

Aline de Almeida Ramos

**Análise de dados
Testes de aderência
Resultados do SAEB
(ANEB - Prova Brasil) - 2017**

Brasília

abril de 2021

Aline de Almeida Ramos

Análise de dados
Testes de aderência
Resultados do SAEB
(ANEB - Prova Brasil) - 2017

Relatório de pesquisa realizado acerca de dados da Prova Brasil, durante o curso da disciplina de Métodos Estatísticos 2.

Universidade de Brasília
Departamento de Estatística

Brasília
abril de 2021

Lista de tabelas

Tabela 1	– Distribuição de frequências da variável nota em matemática	7
Tabela 2	– Resultado dos testes	8
Tabela 3	– Resultado dos testes	8

Sumário

1	SOBRE O SAEB	4
2	TESTES DE ADERÊNCIA	5
2.1	Testes aplicados	5
3	METODOLOGIA	6
4	QUI-QUADRADO	7
4.1	Notas de matemática	7
5	KOLMOGOROV - LILLIEFORS, SHAPIRO-WILK E ANDERSON- DARLING	8
5.1	Notas de matemática	8
5.2	Notas de língua portuguesa	8
6	CONCLUSÃO	10
	REFERÊNCIAS	11

1 Sobre o SAEB

O Sistema de Avaliação da Educação Básica - SAEB tem como objetivo avaliar a qualidade do ensino nas escolas públicas do Brasil. Esse desempenho é medido não somente por meio da aplicação de provas de Língua Portuguesa e de Matemática como também são considerados diversos fatores em que os alunos estão inseridos. A amostra trabalhada no presente relatório corresponde ao desempenho de estudantes do quinto ano do ensino fundamental das escolas públicas brasileiras no ano de 2017.

2 Testes de aderência

Morettin e Bussab (2017) Os testes de aderência tem como objetivo a possibilidade de verificar se uma distribuição se ajusta bem ou não aos dados amostrais. É através da comparação das frequências amostrais com as frequências teóricas esperadas pelo modelo probabilístico que se está julgando válido para descrever os dados observados. Em suma, o intuito é testar:

H0: X tem distribuição f

H1: X tem distribuição f

2.1 Testes aplicados

- Kolmogorov - Lilliefors
- Shapiro-Wilk
- Anderson-Darling
- Qui- Quadrado

3 Metodologia

Para a aplicação dos testes de aderência, foram escolhidas subamostras da amostra inicial. Para a subamostra de tamanho $n=100$ foi construída uma tabela com a distribuição de frequências da variável nota em matemática. Para a segunda subamostra de tamanho $n=30$, foi testada normalidade tanto para a variável nota em língua portuguesa, quanto para nota em matemática. Por fim, há a análise dos resultados dos testes, para verificar se as conclusões obtidas podem se generalizar para a amostra inicial.

4 Qui-Quadrado

4.1 Notas de matemática

H0: Notas em matemática segue distribuição normal
H1: Notas em matemática não segue distribuição normal

Tabela 1 – Distribuição de frequências da variável nota em matemática

Li-Ls	fi	Z(Ls)	P(Z<z)=P(X<x)
122-145	3	-1.66930849522401	0.047528128747078
145-168	8	-1.20112449795384	0.114851455708395
168-191	14	-0.732940500683675	0.231797359852623
191-214	15	-0.264756503413507	0.395598523336993
214-237	16	0.203427493856661	0.580599543966686
237-260	23	0.671611491126829	0.749084469563341
260-283	9	1.139795488397	0.872814243110943
283-306	6	1.60797948566717	0.946080164118353
306-329	1	2.07616348293733	0.981060583302722
329-352	5	2.5443474802075	0.994525894372978

pi	ei	(fi-ei)^2/ei
0.047528128747078	4.7528128747078	183.409075432528
0.0673233269613169	6.73233269613169	934.70364640668
0.116945904144228	11.6945904144228	1648.10552741339
0.163801163484371	16.3801163484371	1343.78041788226
0.185001020629693	18.5001020629693	1351.96115062589
0.168484925596655	16.8484925596655	3093.91524937146
0.123729773547602	12.3729773547602	636.776184696512
0.0732659210074097	7.32659210074097	479.434044643207
0.0349804191843697	3.49804191843697	26.6224022773773
0.0134653110702552	1.34653110702552	1846.63600225527
Total		11545.34

Utilizando $\alpha = 0.05\%$ e com graus de liberdade (gl) = $10-2-1 = 7$, o valor tabelado da estatística Qui-Quadrado é 14.06714. Como a estatística calculada corresponde ao valor de 11545.34 e que é muito maior que o valor tabelado, há então evidências para rejeitar H0, ou seja, a variável notas em matemática não segue distribuição normal.

5 Kolmogorov - Lilliefors, Shapiro-Wilk e Anderson-Darling

5.1 Notas de matemática

Com a subamostra $n=30$, os testes serão utilizados para testar a normalidade da variável.

H0: Notas em matemática segue distribuição normal
H1: Notas em matemática não segue distribuição normal

Tabela 2 – Resultado dos testes

	Kolmogorov	Anderson-Darling	Shapiro-Wilk
estatística	0.187	0.912	0.878
p-valor	0.009	0.017	0.022

Para os três testes dispostos acima, foram obtidos valores pequenos do p-valor, o que permite concluir que a variável notas em matemática não segue distribuição normal, ou seja, rejeitamos H0.

5.2 Notas de língua portuguesa

Com a subamostra $n=30$, os testes serão utilizados para testar a normalidade da variável.

H0: Notas em língua portuguesa segue distribuição normal
H1: Notas em língua portuguesa não segue distribuição normal

Tabela 3 – Resultado dos testes

	Kolmogorov	Anderson-Darling	Shapiro-Wilk
estatística	0.112	0.967	0.296
p-valor	0.434	0.466	0.571

Assim como para as notas de matemática, seguindo o valor apresentado das estatísticas e dos p-valores não é possível afirmar que notas em língua portuguesa seguem distribuição normal, ou seja, aceitamos H_1 .

6 Conclusão

Na análise descritiva realizada com a amostra inicial, em que $n=2000$, concluiu-se com os dados apresentados que para tal tamanho, tanto as notas de português quanto de matemática se aproximavam de uma normal. Ao utilizar os testes de aderência apresentados anteriormente, para subamostras da amostra inicial, o desfecho difere e é possível concluir que para as amostras com $n=100$ e $n=30$ as variáveis não seguem distribuição normal. É válido destacar que mesmo usando uma divisão de classes de tamanho $n=10$ na construção da tabela de distribuição de frequências - pelo teste Qui-Quadrado - o valor calculado era muito maior que o valor encontrado na tabela. Essa é uma característica da Qui-Quadrado em que pode haver certa "explosão" de valores durante a análise dos dados.

Referências

MORETTIN, P. A.; BUSSAB, W. O. *Estatística básica*. [S.l.]: Saraiva Educação SA, 2017.
Citado na página 5.