

# Aplicação do Teste Qui-Quadrado sobre a associação entre proficiência em matemática e fatores socioeconômicos: uma abordagem com dados do SARESP

Camila F. Bassetto<sup>1</sup>  
FCLAR/UNESP, Araraquara, SP

**Resumo.** O presente estudo tem por objetivo apresentar uma aplicação do teste Qui-Quadrado para comprovar a existência de associação entre a proficiência em matemática do aluno com fatores socioeconômicos, tais como escolaridade dos pais e renda familiar. Para tanto, foram consideradas as notas dos alunos matriculados na 3ª série do Ensino Médio da rede estadual de educação, obtidas em matemática no Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (SARESP) de 2019. Os resultados empiricamente obtidos a partir do teste Qui-Quadrado mostram que, ao nível de significância de 1%, a proficiência em matemática está associada às variáveis escolaridade materna e renda familiar.

**Palavras-chave.** Teste Qui-Quadrado, Proficiência em Matemática, Escolaridade da Mãe, Renda Familiar, SARESP.

## 1 Introdução

De acordo com o Relatório Pedagógico do SARESP de 2013, aproximadamente 55% dos alunos da 3ª série do Ensino Médio da rede estadual encontram-se no mais baixo nível de proficiência em matemática, isto é, apresentam habilidades limitadas para solucionar questões propostas para o ano escola [5]. Na avaliação educacional, o desempenho do aluno é considerado a variável indicadora da eficácia deste sistema e é coletado juntamente com informações acerca do contexto escolar e social do indivíduo, características de interesse de muitos estudos, tais como, [7], [1] e [8].

Ao longo da última década, estudos têm revelado evidências empíricas de que o desempenho escolar dos alunos depende não só das habilidades próprias, mas também de fatores que envolvem características sociais, econômicas, culturais e familiares, como mostram os resultados de [12], [9], [1]. Tais autores argumentam que, uma vez que as desigualdades sociais têm implicações diretas sobre a educação, nas avaliações educacionais deve-se considerar que o desempenho do aluno deve ser contextualizado.

Neste sentido, o presente estudo tem por objetivo utilizar o teste Qui-Quadrado sobre a associação entre a proficiência em matemática do aluno com a renda familiar e a escolaridade dos pais. Para a utilização da metodologia proposta, além das notas obtidas pelos alunos da 3ª série do Ensino Médio da rede pública estadual de ensino na edição de 2019 do SARESP, foram consideradas também as respostas dos pais dos alunos às questões sobre renda familiar e sobre o nível de instrução, coletadas mediante aplicação, por esta avaliação, de um questionário socioeconômico.

Este artigo apresenta-se em seções, incluindo esta introdução. A seção dois contém uma breve revisão de literatura sobre estudos dedicados à identificação e compreensão de fatores que atuam

---

<sup>1</sup>camila.bassetto@unesp.br.

sobre o desempenho educacional do aluno. Na seção três, são descritos os níveis de proficiência em matemática, registrados pelo SARESP, as questões sobre renda familiar e escolaridade da mãe e a distribuição dos alunos, por nível de proficiência, conforme as faixas de renda e os níveis de instrução materno. A seção quatro descreve a aplicação do teste Qui-Quadrado para mostrar a associação entre o desempenho em matemática com a escolaridade da mãe e a renda familiar. Ainda nesta seção, os resultados obtidos a partir da análise empírica são apresentados e discutidos. Por fim, na seção cinco, faz-se uma conclusão da presente pesquisa.

## 2 Renda familiar e escolaridade da mãe: impacto no desempenho escolar

A determinante para o aspecto socioeconômico está ligada diretamente com a renda dos pais e, conseqüentemente, com o nível educacional e cultural dos mesmos. Dessa forma, os recursos financeiros que estão à disposição dos pais se convertem em investimento educacional para os filhos, considerando que um menor nível educacional proporciona um menor nível socioeconômico em um longo prazo, como aponta [3].

Considerando a proporção de mães com ensino superior para indicar um elevado status socioeconômico dos alunos que participaram da avaliação, [8] constataram a existência de uma relação positiva forte com o desempenho, isto é, alunos pertencentes a famílias com maior poder aquisitivo exibiram níveis de proficiência mais altos, evidenciando elevada estratificação social no sistema educacional brasileiro. Para a proporção de mães com Ensino Médio, embora não tão acentuada, os resultados mostraram uma relação positiva com o desempenho médio da escola. Autores como [2], caracterizam que a relação entre a escolaridade dos pais e o desempenho do aluno é mais estreita do que a relação entre renda e desempenho. Uma das principais justificativas para isso está no fato de a renda ser variável e o conhecimento que os pais adquiriram é permanente, além do que a escolaridade dos pais é o fator mais relevante para a constituição da renda.

Embora o pai exerça papel significativo na vida do filho, no presente estudo a autora utilizou somente dados associados à mãe, apoiando-se nos estudos de [11] e [10], os quais afirmam que a mãe exerce papel determinante no resultado educacional dos filhos. Tais autores investigaram os determinantes do resultado educacional nos níveis de Ensino Fundamental e Ensino Médio, considerando aspectos familiares dos alunos. Os resultados evidenciaram a importância da educação materna nos indicadores educacionais, o que demonstra a alta estratificação educacional existente no Brasil, pois a trajetória escolar do aluno está bastante relacionada com a sua origem social.

A partir dos estudos descritos anteriormente, é possível observar que tanto a renda familiar como o nível de escolaridade da mãe, além de outros fatores diretamente ligados ao *background* familiar, atuam sobre o desempenho do aluno.

## 3 Distribuição dos alunos de acordo com a renda familiar e a escolaridade materna

O Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (SARESP) passou a vigorar em 1996 e, atualmente, é aplicado no final de cada ciclo, isto é, no 5º, 7º e 9º anos do Ensino Fundamental e na 3ª série do Ensino Médio. Utiliza-se de dois instrumentos de avaliação, sendo o primeiro caracterizado pela aplicação de provas, e o segundo instrumento é o questionário aplicado aos alunos, pais de alunos e escolas (professores, coordenadores e diretores), por meio do qual são extraídas informações sobre características pessoais, contexto socioeconômico e cultural, trajetória escolar, entre outras. O resultado do desempenho dos alunos descreve aquilo que os

mesmos são capazes de fazer em relação às habilidades e competências avaliadas, conforme a Matriz de Referência para Avaliação de Matemática do SARESP<sup>2</sup>. A partir das expectativas de aprendizagem quanto ao conteúdo, competências e habilidades, estabelecidas para cada ano/série da disciplina no Currículo do Estado de São Paulo, os pontos da escala do SARESP são agrupados em quatro níveis: Abaixo do Básico, Básico, Adequado e Avançado, como mostrado na Tabela 1.

Tabela 1: Descrição dos níveis de proficiência em matemática do SARESP.

Níveis de Proficiência	Pontuação	Descrição
Abaixo do Básico (AB)	< 275	Domínio insuficiente dos conteúdos.
Básico (BA)	≥ 275 e < 350	Domínio mínimo dos conteúdos.
Adequado (AD)	≥ 350 e < 400	Domínio pleno dos conteúdos.
Avançado (AV)	> 400	Domínio dos conteúdos acima do requerido.

Uma das variáveis para a qual se pretende mostrar a associação com o desempenho em matemática do aluno é a escolaridade da mãe, presente no questionário socioeconômico aplicado pelo SARESP.

Para a questão *Até que série/nível de ensino a mãe (ou responsável) estudou?* são disponibilizadas as seguintes alternativas: (A) Nunca estudou ou não completou a 4ª série/5º ano (antigo primário); (B) Completou a 4ª série/5º ano, mas não completou a 8ª série/9º ano (antigo ginásio); (C) Completou a 8ª série/9º ano, mas não completou o Ensino Médio (antigo 2º grau); (D) Completou o Ensino Médio, mas não completou o Ensino Superior; (E) Completou o Ensino Superior; (F) Completou a Pós-Graduação (especialização, mestrado ou doutorado); (G) Não sei.

A condição socioeconômica, classificada conforme a renda familiar, também influencia o desempenho do aluno, como apontam os estudos de [1], [9] e [12], entre outros. Presente também no questionário socioeconômico aplicado pelo SARESP, para a questão *Qual é a renda familiar de seu domicílio, ou seja, a soma dos salários e rendimentos?* são dadas as alternativas: (A) Até um salário mínimo (SM) (até R\$ 998,00); (B) De um a dois salários mínimos (de R\$ 998,01 a R\$ 1.996,00); (C) De dois a três salários mínimos (de R\$ 1.996,01 a R\$ 2.994,00); (D) De três a cinco salários mínimos (de R\$ 2.994,01 a R\$ 4.990,00); (E) De cinco a oito salários mínimos (de R\$ 4.990,01 a R\$ 7.984,00); (F) De oito a quinze salários mínimos (de R\$ 7.984,01 a R\$ 14.970,00); (G) Mais de quinze salários mínimos (mais de R\$ 14.970,01); (H) Não sei/não quero responder.

Na edição de 2019, participaram do SARESP 389.660 alunos matriculados na 3ª série do Ensino Médio na rede pública estadual de educação. Deste total, foram descartados os que obtiveram pontuação nula na prova de matemática e, cujos pais não responderam ou optaram pelas alternativas (G) e (H), nas questões sobre a escolaridade da mãe e renda familiar, respectivamente. A amostra contou, então, com 183.008 alunos, distribuídos em 83.048 no nível AB, 85.703 no nível BA, 13.018 no nível AD e 1.239 no nível AV.

Considerando o nível de instrução da mãe, a Tabela 2 apresenta o percentual de alunos conforme a proficiência em matemática, avaliada pelo SARESP no ano de 2019. Observa-se que os maiores percentuais de mães têm o nível de escolaridade correspondente ao Ensino Médio completo (alternativa (D)), em todos os níveis de proficiência. Enquanto para os alunos nos níveis AB e BA, os menores percentuais de mães correspondem à alternativa (F), isto é, não completaram Pós-Graduação (especialização, mestrado ou doutorado), para aqueles classificados nos níveis AD e AV, os menores percentuais estão na alternativa (A) (nunca estudou ou não completou a 4ª série/5º ano (antigo primário)).

Verifica-se também que, à medida que o desempenho do aluno melhora, decrescem os percentuais de mães com os níveis de escolaridade representados pelas alternativas (B) e (C) (aos anos

<sup>2</sup>A Matriz de Referência para Avaliação de Matemática do Ensino Médio é apresentada em [4].

Tabela 2: Valores percentuais, de acordo com o nível de proficiência em mática do SARESP.

Questão	Nível de Proficiência	Respostas						
		A	B	C	D	E	F	G
Escolaridade da mãe	AB	10,60	22,75	18,52	34,59	10,83	2,71	
	BA	7,59	18,78	17,27	41,14	11,70	3,52	
	AD	5,29	15,04	15,24	45,16	13,74	5,52	
	AV	4,76	12,91	14,85	42,70	17,92	6,86	
Renda familiar	AB	22,99	36,46	22,40	12,70	4,03	1,05	0,37
	BA	14,21	34,45	25,66	18,03	5,78	1,52	0,36
	AD	8,54	28,93	27,32	23,92	8,47	2,43	0,40
	AV	7,59	26,23	25,26	25,99	11,14	3,55	0,24

iniciais (antigo ginásio) e anos finais (antigo 2º) do Ensino Fundamental completos, respectivamente). Para o Ensino Superior completo, observa-se que os percentuais aumentam quando o nível de proficiência também aumenta.

Os valores mostrados na Tabela 2 evidenciam um padrão no comportamento do desempenho educacional do aluno conforme o nível de escolaridade da mãe, sugerindo que o nível de proficiência em matemática aumenta à medida que os anos de estudo da mãe também aumentam. Nesse sentido, os valores analisados indicam a existência de uma relação positiva entre proficiência em matemática e escolaridade materna. Assim como mostram os estudos de Riani e Rios-Neto (2008) e Franco e Menezes-Filho (2012), filhos de mães mais escolarizadas possuem melhores desempenhos em matemática, medidos pelos níveis de proficiência, corroborando estudos citados previamente.

No que se refere à renda familiar, a Tabela 2 mostra que o percentual de alunos, com renda familiar de até 1 salário mínimo (SM) e entre 1 e 2 SM, é decrescente conforme os níveis de proficiência aumentam, isto é, as maiores quantidades de alunos nestas faixas de renda pertencem ao nível mais baixo de proficiência, representado por AB, e as menores quantidades concentram-se no nível AV. Nas faixas de renda de 3 a 5 SM, de 5 a 8 SM e de 8 a 15 SM, indicadas, respectivamente, pelas alternativas (D), (E) e (F), os percentuais de alunos exibem comportamento oposto àquele identificado nas faixas de renda de até R\$ 998,00 e entre R\$ 998,01 e R\$ 1.996,00, Tais valores sugerem que, à medida que a renda familiar aumenta, melhora também o desempenho escolar do aluno, uma vez que, faixas de renda mais altas concentram os maiores percentuais de alunos nos níveis AD e AV.

## 4 O Teste Qui-Quadrado aplicado os dados do SARESP

O teste Qui-Quadrado foi desenvolvido por Karl Pearson em 1899, sendo considerado uma técnica ampla, vantajosa e eficaz para a análise de variáveis qualitativas envolvendo duas ou mais categorias. Podem ser utilizados por meio de três testes distintos, isto é, de aderência ou ajustamento, de comparação e de associação [6].

Pra a utilização do teste Qui-Quadrado devem ser estabelecidas duas hipóteses contraditórias acerca da questão a ser respondida ou testada, como apontam [6] e [13]. Tais hipóteses consistem na hipótese de nulidade ( $H_0$ ), a qual afirma que não há distinção entre os grupos testados, e a hipótese alternativa ( $H_1$ ), a qual contradiz a  $H_0$  e que, geralmente, corresponde ao que, de fato, o pesquisador tem por objetivo mostrar ([13]).

No presente estudo, faz-se uso do teste Qui-Quadrado de associação, também denominado teste Qui-Quadrado de independência, dado que o mesmo é empregado sempre que o objetivo é testar a

correlação entre variáveis qualitativas, ou categóricas. Para a realização do teste Qui-Quadrado de associação é necessário separar valores observados (O), coletados via análise descritiva dos dados, e os valores esperados (E), calculados por meio da equação (1)

$$E = \frac{TC \cdot TL}{TG}. \tag{1}$$

Na expressão (1), (TL) e (TC) representam os totais, respectivamente, da linha e da coluna correspondentes, e (TG) é o total geral, dado pela soma dos totais das linhas e das colunas [6].

Após obtido o valor de (E), calcula-se o valor do teste Qui-Quadrado de associação, representado por  $\chi^2$ , dado pela equação (2)

$$\chi^2_{calculado} = \sum \frac{(O - E)^2}{E}. \tag{2}$$

O número de graus de liberdade (gl) é obtido fazendo  $(L - 1) \cdot (C - 1)$ . Considerando um nível de significância de 1% ( $\alpha = 0,01$ ), e o desempenho em matemática associado à escolaridade da mãe, são estabelecidas as seguintes hipóteses  $H_0$  : Não há associação entre o desempenho em matemática e a escolaridade da mãe *versus*  $H_1$  : Há associação entre o desempenho em matemática e a escolaridade da mãe.

Os valores observados e esperados para a escolaridade da mãe são apresentados na Tabela 3.

Tabela 3: Valores observados e valores esperados para a escolaridade da mãe.

Proficiência	Respostas						
	A	B	C	D	E	F	
Observados	AB	8.801	18.893	15.381	28.729	8.995	2.249
	BA	6.509	16.092	14.802	35.257	10.030	3.013
	AD	689	1.958	1.984	5.879	1.789	719
	AV	59	160	184	529	222	85
Esperados	AB	7.287,0	16.837,1	14.680,7	31.944,4	9.546,0	2.752,7
	BA	7.520,0	17.375,4	15.150,0	32.965,6	9.851,2	2.840,7
	AD	1.142,3	2.639,3	2.301,2	5.007,4	1.496,4	431,5
	AV	108,7	251,2	219,0	476,6	142,4	41,1

Para o desempenho associado à renda familiar, são estabelecidas as hipóteses  $H_0$  : Há associação entre o desempenho em matemática e a renda familiar *versus*  $H_1$  : Não há associação entre o desempenho em matemática e a renda familiar.

Os valores observados e esperados para a renda familiar são apresentados na Tabela 4.

A Tabela 5 mostra o resultado do teste Qui-Quadrado de associação para o desempenho em matemática e as variáveis escolaridade da mãe e renda familiar. É possível observar que o p-valor foi muito menor que 0,01 (praticamente igual a 0) considerando as duas variáveis, indicando que a hipótese de nulidade ( $H_0$ ), ou seja, de que o desempenho em matemática é independente da escolaridade da mãe e da renda familiar, deve ser rejeitada.

A independência entre o desempenho em matemática com a escolaridade da mãe e a renda familiar pode ser comprovada também por meio da comparação entre os valores  $\chi^2_{calculado}$  e  $\chi^2_{tabelado}$ . Como mostrado na Tabela 5,  $\chi^2_{calculado} > \chi^2_{tabelado}$ , devendo-se, portanto, rejeitar  $H_0$ .

O resultado do teste Qui-Quadrado sugere, então, considerar que, tanto o nível de instrução mãe quanto a condição socioeconômica familiar influenciam o desempenho em matemática do aluno no SARESP de 2019. A associação existente entre tais variáveis evidencia que, à medida em que

Tabela 4: Valores observados e valores esperados para a renda familiar.

Proficiência		Respostas						
		A	B	C	D	E	F	G
Observados	AB	19.093	30.278	18.600	10.546	3347	875	309
	BA	12.175	29.525	21.991	15.455	4.950	1.301	306
	AD	1.112	3.766	3.556	3.114	1.102	316	52
	AV	94	325	313	322	138	44	3
Esperados	AB	14.736,5	28.994,7	20.175,7	13.358,3	4.327,8	1.150,8	304,0
	BA	15.207,6	29.921,7	20.820,7	13.785,4	4.466,2	1.187,6	313,8
	AD	2.310,0	4.545,0	3.162,6	2.094,0	678,4	180,4	47,7
	AV	219,9	432,6	301,0	199,3	64,6	17,2	4,5

Tabela 5: Teste Qui-Quadrado.

Variável	$\chi^2_{calculado}$	$\chi^2_{tabelado}$	grau de liberdade (gl)	p-valor ( $\alpha = 0,01$ )
Escolaridade da mãe	2.416,9	30,578	15	0,0
Renda familiar	5.258,0	34,805	18	0,0

os anos de estudo da mãe e a renda familiar aumentam, melhores resultados são alcançados no desempenho em matemática pelo aluno da 3ª série do Ensino Médio da rede pública estadual de ensino.

## 5 Conclusões

No presente estudo, o teste Qui-Quadrado foi utilizado para mostrar que o desempenho em matemática do aluno, no SARESP, tem associação com a escolaridade da mãe e com a renda familiar. Para tanto, foram considerados os níveis de proficiência em matemática, no qual o aluno da 3ª série do Ensino Médio da rede pública estadual de ensino se classificou na edição de 2019 do SARESP. Concomitantemente, fez-se uso das respostas das mães dadas às questões sobre o nível de escolaridade e da renda familiar, ambas contidas no questionário socioeconômico aplicado por esta avaliação. A análise empírica realizada, a partir da organização dos percentuais de aluno, de acordo com o nível de proficiência em matemática obtido no SARESP de 2019, considerando as respostas sobre o nível de escolaridade da mãe e da renda familiar, e da empregabilidade do teste Qui-Quadrado, permitiu comprovar que existe associação entre o desempenho do alunos e fatores socioeconômicos, especificamente, com aqueles abordados no presente estudo.

Os resultados obtidos mostraram que, à medida que o nível de proficiência em matemática do aluno aumenta, crescem também os percentuais de mães com mais anos de estudo, isto é, mães com níveis mais altos de instrução tiveram filhos com melhores desempenhos na edição do SARESP de 2019. O mesmo foi observado para a renda familiar. Enquanto alunos classificados no nível mais baixo de proficiência em matemática pertencem a famílias com rendas mensais de até 2 SM, aqueles com proficiências correspondentes aos níveis AD e AV têm melhores condições socioeconômicas, as quais são evidenciadas por rendas mensais de até 15 SM.

Portanto, a aplicação do teste Qui-Quadrado foi útil para comprovar que existe associação entre o desempenho em matemática com a escolaridade da mãe e com a renda familiar. Tal conclusão segue na mesma direção dos resultados de autores como [12] e [1], os quais afirmam que a condição socioeconômica do aluno influencia o seu desempenho, ou seja, que alunos em condições precárias

têm dificuldade de aprendizagem. Este fato indica a necessidade de um refinamento analítico no sentido de se responder como determinantes associados à família favorecem ou interferem no desempenho e por quais caminhos isso se realiza. A continuidade desta pesquisa se torna fundamental para a investigação de outras variáveis relacionadas à condição socioeconômica e, assim, estender a compreensão sobre os elementos que interferem e/ou favorecem um melhor desempenho na escolarização básica.

## Referências

- [1] Alves, M. T. G. e Soares, J. F. O efeito das escolas no aprendizado dos alunos: um estudo com dados longitudinais no Ensino Fundamental. *Educação e Pesquisa*, 34(3):527–544, 2008. DOI: 10.1590/S1517-97022008000300008.
- [2] Barros, R. P. de; Mendonça, R. O impacto de gestão escolar sobre o desempenho educacional. Washington: BID, 1997.
- [3] Barros, R. P. de; Mendonça, R.; Santos, D. D. e Quintaes, G. Determinantes do desempenho educacional no Brasil, Texto para Discussão N<sup>o</sup> 834, Instituto de Pesquisa Econômica e Aplicada (IPEA), 2001.
- [4] Brasil. Secretaria da Educação do Estado de São Paulo Fundação para o Desenvolvimento da Educação – FRD. Matrizes de Referência para o Avaliação: Documento Básico – Ensino Fundamental e Médio, São Paulo, Secretaria da Educação do Estado de São Paulo, 2009.
- [5] BRASIL. Secretaria da Educação do Estado de São Paulo Fundação para o Desenvolvimento da Educação – FRD. Relatório pedagógico SARESP 2013 – Matemática, São Paulo, Secretaria da Educação do Estado de São Paulo, 2013.
- [6] Callegari-Jacques, S. M. Bioestatística: princípios e aplicações. Artmed, 2007.
- [7] Ferrão, M. E.; Beltrão, K. I.; Fernandes, C.; Santos, D.; Suárez, M. e Andrade, A. O SAEB – Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz, *Revista Brasileira de Estudos de População*, 18(1/2):111–130, 2001. DOI: 10.1590/S1413-24782001000200010.
- [8] Franco, A. M. P. e Menezes-Filho, N. Uma análise de rankings de escolas brasileiras com dados do SAEB. *Estudos Econômicos*, 42(2):263–283, 2012. DOI: 10.1590/S0101-41612012000200002.
- [9] Jesus, G. R. e Laros, J. A. Eficácia escolar: Regressão Multinível com Dados de Avaliação em Larga Escala. *Avaliação Psicológica*, 3(2):93–106, 2004. ISSN-e: 2175-3431.
- [10] Riani, J. de L. R. e Rios-Neto, E. L. G. Background familiar versus perfil escolar do município: qual possui maior impacto no resultado educacional dos alunos brasileiros? *Revista Brasileira de Estudos de População*, 25(2):251–269, 2008. DOI: 10.1590/S0102-30982008000200004.
- [11] Rios-Neto, E. L. G.; César, C. C. e Riani, J. de L. R. Estratificação educacional e progressão escolar por série no Brasil. *Pesquisa e Planejamento Econômico*, 32(3):395–415, 2002. ISSN: 2237-2091.
- [12] Soares, T. M. e Mendonça, M. C. M. Construção de um modelo de regressão hierárquico para os dados do Simave-2000. *Pesquisa Operacional, Sociedade Brasileira de Pesquisa Operacional*, 23(3):421–441, 2003. DOI: 10.1590/S0101-74382003000300003 .
- [13] Vieira, S. Introdução à Bioestatística. Campus-1991, 2011.