aprendizaje avanzó significativamente con la Teoría de la Respuesta al Ítem (TRI). El debate sobre la elección de medidas apropiadas se refiere a la comparación de distribuciones de competencias. El T de Theil no es un caso especial de la divergencia de Kullback-Leibler. Sigo defendiendo el uso de medidas básicas de tendencia central, dispersión y desigualdad para comparar distribuciones de competencias. Podría añadirse una medida de asimetría. Las representaciones gráficas pueden resultar esclarecedoras. Es posible experimentar otros métodos, pero no tiene sentido defender el uso de la divergencia de Kullback-Leibler para sustituir las medias y varianzas o evitar el uso de medidas comunes de desigualdad. Para ilustrarlo, se analizan microdatos del Sistema de Avaliação da Educação Básica [Sistema de Evaluación de la Educación Básica].

PALABRAS CLAVE INDICADORES EDUCATIVOS • DESIGUALDADES EDUCATIVAS • DIVERGENCIA DE KULLBACK-LEIBLER.

CONSIDERATIONS ON PROF. DELGADO'S NOTE

ABSTRACT

In previous article I improperly used "grades" as a synonym for "proficiencies". Certainly, the methodology of learning assessment has advanced significantly with Item Response Theory (IRT). The debate about the choice of appropriate measures refers to the comparison of proficiency distributions. Theil's T is not a special case of the Kullback-Leibler divergence. I continue to advocate the use of basic measures of central tendency, dispersion, and inequality in comparing proficiency distributions. A measure of asymmetry could be added. Graphical representations can be enlightening. Other methods could be tried, but it makes no sense to advocate the use of the Kullback-Leibler divergence to replace means and variances or to avoid the use of common measures of inequality. To illustrate, microdata from Sistema de Avaliação da Educação Básica [Basic Education Assessment System] are analysed.

KEYWORDS EDUCATIONAL INDICATORS • EDUCATIONAL INEQUALITIES • KULLBACK-LEIBLER DIVERGENCE.



Este é um texto de acesso aberto distribuído nos termos da licença Creative Commons do tipo BY.

TRATA-SE DE ANALISAR PROFICIÊNCIAS, E NÃO NOTAS

No meu texto inicial (Hoffmann, 2025) não ressalttei que toda análise estatística deve ser feita utilizando as proficiências, obtidas a partir das respostas às questões das provas usando a Teoria da Resposta ao Item (TRI), e não as notas das provas. Pior: achei que podia usar “notas” como sinônimo de “proficiências”. Não usei o termo técnico apropriado, mas não tenho nenhuma dúvida de que devemos analisar as proficiências. A comparação de notas de provas distintas é inapropriada porque há, em geral, diferentes graus de dificuldade nas questões. O desenvolvimento e aplicação da TRI foi um progresso fundamental. Mas os comentários do Prof. Delgado (2025) chamaram minha atenção sobre uma outra questão: o fato de ser razoável admitir que as proficiências tenham as propriedades de uma escala intervalar (na qual é válido comparar grandezas de intervalos), mesmo que as notas das provas sejam apenas ordinais. Sendo válidos os princípios teóricos da TRI, as proficiências obedecerão, sim, a uma escala intervalar. Mas isso torna ainda mais apropriado usar a média, o desvio padrão e o índice de Gini das proficiências como medidas de tendência central, dispersão e desigualdade, respectivamente.

NÃO CONFUNDIR DISPERSÃO COM DESIGUALDADE

O índice de Gini, o T de Theil e o coeficiente de variação são medidas da desigualdade de uma distribuição, e não da sua dispersão. Considero relevante distinguir desigualdade de dispersão. A variância, o desvio padrão e a diferença absoluta média são medidas de dispersão. Note-se que essas medidas mudam quando se altera a unidade de medida da variável. No caso do coeficiente de variação (a razão entre o desvio padrão e a média) e do índice de Gini (que pode ser definido como a razão entre a diferença absoluta média e o dobro da média), a medida de desigualdade se confunde com a medida da dispersão relativa. Uma medida de desigualdade deve ser insensível à mudança na unidade de medida, garantindo que a desigualdade da distribuição da renda em um país não dependa de a renda das pessoas ser medida em reais ou em dólares.

A divergência de Kullback-Leibler não é afetada pela unidade de medida da variável (x_i), mas ela não é uma medida de desigualdade. Já ficou claro que tanto mudanças na média como mudanças positivas ou negativas na desigualdade levam a divergências de Kullback-Leibler positivas.

Consideremos uma distribuição normal com média 300 e desvio padrão 60 e uma outra distribuição normal com média 200 e desvio padrão 40. Uma vez que tanto a média como o desvio padrão da segunda distribuição são iguais a $2/3$ do valor do parâmetro correspondente na primeira distribuição, a desigualdade é a mesma nas duas distribuições. Mas a divergência de Kullback-Leibler da segunda para a primeira distribuição é igual a 3,34, mostrando que essa divergência não pode ser usada para comparar a desigualdade de duas distribuições.

O T DE THEIL NÃO É UM CASO ESPECIAL DA DIVERGÊNCIA DE KULLBACK-LEIBLER

Considerem-se duas distribuições discretas de uma variável com valores x_i ($i = 1, \dots, n$) com probabilidades $Q(x_i)$ e $P(x_i)$ para cada um dos diferentes valores de x_i . Por definição, a divergência de Kullback-Leibler de Q para P , ou ganho de informação quando se passa da distribuição Q para a distribuição P , é dada por

$$K(P \parallel Q) = \sum_{i=1}^n P(x_i) \log \frac{P(x_i)}{Q(x_i)} \quad (1)$$

É certo que Henry Theil, no livro *Economics and information theory*, publicado em 1967, mostra como se inspirou nessa fórmula para criar o que hoje é conhecido como a medida de desigualdade T de Theil. A fórmula matemática do T de Theil é muito semelhante a (1): se $i = 1, \dots, n$ indica as pessoas de uma população, basta substituir $P(x_i)$ pela participação de cada pessoa na renda total e $Q(x_i)$ por $1/n$, que é a participação de cada pessoa na população. Note-se que as variáveis envolvidas são outras. Então não cabe considerarmos o T de Theil um caso particular ou um caso especial de (1), como afirma e tenta demonstrar Delgado (2025). A expressão matemática da Lei de Coulomb é muito semelhante à expressão matemática da Lei de Gravitação de Newton: basta substituir, nesta última, as massas pelas cargas elétricas e ajustar a constante de proporcionalidade. Mas não me parece razoável concluir daí que a Lei de Coulomb seja um caso particular da Lei de Gravitação de Newton.

O T de Theil é, sim, um caso especial de uma medida geral de desigualdade derivada da teoria da informação. Para uma população com valores de uma variável x_i , com $i = 1, 2, \dots, n$, sendo no μ a sua média e sendo ε um parâmetro, essa medida geral é

$$S = \frac{1}{1 - \varepsilon} \left[1 - \frac{1}{n} \sum_{i=1}^n \left(\frac{x_i}{\mu} \right)^{1-\varepsilon} \right] \quad (2)$$

Pode-se demonstrar que o T de Theil é o caso especial dessa medida geral de desigualdade quando ε tende a zero (ver Hoffmann et al., 2019, seção 6.2). Se estamos avaliando a desigualdade de uma distribuição da renda, x_i é, em geral, a renda de uma pessoa. Sendo x_i a proficiência de um aluno em determinada disciplina, a expressão (2) fornece medidas da desigualdade educacional entre alunos. Se a população de alunos estiver dividida em grupos, a medida S pode ser dividida em uma medida da desigualdade *entre* os grupos e uma medida da desigualdade *dentro* dos grupos.

FUI DEMASIADAMENTE ENFÁTICO?

Sim, minha recomendação de abandonar o uso da divergência de Kullback-Leibler na análise de proficiências (Hoffmann, 2025) é, certamente, um exagero. Fui levado a

ela como reação à ideia de que não se deveriam usar, na análise das proficiências, medidas usuais de desigualdade e que o uso da medida de Kullback-Leibler seria uma ótima maneira de fazer as comparações necessárias com clareza.

Para, mais uma vez, tentar deixar clara minha opinião, considere-se a comparação entre Brasil e Sobral no desempenho em matemática apresentada na Figura 2 de Delgado (2025). Basta olhar a Figura 2-a para concluir que o desempenho é substancialmente melhor em Sobral. Imagine que não fosse apresentada a Figura 2 e fosse informado apenas o valor (0,058) da divergência de Kullback-Leibler. Não ficaríamos sabendo nem se havia diferença na proficiência média das duas distribuições e não poderíamos concluir nada sobre o desempenho em Sobral ser pior ou melhor do que no Brasil. Imagine, agora, ainda, que não fosse apresentada a Figura 2, mas fossem informados os valores da média, do desvio padrão e do índice de Gini das duas distribuições. O índice de Gini não é essencial, pois o coeficiente de variação já seria uma medida da desigualdade facilmente calculada. Não teríamos, então, as informações cruciais para comparar as duas distribuições, usando medidas amplamente conhecidas? No que ajudaria acrescentar a divergência de Kullback-Leibler?

É óbvio que a situação é diferente se o analista é um estatístico com experiência na obtenção e interpretação da densidade relativa e nesse caso certamente não tenho nenhuma objeção ao cálculo da divergência de Kullback-Leibler (usual nesse contexto).

DUAS ILUSTRAÇÕES

A Tabela 1 mostra a evolução de medidas de tendência central, dispersão e desigualdade da distribuição da proficiência em matemática de alunos do 5º ano do ensino fundamental, conforme dados do Sistema de Avaliação da Educação Básica (Saeb) para 2019, 2021 e 2023.

TABELA 1

População de alunos e características da distribuição da proficiência em matemática de alunos do 5º ano do ensino fundamental no Brasil

ANO	POPULAÇÃO (SOMA DOS PESOS)	MÉDIA	MEDIANA	DESVIO PADRÃO	CV (%)	ÍNDICE DE GINI	T DE THEIL
2019	2.876.933	227,9	227,1	49,0	21,5	0,123	0,0235
2021	2.821.664	216,9	214,8	48,9	22,5	0,129	0,0256
2023	2.701.761	224,8	223,2	53,4	23,8	0,136	0,0286

Fonte: Elaboração do autor, com base em microdados do Saeb de 2019, 2021 e 2023 (Inep, 2019, 2021, 2023).

A explicação óbvia para a queda da proficiência média de 2019 a 2021 é a pandemia de covid-19 e a consequente interrupção de aulas regulares. Note-se que, apesar da recuperação de 2021 a 2023, a média em 2023 ainda fica abaixo da média em 2019.

As três medidas de desigualdade – o coeficiente de variação (CV), o índice de Gini e o T de Theil – mostram aumento da desigualdade de 2019 a 2023. É claro que há muita coisa para investigar. A proficiência em língua portuguesa mostra o mesmo tipo de evolução? Há diferenças entre regiões do país? Entre escolas públicas e privadas? O que quero deixar claro é que o uso de medidas usuais de tendência central, dispersão e desigualdade da distribuição das proficiências mostra claramente a evolução de aspectos fundamentais do desempenho dos alunos.

Como ilustração adicional, a Tabela 2 apresenta as características da distribuição da proficiência em matemática de alunos do 5º ano do ensino fundamental, conforme dados do Saeb para 2023 em oito estratos de nível socioeconômico da família do aluno. Essa classificação conforme nível socioeconômico está disponível nos microdados do Saeb (Inep, 2023). Note-se como a proficiência média cresce sistematicamente com o nível socioeconômico, mostrando que parte da desigualdade nas proficiências se deve à desigualdade de oportunidades. As proficiências são condicionadas pelo nível socioeconômico da família, que não depende do próprio aluno. Pode-se verificar, também, que as quatro medidas de desigualdade apresentadas na Tabela 2 tendem a decrescer sistematicamente à medida que cresce o nível socioeconômico.

TABELA 2

População de alunos (soma dos pesos) e características da distribuição da proficiência em matemática de alunos do 5º ano do ensino fundamental no Brasil conforme estratos de nível socioeconômico (SE): média, desvio padrão (s), coeficiente de variação (CV), índice de Gini (G) e medidas T e L de Theil

SE	POPULAÇÃO	MÉDIA	s	CV	G	T	L
I	191.620	192,4	54,0	28,1	0,158	0,0384	0,0385
II	247.485	200,2	49,6	24,8	0,139	0,0301	0,0302
III	374.965	210,4	49,4	23,5	0,133	0,0274	0,0279
IV	514.080	218,4	49,6	22,7	0,129	0,0259	0,0266
V	518.416	228,7	49,7	21,7	0,123	0,0240	0,0249
VI	399.975	239,5	50,0	20,9	0,119	0,0223	0,0233
VII	375.812	251,3	50,3	20,0	0,114	0,0207	0,0219
VIII	79.408	265,1	49,6	18,7	0,105	0,0184	0,0197
Total	2.701.761	224,8	53,4	23,8	0,136	0,0286	0,0297

Fonte: Elaboração do autor, com base em microdados do Saeb de 2023 (Inep, 2023).

Vamos dar destaque às medidas T e L de Theil porque são as únicas medidas em que, quando a população é dividida em categorias, a desigualdade total pode ser decomposta em uma medida da desigualdade *entre* as categorias e uma medida da desigualdade *dentro* das categorias, que é uma média ponderada das medidas de desigualdade dentro de cada categoria. No caso da medida L , os fatores de ponderação

são as participações das categorias na população e, para a medida T , esses fatores de ponderação são as participações das categorias no valor total da variável (total das proficiências). Para os oito estratos de nível socioeconômico considerados na Tabela 2, o T geral é 0,0286, com 0,0252 (88,0% do geral) referente à desigualdade dentro dos estratos e 0,0034 (12,0% do geral) referente à desigualdade entre estratos. O L geral é 0,0297, com 0,0263 (88,4% do geral) referente à desigualdade dentro dos estratos e 0,0034 (11,6% do geral) referente à desigualdade entre os estratos. Embora a parcela referente à desigualdade entre estratos seja relativamente pequena, ela merece atenção especial por ser parte evidente da desigualdade de oportunidades.

SÍNTESE E QUESTÕES PARALELAS

Defendo o uso de medidas bem conhecidas de tendência central, dispersão e desigualdade na comparação entre distribuições de proficiências. Não vejo vantagem em usar a divergência de Kullback-Leibler. O uso das curvas de densidade relativa pode ser útil, mas é claramente uma metodologia menos conhecida e há muitos outros tipos de gráficos interessantes, como, por exemplo, as curvas de quantis.

Quais são as comparações relevantes é outro problema. Comparações entre categorias socioeconômicas são fundamentais para avaliar a desigualdade de oportunidades.

Obviamente não contesto a relevância de ter uma distribuição de proficiências de referência (uma meta que direciona os esforços para melhorar).

Comentando as distribuições relativas e a medida de Kullback-Leibler, Delgado (2025, p. 12) afirma que "tal medida é suficiente no sentido estatístico, isto é, carrega toda a informação necessária para a obtenção dos parâmetros descritores de P e Q ". O assunto é complicado, mas essa afirmativa é falsa. A medida de Kullback-Leibler é suficiente apenas para um teste de razão de verossimilhança destinado a verificar se as duas distribuições são ou não diferentes, e nem mesmo o conhecimento da distribuição relativa permite determinar os parâmetros das duas distribuições.

Se soubermos que ambas são distribuições normais, o conhecimento da distribuição relativa permite estabelecer relações entre as médias e as variâncias de P e Q ; contudo, nem mesmo nesse caso especial é possível determinar os valores numéricos das duas médias e das duas variâncias.

AGRADECIMENTOS

O autor agradece a Josimar Gonçalves de Jesus pelas observações sobre uma versão preliminar deste texto.

REFERÊNCIAS

- Delgado, V. M. S. (2025). O índice Theil-T e a divergência KL: Uma nota complementar ao Prof. R. Hoffmann. *Estudos em Avaliação Educacional*, 36, Debate e12362. <https://doi.org/10.18222/eae.v36.12362>
- Hoffmann, R. (2025). Sobre como medir diferenças de resultados no ensino. *Estudos em Avaliação Educacional*, 36, Artigo e10663. <https://doi.org/10.18222/eae.v36.10663>
- Hoffmann, R., Botassio, D. C., & Jesus, J. G. (2019). *Distribuição de renda: Medidas de desigualdade, pobreza, concentração, segregação e polarização* (2ª ed.). Edusp.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). (2019). Microdados do Saeb 2019. *Inep*. <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). (2021). Microdados Saeb 2021. *Inep*. <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). (2023). Microdados Saeb 2023. *Inep*. <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>
- Theil, H. (1967). *Economics and information theory*. Rand McNally.