

NOTA TÉCNICA:

Conversão da escala do NSE médio das escolas e criação de grupos

Maria Teresa Gonzaga Alves¹

Flávia Pereira Xavier

José Francisco Soares

GAME-FAE/UFMG

1. Conversão da escala

A escala original do nível socioeconômico (NSE) médio das escolas obtida pela estimação do modelo da Teoria de Resposta ao Item (TRI) corresponde aos escores que, tipicamente, variam de -3 a 3. Como esta escala não é de fácil interpretação optou-se por convertê-la em um intervalo que contenha valores positivos e que também tenha interpretação substantiva.

A solução mais óbvia de conversão dos valores empíricos observados nos dados num intervalo de 0 a 10 não é a mais adequada, porque supõe que uma escola teria um NSE médio igual à zero. Esta situação tem probabilidade mínima de ocorrência, pois para tal, todos os alunos deveriam ser idênticos com NSE com o valor mínimo. Da mesma maneira, uma escola com NSE médio igual a 10 seria improvável.

É razoável supor que a escala do NSE médio das escolas varie entre valores teoricamente estabelecidos, considerando que há alunos na base e no topo da hierarquia social. Assim, teoricamente é esperado que os alunos da base da hierarquia não possuam nenhum dos itens de bens de consumo

¹ Contato: mtga@ufmg.br

descritos, tenham renda familiar no nível mais baixo, seus pais tenham escolaridade muito baixa (ou nenhuma escolaridade) e exerçam ocupações temporárias, informais ou de baixo prestígio ocupacional (ou ainda estão desempregados). Da mesma forma, teoricamente é esperado que os alunos no topo da hierarquia social possuam todos os itens de consumo nas situações mais vantajosas, tenham renda familiar na categoria mais alta, bem como pais como escolaridade superior e ocupações de alto prestígio, em posições de comando ou liderança.

Todavia a aplicação deste modelo teórico aos dados empíricos enseja dois problemas. Primeiro, as bases de dados utilizadas neste estudo não foram planejadas para serem analisadas em conjunto. Com registrado no Relatório (Cf. p. 14-19), a compatibilização dos itens foi uma das etapas cruciais do trabalho. Além disso, os itens considerados para descrever a hierarquia social não estão igualmente incluídos em todos os questionários das avaliações realizadas pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). Por isto, os casos teóricos acima descritos são impossíveis de serem observados nas bases de dados. O segundo problema é que, mesmo que os itens fossem os mesmos em todos os questionários, os casos extremos podem simplesmente não aparecer nas bases de dados, haja vista que elas não contemplam todos os alunos brasileiros de educação básica.

A solução adotada para estimar os perfis teóricos posicionados mais à direita e à esquerda na escala original do NSE foi adicionar alguns poucos perfis teóricos de alunos à base de dados para a estimação de uma nova escala com amplitude maior e que possa gerar uma interpretação mais substantiva. Os perfis teóricos são casos extremos simulados que possuem

“respostas” para todos os itens dos questionários nas situações menos favoráveis e mais favoráveis referentes, respectivamente, à base e ao topo da hierarquia social. Eles foram adicionados à base de dados com peso equivalente ao dos casos empíricos mais raros observados na base de dados originais, que são aqueles com apenas uma ocorrência. O peso ponderado desses casos raros é igual a zero², mas, mesmo assim, eles são considerados na aplicação do modelo da TRI.

Feito isto, o NSE foi novamente estimado. O resultado não produziu valores diferentes para os casos reais. Como pode ser observado na Tabela 1, a média e o desvio-padrão dos perfis se mantiveram. Mas, como esperado, com a inclusão dos perfis simulados a amplitude da escala aumentou.

Tabela 1: Estatísticas descritivas do índice de NSE por perfil de alunos empírico e com acréscimo de casos simulados

	N	Mínimo	Máximo	Média	Desvio-Padrão
NSE (Perfil original)	1.622.627	-2,31	2,31	0,06	0,57
NSE (Perfil com casos simulados)	1.622.635	-2,47	2,47	0,06	0,57

Fonte: elaboração própria

Estes valores teóricos são assumidos como estimativas plausíveis dos valores mínimos e máximos do NSE na população de referência (Cf. p. 10-14, Relatório NSE). Utilizando-se a regra de conversibilidade, a transformação da escala original do NSE para uma nova escala é feita pela seguinte operação:

² Os pesos empíricos foram computados pela soma das ocorrências de cada perfil nas avaliações do Inep ponderados para a ordem de um milhão de casos. Isto foi feito para garantir que todas as bases de dados tiveram a mesma influência no processo de cálculo do NSE (Cf. Relatório NSE, p. 19-20).

$$NSE10 = \left[\frac{(\text{valor original do NSE} - \text{valor mínimo teórico})}{(\text{valor máximo teórico} - \text{valor mínimo teórico})} \right] \times 10$$

As estatísticas descritivas do NSE médio das escolas na escala convertida (considerando somente as escolas com no mínimo 15 alunos) estão na Tabela 2. Observa-se que os novos valores possibilitam uma interpretação mais substantiva dos resultados. Assim, embora a escala teórica varie de 0 a 10, os valores empíricos das escolas estão contidos em uma amplitude menor. Poder-se-á supor que, caso se dispusesse de dados sobre todas as escolas do país, com questionários planejados para esta finalidade e completamente preenchidos pelos alunos, a amplitude aumentaria um pouco mais. Porém, ela só atingiria os limites caso houvesse, por um lado, uma escola onde todos os alunos tivessem o NSE igual ao mínimo observado, ou, por outro lado, se a escola tivesse todos os alunos com NSE igual ao máximo.

Tabela 2: Estatísticas descritivas do índice de NSE das escolas na escala 0 a 10 (escolas com no mínimo 15 alunos)

	N	Mínimo	Máximo	Média	Desvio- Padrão
NSE médio da escola na escala 0 a 10	69.906	1,30	9,08	4,87	0,96

Fonte: elaboração própria

Os valores da nova escala descrevem as escolas segundo a sua dependência administrativa conforme o esperado (Gráfico 1).

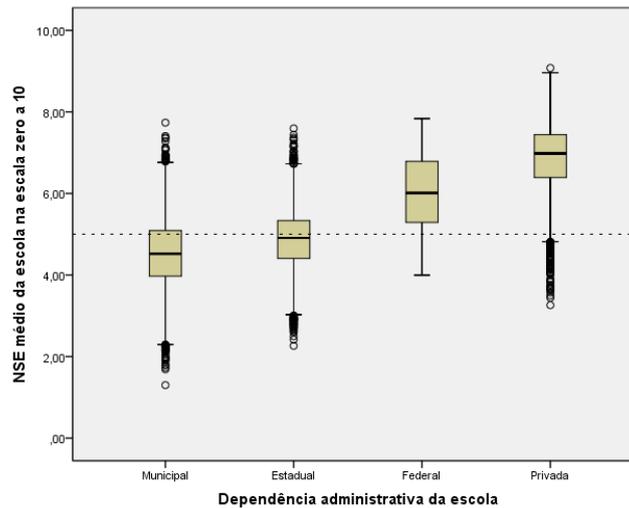


Gráfico 1: Box plot do NSE médio das escolas por rede de ensino

2. Criação de grupos de NSE médio das escolas

Além de converter a escala de NSE médio das escolas numa amplitude teórica a partir de perfis de alunos possíveis, a divisão da escala em grupos torna o uso desta informação também mais fácil. Os agrupamentos de escolas tendem a ser mais estáveis do que a posição específica de uma escola na hierarquia obtida.

A distribuição do NSE médio das escolas não está próxima de uma curva normal (Gráfico 2). De fato, observa-se que há quase uma justaposição de duas curvas normais. Por essa razão a divisão da escala em percentis não se mostra adequada. Em geral, as escolas municipais e estaduais são aquelas com menor NSE enquanto as federais e privadas possuem maior NSE. Além disso, os intervalos dos valores do NSE das redes de ensino se inter cruzam. Ou seja, há escolas privadas com baixo NSE.

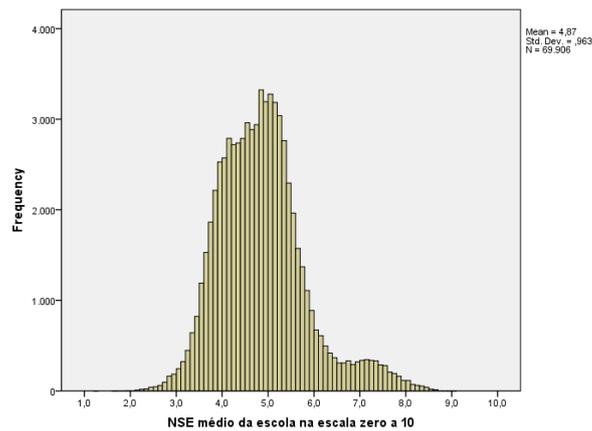


Gráfico 2: Histograma do NSE médio de todas as escolas

Considerando a distribuição do NSE e obedecendo à simetria dos dados, foram criados 7 grupos por meio da análise de *cluster* pelo método de agrupamento não hierárquico designado por *k-means*. O *cluster*, nesse caso, equivale a um conjunto de perfis de escolas no qual cada perfil está mais próximo ao centroide que define o *cluster* do que dos centroides de quaisquer outros *clusters*.

A partir deste procedimento, para tornar os grupos mais interpretáveis, eles foram denominados como: (1) Mais baixo; (2) Baixo; (3) Médio Baixo; (4) Médio; (5) Médio Alto; (6) Alto; e (7) Mais alto

A tabela 3 mostra a frequências dos grupos de escolas por NSE. Como esperado, os grupos extremos possuem percentual menor do que os grupos intermediários.

Tabela 3: Frequência dos Grupos de Escolas por NSE

Grupos	Frequência	Percentual	Percentual acumulado
Mais baixo	2902	4,2	4,2
Baixo	12244	17,5	21,7
Médio baixo	16966	24,3	45,9
Médio	20102	28,8	74,7
Médio alto	11780	16,9	91,5
Alto	3537	5,1	96,6
Mais alto	2375	3,4	100,0
Total	69906	100,0	

Fonte: elaboração própria

A tabela 4 mostra a distribuição dos grupos de NSE por dependência administrativa. Observa-se que as escolas municipais estão concentradas nos grupos “baixo”, “médio baixo” e “médio”. Já as escolas estaduais estão nos grupos “médio baixo”, “médio” e “médio alto”. As federais estão no “médio”, “médio alto” e “alto” e as escolas, enquanto as privadas estão concentradas nos grupos “alto” e “mais alto”. Vale destacar que no grupo “mais baixo” a maioria das escolas é municipal, mas há duas escolas privadas. A princípio, este resultado causou estranhamento. Porém, são duas escolas localizadas no interior da Bahia e que obtiveram o pior desempenho no Enem do estado. Considerando que o desempenho escolar é fortemente correlacionado ao nível socioeconômico, a observação destas duas escolas está dentro de um cenário possível.

Tabela 4: Grupos de NSE por Dependência Administrativa da Escola

	Municipal	Estadual	Federal	Privada	Total
Mais baixo	2522	378	0	2	2902
	6,8%	1,4%	0,0%	0,0%	4,2%
Baixo	9095	3129	1	19	12244
	24,5%	11,8%	0,5%	0,3%	17,5%
Médio baixo	9994	6905	12	55	16966
	26,9%	26,0%	6,5%	0,9%	24,3%
Médio	10151	9679	36	236	20102
	27,3%	36,4%	19,6%	4,0%	28,8%
Médio alto	5046	5875	53	806	11780
	13,6%	22,1%	28,8%	13,5%	16,9%
Alto	383	590	65	2499	3537
	1,0%	2,2%	35,3%	41,9%	5,1%
Mais alto	5	7	17	2346	2375
	0,0%	0,0%	9,2%	39,3%	3,4%
Total	37196	26563	184	5963	69906
	100,0%	100,0%	100,0%	100,0%	100,0%

Fonte: elaboração própria