



Universidade de Brasília

Tailine Nonato

**TESTES PARA NORMALIDADE:
RESULTADOS DO SAEB (ANEB - PROVA
BRASIL) 2017**

Brasília, DF

Abril, 2021

Tailine Nonato

TESTES PARA NORMALIDADE: RESULTADOS DO SAEB (ANEB - PROVA BRASIL) 2017

Relatório de testes para normalidade apresentado à disciplina Métodos Estatísticos 2 da Universidade de Brasília como objeto de avaliação.

Brasília, DF

Abril, 2021

Lista de tabelas

Tabela 1 – Frequências da variável Nota_MT	6
Tabela 2 – Teste χ^2 para Nota em Matemática	7
Tabela 3 – Teste Shapiro-Wilk para Notas em Língua Portuguesa	7
Tabela 4 – Teste Anderson-Darling para Notas em Língua Portuguesa	8
Tabela 5 – Teste Kolmogorov para Notas em Língua Portuguesa	8
Tabela 6 – Teste Shapiro-Wilk para Notas em Matemática	8
Tabela 7 – Teste Anderson-Darling para Notas em Matemática	9
Tabela 8 – Teste Kolmogorov para Notas em Matemática	9

Sumário

	Introdução	4
1	METODOLOGIA	5
2	RESULTADOS	6
2.1	Amostra de tamanho 100	6
2.1.1	Distribuição de frequências de notas em Matemática	6
2.1.2	Verificação para normalidade	6
2.2	Amostras de tamanho 30	7
2.2.1	Notas em Língua Portuguesa	7
2.2.1.1	Shapiro-Wilk	7
2.2.1.2	Anderson-Darling	8
2.2.1.3	Kolmogorov	8
2.2.2	Notas em Matemática	8
2.2.2.1	Shapiro-Wilk	8
2.2.2.2	Anderson-Darling	9
2.2.2.3	Kolmogorov	9
	A – CÓDIGO EM R	10

Introdução

Utilizando os dados do Sistema de Avaliação da Educação Básica (SAEB) de 2017 divulgado pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), esta pesquisa tem por objetivo testar as hipóteses de que as notas em Língua Portuguesa e em Matemática seguem distribuição Normal.

1 Metodologia

O banco de dados em questão dispõe de 2000 observações em 19 variáveis. No relatório anterior (Atividade 2) foram retiradas 50 amostras aleatórias de tamanho 30 e 50 amostras aleatórias de tamanho 100, assim, para este relatório serão utilizadas a primeira amostra de tamanho 30 e a primeira amostra de tamanho 100 retiradas anteriormente.

As variáveis analisadas serão:

- NOTA_LP: Proficiência em Língua Portuguesa transformada na escala única do SAEB;
- NOTA_MT: Proficiência do aluno em Matemática transformada na escala única do SAEB.

O nível de significância considerado foi de 5%.

Para execução das análises, foi utilizado o Software *RStudio* com *R* versão 4.0.3 e com auxílio dos pacotes *tidyverse*, *dplyr*, *infer* e *nortest*.

2 Resultados

A fim de testar se as notas em Língua Portuguesa e em Matemática seguem distribuição Normal, as análises e resultados estão divididos em duas subseções: amostras de 30 observações e amostras de 100 observações.

2.1 Amostra de tamanho 100

Utilizando a amostra de tamanho 100, essa seção tem por objetivo verificar se as notas em Matemática podem ser descritas pela distribuição Normal.

2.1.1 Distribuição de frequências de notas em Matemática

Considerando faixas de 44 pontos, as notas em Matemática estão distribuídas de acordo com a tabela a seguir.

Tabela 1 – Frequências da variável Nota_MT

Nota	Frequência
Entre 133 e 176	17
Entre 177 e 220	27
Entre 221 e 264	33
Entre 265 e 308	20
Entre 309 e 353	3
Total	100

Fonte: INEP, SAEB, 2017.

2.1.2 Verificação para normalidade

$$\begin{cases} H_0 : \text{Nota em Matemática segue distribuição normal.} \\ H_a : \text{Nota em Matemática não segue distribuição normal.} \end{cases}$$

Para testar as hipóteses descritas acima, será feito um teste Qui-Quadrado (χ^2) Goodness of Fit (Aderência),

Utilizando o nível de significância de 0,05, o valor tabelado é de 9,48. Se o valor do teste for maior que o valor tabelado, existem evidências suficientes para rejeitar H_0 , ou

seja, há evidências para dizer que as notas em Matemática **não** segue uma distribuição Normal, caso contrário não rejeita-se a possibilidade.

Tabela 2 – Teste χ^2 para Nota em Matemática

Variável	Valor do teste	Decisão do teste
NOTA_MT	19,47	Rejeita H_0

Fonte: INEP, SAEB, 2017.

Assim, rejeitando H_0 , existem evidências para dizer que as notas em Matemática não seguem uma distribuição Normal.

2.2 Amostras de tamanho 30

Utilizando a amostra de tamanho 30, serão feitos os testes Shapiro-Wilk, Anderson-Darling e Kolmogorov para testar normalidade nas variáveis NOTA_LP e NOTA_MT, que estão divididas em duas subseções.

2.2.1 Notas em Língua Portuguesa

$$\begin{cases} H_0 : \text{Notas em Língua Portuguesa seguem distribuição normal.} \\ H_a : \text{Notas em Língua Portuguesa não seguem distribuição normal.} \end{cases}$$

2.2.1.1 Shapiro-Wilk

Comparando a função de distribuição com a distribuição empírica, obtêm-se os seguintes resultados:

Tabela 3 – Teste Shapiro-Wilk para Notas em Língua Portuguesa

Variável	Valor do teste	Decisão do teste
NOTA_LP	0,976	Não rejeita H_0

Fonte: INEP, SAEB, 2017.

Para $n = 30$ e nível de significância de 5%, o valor tabelado é 0,927. Assim, como o valor observado é maior que o tabelado não há evidências para rejeitar H_0 .

2.2.1.2 Anderson-Darling

Tabela 4 – Teste Anderson-Darling para Notas em Língua Portuguesa

Variável	Valor do teste	Decisão do teste
NOTA_LP	0,278	-

Fonte: INEP, SAEB, 2017.

2.2.1.3 Kolmogorov

Para $n = 30$, o quantil em 95% de confiança é 0,218.

Tabela 5 – Teste Kolmogorov para Notas em Língua Portuguesa

Variável	Valor do teste	Decisão do teste
NOTA_LP	0,111	Não rejeita H_0

Fonte: INEP, SAEB, 2017.

Como o valor do teste não excede o quantil de ordem $(1 - \alpha)$, não há evidências para rejeitar H_0 .

2.2.2 Notas em Matemática

$$\begin{cases} H_0 : \text{Notas em Matemática seguem distribuição normal.} \\ H_a : \text{Notas em Matemática não seguem distribuição normal.} \end{cases}$$

2.2.2.1 Shapiro-Wilk

Comparando a função de distribuição com a distribuição empírica, obtêm-se os seguintes resultados:

Tabela 6 – Teste Shapiro-Wilk para Notas em Matemática

Variável	Valor do teste	Decisão do teste
NOTA_MT	0,977	Não rejeita H_0

Fonte: INEP, SAEB, 2017.

Para $n = 30$ e nível de significância de 5%, o valor tabelado é 0,927. Assim, como o valor observado é maior que o tabelado não há evidências para rejeitar H_0 .

2.2.2.2 Anderson-Darling

Tabela 7 – Teste Anderson-Darling para Notas em Matemática

Variável	Valor do teste	Decisão do teste
NOTA_MT	0,302	-

Fonte: INEP, SAEB, 2017.

2.2.2.3 Kolmogorov

Para $n = 30$, o quantil em 95% de confiança é 0,218.

Tabela 8 – Teste Kolmogorov para Notas em Matemática

Variável	Valor do teste	Decisão do teste
NOTA_MT	0,091	Não rejeita H_0

Fonte: INEP, SAEB, 2017.

Como o valor do teste não excede o quantil de ordem $(1 - \alpha)$, não há evidências para rejeitar H_0 .

ANEXO A – Código em R

Link para o diretório no GitHub:

<<https://github.com/tailineju/SAEB-2017>>