



Allan Victor Almeida Faria, Hermes Jasper Winarski, Kassyano Kevyn Andrade
de Souza

Relatório de Pesquisa

Brasília, DF

Dezembro, 2020



Allan Victor Almeida Faria, Hermes Jasper Winarski, Kassyano Kevyn Andrade
de Souza

Relatório de Pesquisa

Relatório de Pesquisa do projeto Saeb 2017.

Universidade de Brasília (UnB)
Instituto de Ciências Exatas (IE)
Departamento de Estatística (DE)

Brasília, DF
Dezembro, 2020

Resumo

A fim de compreender o comportamento do tempo de afazeres domésticos e o desempenho de estudantes que realizaram a Prova Brasil, foi realizado uma série de testes estatísticos para saber como outras variáveis demográficas influenciam nesses fatores. Foram utilizadas como variáveis explicativas: Raça/Cor, Escolaridade da Mãe, Sexo e Localização.

Diante disso, foi possível observar uma diferença significativa entre o tempo de afazeres domésticos de acordo com o sexo, em que estudantes do sexo feminino realizam mais tempo de atividades domiciliar. A escolaridade da mãe também foi um fator existente para decretar uma diferença no tempo de atividades domésticas, já que estudantes que não sabem o nível de escolaridade da mãe trabalham menos em casa do que os outros grupos.

Como fatores responsáveis pelo desempenho, foi observada a diferença entre estudantes de fenótipo branco com o restante das raças/cores, em que esse grupo demonstrou um desempenho maior nas provas. O nível de Escolaridade da Mãe também indica certa tendência a aumento nas notas de acordo com a o grau de formação da figura materna dos estudantes. O último elemento entre os estudados que reflete nas notas é a localização, em que estudantes da região urbana tiveram desempenho superior aos da zona rural.

Palavras-chaves: 1. Saeb 2017. 2. Amostra. 3. Pesquisa.

Lista de ilustrações

Figura 1 – Proporção dos sexos por período de tempo em afazeres domésticos por parte dos alunos.	12
Figura 2 – Proporção total do nível de escolaridade da mãe com base nos períodos de tempo de afazeres domésticos por parte dos alunos.	13
Figura 3 – Distribuições das somas das notas com base na raça/cor dos alunos. . .	16
Figura 4 – Distribuições empíricas das somas das notas com base nas localizações das das escolas dos alunos.	17
Figura 5 – Distribuições das somas das notas com base nas escolaridades das mães dos alunos	18

Lista de tabelas

Tabela 1 – Testes de igualdade na variabilidade sobre as relações com o tempo de afazeres domésticos por parte dos alunos.	14
Tabela 2 – Comparações dois a dois entre as ordens das posições sobre os tempos de afazeres domésticos com base na escolaridade das mães dos alunos .	15
Tabela 3 – Testes para as relações com soma das notas dos alunos.	19
Tabela 4 – Comparações dois a dois entre as médias sobre a soma das notas com base na raça/cor dos alunos.	20
Tabela 5 – Comparações dois a dois das notas entre os alunos com base na escolaridade da mãe	21

Lista de abreviaturas e siglas

INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
SAEB	Sistema de Avaliação da Educação Básica

Lista de símbolos

σ Letra grega minúscula sigma

μ Letra grega minúscula mu

Sumário

	Introdução	8
I	METODOLOGIA	9
1	METODOLOGIA	10
II	RESULTADOS	11
2	COMPARAÇÕES	12
2.1	Tempo em afazeres domésticos	12
2.2	Notas	16
3	CONCLUSÃO	22
	REFERÊNCIAS	23
	ANEXOS	24
	ANEXO A – AMOSTRA	25



Introdução

Este documento tem como objetivo analisar os fatores sociais de alunos brasileiros do 9º ano de 2017, sobre os quais avalia possibilidades de diferenças no desempenho do aprendizado básico destes. Para a análise, serão utilizados dados do Sistema de Avaliação da Educação Básica (SAEB) de 2017 divulgado pelo [INEP \(2018\)](#). O banco de dados utilizado no estudo é constituído por dados de estudantes do 9º ano que realizaram a Prova Brasil no ano de 2017.

O SAEB avalia o desempenho em Matemática e Língua Portuguesa por meio de testes e questionários sobre fatores sociais de alunos, aplicados bienalmente em redes públicas e amostras de escolas privadas sobre turmas do 5º e 9º anos do ensino fundamental e 3º ano do ensino médio.

Esta análise parte de uma amostragem aleatória simples de 5.271 alunos deste banco de dados, e relaciona fatores como raça/cor, sexo dos alunos, localizações das escolas e escolaridade da mãe destes, com as variáveis a serem explicadas como a soma destas notas e o tempo de afazeres domésticos realizados por dia com base nos alunos.

A aplicação dos testes de hipóteses para a análise da amostra, bem como dos gráficos (para a análise descritiva), utiliza ferramentas do software R nas versões 3.6.3 e 4.0.3. Parte do processamento dos dados foi feito com o software PYTHON versão 3.7.3.¹

¹ Pacotes externos usados para a manipulação dos dados:
R: *tidyverse*, *data.table*, *reshape2*, *patchwork*, *EnvStats*, *PMCMR*;
Python: *pandas*

Parte I

Metodologia



1 Metodologia

Como variáveis explicativas para a soma das notas, o estudo relaciona o sexo do aluno, a raça/cor, a escolaridade da mãe e as localizações das escolas, e avalia o quanto essas variáveis influenciam no desempenho na Prova Brasil. Para avaliar estas relações, um estudo prévio foi realizado com amostras de tamanhos 30, 50 e 100 dos 5.271 alunos em que, com os testes [Anderson e Darling \(1954\)](#), [Shapiro e Wilk \(1965\)](#) e [Shapiro e Francia \(1972\)](#), se concluiu que as distribuições das notas são normais. Este resultado será utilizado para atender os testes paramétricos que pressupõem normalidade dos dados. Os testes a serem aplicados serão o teste ANOVA Fisher por [Fisher \(1928\)](#) para avaliar os fatores com mais de 2 categorias e o teste t-Student para as comparações dois a dois ([O'CONNOR; ROBERTSON; EDMUND, 1908](#)). Estes testes utilizam das médias amostrais (\bar{x}) de cada grupo para avaliar se as distâncias entre as médias populacionais (μ) são significativas entre eles, no qual previamente avalia as variâncias amostrais (S^2) para verificar se as variâncias populacionais (σ^2) destes são iguais (Teste B), através do teste proposto por [Bartlett \(1954\)](#).

Quanto ao tempo gasto com afazeres domésticos, estas mesmas variáveis são exploradas, com exceção das localizações das escolas, para observar se há indícios de diferenças sociais entre a disponibilidade de tempo em casa para outras possíveis tarefas na formação do aprendizado básico. Para estas relações, os testes estatísticos não paramétricos são apropriados, e utiliza-se o teste de [Kruskal e Wallis \(1952\)](#) para fatores com mais de duas categorias (Teste K), e o teste de [Mann e Whitney \(1947\)](#) para a comparação dois a dois das categorias (Teste W). Estes testes avaliam as distribuições das informações com base na posição, para verificar se as distâncias entre as categorias são significativas.

Para avaliar os resultados dos testes, foi proposto o uso da correção de [Bonferroni \(1936\)](#) para os testes com mais de duas categorias. Se houver evidências para rejeitar igualdade destas, a comparação dois a dois é efetuada e a correção é utilizada. Esta correção é sobre o p-valor, que é avaliado em uma escala de significância, adotado por este estudo como uma confiança de 95%, no que diz respeito a aceitar a hipótese nula H_0 .

Parte II

Resultados

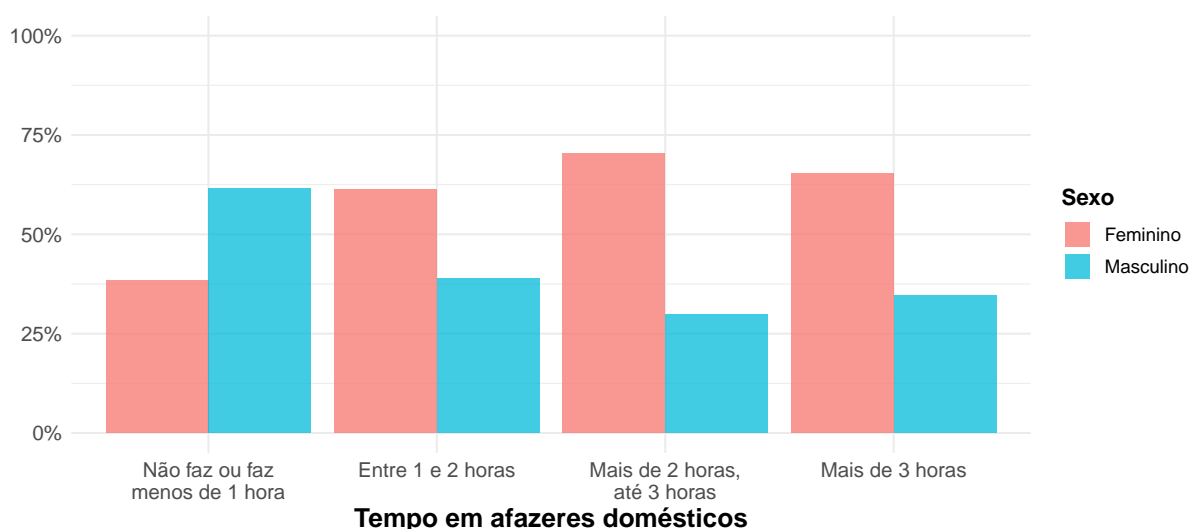


2 Comparações

2.1 Tempo em afazeres domésticos

O tempo gasto com afazeres domésticos pode privar o aluno de exercer o estudo do ensino básico. Foram feitas análises sociais com base nas diferenças entre os períodos de tempo gastos diariamente nestas atividades.

Figura 1 – Proporção dos sexos por período de tempo em afazeres domésticos por parte dos alunos.



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

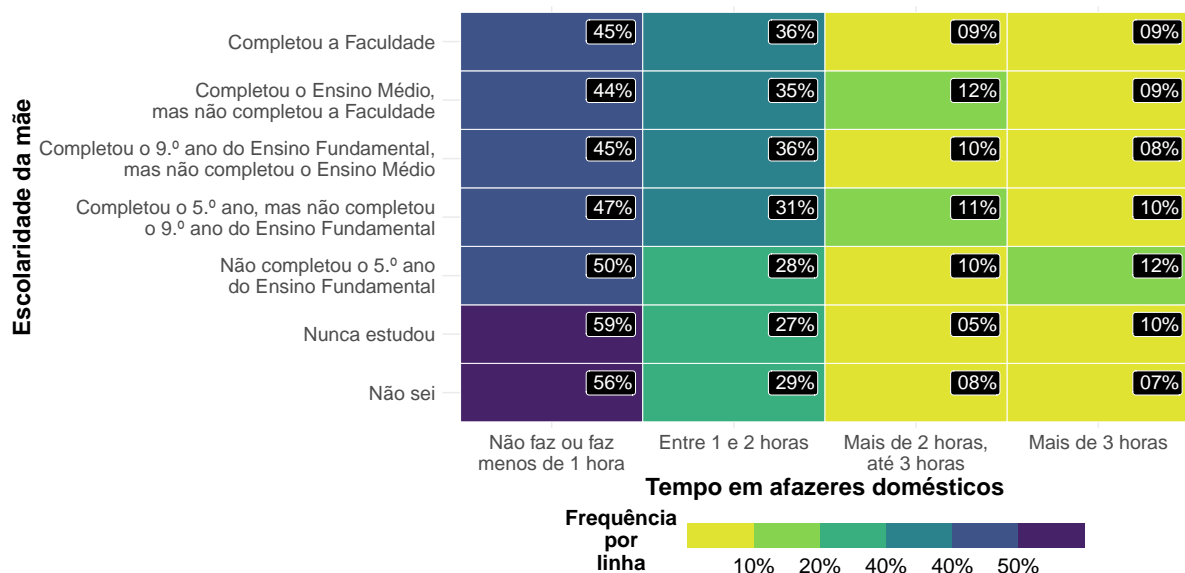
A [Figura 1](#) apresenta como se distribuem estes períodos em relação ao sexo do aluno, no qual estudantes do sexo feminino tendem a gastar mais tempo com atividades domésticas. O único período em que o sexo feminino teve menos representatividade foi a categoria que menos tempo diário é gasto com atividades domésticas, “Não faz ou faz menos de 1 hora”.

Para os períodos em que pelo menos uma hora por dia é gasta, o sexo feminino representa no mínimo 60% de todos os indivíduos, o que pode indicar que a proporção de



estudantes do sexo masculino possui mais disponibilidade de tempo em casa para estudos.

Figura 2 – Proporção total do nível de escolaridade da mãe com base nos períodos de tempo de afazeres domésticos por parte dos alunos.



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

Outras observações sobre tempo gasto podem ser feitas ao observar a [Figura 2](#), que mostra o tempo gasto pelos alunos com atividades domésticas pra cada nível de escolaridade da mãe. Ao analisar a proporção de alunos que exercem estas atividade, sobre o total de alunos para cada nível desta escolaridade, pelo menos 70% se localiza entre o período de tempo que não fazem ou fazem até 2 horas.

Nos testes estatísticos, que avaliam a hipótese de igualdade das distribuições com base nesses afazeres ([Tabela 1](#)) e com a confiança de 95%, a variável sexo obteve evidências significativas para afirmar que o tempo médio não é igual entre os sexos, no qual o sexo feminino possui proporções superiores em períodos de tempo maiores nestas atividades.

Ao testar a variável raça/cor dos alunos sob a mesma hipótese, não foram obtidas evidências significativas de diferenças, em média, entre as raças/cores no tempo gasto nestes afazeres.

O teste realizado com base na escolaridade da mãe obteve evidências significativas



Tabela 1 – Testes de igualdade na variabilidade sobre as relações com o tempo de afazeres domésticos por parte dos alunos.

Teste	H_0	P-valor	Decisão de H_0 (95%)
K	$\mu_{Raça/Cor}$ iguais	0.369	Aceita
K	$\mu_{Esc(mãe)}$ iguais	Aprox. 0	Rejeita
W	$\mu_M = \mu_F$ iguais	Aprox. 0	Rejeita

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples.

Anotações: Os subíndices M e F referem-se, respectivamente, aos sexos Masculino e Feminino dos alunos. Esc (mãe) diz respeito à escolaridade da mãe destes. O Aprox. 0 refere-se a um número muito pequeno considerado por este estudo aproximadamente zero.

sobre a existência de diferença no tempo gasto com afazeres domésticos. Ao efetuar testes pareados para cada nível escolar da mãe com a confiança de 95% ([Tabela 2](#)), a proporção de alunos que gastam tempo nestes afazeres obteve diferenças nas distribuições apenas para aqueles que responderam não saber a escolaridade mãe com os outros níveis desta escolaridade. Apenas aqueles cuja mãe nunca estudou obtiveram igualdade com os que não souberam responder.



Tabela 2 – Comparações dois a dois entre as ordens das posições sobre os tempos de afazeres domésticos com base na escolaridade das mães dos alunos

Comparações	P-valor	Evidência (RA 95%)
Não sabe = Nunca estudou	1.0000	Iguais
Não sabe = Incompleto 5.º ano do EF	0.0078	Desiguais
Não sabe = Completou 5.º ano do EF	0.0005	Desiguais
Não sabe = Completou 9.º ano do EF	0.0001	Desiguais
Não sabe = Completou EM	Aprox. 0	Desiguais
Não sabe = Completou Faculdade	0.0011	Desiguais
Nunca estudou = Incompleto 5.º ano do EF	1.0000	Iguais
Nunca estudou = Completou 5.º ano do EF	0.5598	Iguais
Nunca estudou = Completou 9.º ano do EF	0.4165	Iguais
Nunca estudou = Completou EM	0.1114	Iguais
Nunca estudou = Completou Faculdade	0.4707	Iguais
Incompleto 5.º ano do EF = Completou 5.º ano do EF	1.0000	Iguais
Incompleto 5.º ano do EF = Completou 9.º ano do EF	1.0000	Iguais
Incompleto 5.º ano do EF = Completou EM	1.0000	Iguais
Incompleto 5.º ano do EF = Completou Faculdade	1.0000	Iguais
Completo 5.º ano do EF = Completou 9.º ano do EF	1.0000	Iguais
Completo 5.º ano do EF = Completou EM	1.0000	Iguais
Completo 5.º ano do EF = Completou Faculdade	1.0000	Iguais
Completo 9.º ano do EF = Completou EM	1.0000	Iguais
Completo 9.º ano do EF = Completou Faculdade	1.0000	Iguais
Completou EM = Completou Faculdade	1.0000	Iguais

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples. Teste de Wilcoxon dois a dois efetuado com a correção de Bonferroni no P-valor.

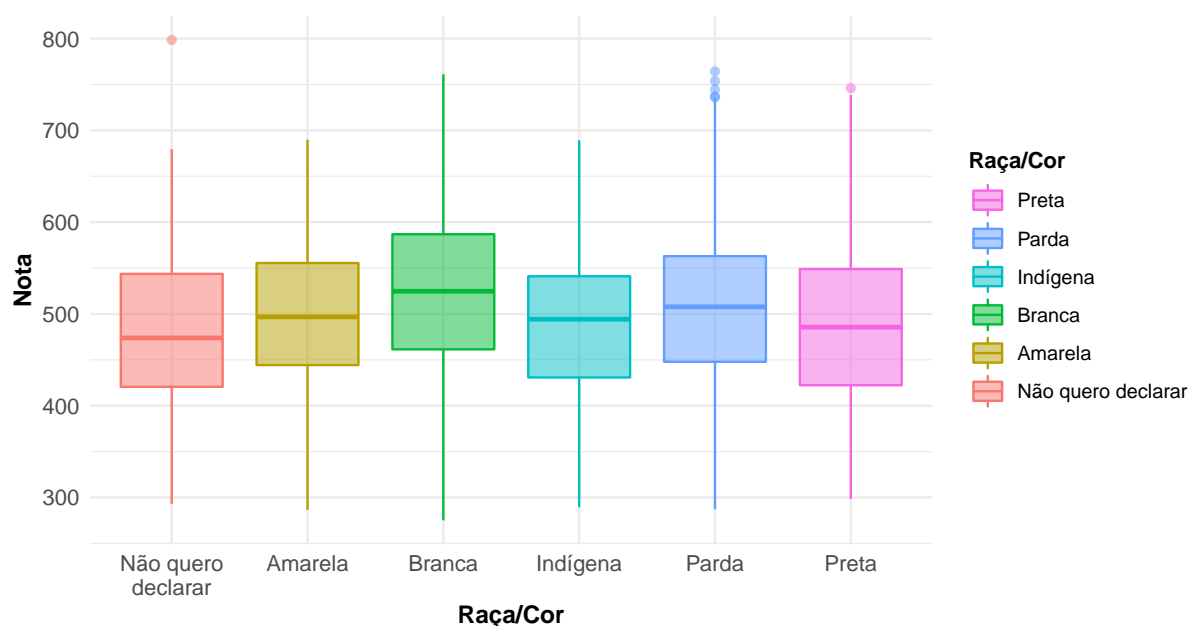
Anotações: Aprox. 0 refere-se à algum número muito pequeno considerando aproximadamente zero. O EF e EM remete ao ensino fundamental e ensino médio respectivamente.



2.2 Notas

Para medir o desempenho dos alunos no aprendizado básico, a soma das notas em Língua Portuguesa e Matemática da Prova Brasil foi utilizado para verificar se há indícios de desigualdade com algum grupo das relações abordadas pelo estudo.

Figura 3 – Distribuições das somas das notas com base na raça/cor dos alunos.



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

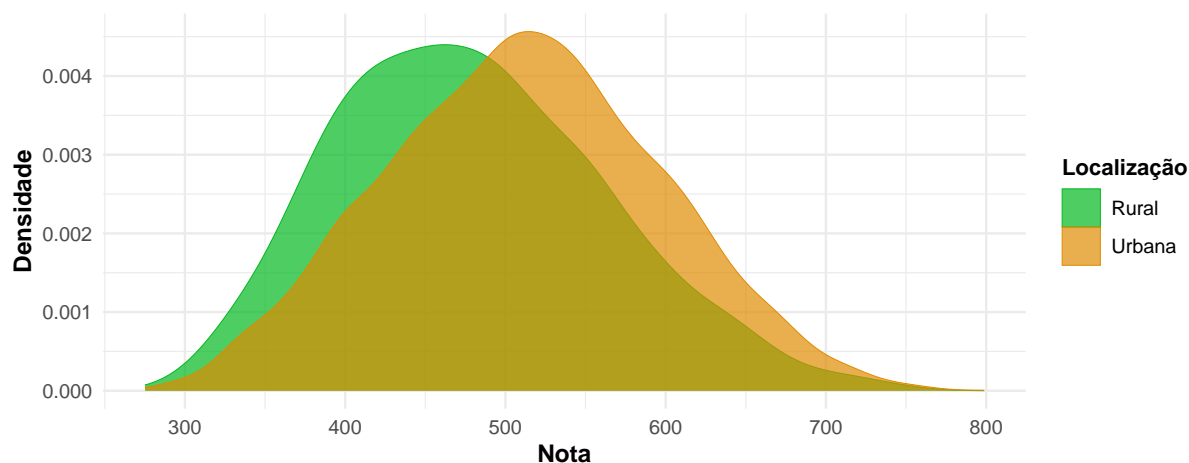
Nota: Amostra retirada de uma amostragem aleatórias simples.

Sobre a [Figura 3](#), nota-se que a soma das notas dos grupos tem uma tendência da medida central (50% dos alunos) ser localizada na nota 500, que varia pouco de acordo com a raça, sendo o fenótipo Branco aquele que possui as maiores valores nas medidas de posição sobre as estas notas.

A raça Parda é a que possui o maior número de outliers entre os estudantes, mas a maior nota observada esteve presente no grupo que não quis declarar a raça/cor.



Figura 4 – Distribuições empíricas das somas das notas com base nas localizações das das escolas dos alunos.



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

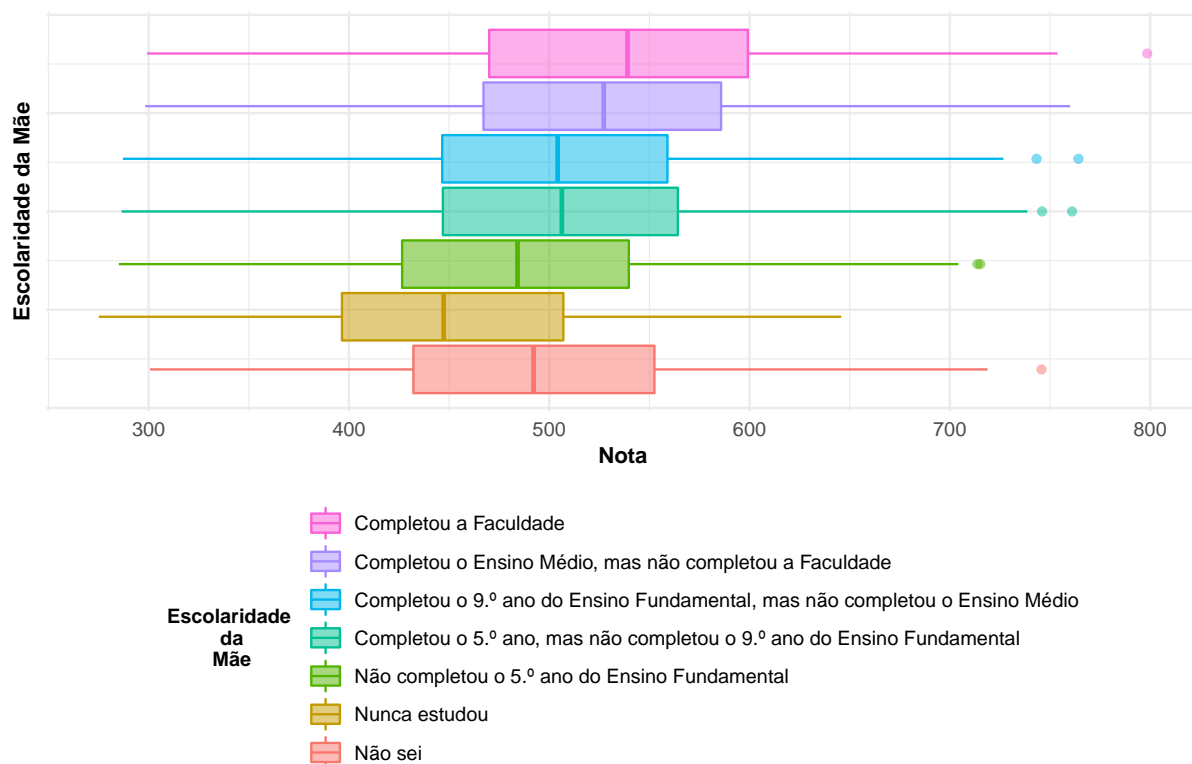
Na [Figura 4](#), é possível observar um grau de assimetria um pouco maior nestas notas para as escolas das regiões rurais em relação às regiões urbanas. A distribuição das notas nas regiões rurais foi um pouco mais inclinada para a esquerda e com moda inferior à moda das zonas urbanas, que possui uma distribuição mais centralizada.

As escolas rurais apresentaram uma nota média de 479, que também é inferior quando comparada com a nota média das escolas urbanas, que foi de 512.

Para a [Figura 5](#), observa-se um crescimento das notas em geral, à medida que o grau de escolaridade das mães é maior, de modo que apenas aqueles alunos que responderam que desconhecem a escolaridade da mãe tiveram comportamento independente a essa observação. Esse comportamento é observado de forma equivalente aos valores extremos, onde a maior nota registrada vem por parte do aluno cuja mãe completou a faculdade.



Figura 5 – Distribuições das somas das notas com base nas escolaridades das mães dos alunos



Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatórias simples.

Ao realizar testes sobre a soma das notas, foi verificado se havia indícios de diferenças significativas entre as médias da soma das notas com base nas relação efetuada pelo estudo, com a confiança 95%. Para efetuar estes testes, foi avaliado primeiramente a hipótese de igualdade sobre as variâncias das notas (Teste B) para cada relação abordada pelo estudo como as localizações das escolas, escolaridade da mãe, sexo e a raça/cor do aluno, no qual apenas a relação com o sexo rejeita esta hipótese como demonstrado na [Tabela 3](#).



Tabela 3 – Testes para as relações com soma das notas dos alunos.

Teste	H_0	P-valor	Decisão de H_0 (95%)
B	$\sigma_R^2 = \sigma_U^2$	0.503	Aceita
B	$\sigma_M^2 = \sigma_F^2$	0.002	Rejeita
B	$\sigma_{Raça/Cor}^2$ iguais	0.265	Aceita
B	$\sigma_{Esc(mãe)}^2$ iguais	0.132	Aceita
T	$\mu_R = \mu_U$	Aprox. 0	Rejeita
T	$\mu_M = \mu_F$	0.905	Aceita
ANOVA	$\mu_{Raça/Cor}$ iguais	Aprox. 0	Rejeita
ANOVA	$\mu_{Esc(mãe)}$ iguais	Aprox. 0	Rejeita

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples. Teste T dois a dois efetuado com a correção de Bonferroni no P-valor.

Anotações: Os subíndices, com base nos alunos, R e U referem-se às localizações das escolas rurais e urbanas, M e F sobre os sexos Masculino e Feminino respectivamente e $Esc(mãe)$ diz respeito à escolaridade da mãe. O Aprox. 0 refere-se à algum número muito pequeno considerado por este estudo aproximadamente zero.

Ao concluir sobre estas relações com as variâncias na [Tabela 3](#), foram avaliadas as hipóteses de igualdades destas médias, que pelo mesmo nível de confiança. Houveram indícios de igualdade com base nos sexos (Teste T), que diz sobre não existir diferença em média entre as notas dos sexos masculino e feminino.

Sobre a hipótese de igualdade de notas por raça/cor do aluno, e a escolaridade da mãe (Teste ANOVA) sobre essas notas, houveram indícios significativos de diferença nas médias em alguma destas relações, onde os testes dois a dois (Teste T) são apropriados para investigar que diferenças foram essas.

Para a relação com as localizações das escolas, a hipótese de igualdade também foi rejeitada, e as escolas urbanas obtiveram, em média, um valor superior às escolas rurais.



Tabela 4 – Comparações dois a dois entre as médias sobre a soma das notas com base na raça/cor dos alunos.

Comparações	P-valor	Evidência (RA 95%)
Amarela = Não quero declarar	0.4113	Iguais
Amarela = Branca	0.0005	Desiguais
Amarela = Indígena	1.0000	Iguais
Amarela = Parda	1.0000	Iguais
Amarela = Preta	1.0000	Iguais
Branca = Não quero declarar	Aprox. 0	Desiguais
Branca = Indígena	0.0010	Desiguais
Branca = Parda	Aprox. 0	Desiguais
Branca = Preta	Aprox. 0	Desiguais
Indígena = Não quero declarar	1.0000	Iguais
Indígena = Parda	0.7758	Iguais
Indígena = Preta	1.0000	Iguais
Parda = Não quero declarar	Aprox. 0	Desiguais
Parda = Preta	Aprox. 0	Desiguais

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples. Teste T dois a dois efetuado com a correção de Bonferroni no P-valor.

Anotações: Aprox. 0 refere-se à algum número muito pequeno considerando aproximadamente zero.

Na [Tabela 4](#), as comparações dois a dois apresentaram que a raça/cor branca obteve diferença, nestes testes de hipótese de igualdade das médias, entre todas as outras raças, confirmando o que a [Figura 3](#) apontava.

As únicas desigualdades dessa hipótese que não foram entre as comparações com os indivíduos da raça/cor branca foram entre os alunos pardos com os pretos e aqueles que não quiseram declarar.

Todos os outros testes não só levam à aceitação da hipótese de igualdade das médias (H_0), como apresentaram um p-valor significativamente alto, que diz respeito a níveis de confiabilidade elevados.

As hipóteses iniciais para a variável do nível de escolaridade da mãe são passadas como comparações dois a dois na [Tabela 5](#), que supõe esta hipóteses de igualdade entre as médias da soma das nota dos alunos, com cada grupo de nível escolar das mães.



Tabela 5 – Comparações dois a dois das notas entre os alunos com base na escolaridade da mãe

Comparações	P-valor	Evidência (RA 95%)
Não sabe = Nunca estudou	Aprox. 0	Desiguais
Não sabe = Incompleto 5.º ano do EF	1.0000	Iguais
Não sabe = Completou 5.º ano do EF	0.0084	Desiguais
Não sabe = Completou 9.º ano do EF	0.1927	Iguais
Não sabe = Completou EM	Aprox. 0	Desiguais
Não sabe = Completou Faculdade	Aprox. 0	Desiguais
Nunca estudou = Incompleto 5.º ano do EF	0.0038	Desiguais
Nunca estudou = Completou 5.º ano do EF	Aprox. 0	Desiguais
Nunca estudou = Completou 9.º ano do EF	Aprox. 0	Desiguais
Nunca estudou = Completou EM	Aprox. 0	Desiguais
Nunca estudou = Completou Faculdade	Aprox. 0	Desiguais
Incompleto 5.º ano do EF = Completou 5.º ano do EF	0.0002	Desiguais
Incompleto 5.º ano do EF = Completou 9.º ano do EF	0.0048	Desiguais
Incompleto 5.º ano do EF = Completou EM	Aprox. 0	Desiguais
Incompleto 5.º ano do EF = Completou Faculdade	Aprox. 0	Desiguais
Completo 5.º ano do EF = Completou 9.º ano do EF	1.0000	Iguais
Completo 5.º ano do EF = Completou EM	Aprox. 0	Desiguais
Completo 5.º ano do EF = Completou Faculdade	Aprox. 0	Desiguais
Completo 9.º ano do EF = Completou EM	Aprox. 0	Desiguais
Completo 9.º ano do EF = Completou Faculdade	Aprox. 0	Desiguais
Completou EM = Completou Faculdade	1.0000	Iguais

Fonte: Amostra de 5.271 alunos do 9º ano do SAEB 2017.

Nota: Amostra retirada de uma amostragem aleatória simples. Teste T dois a dois efetuado com a correção de Bonferroni no P-valor.

Anotações: Aprox. 0 refere-se à algum número muito pequeno considerando aproximadamente zero e EF e EM remete ao ensino fundamental e ensino médio respectivamente.

E dentre esses grupos, a hipótese de igualdade só foi aceita entre as seguintes comparações: alunos que não sabiam a escolaridade da mãe com aqueles cujas mães têm o 5º ano incompleto ou 9º ano completo, mães que têm completo o 5º ano com as que completaram o 9º ano, com base no ensino fundamental, e aquelas que têm o ensino médio completo com as que possuem a faculdade completa, com um nível de significância de 5%.

Ao considerar estas igualdades e as outras desigualdades das médias sobre a soma destas notas, há maiores valores para aquelas mães que possuem escolaridade superior.



3 Conclusão

De acordo com os resultados obtidos utilizando testes estatísticos com uma confiança de 95%, houve indícios significativos que para alguns fatores sociais, há desigualdades entre as avaliações das relações neste estudo com a soma das notas em Língua Portuguesa e Matemática ou o tempo de afazeres diários em casa por parte dos alunos do 9º ano do ensino fundamental.

A raça/cor não diz sobre os tempos destes afazeres, mas ao analisar os resultados destes testes sobre a soma destas notas, os alunos que se autodeclararam com raça/cor Branca, em média, obtiveram maiores notas.

Sobre o sexo do aluno, aqueles que são do sexo feminino acabam por fazer mais destes afazeres, não havendo o mesmo tempo disponível em casa hipoteticamente que comparado ao outro sexo. Já para o desempenho médio entre os sexos, não houve evidência de desigualdade entre esta soma.

O nível de escolaridade da mãe do aluno diz sobre o período de tempo gasto por ele nestas tarefas diárias, no qual apenas os alunos não sabem a escolaridade da mãe resultam em exercer menos tempo nestas tarefas. Com base nesta soma das notas, estes alunos que não sabem, obtêm a média equivalente as mães que tem o 5º incompleto ou as que tem o 9º completo sobre o ensino fundamental, mas de forma geral, o nível de escolaridade da mãe diz em média que, quanto maior este nível, maior tende a ser a nota.

Para a localização da escola do aluno, foi analisado apenas esta soma das notas, no qual houve evidência significativa para afirmar que aqueles que estudam em uma zona urbana tendem a ter notas superiores àqueles que estudam em uma zona rural.

De forma geral, fatores sociais influenciam o alunos do 9º no desenvolvimento da educação básica, pelos quais existem desigualdades que prejudicam a formação do conhecimento destas turmas, ao passo que alguns alunos dispõem de mais tempo em casa ou exemplos maternos, que acabam por ter influência no aluno sobre as notas avaliadas pela Prova Brasil efetuada de forma igualitária para toda estas turmas.



Referências

ANDERSON, T. W.; DARLING, D. A. A test of goodness of fit. *Journal of the American statistical association*, Taylor & Francis, v. 49, n. 268, p. 765–769, 1954. Citado na página 10.

BARTLETT, M. S. A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 296–298, 1954. Citado na página 10.

BONFERRONI, C. Teoria statistica delle classi e calcolo delle probabilit . *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, v. 8, p. 3–62, 1936. Citado na p gina 10.

FISHER, R. A. The general sampling distribution of the multiple correlation coefficient. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, The Royal Society London, v. 121, n. 788, p. 654–673, 1928. Citado na p gina 10.

INEP. *SAEB 2017 MICRODADOS DA ANEB E DA ANRESC (PROVA BRASIL)*: Saeb (aneb/prova brasil 2017). [S.l.], 2018. 33 p. Dispon vel em: <<http://portal.inep.gov.br/web/guest/microdados>>. Citado na p gina 8.

KRUSKAL, W. H.; WALLIS, W. A. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, Taylor & Francis Group, v. 47, n. 260, p. 583–621, 1952. Citado na p gina 10.

MANN, H. B.; WHITNEY, D. R. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, JSTOR, p. 50–60, 1947. Citado na p gina 10.

O’CONNOR, J. J.; ROBERTSON, E. F.; EDMUND, F. Student’s t-test. *MacTutor History of Mathematics archive*, 1908. Citado na p gina 10.

SHAPIRO, S. S.; FRANCIA, R. An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, Taylor & Francis Group, v. 67, n. 337, p. 215–216, 1972. Citado na p gina 10.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, JSTOR, v. 52, n. 3/4, p. 591–611, 1965. Citado na p gina 10.

Anexos



ANEXO A – Amostra

Os bancos de dados dos alunos participantes do SAEB de 2017 foram disponibilizados como amostras de tamanho 2.000 em formato CSV. Cada banco de dados era uma amostra aleatória diferente do banco de dados original do SAEB, para cada aluno da Turma A de Métodos Estatísticos 2 do semestre 1/2020 da Universidade de Brasília (UnB).

Os dados utilizados nesse documento foi a junção dos três bancos de dados dos integrantes com a retirada dos alunos que detinham a *NA's*, nestas amostras com os seguintes nomes: **amostra_190015853.csv**, **amostra_190029498.csv** e **amostra_190127180.csv**