



UNIVERSIDADE FEDERAL DE  
PERNAMBUCO CENTRO DE EDUCAÇÃO  
PROGRAMA DE PÓS- GRADUAÇÃO EM  
EDUCAÇÃO MATEMÁTICA E  
TECNOLÓGICA  
CURSO DE DOUTORADO

ANDRÉ FELLIPE QUEIROZ ARAÚJO

**ENSINO E APRENDIZAGEM DE AMOSTRAGEM, CURVA  
NORMAL E SUAS RELAÇÕES NO 3º ANO DO ENSINO MÉDIO**

RECIFE  
2024

**ANDRÉ FELLIPE QUEIROZ ARAÚJO**

**ENSINO E APRENDIZAGEM DE AMOSTRAGEM, CURVA  
NORMAL E SUAS RELAÇÕES NO 3º ANO DO ENSINO MÉDIO**

Tese apresentada ao Programa de Pós-Graduação em Educação Matemática e Tecnológica do Centro de Educação da Universidade Federal de Pernambuco-UFPE, como requisito parcial para obtenção do título de Doutor em Educação Matemática e Tecnológica. Área de concentração: Ensino de Ciências e Matemática.

Orientadora: Prof.<sup>a</sup> Dr<sup>a</sup>. Gilda Lisbôa Guimarães

RECIFE

2024

Catalogação na fonte  
Bibliotecária Katia Tavares, CRB-4/1431

A663e Araújo, André Fellipe Queiroz  
Ensino e aprendizagem de amostragem, curva normal e suas relações no 3º ano do ensino médio / André Fellipe Queiroz de Araújo. –2024.  
309 f.

Orientação de: Gilda Lisbôa Guimarães.

Tese (doutorado) - Universidade Federal de Pernambuco, Centro de Educação. Programa de Pós-graduação em Educação Matemática e Tecnológica, 2024.

Inclui Referências.

1. Matemática – Estudo e ensino. 2. Amostragem. 3. Curva normal. 4. Estatística. 5. Probabilidade. 6. Geogebra. I. Guimarães, Gilda Lisbôa (Orientação). II. Título.

372.7 (22. ed.)

UFPE (CE2024-038)

ANDRÉ FELLIPE QUEIROZ ARAÚJO

**ENSINO E APRENDIZAGEM DE AMOSTRAGEM, CURVA  
NORMAL E SUAS RELAÇÕES NO 3º ANO DO ENSINO MÉDIO**

Tese apresentada ao Programa de Pós-Graduação em Educação Matemática e Tecnológica do Centro de Educação da Universidade Federal de Pernambuco-UFPE, como requisito parcial para obtenção do título de Doutor em Educação Matemática e Tecnológica. Área de concentração: Ensino de Ciências e Matemática.

Aprovado em: 05/04/2024

**BANCA EXAMINADORA**

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Gilda Lisbôa Guimarães (Orientadora)  
Universidade Federal de Pernambuco - UFPE

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Jaqueline Aparecida Foratto Lixandrão Santos (Examinadora Interna)  
Universidade Federal de Pernambuco - UFPE

---

Prof<sup>o</sup>. Dr<sup>o</sup>. Humberto José Bortolossi (Examinador Externo)  
Universidade Federal Fluminense - UFF

---

Prof<sup>o</sup>. Dr<sup>o</sup>. Diego Bezerra de Melo Maciel (Examinador Externo)  
Universidade Federal de Pernambuco - UFPE

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Liliana Mabel Tauber (Examinadora Externa)  
Universidad Nacional Del Litoral – UNL

## AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus pelo dom da vida, pelo seu infinito amor e bondade, por ter me guiado e concedido saúde, graça e sabedoria para concluir mais uma etapa na minha trajetória acadêmica.

Aos meus pais Joás Araújo e Léa Maria, e aos meus irmãos, Débora Raquel e Samuel Marcos pelo apoio inabalável e investimento na minha educação que foram fundamentais em toda a minha jornada. Sou imensamente grato por todo o incentivo e encorajamento proporcionado.

À minha amada esposa Stephany Chiappetta por todo amor, companheirismo, apoio e conselhos que muito me ajudaram durante todo o desenvolvimento desse estudo. Sua dedicação e apoio incondicionais são uma fonte constante de força e inspiração para mim. Minha imensa gratidão!

À minha orientadora e amiga, professora Gilda Guimarães, por todo acolhimento e ricas orientações que ao longo dessa trajetória me proporcionaram muito aprendizado e foram fundamentais para a realização da pesquisa. Guardarei profunda admiração pela sua inteligência, profissionalismo e por sua forma brilhante e dedicada de conduzir pesquisa científica. Muito obrigado por tudo!

Aos professores Jaqueline Lixandrão, Humberto Bortolossi, Diego Maciel e Liliana Tauber pelas significativas contribuições durante a qualificação desta pesquisa e por também fazerem parte da banca examinadora.

A todos que fazem o grupo de Estudos em Educação Estatística no Ensino Fundamental - GREF pelas importantes contribuições e sugestões durante o andamento desse estudo.

Ao Luan Luna, um amigo que o Doutorado me presenteou. Sempre disposto a ajudar. Obrigado pela parceria, trocas de experiências e valiosas sugestões e apontamentos durante todo o percurso da pesquisa.

Ao Diego Maciel, outro amigo que a academia proporcionou. Sempre solícito para responder as minhas interrogações e colaborar com a pesquisa. Grato pela parceria e contribuições.

À secretaria de Educação do Governo de Pernambuco e a todos que fazem a Escola de Aplicação Professor Chaves e a Escola Dom Carlos Coelho. Agradeço todo o apoio e incentivo.

Aos amigos de turma e professores do programa de Pós-Graduação em Educação Matemática e Tecnológica- EDUMATEC da Universidade Federal de Pernambuco- UFPE que direta ou indiretamente contribuíram durante o meu curso de Doutorado e o desenvolvimento da pesquisa. Obrigado pela amizade, atenção e consideração.

## RESUMO

A presente pesquisa teve o objetivo de investigar o que estudantes do 3º ano do Ensino Médio (16/17 anos de idade) sabem e podem aprender sobre Amostragem, Curva Normal e suas relações. Para tanto, nos apoiamos no modelo de Letramento Estatístico e Probabilístico proposto por Gal (2002, 2005) e no modelo de Ciclo Investigativo elaborado por Guimarães e Gitirana (2013). Em termos metodológicos, para alcançar o referido objetivo, traçamos 4 estudos inter-relacionados, dos quais os três primeiros deram base para a realização do último. No estudo 1, analisamos o que tem sido proposto sobre Amostragem, Curva Normal e suas relações nos livros didáticos aprovados pelo Programa Nacional do Livro e do Material Didático 2021(Livro de Projetos e Livros de Conhecimento). No estudo 2, investigamos o que foi proposto nas questões da prova de Matemática e suas Tecnologias do Exame Nacional do Ensino Médio (ENEM) de 2022 referente a esses dois conceitos e suas relações. Já no estudo 3, identificamos os conhecimentos iniciais de estudantes do 3º ano do Ensino Médio sobre a referida temática. Na sequência, realizamos um estudo experimental que contemplou uma intervenção de ensino para analisar a contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada. Os resultados do estudo 1 revelaram que tanto os livros de projetos quanto os de conhecimento oferecem propostas interessantes para o ensino de Estatística e Probabilidade, principalmente na análise e realização de pesquisas estatísticas. Os livros de projetos exploram dados reais na maioria das situações, promovendo uma compreensão mais profunda do contexto social. No entanto, há uma ausência de discussão sobre seleção de amostra e representatividade, priorizando análises descritivas em detrimento da abordagem de tópicos da inferência estatística. Os livros de conhecimento, por outro lado, exploram mais dados fictícios e enfatizam o cálculo de medidas estatísticas carecendo, também, de uma abordagem sobre Amostragem e Curva Normal de forma articulada. No estudo 2 verificamos que nenhuma questão analisada do ENEM explora o conceito de Amostragem e nem da Curva Normal. As questões que envolvem a Estatística têm uma abordagem voltada apenas para o cálculo de medidas de centralidade e as que exploram a probabilidade abordam unicamente o significado de probabilidade enquanto razão clássica, com predominância nesses dois casos de dados fictícios. Por meio do estudo 3, constatamos inicialmente que os estudantes apresentaram lacunas sobre a compreensão da representatividade da amostra, incluindo a variabilidade amostral e os métodos de Amostragem, como também, no modelo da Curva Normal, contemplando sua representação gráfica e propriedades. Em acréscimo, evidenciou-se fragilidades na compreensão da relação entre a distribuição das médias amostrais com os parâmetros populacionais, pautada no Teorema Central do Limite que relaciona a Amostragem com a Curva Normal. Entretanto, o estudo experimental revelou que a intervenção de ensino realizada desempenhou um papel significativo na aprendizagem dos conceitos relacionados à Amostragem e à Curva Normal, bem como na articulação entre esses conceitos. Os recursos, tanto o lápis e papel, quanto o software Geogebra contribuíram significativamente para o desenvolvimento das atividades e a construção do conhecimento. Dessa forma, acreditamos que a realização dessa pesquisa possibilitará importantes discussões e reflexões sobre o processo de ensino e aprendizagem da Amostragem e da Curva Normal no 3º ano do Ensino Médio, propiciando a ampliação das possibilidades didático-pedagógicas para a realização desse processo com vistas a favorecer, aos estudantes, o desenvolvimento do

Letramento Estatístico e Probabilístico (Gal,2002,2005) relativo a esses conceitos e dos fenômenos e situações do nosso cotidiano que comumente explicitam seus fundamentos e relações.

**Palavras - Chave:** amostragem; curva normal; estatística e probabilidade; ensino médio; lápis e papel; geogebra.

## **ABSTRACT**

The present research aimed to investigate what 3rd year high school students (16/17 years old) know and can learn about Sampling, Normal Curve and their relationships. To this end, we rely on the Statistical and Probabilistic Literacy model proposed by Gal (2002, 2005) and the Investigative Cycle model developed by Guimarães and Gitirana (2013). In methodological terms, to achieve this objective, we outlined 4 interrelated studies, of which the first three provided the basis for carrying out the last. In study 1, we analyzed what has been proposed about Sampling, Normal Curve and their relationships in textbooks approved by the National Book and Teaching Material Program 2021 (Project Book and Knowledge Books). In study 2, we investigated what was proposed in the questions of the Mathematics and its Technologies test of the 2022 National High School Exam (ENEM) regarding these two concepts and their relationships. In study 3, we identified the initial knowledge of 3rd year high school students on the aforementioned topic. Next, we carried out an experimental study that included a teaching intervention to analyze the contribution of different resources (Pencil and paper and Geogebra software) to learning Sampling and Normal Curve in an articulated way. The results of study 1 revealed that both project and knowledge books offer interesting proposals for teaching Statistics and Probability, especially in the analysis and conduct of statistical research. Project books explore real data in most situations, promoting a deeper understanding of the social context. However, there is an absence of discussion about sample selection and representativeness, prioritizing descriptive analyzes to the detriment of addressing statistical inference topics. Knowledge books, on the other hand, explore more fictitious data and emphasize the calculation of statistical measures, also lacking an approach to Sampling and the Normal Curve in an articulated way. In study 2 we found that none of the ENEM questions analyzed explored the concept of Sampling or the Normal Curve. The questions involving Statistics have an approach focused only on calculating centrality measures and those that explore probability only address the meaning of probability as a classical ratio, with predominance in these two cases of fictitious data. Through study 3, we initially found that students had gaps in understanding the representativeness of the sample, including Sampling variability and Sampling methods, as well as in the Normal Curve model, including its graphical representation and properties. In addition, weaknesses were highlighted in understanding the relationship between the distribution of sample means and population parameters, based on the Central Limit Theorem that relates Sampling to the Normal Curve. However, the experimental study revealed that the teaching intervention carried out played a significant role in learning concepts related to Sampling and the Normal Curve, as well as in the articulation between these concepts. The resources, both pencil and paper, and the Geogebra software contributed significantly to the development of activities and the construction of knowledge. Therefore, we believe that carrying out this research will enable important discussions and reflections on the teaching and learning process of Sampling and the Normal Curve in the 3rd year of High School, providing the expansion of didactic-pedagogical possibilities for carrying out this process with a view to favor, for students, the development of Statistical and Probabilistic Literacy (Gal, 2002, 2005) related to these concepts and the phenomena

and situations of our daily lives that commonly explain their foundations and relationships.

**Keywords:** sampling; normal curve; statistics and probability; high school; pencil and paper; geogebra.

## LISTA DE FIGURAS

Figura 1 – Representação Gráfica da Curva Normal	37
Figura 2 – Representação de Curvas Normais com parâmetros diferentes	40
Figura 3 – Área sob a curva entre os pontos a e b.	40
Figura 4 – Curva Normal padronizada	42
Figura 5 – Tabela Z com a Área ou Probabilidade para a Curva Normal padronizada	43
Figura 6 – Área sob a curva entre os pontos 7 e 9	44
Figura 7 – Área sob a Curva Normal padronizada entre os pontos 0 e 1,25.	44
Figura 8 – Cálculo da área de um intervalo sob a Curva Normal com a utilização do Geogebra	46
Figura 9 – Distribuições populacionais não normal (1), normal (2) e suas respectivas distribuições amostrais das médias	48
Figura 10 – Análise do gráfico pela apresentadora	56
Figura 11 – Notícia sobre os casos de Dengue no Rio de Janeiro	58
Figura 12 – Distribuição das alturas de um conjunto de pessoas	59
Figura 13 – Notícia sobre a chance de Fakes News viralizarem	60
Figura 14 – Resultado de uma pesquisa amostral	61
Figura 15 – Noticia do mercado financeiro com menção a incerteza e risco	68
Figura 16 – Noticia em contexto de saúde com menção à chance	72
Figura 17 – Ciclo Investigativo	76
Figura 18 – Ferramenta Planilha no Geogebra	128
Figura 19 – Ferramenta de Análise Univariada no Geogebra	129
Figura 20 – Ferramenta de Análise Multivariada	130
Figura 21 – O efeito do tamanho da amostra para a determinação da proporção de elementos amostrais	131
Figura 22 – Construção de um Histograma no Geogebra	134
Figura 23 – Cálculo de Probabilidade na Curva Normal	135
Figura 24 – Gráfico da distribuição das médias Amostrais	138
Figura 25 – Gráfico da distribuição das médias Amostrais no Geogebra	141
Figura 26 – Ferramenta didática no Geogebra sobre Amostragem e Curva Normal	141

Figura 27 – Situação de explicação com abordagem explícita ao conceito de Amostra	149
Figura 28 – Situação de resolução com abordagem implícita ao conceito de Amostra	149
Figura 29 – Situação de resolução envolvendo o conceito de população/amostra com pessoas	151
Figura 30 – Situação de resolução envolvendo o conceito de população/amostra com objetos	152
Figura 31 – Situação de explicação - Análise dos dados de pesquisa amostral	153
Figura 32 – Situação de explicação – Realização de Pesquisa Estatística	154
Figura 33 – Situação de Explicação com Análise da representação gráfica	156
Figura 34 – Situação de Explicação com Análise da representação gráfica	156
Figura 35 – Situação de Explicação com o conceito de amostra explícito	159
Figura 36 – Situação de Resolução com o conceito de amostra explícito	159
Figura 37 – Situação de Resolução com a realização de pesquisa estatística	162
Figura 38 – Situação de explicação com o conceito de amostra explícito	163
Figura 39 – Situação de resolução com o conceito de amostra explícito	164
Figura 40 – Situação de Explicação com dados fictícios	165
Figura 41 – Situação de Resolução com dados fictícios	165
Figura 42 – Situação de explicação sobre técnicas de Amostragem	168
Figura 43 – Situação de Resolução com cálculos de medidas estatísticas	168
Figura 44 – Situação de explicação com abordagem explícita	170
Figura 45 – Situação de Resolução com abordagem explícita	170
Figura 46 – Situação de explicação com análise gráfica	173
Figura 47 – Situação de explicação com o cálculo ou identificação das medidas de dispersão	173
Figura 48 – Situação de explicação com o cálculo de probabilidades	174
Figura 49 – Situação de resolução com o cálculo ou identificação das medidas de dispersão	174
Figura 50 – Situação de explicação com o conceito de Amostra explícito	176
Figura 51 – Situação de Resolução com o conceito de Amostra explícito	176

Figura 52 – Situação de explicação com Dados Reais	177
Figura 53 – Situação de explicação com técnicas de Amostragem	178
Figura 54 – Situação de explicação com a realização de pesquisa amostral	178
Figura 55 – Situação de explicação com a análise dos dados de pesquisa amostral	180
Figura 56 – Situação de resolução com a realização de pesquisa amostral	180
Figura 57 – Número de inscritos no ENEM em todas as edições até 2022	184
Figura 58 – Questão do ENEM que aborda o conceito de Média aritmética	188
Figura 59 – Questão do ENEM que aborda o conceito de Média aritmética	189
Figura 60 – Questão do ENEM que aborda o conceito de Mediana	190
Figura 61 – Questão do ENEM que aborda o conceito de Mediana	191
Figura 62 – Questão do ENEM que aborda a Probabilidade	192
Figura 63 – Questão do ENEM que aborda a Probabilidade	192
Figura 64 – Resposta adequada de um estudante para a 1º Questão	196
Figura 65 – Resposta Inadequada de um estudante para a 2º (A) Questão	197
Figura 66 – Resposta adequada de um estudante para a 2º (A) Questão	197
Figura 67 – Resposta Inadequada de um estudante para a 2º (B) Questão	198
Figura 68 – Resposta adequada de um estudante para a 2º (B) Questão	198
Figura 69 – Resposta adequada de um estudante para a 3º (A) Questão	198
Figura 70 – Resposta Parcialmente Adequada de um estudante para a 3º (A) Questão	200
Figura 71 – Resposta Inadequada de um estudante para a 3º (A) Questão	200
Figura 72 – Resposta adequada de um estudante para a 3º (B) Questão	201
Figura 73 – Resposta Parcialmente adequada de um estudante para a 3º (B) Questão	201
Figura 74 – Resposta Inadequada de um estudante para a 3º (B) Questão	202
Figura 75 – Resposta adequada de um estudante para a 3º (C) Questão	202
Figura 76 – Resposta Parcialmente adequada de um estudante para a 3º (C) Questão	202
Figura 77 – Resposta inadequada de um estudante para a 3º (C) Questão	203
Figura 78 – Resposta adequada de um estudante para a 3º (D) Questão	203
Figura 79 – Resposta parcialmente adequada de um estudante para a 3º (D) Questão	204

Figura 80 – Resposta inadequada de um estudante para a 3º (D) Questão	204
Figura 81 – Resposta adequada de um estudante para a 4º Questão	205
Figura 82 – Resposta parcialmente adequada de um estudante para a 4º Questão	206
Figura 83 – Resposta inadequada de um estudante para a 4º Questão	206
Figura 84 – Resposta adequada de um estudante para a 5º Questão	207
Figura 85 – Resposta inadequada de um estudante para a 5º Questão	208
Figura 86 – Resposta adequada de um estudante para a 6º (A) Questão	209
Figura 87 – Resposta inadequada de um estudante para a 6º (A) Questão	209
Figura 88 – Resposta adequada de um estudante para a 6º (B) Questão	210
Figura 89 – Resposta inadequada de um estudante para a 6º (B) Questão	210
Gráfico 1 – Desempenho das turmas no pré-teste e pós-teste	213
Figura 90 – Resposta adequada de um estudante para a 1ºquestão do pré-teste – turma Lápis e Papel	216
Figura 91 – Resposta adequada de um estudante para a 1º questão do pós-teste – turma Geogebra	216
Figura 92 – Resposta inadequada de um estudante para a 2º (A) questão do pré-teste – turma Lápis e Papel	218
Figura 93 – Resposta inadequada de um estudante para a 2º (A) questão do pré-teste – turma Geogebra	218
Figura 94 – Resposta adequada de um estudante para a 2º (A) questão do pós-teste – turma Lápis e Papel	219
Figura 95 – Resposta adequada de um estudante para a 2º (A) questão do pós-teste – turma Geogebra	219
Figura 96 – Resposta inadequada de um estudante para a 2º (B) questão do pré-teste – turma Lápis e Papel	220
Figura 97 – Resposta inadequada de um estudante para a 2º (B) questão do pré-teste – turma Geogebra	220
Figura 98 – Resposta adequada de um estudante para a 2º (B) questão do pré-teste – turma Lápis e Papel	221
Figura 99 – Resposta inadequada de um estudante para a 2º (B) questão do pré-teste – turma Geogebra	221

- Figura 100 – Resposta parcialmente adequada de um estudante para a 3º questão do pré-teste – turma Lápis e Papel 223
- Figura 101 – Resposta adequada de um estudante para a 3º questão do pós-teste – turma Lápis e Papel 224
- Figura 102 – Resposta parcialmente adequada de um estudante para a 3º questão do pré-teste – Geogebra 225
- Figura 103 – Resposta adequada de um estudante para a 3º questão do pós-teste – turma Geogebra 226
- Figura 104 – Resposta parcialmente adequada de um estudante para a 4º questão do pré-teste – turma Lápis e Papel 228
- Figura 105 – Resposta parcialmente adequada de um estudante para a 4º questão do pré-teste – turma Geogebra 229
- Figura 106 – Resposta adequada de um estudante para a 4º questão do pré-teste – turma Lápis e Papel 229
- Figura 107 – Resposta adequada de um estudante para a 4º questão do pré-teste – turma Lápis e Papel 230
- Figura 108 – Resposta inadequada de um estudante para a 5º questão do pré-teste – turma Lápis e Papel 231
- Figura 109 – Resposta inadequada de um estudante para a 5º questão do pré-teste – turma Geogebra 232
- Figura 110 – Resposta adequada de um estudante para a 5º questão do pós-teste – turma Lápis e Papel 233
- Figura 111 – Resposta adequada de um estudante para a 5º questão do pós-teste – turma Lápis e Papel 233
- Figura 112 – Resposta inadequada de um estudante para a 6º(A) questão do pré-teste – turma Lápis e Papel 235
- Figura 113 – Resposta inadequada de um estudante para a 6º(A) questão do pré-teste – turma Geogebra 236
- Figura 114 – Resposta inadequada de um estudante para a 6º(A) questão do pré-teste – turma Lápis e Papel 236
- Figura 115 – Resposta adequada de um estudante para a 6º(A) questão do pré-teste – turma Geogebra 237

Figura 116 – Resposta inadequada de um estudante para a 6º(B) questão do pré-teste – turma Lápis e Papel	238
Figura 117 – Resposta inadequada de um estudante para a 6º(B) questão do pré-teste – turma Geogebra	238
Figura 118 – Resposta adequada de um estudante para a 6º(B) questão do pré-teste – turma Lápis e Papel	239
Figura 119 – Resposta adequada de um estudante para a 6º(B) questão do pré-teste – turma Geogebra	239
Figura 120 – Primeira questão da atividade 1	244
Figura 121 – Segunda questão da atividade 1	248
Figura 122 – Terceira questão da atividade 1	251
Figura 123 – quarta questão da atividade 1	256
Figura 124 – Quinta questão da atividade 1	259
Figura 125 – Primeira e Segunda questão da atividade 2	264
Figura 126 – Terceira questão da atividade 2	265
Figura 127 – Reposta de uma dupla para a 3º questão da atividade 2	270
Figura 128 – Primeira e Segunda questão da atividade 2 – Turma Geogebra	271
Figura 129 – Terceira questão da atividade 2 – Turma Geogebra	271
Figura 130 – Simulação realizada por uma Dupla na 3º questão – Turma Geogebra	275
Figura 131 – Simulação realizada por uma Dupla – Turma Geogebra	276
Figura 132 – Curva Normal das Notas de Alunos do Ensino Médio em Matemática	278
Figura 133 – Primeira questão da terceira atividade – Turma Lápis e Papel	280
Figura 134 – Segunda questão da terceira atividade – Turma Lápis e Papel	280
Figura 135 - Reposta de uma dupla para a 1º questão da atividade 3	282
Figura 136 – Reposta de uma dupla para a 2º questão da atividade 3	283
Figura 137 – Primeira questão da terceira atividade – Turma Geogebra	284
Figura 138 – Segunda questão da terceira atividade – Turma Geogebra	285
Figura 139 – Ferramenta análise univariada – Turma Geogebra	287
Figura 140 – Ferramenta Calculadora de Probabilidade – Turma Geogebra	288
Figura 141 – Resposta de um trio para a quarta atividade	291

Figura 142 – Resolução da quarta atividade – Turma Geogebra	293
Figura 143 – Simulação realizada por uma Dupla – Turma Geogebra	294

## LISTA DE QUADROS

Quadro 1 – Amostras possíveis de uma população e respectivas médias	49
Quadro 2 – Modelo do Letramento Estatístico	55
Quadro 3 – Modelo do Letramento Probabilístico	67
Quadro 4 – Critérios de correção por item do teste diagnóstico	114
Quadro 5 – Média das nove amostras	137
Quadro 6 – Critérios de correção por item do pré-teste e pós-teste	147
Quadro 7 – Desempenho médio por item da Questão 3	199
Quadro 8 – Percentual de resposta adequada por item da questão 1	215
Quadro 9 – Percentual de resposta adequada por item da questão 2	217
Quadro 10 – Percentual de resposta adequada por item da questão 3	222
Quadro 11 – Percentual por tipo de resposta da Questão 4	227
Quadro 12 – Percentual por tipo de resposta da Questão 5	231
Quadro 13 – Percentual por tipo de resposta da Questão 6º A	235
Quadro 14 – Percentual por tipo de resposta da Questão 6º B	237
Quadro 15 – Recorte do diálogo preliminar à Atividade 1 da turma Lápis e Papel	241
Quadro 16 – Recorte do diálogo preliminar à Atividade 1 da turma Geogebra	242
Quadro 17 – Recorte do diálogo da 1º questão da Atividade 1 - turma Lápis e Papel	244
Quadro 18 – Recorte do diálogo da 1º questão da Atividade 1 – turma Geogebra	246
Quadro 19 – Recorte do diálogo da 2º questão da Atividade 1 - turma Lápis e Papel	248
Quadro 20 – Recorte do diálogo da 2º questão da Atividade 1 - turma Geogebra	250
Quadro 21 – Recorte do diálogo da 3º questão da Atividade 1 - turma Lápis e Papel	253
Quadro 22 –Recorte do diálogo da 3º questão da Atividade 1 - Geogebra	254
Quadro 23 – Recorte do diálogo da 4º questão da Atividade 1 - turma Lápis e Papel	257

Quadro 24 – Recorte do diálogo da 4º questão da Atividade 1 - Geogebra	258
Quadro 25 – Recorte do diálogo da 5º questão da Atividade 1 - turma Lápis e Papel	260
Quadro 26 – Recorte do diálogo da 5º questão da Atividade 1 - turma Geogebra	262
Quadro 27 – Recorte do diálogo da 2º e 3º questão da Atividade 2 - turma Lápis e Papel	266
Quadro 28 – Recorte do diálogo da 2º e 3º questão da Atividade 2 - turma Geogebra	272
Quadro 29 – Frequências Relativas dos intervalos da Curva Normal	279
Quadro 30 – Recorte do diálogo da 2º e 3º questão da Atividade 3 - turma Lápis e Papel	280
Quadro 31 – Recorte do diálogo da 2º e 3º questão da Atividade 3 - turma Geogebra	285
Quadro 32 – Recorte do diálogo da atividade 4 - turma Lápis e Papel	289
Quadro 33 – Recorte do diálogo da atividade 4 - turma Geogebra	292

## LISTA DE TABELAS

Tabela 1 – Quantitativo das situações de Amostra por tipo de abordagem	148
Tabela 2 – Quantitativo das situações de Amostragem por tipo de contexto	150
Tabela 3 – Quantitativo das situações de Amostragem por tipo de população/amostra	151
Tabela 4 – Quantitativo das situações de Amostragem por habilidades	153
Tabela 5 – Quantitativo das situações de Amostra por tipo de abordagem	158
Tabela 6 – Quantitativo das situações de Amostragem por tipo de contexto	160
Tabela 7 – Quantitativo das situações de Amostragem por tipo de população/amostra	160
Tabela 8 – Quantitativo das situações de Amostragem por habilidades	161
Tabela 9 – Quantitativo das situações de Amostra por tipo de abordagem	163
Tabela 10 – Quantitativo das situações de Amostragem por tipo de contexto	164
Tabela 11 – Quantitativo das situações de Amostragem por tipo de população/amostra	166
Tabela 12 – Quantitativo das situações de Amostragem por habilidades	166
Tabela 13 – Quantitativo das situações de Curva Normal por tipo de abordagem	169
Tabela 14 – Quantitativo das situações de Curva Normal por tipo de contexto	171
Tabela 15 – Quantitativo das situações de Curva Normal por tipo de representação	171
Tabela 16 – Quantitativo das situações de Curva Normal por habilidades	172
Tabela 17 – Quantitativo das situações de Amostra por tipo de abordagem	175
Tabela 18 – Quantitativo das situações de Amostragem por tipo de contexto	177
Tabela 19 – Quantitativo das situações de Amostragem por tipo de população/amostra	178
Tabela 20 – Quantitativo das situações de Amostragem por habilidades	179
Tabela 21 – Percentual de resposta adequada por item da Questão 1	195

## SUMÁRIO

<b>1 INTRODUÇÃO</b>	22
<b>2 AMOSTRAGEM, CURVA NORMAL E SUAS RELAÇÕES</b>	26
2.1 A AMOSTRAGEM	26
2.2 A CURVA NORMAL	35
2.3 RELACIONANDO A AMOSTRAGEM COM A CURVA NORMAL: UMA ABORDAGEM INFORMAL DO TEOREMA CENTRAL DO LIMITE	47
<b>3 FUNDAMENTAÇÃO TEÓRICA</b>	53
3.1 O LETRAMENTO ESTATÍSTICO E PROBABILÍSTICO	53
3.2 O CICLO INVESTIGATIVO NO ENSINO DE ESTATÍSTICA E PROBABILIDADE	75
<b>4 ORIENTAÇÕES CURRICULARES E REVISÃO DE LITERATURA</b>	81
4.1 RECOMENDAÇÕES DOS DOCUMENTOS OFICIAIS SOBRE O ENSINO DA ESTATÍSTICA E DA PROBABILIDADE	81
4.2 ALGUNS ESTUDOS ANTECEDENTES SOBRE O ENSINO E APRENDIZAGEM DE AMOSTRAGEM	91
4.3 ALGUNS ESTUDOS ANTECEDENTES SOBRE O ENSINO E APRENDIZAGEM DA CURVA NORMAL	99
<b>5 MÉTODO</b>	105
5.1 OBJETIVOS	105
5.2 PROCEDIMENTOS METODOLÓGICOS	105
<b>6 RESULTADOS DO ESTUDO 1</b>	148
6.1 LIVROS DE PROJETOS INTEGRADORES	148
6.2 LIVROS DE PROJETOS DE VIDA	157
6.3 LIVROS DE CONHECIMENTO	162
6.4 LIVROS DE CIÊNCIAS HUMANAS E MATEMÁTICA	175
<b>7 RESULTADOS DO ESTUDO 2</b>	182
7.1 UM BREVE HISTÓRICO DO ENEM	182
7.2 MATRIZ DE REFERÊNCIA E A ANÁLISE DAS QUESTÕES	185
<b>8 RESULTADOS DO ESTUDO 3</b>	195
<b>9 RESULTADOS DO ESTUDO 4</b>	212

9.1 DESEMPENHO GERAL DAS TURMAS NO PRÉ-TESTE E NO PÓS- TESTE	212
9.2 INTERVENÇÃO DE ENSINO	240
<b>CONSIDERAÇÕES FINAIS</b>	296
<b>REFERÊNCIAS</b>	302

## 1 INTRODUÇÃO

A Estatística e a Probabilidade, enquanto áreas de conhecimento, exercem um relevante papel em nossa sociedade. As pesquisas estatísticas, por exemplo, permitem a análise e interpretação de uma variedade de dados e informações do nosso cotidiano. Consequentemente, seus resultados possibilitam uma melhor compreensão da realidade e subsidiam as tomadas de decisões do poder público, empresas e dos cidadãos.

Nesse contexto, Gal (2002,2005) aponta que, para o pleno exercício da cidadania, torna-se necessário que as pessoas sejam letradas em Estatística e Probabilidade para analisar e interpretar de modo adequado diversos fenômenos sociais. O autor destaca que a maioria dos indivíduos se portam como consumidores de dados e, para isso, necessitam de uma postura crítica frente às informações veiculadas. Além disso, há aqueles que são produtores de dados, inseridos em contextos de investigação, os quais devem compreender como são realizadas as pesquisas estatísticas.

Nesse sentido, é essencial que na escolaridade básica sejam sistematizadas possibilidades para o desenvolvimento do Letramento Estatístico e Probabilístico (Gal, 2002; 2005), para possibilitar, ao estudante, a construção de competências e habilidades voltadas para a compreensão dos conceitos e fundamentos relativos a essas duas áreas de conhecimento, como também, a realização de pesquisas estatísticas e o olhar crítico para veiculação de dados.

No que diz respeito aos processos de investigação, Guimarães e Gitirana (2013) formularam um modelo de ciclo investigativo, o qual contempla as fases que uma pesquisa estatística deve percorrer. Entre elas, destaca-se a definição se a pesquisa deve ser censitária ou amostral, o que implica em reflexões e estudos sobre conceitos relativos à Amostragem. Outra fase pertinente é análise e interpretação dos dados que abrange conceitos e modelos probabilísticos que permitem o tratamento dos dados.

A partir desse contexto, esta pesquisa de Doutorado tem como foco os conceitos de Amostragem e Curva Normal, importantes tópicos de conhecimento estatístico e probabilístico que estão presentes em muitos fenômenos do nosso cotidiano. Esses conceitos ainda são pouco explorados na Educação Básica no Brasil, entretanto, são fundamentais para a compreensão real dos dados e a construção de

dados estatísticos para tomadas de decisão (Luna e Guimarães, 2021; Araújo, 2020; Gonçalves, 2014). A Amostragem envolve as técnicas para seleção de amostras, de maneira que as mesmas sejam representativas de suas respectivas populações (Bayer, Echevest, Bittencourt e Rocha, 2005). Assim, refletir sobre os fatores que implicam na representatividade de uma amostra é fundamental. Da mesma forma, a Curva Normal que é considerada o principal modelo probabilístico de análise de dados presente na Inferência Estatística (Batanero, Tauber e Sánchez, 2004) precisa ser compreendida pelos cidadãos.

Com relação a essa temática, a Base Nacional Comum Curricular – BNCC (Brasil, 2018), principal documento orientador para o ensino no Brasil, propõe que na etapa do Ensino Médio, o estudante deve realizar pesquisas estatísticas amostrais, como também analisar criticamente as amostras de pesquisas estatísticas. Além disso, recomenda a elaboração e resolução de problemas que envolvem as medidas de tendência central e de dispersão e a identificação de espaços amostrais em eventos aleatórios para o cálculo de probabilidades em diferentes contextos, o que contempla implicitamente a Curva Normal.

Diante do exposto, dada a pertinência dessa temática para a formação dos cidadãos e reconhecendo que na Educação Básica devem ser sistematizadas possibilidades para a realização do seu ensino, esta pesquisa tem como objetivo geral investigar o que estudantes do 3º ano do Ensino Médio sabem e podem aprender sobre Amostragem, Curva Normal e suas relações.

Com a finalidade de alcançarmos esse objetivo geral, buscamos, enquanto objetivos específicos:

- Analisar o que tem sido proposto sobre Amostragem, Curva Normal e suas relações nos materiais para o ensino de Matemática (Livro de Projetos e Livros de Conhecimento) do PNLD 2021.
- Investigar o que foi proposto nas questões da prova de Matemática e suas Tecnologias do ENEM de 2022 sobre a Amostragem, Curva Normal e suas relações.
- Identificar os conhecimentos iniciais de estudantes do 3º ano do Ensino Médio sobre Amostragem, Curva Normal e suas relações.
- Analisar a contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada a partir de um processo interventivo.

Para tanto, fundamentamos a pesquisa no modelo de Letramento Estatístico e Probabilístico proposto por Gal (2002, 2005), os quais, em linhas gerais, são definidos como a capacidade de interpretar e avaliar criticamente uma informação estatística e probabilística em qualquer contexto e discutir ou comunicar suas reações, crenças e opiniões frente às informações e dados. Além disso, nosso estudo também está fundamento no modelo de ciclo investigativo elaborado por Guimarães e Gitirana (2013) que apresenta a pesquisa e suas fases como um eixo estruturador para a aprendizagem da Estatística.

Em termos metodológicos, para alcançarmos o objetivo geral da pesquisa, quatro estudos foram realizados. No estudo 1, analisamos os materiais para o ensino de Matemática (Livro de Projetos e Livros de Conhecimento) aprovados pelo Programa Nacional do Livro e do Material Didático - PNLD (Brasil, 2021). Assim, buscamos verificar o que é proposto nesses livros sobre Amostragem e Curva Normal e suas relações, reconhecendo que o Livro Didático, por vezes, é a principal fonte de informação no âmbito educacional e se constitui de um importante recurso didático, tanto para as ações docentes, auxiliando o planejamento das aulas e a avaliação da aprendizagem, como também para os discentes, propiciando a construção de conhecimentos e o desenvolvimento de competências e habilidades. (Thomaz, 2013; Bittar, 2017).

No estudo 2, investigamos o que foi proposto nas questões da prova de Matemática e suas Tecnologias do Exame Nacional do Ensino Médio (ENEM) de 2022 sobre os conceitos de Amostragem, Curva Normal e suas relações. Essa investigação foi conduzida a partir do reconhecimento de que esse exame se constitui como a principal porta de entrada para as universidades no Brasil e, por conseguinte, influencia diretamente as práticas educacionais nas escolas, especialmente na preparação dos estudantes para especificidades da prova, sobretudo nas turmas do 3º ano (Andrade, Batista, Santos e Mocrosky, 2020). Nas aulas de Matemática, por exemplo, é comum os professores utilizarem questões de edições anteriores do ENEM para ensinar os conteúdos planejados e, simultaneamente, familiarizarem os estudantes com o nível de exigência da prova. (Melão e Soares, 2012; Otobelli, 2018).

Já no estudo 3, identificamos os conhecimentos iniciais de estudantes do 3º ano do Ensino Médio sobre a referida temática. Esse estudo possibilitou a realização de uma diagnose com vistas a revelar o nível de aprendizagem nessa etapa de escolarização sobre a temática da presente pesquisa, como também, validar o

instrumento de análise para o estudo seguinte. A partir dos resultados desses três estudos, realizamos no estudo 4 uma intervenção de ensino que permitiu analisar a contribuição dos diferentes recursos (Lápis e Papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada no 3º ano do Ensino Médio. Acreditamos que esse estudo pode ampliar as possibilidades didático-pedagógicas para o processo de ensino e aprendizagem dessa temática, explicitando as potencialidades e desafios de cada recurso empregado.

Diante disso, a fim de cumprir com as propostas mencionadas anteriormente, estruturamos o presente estudo da seguinte maneira: No capítulo 2, abordamos a definição teórica de Amostragem e dos seus conceitos subjacentes. Além disso, apresentaremos o modelo teórico da Curva Normal e o Teorema Central do Limite que estabelece a relação entre a Amostragem e esse modelo, através da distribuição de probabilidade da média amostral.

O capítulo 3 é destinado para a apresentação da fundamentação teórica do nosso estudo que contempla a perspectiva do Letramento Estatístico e Probabilístico proposto por Gal (2002, 2005) e do Ciclo Investigativo como eixo estruturador para o ensino de Estatística elaborado por Guimarães e Gitirana (2013).

O capítulo 4 foi direcionado para a apresentação sobre as recomendações dos documentos oficiais sobre o ensino da Estatística e da Probabilidade com foco para a Amostragem e Curva Normal, como também, apresentamos uma revisão de literatura que contempla estudos anteriores que versam sobre a compreensão de estudantes e professores sobre estes conceitos e a importância e possibilidades de seu ensino e abordagem em sala de aula.

No capítulo 5 apresentamos nossa metodologia, destacando, os procedimentos metodológicos para cada estudo, os participantes da pesquisa, instrumento de coletadas de dados e os critérios de análise de dados.

Por fim, nos capítulos 6,7, 8, 9 apresentamos os resultados do estudo 1, 2,3 e 4, respectivamente. Em seguida, apresentaremos as considerações finais e as Referências.

## 2 AMOSTRAGEM, CURVA NORMAL E SUAS RELAÇÕES

Este capítulo tem por objetivo discorrer sobre a definição teórica de Amostragem e de seus conceitos subjacentes, tais como, população, censo, amostra, representatividade, variabilidade amostral e técnicas de Amostragem. Além disso, apresentaremos o modelo teórico da Curva Normal e o Teorema Central do Limite que estabelece a relação entre a Amostragem e esse modelo.

### 2.1 A AMOSTRAGEM

De modo formal, a Estatística é classificada como uma ciência que trata da coleta, organização, análise e interpretação de dados a fim de propiciar a melhor tomada de decisões a partir do que os dados têm a revelar (Bayer *et al*, 2005). Nesse sentido, essa ciência exerce um papel relevante por ter um vasto campo de aplicações em diversas áreas do conhecimento e, sobretudo, por servir de base para a realização de pesquisas estatísticas.

Em nosso cotidiano, por exemplo, frequentemente são divulgados nos diversos meios de comunicação (TV, jornal, rádio, internet, etc.) dados estatísticos de pesquisas sobre diferentes aspectos relacionados a nossa sociedade, tais como: econômico, político, social, biológico, entre outros. Esses resultados viabilizam as tomadas de decisões porque permitem o poder público, as empresas e os cidadãos planejarem suas ações visando uma maior assertividade em suas escolhas.

Em sua composição, a Estatística é dividida em duas grandes áreas: a Descritiva e a Inferencial. Em síntese, a Estatística Descritiva reúne métodos e conceitos para organização, resumo e descrição de dados e informações, como por exemplo: tabelas, gráficos, medidas de tendência central (médias, moda e mediana) e medidas de variabilidade ou dispersão (amplitude, desvio padrão e variância), entre outros.

A Estatística Inferencial, por sua vez, é responsável por dar suporte ao pesquisador, apresentando um conjunto de métodos e modelos que buscam caracterizar e projetar o comportamento de uma população a partir das estimativas de um subconjunto dela, ou seja, uma amostra dessa população. Para que isso ocorra, a Estatística Inferencial está apoiada na Amostragem, a qual é definida como a área responsável pelo desenvolvimento de “técnicas para seleção das unidades

*populacionais que formarão a amostra, de maneira que as mesmas sejam representativas de suas respectivas populações*" (Bayer et al, 2005, p.2). Deste modo, por meio das amostras representativas é possível obter conclusões verídicas que possam caracterizar e representar as suas respectivas populações. Para tanto, a Inferência Estatística também está alicerçada na Probabilidade, a qual permite estimar o grau de confiabilidade nos resultados encontrados nas pesquisas amostrais.

Diante dessas definições, para a realização de uma pesquisa estatística é pertinente compreendermos mais a fundo os conceitos de população e amostra. No âmbito da Estatística, a população diz respeito ao conjunto de todos os elementos (indivíduos, objetos, animais, etc.) que possuem ao menos uma característica em comum (Triola, 2008). Como por exemplo, podemos citar a população de brasileiros, a população de felinos de uma cidade e a população de livros de uma biblioteca.

Dessa forma, ao se realizar uma pesquisa com toda a população de interesse, a classificamos como censitária ou simplesmente censo. No Brasil, por exemplo, o Instituto Brasileiro de Geografia e Estatística (IBGE) realiza a cada 10 anos o censo demográfico, o qual levanta características de todos os brasileiros, como idade, sexo, renda, entre outros. Esse tipo de pesquisa possibilita o levantamento de dados e resultados com mais precisão, tendo em vista que a coleta é realizada com todos os elementos da população.

No entanto, em situações em que o tamanho da população é grande, nas quais o custo e tempo para desenvolver a pesquisa são enormes ou ainda quando a pesquisa resulta na inutilização de um produto, nem sempre é possível/recomendável realizar um censo. Logo, torna-se mais viável realizar uma pesquisa com uma parte representativa da população, ou seja, uma pesquisa amostral. Assim, o termo amostra, na Estatística, é definido como um subconjunto (grupo) da população (Triola, 2008).

Nesse contexto, no processo de seleção de uma amostra é imprescindível que ela seja representativa para que sejam garantidas generalizações adequadas e confiáveis para a população. Deste modo, a representatividade está ligada ao quanto as características da população estão contempladas na amostra. Em uma pesquisa com a população do estado de Pernambuco, por exemplo, a amostra deve levar em conta algumas características desse público, como o sexo, classe social, nível de escolarização, idade, dentre outros, para que de fato ela seja representativa. Assim, se a referida população tiver mais mulheres do que homens, a amostra deve ter,

proporcionalmente, mesmo percentual do sexo de sua população. Em síntese, uma amostra representativa reflete, com a maior precisão possível, as características de sua respectiva população, a partir das variáveis em estudo e obedecendo às devidas proporções.

Nesse viés, a representatividade de uma amostra depende de outros conceitos que estão subjacentes à Amostragem, tais como: a variabilidade, o tamanho da amostra e o método de Amostragem. Se selecionarmos mais de uma amostra de mesmo tamanho de uma população, é natural que os dados amostrais não apresentem os mesmos resultados, ou seja, as estimativas variem entre si e também em relação à população. Além disso, a depender da característica populacional, seus dados podem ter maior ou menor grau de variabilidade. Logo, em um processo amostral, a variabilidade é um fator importante a ser considerado e está associada ao grau dispersão dos dados e a quantidade de fatores para caracterizar uma população.

Em um lote de peças produzidas em uma fábrica, por exemplo, é comum termos uma variabilidade pequena, ou seja, suas características são mais homogêneas do que heterogêneas, o que nos permite a seleção de poucas peças para caracterizarmos toda a população. Diante disso, podemos inferir que o tamanho da amostra está estritamente relacionado com a variabilidade dessa população. Em uma população mais homogênea, isto é, que tenha uma menor variabilidade (lote de peças), menor é o tamanho necessário da amostra para representá-la. Do contrário, uma população mais heterogênea (população pernambucana), maior é esse tamanho, pois será necessário conter todas as características da população na amostra para que a mesma seja, de fato, representativa. Logo, a representatividade, variabilidade e tamanho da amostra são conceitos interligados no processo de Amostragem.

De acordo com Reyes (2019), o processo de pesquisa amostral implica em 3 tipos de variabilidade, a saber: 1) A *variabilidade natural* está associada à natureza dos elementos, sejam eles populacionais ou amostrais, tendo em vista que os mesmos são essencialmente diferentes. Nesse sentido, qualquer variável estática, como exemplo da altura de pessoas, vai apresentar variabilidade porque os elementos (pessoas) se diferem entre si. 2) A *Variabilidade Amostral* diz respeito a variação das estatísticas amostrais. Como já dito anteriormente, se selecionarmos mais de uma amostra de mesmo tamanho de uma população e para cada uma delas calcularmos uma estatística, a exemplo da média, os valores dessas estatísticas irão variar de uma amostra para outra. 3) A *Variabilidade de medição* representa a diferença entre o

resultado amostral e o verdadeiro resultado da população, tendo em vista que dificilmente uma amostra vai apresentar resultados idênticos a sua população. Por exemplo, quando se estabelece um intervalo de confiança para uma estatística amostral, se considera um intervalo de valores no qual a estimativa pontual é o centro e a variabilidade de medição (margem de erro) determina a estimativa máxima de erros do resultado, os quais têm relação com a aleatoriedade do processo de seleção da amostra. Em pesquisas, é comum a utilização do intervalo de 95% de confiança, ou seja, se fossem realizadas 100 pesquisas idênticas, 95% delas apresentariam o mesmo intervalo que conteria o verdadeiro parâmetro da população. Assim, quando se anuncia que o Candidato X tem 48% de intenção de votos com margem de erro de 2%, entende-se que sua intenção de voto está no intervalo entre 46 a 50%. A partir disso, comprehende-se que quanto maior for o tamanho da amostra, menor será a variabilidade de medição, ou seja, as estimativas estarão cada vez mais próximas dos parâmetros populacionais.

Em continuidade, outro fator já mencionado anteriormente como pertinente para a representatividade amostral são os métodos de Amostragem. Esses métodos se ocupam do processo de seleção dos elementos que irão compor a amostra e são divididos em dois grupos: não-probabilísticos (não-aleatórios) e probabilísticos (aleatórios).

### 2.1.1 Amostragem Não-probabilística

Os métodos não probabilísticos são aqueles em que os elementos da população não têm a mesma probabilidade ser escolhido para participar da amostra ou essa probabilidade é desconhecida (Moore, 1995). A seleção desses elementos fica a critério e julgamento do pesquisador, não sendo possível medir o erro amostral e as suas escolhas pessoais podem vir a resultar em tendenciosidade. Por conta desse caráter não aleatório, não é possível, através dos métodos não-probabilísticos, se obter representatividade para a população. Porém, eles têm suas utilidades por serem mais baratos e convenientes, sendo bastante úteis para pesquisas exploratórias e de geração de hipóteses quando não há intenção de generalizar os resultados obtidos na amostra para a população.

A escolha desse tipo de Amostragem ainda se justifica nas ocasiões em que há a inviabilidade de identificação de todos os elementos da população, como

também, limitações de tempo, recursos financeiros e materiais necessários para a realização de uma pesquisa (Mattar, 1996). Os métodos não-probabilísticos são classificados em: Amostragem por Resposta Voluntária, Amostragem por Conveniência, Amostragem por Julgamento e Amostragem por Cotas (Triola, 2008).

### 2.1.2 Amostragem por resposta voluntária

Essa técnica se caracteriza pelo julgamento, por parte dos respondentes, se irão participar da pesquisa e compor a amostra ou não. Triola (2008) indica três tipos de situações que utilizam esse método de Amostragem para realizar uma pesquisa: 1) Pesquisas realizadas pela internet, através de sites ou redes sociais, nas quais os internautas decidem se respondem ou não. 2) Pesquisas feitas por empresas de transporte e distribuição da correspondência, nas quais as pessoas possuem a prerrogativa de decidir se enviam ou não as respostas; 3) Pesquisas por telefone, divulgadas em anúncios de veículos de comunicação, como jornais, rádio e televisão, através dos quais pedem que o consumidor ligue para o contato telefônico para registrar sua opinião.

Em nosso cotidiano, é comum observarmos esse tipo de Amostragem. Sites de notícia, por exemplo, utilizam enquetes para medir a opinião pública sobre determinado assunto. Como também, alguns canais de televisão disponibilizam um número para os telespectadores ligarem e deixarem suas opiniões sobre determinada decisão do poder público ou tema abordado. Em situações desse tipo, apenas os telespectadores que assistem ao canal no momento da enquete optam em responder ou não a pesquisa, o que caracteriza uma Amostragem por Resposta Voluntária e também não probabilística, ou seja, não se pode generalizar os resultados observados para toda a população.

### 2.1.3 Amostragem por Conveniência

Como o próprio nome sugere, esse tipo de Amostragem seleciona os elementos para compor a amostra com base na conveniência do pesquisador (Triola, 2008). O principal critério utilizado para seleção amostral é a facilidade de acesso a esses elementos. Esse tipo de Amostragem se caracteriza por ser mais barata e de

fácil execução, sendo bastante utilizado em pesquisas exploratórias, nas quais o objetivo é gerar ideias e desenvolver hipóteses.

Em uma pesquisa com pessoas, por exemplo, o entrevistador pode se localizar em centros urbanos populosos, como uma avenida ou um parque e ir entrevistando pessoas que estão próximas a ele. Pode-se ainda, entrevistar amigos, colegas e familiares, ou seja, pessoas as quais ele tem fácil acesso. Um outro exemplo se verifica em pesquisas na área da saúde, nas quais é comum a seleção de animais já disponíveis nos laboratórios para se realizar teste de medicamentos, assim o critério de escolha se dá por conveniência, tornando o estudo mais prático para os pesquisadores.

#### 2.1.4 Amostragem por Julgamento

Nesse tipo de técnica, o pesquisador intencionalmente escolhe os elementos para compor a amostra com base no seu julgamento, selecionado o que acredita serem mais apropriados e representativos para o estudo em questão (Triola, 2008). Essa técnica se caracteriza por ser mais rápida e de fácil execução e é normalmente utilizada por empresas que produzem materiais de consumo.

Como exemplo, podemos citar uma empresa fabricante de computadores que pretende lançar uma nova linha de notebooks e, para tanto, deseja testar a opinião de usuários sobre o modelo, desempenho e funcionalidades. Nesse caso, não faz sentido o entrevistador selecionar qualquer pessoa para compor a amostra, mas sim, usuários de notebook ou técnicos de informática, porque estes terão uma análise mais criteriosa em relação ao novo produto do que o público em geral.

#### 2.1.5 Amostragem por Cotas

Nessa técnica, os elementos que fazem parte da amostra são selecionados segundo categorias (quotas) estabelecidas de acordo com a proporção desses elementos na população. Em se tratando de uma pesquisa com pessoas, por exemplo, a população pode ser dividida em cotas como idade, o sexo, a classe social, dentro outros, que serão contempladas na amostra na mesma proporção que ocorrem na população (Triola, 2008).

Esse tipo de técnica é frequentemente utilizado em pesquisas de mercado e de opinião pública. Para realiza-la, inicialmente, o pesquisador estabelece as categorias (cotas) de controle dos elementos da população que julga como importantes para o estudo e, posteriormente, faz seleção de elementos da amostra com base em um processo de conveniência ou julgamento. Se imaginarmos, por exemplo, uma pesquisa de opinião realizada em uma determinada cidade na qual, do total de habitantes, há mais mulheres do que homens, consequentemente, de acordo com a cota do sexo, os elementos da amostra serão compostos por mais mulheres do que homens, na mesma proporção do que é observado na população.

Apresentados os tipos de técnicas não probabilísticas, abordaremos na seção a seguir as técnicas probabilísticas, também conhecidas no campo estatístico como Amostragem Aleatória.

#### 2.1.6 Amostragem Probabilística

Os métodos probabilísticos, por sua vez, são aqueles em que a probabilidade de cada elemento da população pertencer à amostra é conhecida e diferente de zero (Moore, 1995). Assim, nesses métodos, o pesquisador tem a listagem de todos os elementos da população e os elementos que irão compor a amostra são selecionados de modo aleatório e imparcial, como por exemplo, por um sorteio manual ou eletrônico.

Consequentemente, as amostragens probabilísticas garantem a representatividade e generalizações para a população. Além disso, é possível, estimar as margens de erro das estimativas, possibilitando um maior controle acerca dos resultados gerados e como eles podem ser generalizados. Isso significa que podemos associar aos resultados uma certa probabilidade de que estejam, de fato, corretos, isto é, uma medida que indica o grau da confiabilidade das conclusões obtidas. No presente estudo, abordaremos os métodos probabilísticos mais comuns em pesquisas (Triola, 2008): Amostragem Aleatória Simples, Amostragem Sistemática, Amostragem Estratificada e Amostragem por Conglomerados.

#### 2.1.7 Amostragem Aleatória Simples

Nesse tipo de técnica, todos os elementos da população têm a mesma chance de pertencer à amostra e são escolhidos mediante a um sorteio (Moore, 1995; Triola,

2008). O primeiro passo para o pesquisador é ter uma lista com a identificação de toda a população para, em seguida, escolher aleatoriamente os elementos amostrais mediante a um sorteio manual ou eletrônico.

Assim, se a população tem tamanho  $N$ , cada elemento tem  $1/N$  probabilidade de pertencer a amostra, o que garante a representatividade. Esse tipo de Amostragem é recomendado para população não muito grande e mais homogênea, tendo em vista que o pesquisador conhecer as unidades da população e estas tendem a ter características muito próximas. Como exemplo, podemos citar um sorteio de estudantes de uma escola para a composição de uma comissão de representação estudantil, no qual todos os elementos da população são conhecidos e essa população é homogênea porque o único critério para pertencer à comissão é ser estudante matriculado na escola.

### 2.1.8 Amostragem Sistemática

Consiste na técnica em que os elementos da amostra são escolhidos conforme algum critério ou fator de repetição. Deste modo, as unidades da população são listadas e ordenadas (lista de funcionários, casas de rua, peças de um lote) e a retirada dos elementos amostrais é feita periodicamente, a partir de um intervalo pré-estabelecido (Moore, 1995; Triola, 2008).

Nesse sentido, os elementos da amostra são selecionados a partir de intervalos fixos ( $k$ ). Logo, divide-se o tamanho da população ( $N$ ) pelo tamanho da amostra ( $n$ ) a ser selecionada e, consequentemente, se obtém o intervalo de Amostragem ( $i$ ). Na sequência, um número aleatório entre 1 que representa o primeiro elemento e  $i$  que indica o último elemento do intervalo é escolhido, o qual será o primeiro elemento da amostra. A partir deste, os próximos são extraídos pela periodicidade do intervalo ( $k$ ), utilizando o mesmo número de escolha do primeiro intervalo. Dessa forma, as unidades da amostra são escolhidas a partir do mesmo número ordinal selecionado em intervalos de tamanho  $k$  (Triola, 2008).

Para exemplificar, consideremos que uma fábrica de peças automotivas deseja selecionar uma amostra sistemática de 10 peças em um lote de 200 peças produzidas. Para tanto, inicialmente divide-se o tamanho da população pelo tamanho da amostra, e nesse caso, obtém-se o tamanho de Amostragem igual a 20. Na sequência, um número aleatório é escolhido entre 1 e 20. Suponhamos que o número escolhido

aleatoriamente foi o 8. Consequentemente, a amostra começará a ser formada pelo número 8 do primeiro intervalo de 20 números e pelos próximos oitavos números nos intervalos de 20 números subsequentes, ou seja, a amostra será composta pelas peças com os seguintes números: 8, 28, 48, 68, 88, 108, 128, 148, 168, 188.

Essa técnica é frequentemente utilizada em nosso cotidiano para se medir a qualidade de produtos industrializados. Como no exemplo citado, alguns produtos são selecionados dentre um lote de produção a fim de se verificar possíveis falhas ou adulteração.

### 2.1.9 Amostragem Estratificada

Por meio dessa técnica, a população é dividida em subgrupos mais homogêneos denominados de estratos e, em seguida, os elementos de cada estrato são escolhidos de forma aleatória (Triola, 2008). Assim, a partir do conhecimento de toda a população, as variáveis de estratificação são elaboradas de acordo com as características de interesse da pesquisa, de modo que todo elemento populacional tenha a chance de pertencer somente a um estrato e nenhum desse elemento deve ser descartado.

A Amostragem Estratificada se assemelha à Amostragem por Cotas apresentadas anteriormente. No entanto, a diferença determinante reside no fato da estratificada ser probabilística, o que garante o conhecimento sobre a probabilidade de cada elemento da população pertencer aos estratos das pesquisas.

Esse tipo de Amostragem é frequentemente utilizado em populações mais heterogêneas e podemos citar o exemplo da pesquisa de opinião pública sobre a aprovação da gestão do governo federal no Brasil, na qual a população brasileira é dividida em estratos como, idade, sexo, classe social, região, religião, dentre outros para compor uma única amostra e representar toda a população.

### 2.1.10 Amostragem por Conglomerado

Nesse tipo de técnica, a população é dividida em subgrupos, denominados conglomerados, de forma que cada grupo seja o mais representativo possível da população. Ou seja, carregue consigo o máximo de características de interesse do estudo e, consequentemente, representam o universo pesquisado.

No processo dessa técnica, inicialmente são determinados os conglomerados. Em seguida, uma amostra aleatória simples desses conglomerados é obtida, e por fim, os elementos dentro de cada grupo são analisados. Podemos citar, como exemplo, pesquisas de opinião que são realizadas em grandes cidades. O pesquisador para tornar o processo mais simples sem deixar de ser representativo, divide a área da cidade em conglomerados, como quarteirões ou bairros. Posteriormente, seleciona aleatoriamente alguns desses bairros para, na sequência, pesquisar todos os moradores que estão localizados neles.

Esse tipo de Amostragem é recomendado quando a população de interesse está espalhada por uma extensa área geográfica. Dessa forma, realizar estudos por agrupamentos torna-se mais vantajoso por possibilitar a seleção de uma amostra de tamanho maior em um espaço geográfico menor com economia de tempo e custo.

Apresentamos nessa seção o conceito de Amostragem e as principais técnicas que estão subjacentes a ela. Dando continuidade ao estudo, passaremos a abordar na seção seguinte os fundamentos teóricos da Curva Normal, classificada como o principal modelo de análise de dados presente na Inferência Estatística (Batanero, Tauber e Sánchez, 2004).

## 2.2 A CURVA NORMAL

Na seção anterior, vimos que a Inferência Estatística é responsável por fornecer técnicas e modelos que possibilitam o pesquisador inferir sobre uma população a partir de uma amostra representativa da mesma. Nesse viés, antes de apresentarmos o modelo da Curva Normal, é pertinente compreendermos, inicialmente, o conceito de variável aleatória.

Nos processos inferenciais, ao se realizar um experimento aleatório ou uma pesquisa amostral, é possível atribuir um valor numérico para a variável de interesse a fim de uma ampla caracterização dos dados. Nesse sentido, em termos de conceitualização, uma variável aleatória  $X$  é aquela que representa um valor numérico associado a cada um dos resultados de um experimento aleatório. Como exemplo, em uma pesquisa amostral probabilística sobre a renda mensal dos moradores de uma cidade, ao entrevistar um munícipe se obtém um valor numérico aleatório para a variável da pesquisa. Em uma pesquisa pautada em um questionário no qual se apresenta uma proposição e as respostas possíveis são “sim” ou “não”, é possível

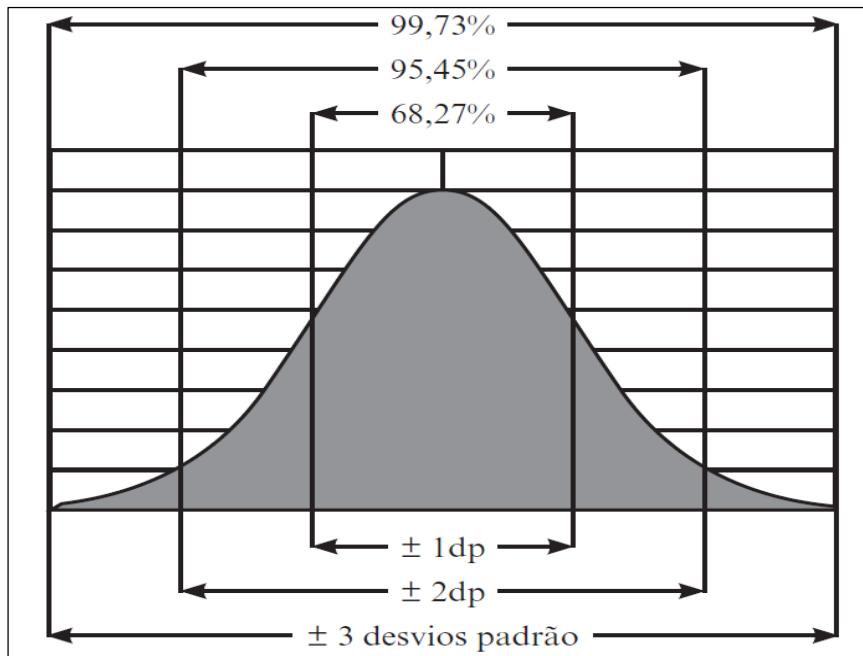
associar a questão problema uma variável que toma dois valores, 0 ou 1, por exemplo, correspondentes às respostas “sim” ou “não”, respectivamente. Em síntese, uma variável aleatória associa um número a cada resultado possível, possibilitando uma análise quantitativa para os estudos.

Nessa direção, as variáveis aleatórias são classificadas em dois tipos: Discreta ou Contínua. Ela será discreta quando for resultado de uma contagem, ou seja, assumir respostas cujos os valores numéricos são sempre inteiros, a exemplo do número de pessoas de uma população, o número de irmãos e o número de letras do nome. Já as do tipo Contínua são aquelas em que os valores são resultados de uma mensuração, isto é, podem assumir valores que pertencem a um intervalo real, ou seja, número decimais. Como exemplo, podemos citar a altura, a massa, temperatura, renda e pressão arterial (Cazorla, Magina, Guimarães e Gitirana, 2017; Ferreira, 2015).

Para a representação e análise de dados dessas variáveis, a inferência dispõe ainda de modelos de distribuição de probabilidades que podem ser representados em tabelas ou gráfico, os quais apresentam a probabilidade para cada valor que a variável aleatória pode assumir, o que possibilita inferir sobre o comportamento dos dados (Triola, 2008).

Em se tratando de variáveis contínuas, a Distribuição Normal ou Curva Normal é considerada o principal modelo de análise do comportamento dos dados, abarcando uma variedade de fenômenos do nosso cotidiano que apresentam uma Distribuição Normal ou aproximadamente normal dos dados (Batanero *et al*, 2004; Gonçalves, 2014). Graficamente, a Curva Normal tem a forma de um sino (Figura 1) e a importância desse modelo se verifica ainda por ser uma aproximação para outros tipos de distribuições.

Figura 1 – Representação Gráfica da Curva Normal



Fonte: São Paulo (2014).

Walpole, Myers, Myers e Ye (2008) e Tauber (2001) destacam que diversos fenômenos físicos, sociais, biológicos, geográficos que ocorrem no nosso cotidiano podem ser modelados pela Curva Normal. Se pensarmos na variável altura (cm), por exemplo, para uma população suficiente grande, haverá uma média como centro da distribuição, uma concentração maior de valores em torno da média, e quanto maior o desvio-padrão em relação a esse valor, menor será a frequência dos dados, sendo pouco provável terem pessoas com a estatura muito alta ou muito baixa. Assim, a Curva Normal contribui amplamente para a inferência estatística porque diversos fenômenos na natureza seguem um determinado padrão, assemelhando-se à Curva Normal, em que os valores extremos têm menor probabilidade de ocorrência e os valores mais próximos da média ocorrem em maior número de vezes, ou seja, com uma maior probabilidade.

Historicamente, de acordo com Duarte (2010), a Curva Normal se originou no século XVIII através da observação do erro de mensuração, também conhecido por “lei do erro”. Segundo o autor, alguns registros históricos apontam que os erros de mensuração eram observados por astrônomos da época, na tentativa de estimar as órbitas dos corpos celestes. Posteriormente, através do matemático francês Abraham de Moivre (1667 – 1754), a Curva Normal foi introduzida a um formalismo matemático. Em um artigo impresso no livro “A doutrina do acaso”, Moivre foi o primeiro matemático

a trabalhar com a “lei do erro” e procurou apresentar de forma algébrica a teoria das probabilidades. (Bittencourt e Viali, 2006).

Estudos iniciais com a Curva Normal também foram realizados pelo matemático belga Adolph Quetelet (1796 –1874), ao coletar dados relativos a medidas do tórax de soldados escoceses e da altura de soldados franceses. Ele verificou que os dados dessas medidas obedeciam a uma Distribuição Normal. Além disso, a Curva Normal também é denominada de Distribuição Gaussiana. Este fato deve-se ao matemático alemão Karl F. Gauss (1777 – 1855) que realizou relevantes estudos formais sobre a lei dos erros. Gauss contribuiu para a formulação de fundamentos teóricos do modelo normal, o qual se constitui de uma representação adequada para modelar fenômenos sociais.

### 2.2.1 O modelo da Curva Normal

Como visto, a distribuição Normal possui uma representação gráfica na forma de curva, se assemelhando ao formato de um sino. Esse modelo de normalidade para a distribuição de dados, é definido, no campo algébrico, pela seguinte função de densidade de probabilidade: (Duarte, 2010).

$$f(x|\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ para } x \in \mathbb{R}, -\infty < \mu < \infty \text{ e } \sigma > 0$$

A partir dessas informações, destaca-se que essa função densidade é uma função sempre não negativa e a área sob seu gráfico e o eixo das abscissas, ou seja, a área da curva, é sempre igual a 1. Assim, em termos matemáticos, tal função satisfaz as seguintes condições:

- 1)  $f(x) \geq 0$ , para todo  $x \in \mathbb{R}$

- 2) A área definida por  $f(x)$  é igual a 1.

Além disso, a Curva Normal é uma distribuição continua, ou seja, sua variável pode assumir qualquer valor real dentro do intervalo  $(-\infty, +\infty)$ . Em termos matemáticos, trata-se de uma função exponencial de uma variável real, que fica determinada a partir dois parâmetros, a  $\mu$  (média) e  $\sigma$  (desvio padrão) da população. Cazorla e Santana (2006, p.18) definem a média como uma medida que “resume e representa um conjunto de dados em um único valor”. E o seu cálculo “remete à divisão em partes

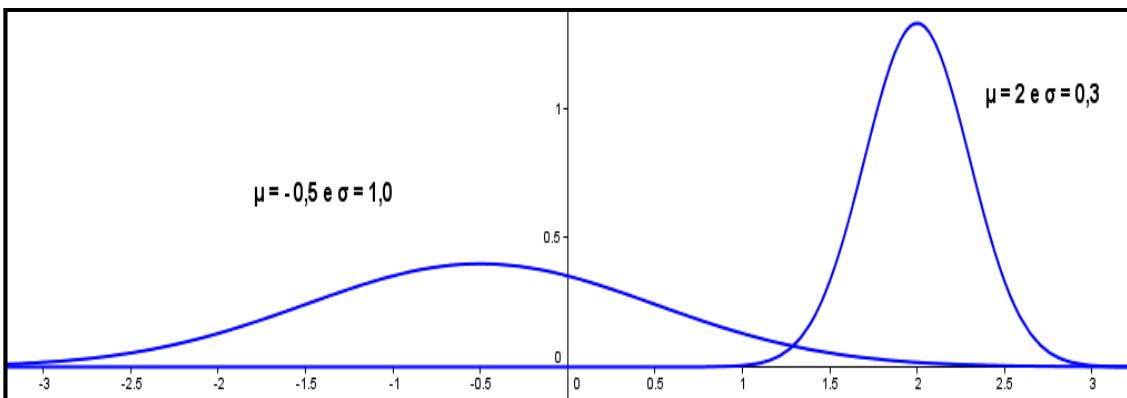
*iguais do todo entre seus componentes. Assim, seu algoritmo consiste em somar todos os valores da variável e dividir pelo número de dados".* Com relação ao desvio padrão, Martins (2013, p. 01) o conceitua como “*uma medida de dispersão dos dados relativamente à média*”, ou seja, representa a distância de um dado em relação à média do conjunto ao qual esse dado pertence.

Com base nessas definições, para indicar que uma determinada variável ( $X$ ) possui uma distribuição normal a partir da média e do Desvio Padrão, utiliza-se a notação  $X \sim N(\mu, \sigma^2)$ . Além das informações abordadas até aqui, cabe destacarmos ainda que a Curva Normal possui as seguintes propriedades (Duarte, 2010).

- 1) A Curva Normal é simétrica em relação a  $\mu$  (média), ou seja, ao centro.
- 2) Em uma Curva Normal, os valores de média, moda e mediana coincidem.
- 3) o valor máximo de  $f(x)$  é obtido quando  $x = \mu$ .
- 4) A área total sob curva e o eixo da abscissa é igual a 1 ou 100%.
- 5) A Curva Normal possibilita determinar probabilidades associadas aos valores da área de intervalos da distribuição.
- 6) A probabilidade de uma variável assumir um valor entre dois pontos é determinada pela área sob a curva que os contêm.
- 7) A Curva Normal é assintótica em relação ao eixo das abscissas.

Em acréscimo, o formato da Curva Normal depende dos parâmetros média ( $\mu$ ) e desvio padrão ( $\sigma$ ) estabelecidos em cada situação. Consequentemente, ao variá-los, o gráfico da Curva apresentará movimentos de translação e contração/expansão. Isso significa que o valor da média ( $\mu$ ) estabelece o centro da Curva Normal e o desvio padrão ( $\sigma$ ) indica a dispersão do conjunto. Deste modo, quanto maior o desvio padrão, maior será a dispersão e a amplitude dos dados, e, consequentemente, a Curva Normal será mais comprimida verticalmente e mais expandida horizontalmente. Em contrapartida, quanto menor o desvio padrão, menor será a dispersão do conjunto e a amplitude dos dados e, logo, a Curva Normal será mais expandida verticalmente e mais comprimida horizontalmente. Essas informações podem ser observadas na Figura 2, com dois exemplos distintos de Curvas Normais com seus respectivos parâmetros.

Figura 2 – Representação de Curvas Normais com parâmetros diferentes

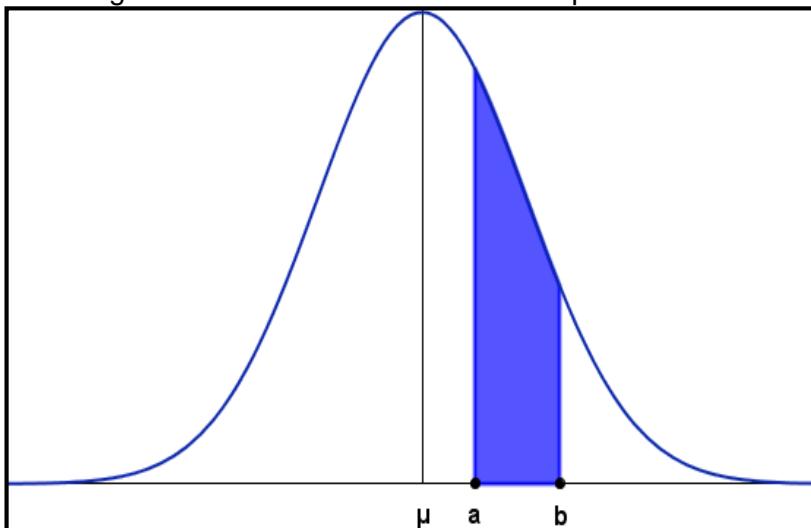


Fonte: Elaborado pelo autor a partir do Geogebra (2020).

### 2.2.2 O Cálculo de probabilidades na Curva Normal

Na seção anterior, vimos que a Curva Normal possibilita determinar probabilidades associadas da área sob a curva e o eixo das abscissas. Deste modo, para quaisquer que sejam os valores da média ( $\mu$ ) e do desvio padrão ( $\sigma$ ) de uma Curva Normal, a área total da curva será sempre igual a 1 ou 100% de probabilidade. Neste viés, a área sob a Curva Normal entre dois pontos quaisquer se configura a probabilidade da variável tomar um valor entre esses pontos, como podemos observar na Figura 3 adiante, a área sob a curva determinada pelo intervalo entre dois pontos  $a$  e  $b$  em uma Curva Normal.

Figura 3 – Área sob a curva entre os pontos  $a$  e  $b$ .



Fonte: Elaborado pelo autor a partir do Geogebra (2020).

Diante disso, o cálculo da área de um dado intervalo  $(a, b) \in \mathbb{R}$  de uma Curva Normal, que corresponde a probabilidade da variável em questão assumir valor nesse intervalo, é determinado formalmente por:

$$P(a < x < b) = \int_a^b f(x)dx$$

Conhecidos os parâmetros média ( $\mu$ ) e o desvio padrão ( $\sigma$ ) da Curva Normal associada a Função de Densidade e o intervalo de interesse sob a Curva, utiliza-se recursos do cálculo diferencial integral para determinar a área e, consequentemente, a probabilidade da variável assumir valor no intervalo estabelecido.

Em acréscimo, tendo em vista que a Curva Normal fica completamente especificada por sua média e seu desvio padrão, devemos compreender ainda que em todo modelo normal, aproximadamente 68% da área total da curva está localizada no intervalo  $\mu \pm \sigma$ . Por sua vez, o intervalo  $\mu \pm 2\sigma$ , isto é, a área sob a curva compreendida entre a média e duas vezes o valor do desvio padrão tanto para mais como para menos, contempla aproximadamente 95% da área total. Por fim, o intervalo  $\mu \pm 3\sigma$ , ou seja, a área sob a curva compreendida entre a média e três vezes o valor do desvio padrão tanto para mais como para menos, abrange aproximadamente 99,7% da área total compreendida sob a curva e o eixo ( $x$ ). Essas informações podem ser visualizadas graficamente na Figura 1, apresentada anteriormente.

No entanto, de modo alternativo, mais simples e prático, o cálculo da área entre dois pontos quaisquer de uma Curva Normal e consequentemente a probabilidade da variável tomar valor no intervalo estabelecido por esses pontos, pode ser obtido através de dois métodos. O primeiro é a Curva Normal Padronizada ( $Z$ ), que consiste em efetuar a mudança da variável, transformando a variável aleatória ( $X$ ), normalmente distribuída, em uma variável aleatória ( $Z$ ), também normalmente distribuída, através da seguinte expressão:

$$Z = \frac{x - \mu}{\sigma}$$

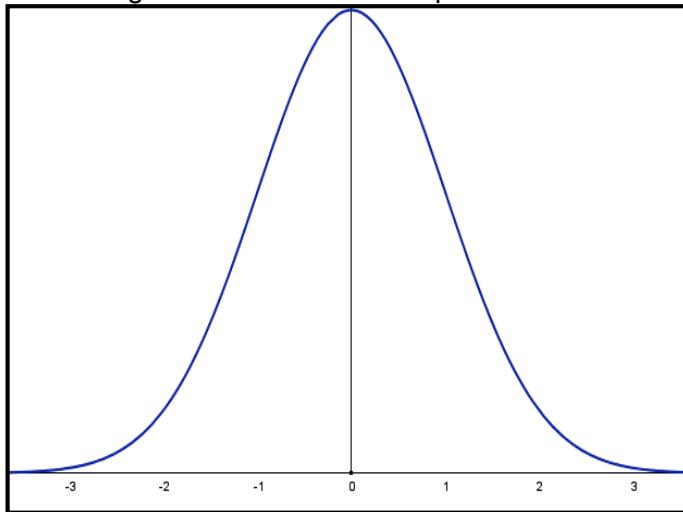
Assim, padronizar uma variável aleatória ( $X$ ) de média ( $\mu$ ) e desvio padrão ( $\sigma$ ) consiste em determinar a partir desta, uma nova variável aleatória ( $Z$ ), cujos valores são obtidos subtraindo-se de cada valor de ( $X$ ) da média populacional e posteriormente, divide-se o resultado pelo desvio padrão.

Desta forma, para o cálculo de probabilidades, devemos inicialmente transformar a variável aleatória ( $X$ ), através da expressão indicada, em uma distribuição equivalente e padronizada de variável ( $Z$ ). Com essa transformação, o valor encontrado ( $Z$ ) indica a intensidade de afastamento do ponto de variável ( $X$ ) até a média ( $\mu$ ), em relação ao desvio padrão ( $\sigma$ ), isto é, o comprimento relativo em termos de desvios padrões, entre a média ( $\mu$ ) até o ponto ( $X$ ). A distribuição normal padronizada (Figura 4) possui sempre parâmetros  $Z \sim N(0,1)$ , isto é, média  $\mu = 0$  e desvio padrão  $\sigma = 1$ .

$$f(z) = \frac{f(x) - f(\mu)}{f(\sigma)} = \frac{\mu - \mu}{\sigma} = 0$$

$$\sigma(z) = \frac{1}{\sigma^2} \cdot \sigma(x - \mu) = 1$$

Figura 4 – Curva Normal padronizada



Fonte: Elaborado pelo autor a partir do Geogebra (2020).

Após a realização dessa transformação e determinação do valor do ponto da variável  $Z$ , a segunda etapa do cálculo consiste em consultar a tabela da Distribuição Normal padronizada ( $Z$ ) (Figura 5), que contempla os valores da área ou probabilidade entre a média ( $u = 0$ ) e qualquer ponto ( $z$ ) e assim, encontrar o valor da probabilidade desse ponto determinado. Deste modo, esse método possibilita, através de qualquer Distribuição Normal, com seus respectivos valores da média e do desvio padrão, a transformação da variável aleatória ( $X$ ) em questão em uma variável padronizada ( $Z$ ),

e com o respectivo valor de ( $Z$ ), encontrar a área e a probabilidade desejada entre qualquer intervalo da Distribuição Normal.

Figura 5 – Tabela Z com a Área ou Probabilidade para a Curva Normal padronizada

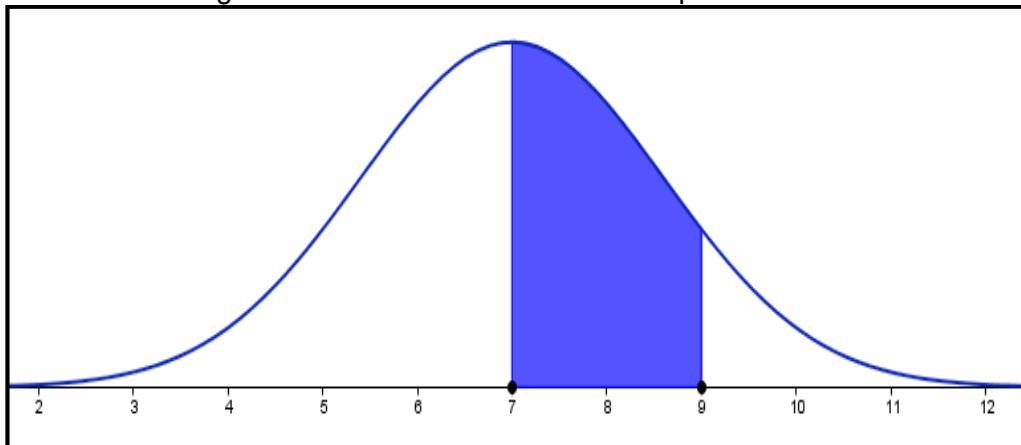
$z$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2518	0,2549
0,7	0,2580	0,2612	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4986	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

Fonte: Anderson *et al* (2002).

Diante do exposto, para uma melhor compreensão de todo esse processo de cálculo de probabilidades, através da Distribuição Normal padronizada, exemplificaremos a aplicação do referido método na seguinte situação problema: Em um Banco de crédito financeiro, o tempo de duração necessário para o atendimento aos clientes é normalmente distribuído, com média igual a 7 minutos e desvio padrão de 1,6 minutos. Qual a probabilidade de que um atendimento dure de 7 a 9 minutos?

Interpretando os dados do problema, podemos compreender que a variável em questão, normalmente distribuída, possui média  $\mu = 7$  e desvio padrão  $\sigma = 1,6$  e pergunta-se a probabilidade de que um atendimento dure de 7 a 9 minutos, ou seja, a área compreendida entre a média e o ponto  $x = 9$ , representada na Figura abaixo.

Figura 6 – Área sob a curva entre os pontos 7 e 9



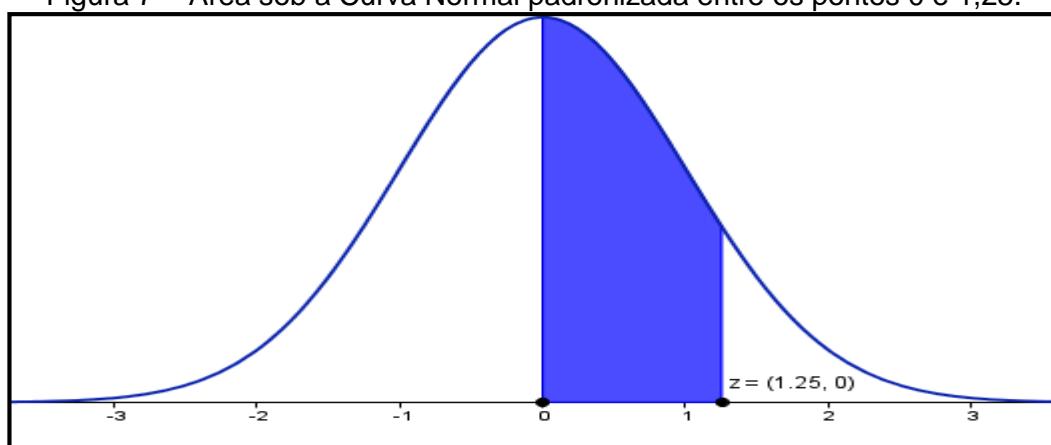
Fonte: Elaborado pelo autor a partir do Geogebra (2020).

Inicialmente, aplicando o método da transformação da variável, obtemos o valor Z para o ponto  $x = 9$ , e como sabemos, o valor da média na distribuição normal padronizada é igual a 0. Deste modo, o valor de Z é igual a:

$$Z = \frac{9 - 7}{1,6} = 1,25$$

Desta forma, após a padronização da variável, a área que corresponde a probabilidade desejada está compreendida entre a média  $\mu = 0$  e o ponto  $z = 1,25$ , como representado na Figura 7 .

Figura 7 – Área sob a Curva Normal padronizada entre os pontos 0 e 1,25.



Fonte: Elaborado pelo autor a partir do Geogebra (2020).

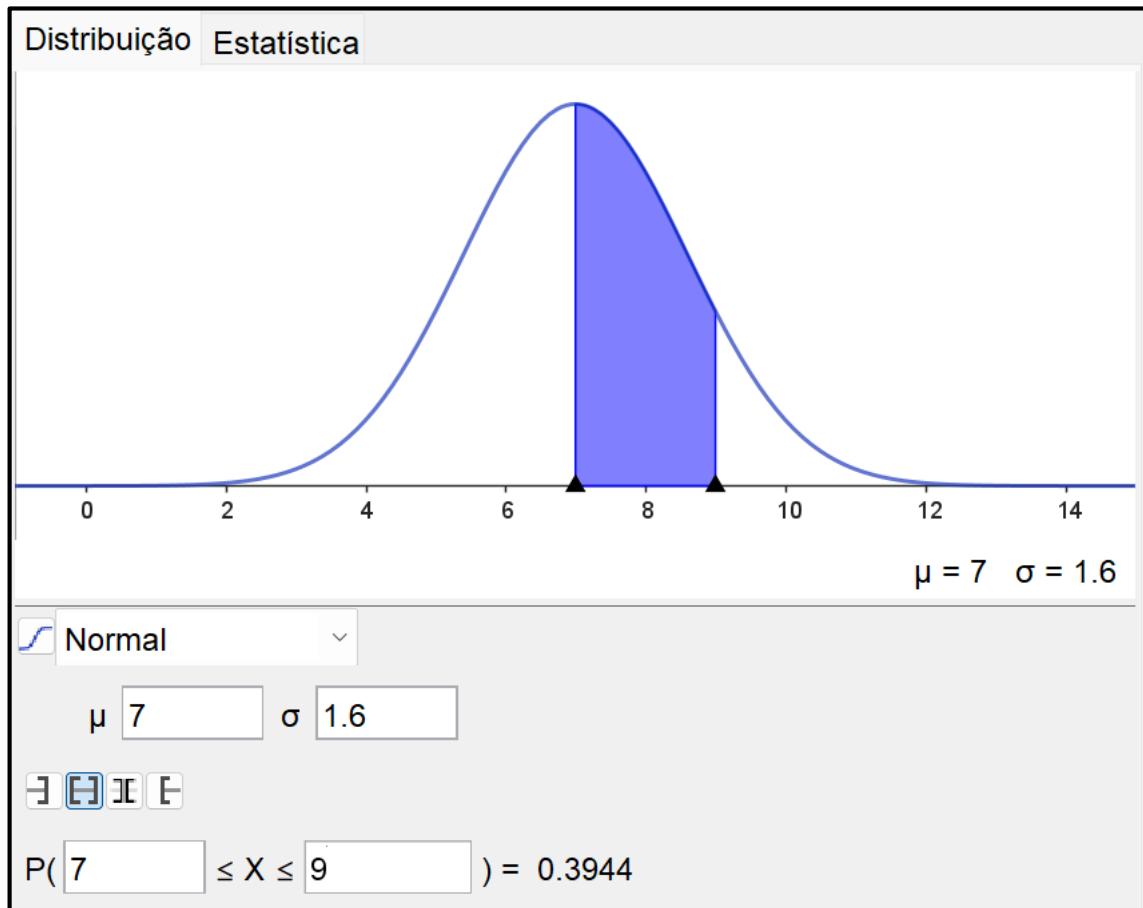
Como escala padronizada ( $Z$ ) permite associar os valores de qualquer variável aleatória ( $X$ ), com os valores de sua escala, a partir dos parâmetros média ( $\mu$ ) e desvio padrão ( $\sigma$ ), comprehende-se que as duas áreas apresentadas anteriormente da referida situação problema são equivalentes, ou seja, possuem a mesma medida. Em continuidade, para encontrarmos o valor da área determinada pela Distribuição Normal padronizada ( $Z$ ), devemos consultar na tabela dessa Distribuição, apresentada anteriormente e que contempla os valores da área ou probabilidade entre a média ( $u = 0$ ) e qualquer ponto ( $Z$ ).

Nesta direção, ao analisar a tabela e buscarmos a área desejada, devemos observar a primeira coluna, que contempla a parte inteira e a primeira decimal do ponto  $Z$  e a primeira linha que contempla segunda decimal desse ponto. Assim, a área ou a probabilidade desejada estará na interseção entre a coluna e a linha relativas ao ponto  $Z$  analisado.

Retornado a nossa situação problema, desejamos encontrar a área correspondente ao ponto  $z = 1,25$ . Como este ponto,  $(1,25 = 1,2 + 0,05)$ , possui a parte inteira e primeira decimal igual a 1,2 e a segunda decimal igual a 0,05, ao analisarmos a tabela, podemos compreender que a área correspondente a esse ponto está na interseção entre a décima terceira linha (1,2) e a sexta coluna (0,05). Assim, a área procurada é igual a 0,3944, ou seja, a probabilidade de que o atendimento dure de 7 a 9 minutos é de 39,44%.

A segunda possibilidade para o cálculo de probabilidades associada a intervalos sob a Curva Normal se dá através da utilização de um Software. Nesse estudo recorremos ao software Geogebra, ferramenta gratuita, de livre acesso e frequentemente utilizada no processo de ensino e aprendizagem de Matemática, dentre outras disciplinas. Esse programa computacional possui, dentre inúmeras funcionalidades, a “Calculadora de Probabilidades” que contempla, em seu layout, opções para o preenchimento dos dois parâmetros da Curva Normal, a média e o desvio padrão, estabelecendo a Curva e o intervalo de interesse para o cálculo da área. Assim, de forma mais prática e dinâmica é possível determinar a área que corresponde a probabilidade associada a esse intervalo (Figura 8).

Figura 8 – Cálculo da área de um intervalo sob a Curva Normal com a utilização do Geogebra



Fonte: Elaborado pelo autor a partir do Geogebra (2020).

Como podemos observar, a utilização no Geogebra dos campos média ( $\mu$ ) o desvio padrão ( $\sigma$ ) e o intervalo de interesse ( $7 \leq x \leq 9$ ), o programa determina, assim como no método da Curva Normal Padronizada, a área igual a 0,3944, ou seja, a probabilidade de que o atendimento dure de 7 a 9 minutos é de 39,44%.

Diante desse contexto, acreditamos que o conceito e representação da Curva Normal articulado com a de Amostragem, contemplando os conceitos estatísticos das medidas de tendência central e dispersão, e o cálculo de probabilidades, pode ser amplamente abordado no Ensino Médio, através de diferentes recursos didáticos, tanto com lápis e papel, através da tabela da Curva Normal padronizada, como também pela utilização do software Geogebra. No entanto, para compreendermos essa referida articulação, passaremos a apresentar na próxima seção o Teorema Central do Limite.

### 2.3. RELACIONANDO A AMOSTRAGEM COM A CURVA NORMAL: UMA ABORDAGEM INFORMAL DO TEOREMA CENTRAL DO LIMITE

Discutimos, anteriormente, que a Estatística Inferencial se fundamenta nos processos de Amostragem para realizar inferências sobre a população. Esses processos reúnem desde a seleção de amostras representativas, até modelos de distribuição de probabilidade, a exemplo da Distribuição Normal, que possibilitam um olhar analítico aprimorado sobre o comportamento dos dados.

Além disso, com base na variabilidade (Reyes, 2019) vimos que por mais que a seleção de uma amostra representativa seja de modo adequado, haverá sempre a tendência das estatísticas amostrais apresentarem variação com relação aos parâmetros ou com os estimadores de outras amostras de mesmo tamanho extraídas da mesma população.

Para fundamentar essa variabilidade, a inferência estatística se pauta em alguns fundamentos teóricos. Dentre eles, focaremos nesse estudo na Distribuição Amostral e no Teorema Central do Limite.

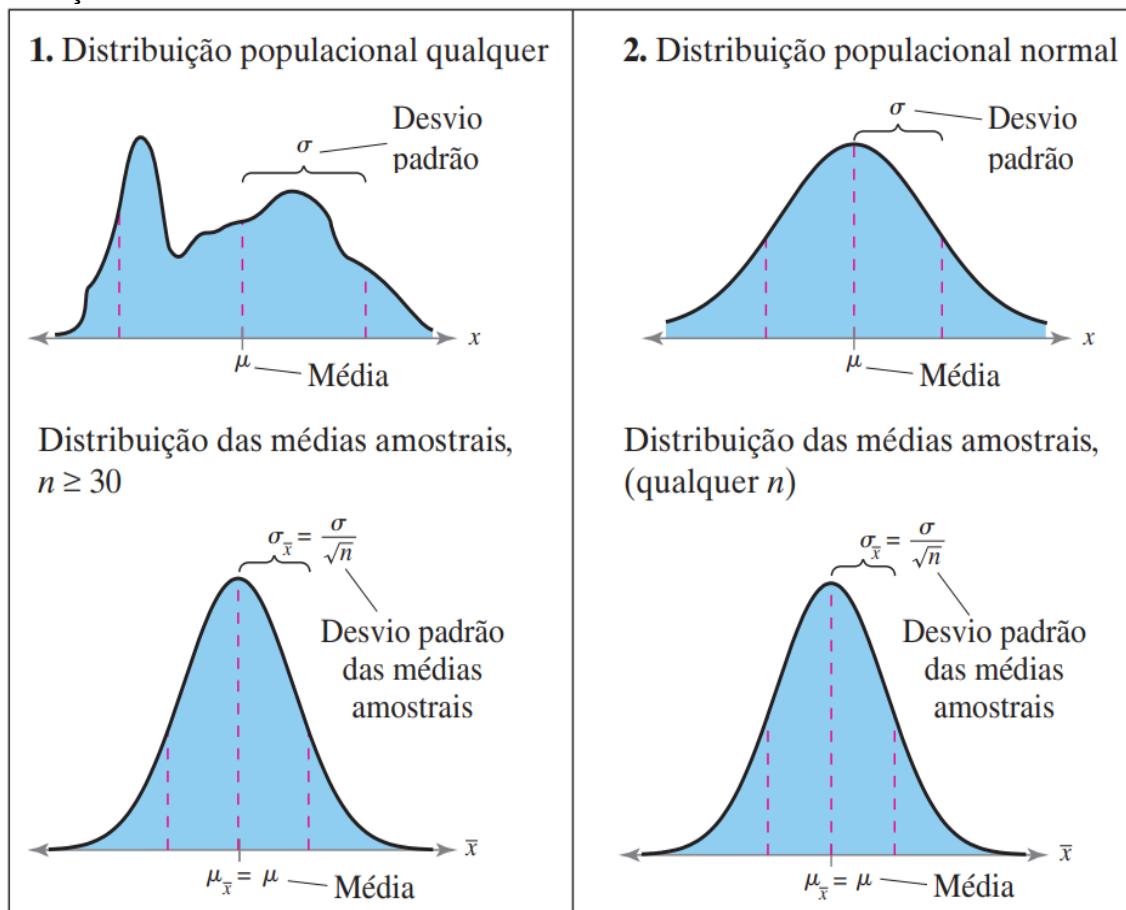
Nesse sentido, a Distribuição Amostral diz respeito à distribuição de probabilidade das medidas estatísticas de todas as amostras possíveis, de mesmo tamanho, extraídas de uma mesma população. Assim, se de uma população de tamanho N, retirarmos todas as amostras possíveis de tamanho n, obteríamos para cada uma delas estatísticas amostrais. Em se tratando da média, por exemplo, as médias extraídas de cada amostra determinam uma distribuição amostral das médias. Em síntese e considerando as estatísticas amostrais (média, desvio-padrão, etc.), uma Distribuição Amostral de uma estatística é a distribuição de todos os valores da estatística de interesse quando todas as amostras possíveis de mesmo tamanho são extraídas da população. (Triola, 2008; Larson e Farber, 2015).

O entendimento sobre as características da Distribuição Amostral de uma estatística é de fundamental importância para se estimar parâmetros para a população. No que diz respeito a Distribuição Amostral das Médias, o Teorema Central do Limite apresenta importantes conclusões: 1) Se a população obedece a uma distribuição Normal, então a distribuição amostral das médias extraídas dessa população também obedecerá a uma distribuição normal, para qualquer tamanho da amostra. 2) Se a distribuição da população é desconhecida ou não segue uma distribuição normal, à medida que o tamanho da amostra aumenta, a distribuição das

médias amostrais tende a se aproximar de uma distribuição normal ou tornar-se aproximadamente normal. Logo, mesmo que a população tenha uma distribuição desconhecida ou não seja normal, haverá a tendência da aproximação da distribuição das médias amostrais para uma distribuição normal, desde que o tamanho  $n$  da amostra seja grande o suficiente, com  $n \geq 30$  (Larson e Farber, 2015).

A partir desse teorema se conclui que se a representação gráfica dos dados de uma população é uma distribuição normal, ou seja, uma Curva Normal, a distribuição amostral das médias também será. Além disso, se o comportamento da população for desconhecido ou aproximadamente normal, a distribuição das médias amostrais (para amostras com tamanho maior ou igual a 30) poderá ser uma Curva Normal ou se aproximar da normalidade à proporção que o tamanho da amostra aumenta. Para ambos os casos, a mais importante conclusão é que a distribuição amostral das médias tem média igual à da população. (Triola, 2008, Larson e Farber, 2015). Essas conclusões podem ser representadas graficamente através da Figura 9.

Figura 9 – Distribuições populacionais não normal (1), normal (2) e suas respectivas distribuições amostrais das médias



Fonte: Larson e Farber (2015, p. 247).

Como podemos observar, a distribuição amostral das médias tem a mesma média que a população. Isso implica que distribuição amostral tem o mesmo centro da Curva Normal populacional. No entanto, o seu desvio-padrão é menor do que o desvio-padrão populacional, ou seja, é mais concentrada em torno da média. Isso é justificado, estatisticamente, pelo fato de que as médias amostrais são menos dispersas que os elementos populacionais, os quais podem ser muito grandes e/ou muitos pequenos. Nesse sentido, mesmo que uma amostra tenha um valor extremo, o mesmo será diluído entre os demais valores amostrais, o que diminui sua interferência no valor da média. Quanto maior o tamanho da amostra menor será a variabilidade, neste caso, o desvio-padrão, porque mais diluídos estão os valores extremos da população. Desta forma, a dispersão da distribuição amostral está diretamente ligada a dispersão da população e ao tamanho amostral (Triola, 2008; Larson e Farber, 2015; Silva e Pinto, 2021).

Para mostrar, de forma prática, a conclusão do referido teorema, de que a média da distribuição amostral das médias é igual à média da população, apresentamos um exemplo abordado por Larson e Farber (2015). Considere uma população formada pelos elementos 1,3,5 e 7. A média dessa população é igual a 4, ou seja, corresponde a soma desses elementos dividida por 4, isto é, a quantidade deles. O Quadro 1, lista todas as amostras possíveis de tamanho  $n = 2$ , retiradas dessa população, com reposição dos elementos, e suas respectivas médias.

Quadro 1 – Amostras possíveis de uma população e respectivas médias

Amostra	Elementos da Amostra	Média da Amostra ( $\bar{x}$ )
1	1,1	1
2	1,3	2
3	1,5	3
4	1,7	4
5	3,1	2
6	3,3	3
7	3,5	4
8	3,7	5
9	5,1	3
10	5,3	4
11	5,5	5
12	5,7	6
13	7,1	4
14	7,3	5
15	7,5	6
16	7,7	7

Fonte: Larson e Farber (2015).

Considerando as médias determinadas por cada amostra, a média desses valores é igual ao somatório de cada média, dividido por 16 que é a quantidade de amostras geradas, ou seja, a média da distribuição amostral das médias  $\mu_{\bar{x}} = \frac{64}{4} = 4$ . Diante disso, podemos constatar que a média dessa distribuição é igual a média população, indo ao encontro do que determina o referido teorema.

As definições fundamentadas no Teorema Central do Limite são bases para os procedimentos realizados em toda inferência estatística. A estimação dos parâmetros populacionais através de amostras tem relação direta com as conclusões sobre a distribuição amostral da estatística que é utilizada para a estimação do parâmetro de interesse. Assim, nem sempre o pesquisador necessita conhecer a distribuição populacional para se fazer inferências sobre ela a partir dos dados amostrais. Na prática, em uma pesquisa amostral, apenas uma amostra representativa é selecionada, obtendo-se estimativas pontuais, a exemplo da média. Porém, a média obtida na única amostra, dificilmente, será igual à média populacional, mas sim, próxima desse valor, apresentando variações a depender da natureza da amostra. Logo, a média amostral, nesse contexto, é apresentada como uma variável aleatória, possuindo, consequentemente, uma distribuição de probabilidade específica. Nessa direção, independentemente da distribuição de probabilidade associada ao fenômeno de interesse, a sua média possui distribuição amostral aproximadamente Normal, com média igual à média populacional ( $\mu$ ). Com base nesses princípios, todas as propriedades da Curva Normal, discutidas neste capítulo, se aplicam também para a média amostral, enquanto variável aleatória.

Considerando que a média amostral dificilmente será igual a média populacional, adota-se, nos processos de inferenciais, uma estimativa intervalar para essa média, ou seja, um intervalo de valores, ou amplitude de valores, usado para estimar um parâmetro populacional. A partir disso, torna-se necessário definir o grau de confiança de que uma determinada estimativa intervalar contenha a média populacional. Assim, o nível de confiança corresponde à probabilidade de que a estimativa intervalar contenha o parâmetro populacional, supondo que o processo de estimação é repetido um grande número de vezes (Larson e Farber, 2015). Com base na estimativa intervalar, define-se o intervalo de confiança, o qual é uma maneira específica de apresentar essa estimativa fornecendo uma faixa de valores, calculada a partir de uma amostra de dados, dentro da qual espera-se que esteja o valor

verdadeiro de um parâmetro populacional uma certa probabilidade específica, ou seja, nível de confiança. Assim, a informação de que uma pesquisa amostral tem 95% de nível de confiança, por exemplo, significa que se a seleção da amostra for repetida muitas vezes, em 95% delas o intervalo resultante irá conter o verdadeiro valor do parâmetro populacional que está sendo estimado (Larson e Farber, 2015).

No intervalo de confiança, a estimativa pontual, como por exemplo, a média da amostra, corresponde ao centro do intervalo e a ela é adicionado e subtraído o erro máximo de estimativa, também chamada de margem de erro. Dessa forma, fica gerado um intervalo que contém a média da amostra e também contenha a média populacional, considerando a variabilidade estabelecida através da margem de erro (Larson e Farber, 2015). Nesse viés, o nível de confiança, intervalo de confiança e a margem de erro estão intrinsecamente ligados ao tamanho da amostra e a relação entre precisão e confiança nos processos de pesquisa amostral. Se os pesquisadores priorizarem a precisão dos resultados e desejarem estimativas mais precisas dos parâmetros populacionais, certamente optarão por usar um tamanho de amostra maior. Consequentemente, aumentará a precisão dos resultados, mas também pode aumentar os custos e tempo da pesquisa. Em outra perspectiva, se os pesquisadores estiverem mais interessados em um maior grau de confiança na precisão de suas estimativas, eles poderão optar por um nível de confiança mais alto. Logo, resultará em intervalos de confiança mais amplos e uma menor chance de erro amostral, mas também pode implicar em uma precisão relativamente menor, tendo em vista que os intervalos de confiança mais amplos têm uma margem de erro maior.

Diante do exposto nesse capítulo, acreditamos que é possível na etapa do Ensino Médio da escolarização básica, a abordagem dos conceitos da Amostragem articulada com o modelo da Curva Normal, contemplando os conceitos estatísticos das medidas de tendência e dispersão e o cálculo de probabilidades, através de diferentes recursos didáticos, a exemplo do lápis e papel com utilização da tabela de Curva Normal padronizada e o software Geogebra. Acreditamos que essa abordagem, a qual notadamente não tem um rigor estatístico e matemático, possibilitará ao estudante, o entendimento sobre a importância da Amostragem e do modelo da Curva Normal para a inferência Estatística. Por exemplo, ao compreenderem os princípios subjacentes ao intervalo de confiança e ao teorema central do limite, o estudante pode reconhecer como esses conceitos são aplicáveis na tomada de decisões fundamentadas em evidências, em diversas áreas do cotidiano. Além disso, ao

explorarem o papel da média como uma variável aleatória, podem identificar inúmeros fenômenos e situações cotidianas que contemplam a distribuição de médias amostrais, e que esta se aproxima de uma distribuição normal, independentemente da distribuição da população subjacente. Por conseguinte, permite a realização de inferências sobre a população com base em uma amostra, mesmo que os pesquisadores não tenham conhecimento sobre a distribuição populacional.

### 3 FUNDAMENTAÇÃO TEÓRICA

Este capítulo é destinado para a apresentação da concepção de Letramento Estatístico e Probabilístico proposto por Gal (2002, 2005) e do Ciclo Investigativo como eixo estruturador para o ensino de Estatística elaborado por Guimarães e Gitirana (2013), os quais fundamentam teoricamente o presente estudo.

#### 3.1 O LETRAMENTO ESTATÍSTICO E PROBABILÍSTICO

Iniciaremos essa seção apresentando uma concepção geral para o termo Letramento. Na ótica de Cope e Kalantzis (2000), o "letramento" refere-se ao processo pelo qual um indivíduo adquire e desenvolve habilidades relacionadas à leitura e à escrita. Esse desenvolvimento vai além da simples técnica de decifrar letras e palavras, envolvendo também a interpretação e produção de textos. Além disso, o letramento abrange a compreensão dos diferentes usos sociais da leitura e da escrita em contextos variados. Nessa direção, no âmbito da Educação Básica, o letramento está relacionado com a capacidade de ler e interpretar diversos tipos de informações que levem a compreensão dos significados dos conceitos em diferentes situações, possibilitando uma aprendizagem que responda às demandas sociais e propicie o senso crítico.

De modo mais específico, no campo da Educação Estatística, o educador Iddo Gal (2002) enfatiza que os conhecimentos básicos de Estatística são essenciais para que os cidadãos possam se portar de maneira crítica e emancipada em sociedades que constantemente estão imersas em cadeias de informações. Nessa perspectiva, o autor concebe o Letramento Estatístico como:

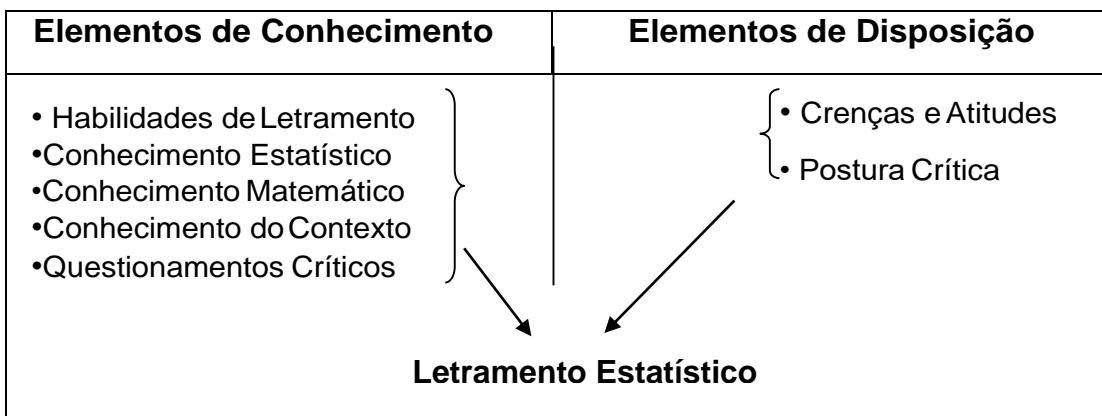
- a) competência da pessoa para interpretar e avaliar criticamente a informação estatística, os argumentos relacionados aos dados ou aos fenômenos estocásticos, que podem se apresentar em qualquer contexto e quando relevante; b) competência da pessoa para discutir ou comunicar suas reações para tais informações estatísticas, tais como seus entendimentos do significado da informação, suas opiniões sobre as implicações desta informação ou suas considerações acerca da aceitação das conclusões fornecidas (GAL, 2002, p. 2-3).

Assim, para Gal (2002,2019) o indivíduo é considerado letrado estatisticamente quando detém as competências para interpretar e avaliar criticamente uma informação estatística em qualquer contexto e discutir ou comunicar suas reações, crenças e opiniões frente às informações e dados do campo estatístico. O autor enfatiza que o Letramento Estatístico pode auxiliar os cidadãos a responder demandas sociais em várias perspectivas. Por exemplo, possibilita a compreensão das tendências de fenômenos sociais, políticos e econômicos, como também, contribui para a tomada de decisão em contextos que contemplam situações aleatórias. Além disso, pode apoiar o debate público e o planejamento de políticas públicas. Por fim, cita que em muitos ambientes de trabalho a estatística colabora com o aperfeiçoamento da qualidade dos processos de produção.

Gal (2002) também destaca que há cidadãos que se colocam como produtores de dados, isto é, estão envolvidos em processos de investigação estatísticas e, consequentemente, necessitam desenvolver habilidades para compreender as pesquisas estatísticas. No entanto, a grande maioria cidadãos se portam como consumidores de dados, ou seja, estão inseridos em contextos de leituras de dados e informações estatísticas. Como exemplo desses contextos, o autor cita programas de televisão, jornais, sites de internet, eventos cívicos ou políticos e relatórios de dados em ambientes de trabalho. Nesses contextos, Gal (2002) aponta que a informação estatística pode ser apresentada de três maneiras: através de textos (escrito ou oral), números/símbolos e apresentações gráficas e tabulares. Logo, esse público necessita compreender, interpretar, avaliar criticamente as informações estatísticas encontradas nesses contextos.

Diante dessas considerações, Gal (2002) elaborou um modelo de Letramento Estatístico (Quadro 2) que contempla dois componentes que se relacionam entre si, a saber: *Elementos de Conhecimento* relacionado a aspectos cognitivos e *Elementos de Disposição* relativos a facetas atitudinais e emocionais.

Quadro 2 – Modelo do Letramento Estatístico



Fonte: Gal (2002, p. 4).

Explorando cada componente desse modelo, destacamos que os *Elementos de Conhecimento* estão relacionados com a capacidade de o indivíduo ler, interpretar e analisar de forma crítica uma informação ou dado estatístico. Esse campo contempla cinco elementos que não são independentes e se inter-relacionam: habilidades de letramento, conhecimento estatístico, conhecimento matemático, conhecimento do contexto e questionamentos críticos.

As *Habilidades de Letramento*, por sua vez, dizem respeito às habilidades gerais que os indivíduos devem possuir para se comunicar de forma oral ou escrita, compreender as informações estatísticas de diferentes níveis de complexidade, veiculadas por diversos instrumentos, como textos virtuais, orais e escritos, gráficos e tabelas que abarcam uma gama de conceitos estatísticos relativos a um dado contexto.

Gal (2002) aponta que é comum as informações estatísticas serem veiculadas através de textos de forma escrita ou oral, como também tabelas e gráficos, o que exige habilidades para a adequada leitura e interpretação dos dados. Porém, essas informações são transmitidas por diferentes agentes, a exemplo de jornalistas, políticos, funcionários, anunciantes, dentre outros. Consequentemente, a maneira de abordar a informação por cada um deles é distinta porque os objetivos são diferentes, a fim de convencer o leitor ou ouvinte a aceitar um argumento. O estudo de Luna (2023) apresenta (Figura 10) uma forma inadequada de uma apresentadora de televisão ao fazer a leitura de um gráfico.

Figura 10 – Análise do gráfico pela apresentadora



Fonte: G1 – Bahia (26/03/2020)

Nessa situação, para explicar a diferença entre a projeção e o número de casos confirmados da COVID-19, a apresentadora recorreu à técnica visual ao utilizar o palmo da mão para argumentar sobre o distanciamento entre as duas linhas do gráfico. Esse é um exemplo de uma forma de comunicar dados estatísticos que pode induzir os cidadãos a aceitarem conclusões equivocadas.

Maciel (2023) apontam que em algumas ocasiões, os conceitos estatísticos acabam sendo abordados na mídia de maneira rasa e sutil. A exemplo de manchetes que divulgam a “projeção do PIB” ou a “margem de erro da pesquisa”, mas ocultam importantes conceitos ou procedimentos estatísticos necessários para uma melhor compreensão das informações por parte do leitor.

Assim, Gal (2002,2019) chama a atenção que os veículos de comunicação podem usar termos técnicos estatísticos ou jargões de forma proposital, diferente do significado clássico do conceito, com o intuito de convencer o telespectador/leitor. Logo, a habilidade de letramento combinada com os outros elementos do Letramento Estatístico contribui para a compreensão adequada das informações em textos simples ou complexos, como também, em gráficos com poucos dados ou aqueles que demandam mais habilidades para interpretá-los.

O segundo elemento, o *Conhecimento Estatístico*, envolve o domínio básico sobre o significado dos conceitos e procedimentos estatísticos. O autor destaca inicialmente que esse conhecimento básico não deve ser encarado como algo estagnado, tendo em vista que a depender do contexto os níveis de complexidade dos conceituais estatísticos exigidos para a interpretação dos dados podem ser diferentes.

O conhecimento estatístico proposto por Gal (2002) em seu modelo está fundamentado em cinco capacidades: 1- Saber por que os dados são necessários e como podem ser produzidos; 2 - Familiaridade com conceitos e ideias básicas relacionadas à estatística descritiva; 3 - Familiaridade com conceitos básicos e ideias relacionadas a representações gráficas e tabulares; 4 - Compreensão de noções básicas de Probabilidade e sua relação com a Estatística para determinar amostras aleatórias, teste de significância, raciocínio inferencial, como intervalos de confiança ou teste de hipóteses, e por fim, 5 - Saber como as conclusões ou inferências estatísticas são obtidas.

A primeira capacidade “saber por que os dados são necessários e como podem ser produzidos” está relacionada com a compreensão sobre a origem dos dados relativos a uma determinada conclusão estatística. Além disso, demanda o conhecimento do processo de coleta até as conclusões. Ou seja, leva ao reconhecimento sobre o porquê da realização da coleta dos dados e o que fundamenta as conclusões que são veiculadas, tendo em vista que os dados brutos são resumidos para os agentes apresentarem apenas suas principais características e tendências. Um exemplo dessa capacidade é a realização de pesquisas amostrais. Os cidadãos devem conhecer qual o objetivo da pesquisa, método de Amostragem utilizado, o tamanho da amostra, o nível de confiança, a margem de erro, dentre outros aspectos. São termos essenciais que dão credibilidade à informação estatística.

A segunda capacidade “2 - Familiaridade com conceitos e ideias básicas relacionadas à estatística descritiva” diz respeito à familiaridade com conceitos e formas de apresentação de dados que são comumente utilizados para abordar os resultados aos cidadãos. Gal (2002) argumenta que a porcentagem, a média aritmética e a mediana são os principais conceitos utilizados para transmitir as informações estatísticas. Diante disso, os cidadãos devem reconhecer os significados desses conceitos e suas implicações. Por exemplo, a Figura 11 aborda uma notícia a qual utiliza a porcentagem para indicar a principal conclusão do estudo.

Figura 11 – Notícia sobre os casos de Dengue no Rio de Janeiro

## Casos de dengue aumentam quase 1.000% no Rio de Janeiro em relação a 2023

Estado já tem 17 mil casos da doença em janeiro, segundo dados da Central de Inteligência em Saúde (CIS)

Redação  
Brasil de Fato | Rio de Janeiro (RJ) | 30 de janeiro de 2024 às 20:23

Fonte: Portal Brasil de Fato (30/01/2024)

A compreensão dessa informação implica no reconhecimento que a porcentagem abordada indica a intensidade do aumento dos casos de dengue no estado do Rio de Janeiro. Logo, o leitor familiarizado com o conceito de porcentagem irá assimilar que o número de casos aumentou em 10 vezes o valor inicial e que esse tipo de aumento é extremamente significativo e geralmente indica um crescimento muito rápido ou uma mudança drástica na situação. O teórico ainda enfatiza que essa familiaridade leva o cidadão a reconhecer que a média indica a concentração dos dados em torno de um valor, porém ela é afetada por valores extremos e que a porcentagem pode indicar um aumento com valores acima de 100%, como abordado anteriormente. Diante disso, os cidadãos devem estar conscientes que os diferentes valores utilizados para esses conceitos produzem diferentes significados a depender dos contextos que estão sendo utilizados.

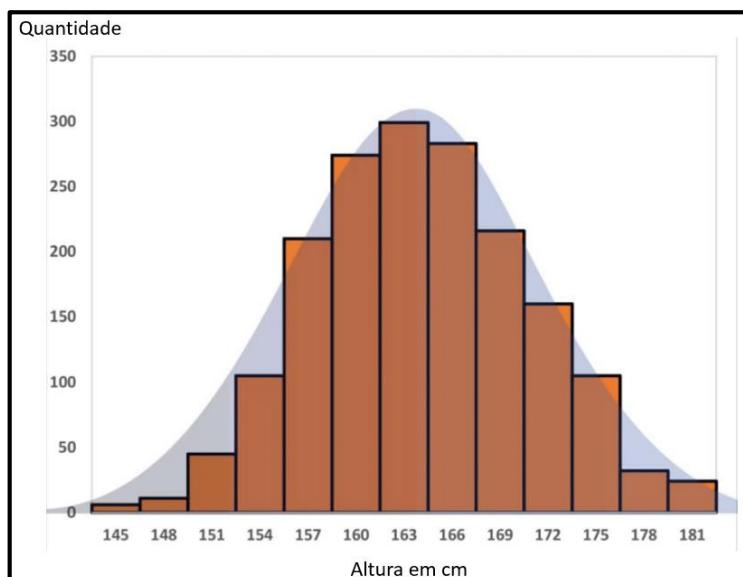
Em continuidade, a terceira capacidade “Familiaridade com conceitos básicos e ideias relacionadas a representações gráficas e tabulares” é relativa ao conhecimento sobre os recursos para a apresentação dos dados de uma pesquisa. Mais especificamente, o autor cita que a compreensão dos gráficos e tabelas é fundamental para a análise e interpretação do comportamento dos dados em variados contextos. A partir disso, essa capacidade envolve a leitura dos dados, entre os dados e para além dos dados, possibilitando uma análise de um ponto específico da representação gráfica, mas também uma visão global das tendências dos dados. Os cidadãos também devem ficar atentos a gráficos que podem ser criados, propositadamente, para explicitar ou ocultar uma tendência, ou ainda, direcionar o leitor para a assimilação de uma conclusão específica. Isso demanda compreensão de tabelas em diferentes níveis de complexidade e conhecer os diferentes tipos de

representação gráfica e seus principais elementos, a exemplo da escala, eixos, legenda e marcadores de dados.

Sobre esse tema, Guimarães, Cavalcanti e Evangelista (2020) argumentam que a maior dificuldade em construir e interpretar gráficos, tanto para crianças quanto para adultos, é a escala. As autoras apontam que essa dificuldade diz respeito a compreensão dos valores não explícitos na escala, e, por conseguinte, gera incompreensões com suas subunidades. Além disso, destacam que a experiência de vida dos adultos não é suficiente para superar esse obstáculo, o que reforça a importância da escola oferecer um ensino sistematizado sobre esse tópico conceitual.

Como exemplo dessa capacidade, apresentamos o gráfico da Curva Normal, principal modelo de análise de dados da Inferência Estatística (Batanero, Tauber e Sánchez, 2004). A Figura 12 aborda a distribuição das medidas da altura em centímetros de um conjunto de pessoas, na qual a representação gráfica dos dados se aproxima de uma Curva Normal.

Figura 12 – Distribuição das alturas de um conjunto de pessoas



Fonte: CONRE-3 (2017).

Compreender essa representação gráfica exige, além do domínio sobre a escala e eixos, a identificação da média como centro da distribuição, a qual indica o eixo de simetria e a concentração dos dados, como também, o desvio-padrão, enquanto medida de variabilidade, que denota o grau de dispersão. Assim, a Curva pode apresentar movimentos em função dessas duas medidas. Quanto maior o desvio padrão, maior será a dispersão do conjunto e a amplitude dos dados, e

consequentemente a Curva Normal será mais espalhada horizontalmente e mais comprimida verticalmente. Isso acontece porque a distribuição dos valores se estende por uma faixa maior em relação à média, resultando em uma curva mais ampla. Por outro lado, quanto menor o desvio padrão, menor será a dispersão do conjunto e a amplitude dos dados e, consequentemente, a Curva Normal será mais expandida verticalmente e mais comprimida horizontalmente. Isso ocorre porque os valores se concentram mais próximos da média, resultando em uma curva mais alta e afunilada. Portanto, o desvio padrão influencia tanto a dispersão horizontal dos dados (amplitude) quanto a dispersão vertical (altura) da Curva Normal.

Na sequência, através da quarta capacidade, “compreender noções básicas de Probabilidade”, Gal (2002) argumenta que todo cidadão pode se deparar com situações ou ler notícias relacionadas ao acaso e a acontecimentos aleatórios. Logo, necessitam interpretá-las adequadamente para uma melhor tomada de decisão. O autor cita como exemplo, situações de previsão do tempo, fatores genéticos e riscos de diferentes naturezas. Em tais situações, a compreensão se restringe às avaliações de chance relacionadas a esses eventos. Essas avaliações podem ser obtidas através de julgamentos pessoais ou de modelos matemáticos e probabilísticos elaborados. Qualquer que seja a abordagem empregada para estimar essas probabilidades, é crucial que a pessoa seja capaz de interpretá-las e usá-las a seu favor. O exemplo dessa capacidade é retrata da Figura 13, a qual aborda uma situação bastante atual que é o compartilhamento de “Fake News”, amparado pelo o aumento do uso das mídias sociais e da disseminação rápida de informações online.

Figura 13 – Notícia sobre a chance de Fakes News viralizarem

Notícia 1 • Estadão / Ciência

## 'Fake news' têm 70% mais chance de viralizar que as notícias verdadeiras, segundo novo estudo

Fonte: Portal Estadão (2018).

Ao se deparar com essa notícia, o cidadão letrado estatisticamente deve compreender que o termo “chance” mencionado na manchete remete a uma

probabilidade significativamente maior - 70% mais chance – das notícias fakes se tornarem virais em comparação com notícias verdadeiras. Isso significa que as informações falsas estão sendo compartilhadas e difundidas mais rapidamente e em maior escala do que informações verdadeiras. Logo, o cidadão, após interpretar o conceito probabilístico e consequentemente a informação, deve estar ciente dos possíveis impactos negativos dessa tendência e adotar uma postura crítica em relação às informações que consomem e compartilham, verificando a credibilidade das fontes e buscando por evidências antes de aceitar uma notícia como verdadeira.

A última competência, “Saber como as conclusões ou inferências estatísticas são obtidas” remete ao entendimento do processo de Amostragem para a inferência estatística. O autor destaca que os cidadãos necessitam compreender como os dados são coletados, analisados e como as conclusões são formuladas, tendo em vista que a maioria deles são consumidores de dados e estão envolvidos em contextos de leitura de informações estatísticas.

Gal (2002) aponta que essa competência está relacionada com quatro aspectos importantes. O primeiro diz respeito na identificação da representatividade e generalização da amostra, reconhecendo que a pesquisa amostral não é exata, mas possui uma probabilidade de representar a população. Outro aspecto é estar atento a possíveis vieses ou erros no processo amostral que inviabiliza a representação. O terceiro aspecto tem relação com o entendimento sobre a variabilidade, mais precisamente a margem de erro que permite controlar erros de medição nas pesquisas inferenciais. O último aspecto se refere ao conhecimento de que pesquisas amostrais podem apresentar diferenças, porém, em muitos casos, são tão ínfimas que se tornam irrelevantes, podendo ser pontuais e sem nenhuma intencionalidade. Para exemplificar essa capacidade, apresentamos na Figura 14 o resultado de uma pesquisa amostral com a população brasileira sobre o mercado de Pets.

Figura 14 – Resultado de uma pesquisa amostral



Fonte: OpinionBox (2023).

O principal resultado dessa pesquisa aponta que 70% dos brasileiros que possuem animais de estimação preferem comprar produtos de empresas que apoiam a causa animal. No texto vinculado a essa imagem, é informado que foram entrevistadas 2.000 pessoas que têm pets ao redor do país, respeitando as proporções de gênero, região, faixa etária e classe social. A coleta foi feita em agosto de 2023 e a margem de erro é de 2,2 pontos percentuais. Essa situação exemplifica o que o cidadão a partir do letramento estatístico não deve se ater apenas ao resultado de uma pesquisa amostral, mas analisar e julgar o processo amostral, incluindo o tamanho da amostra e o método de Amostragem para reconhecer sua representatividade.

Outro elemento de conhecimento presente no modelo de letramento estatístico proposto por Gal (2002,2019) é o *conhecimento matemático*. A construção do conhecimento estatístico e, por conseguinte, o Letramento Estatístico, também requer o domínio de ferramentas e conceitos matemáticos. Logo, o conhecimento matemático é responsável por dar suporte necessário ao conhecimento estatístico e está direcionado para compreensão da ideia de número, seja ele natural, fracionário, decimal ou percentual com diferentes significados e quantidades, como, por exemplo, a ideia de média, quantificação de variância, e probabilidade, possibilitando ao indivíduo, o desenvolvimento de habilidades e competências matemáticas que corroborem com a formação de seu Letramento Estatístico.

O autor argumenta que os meios de comunicação, para divulgar informações estatísticas, costumam utilizar medidas de tendência central, a exemplo da média e também a porcentagem, como exemplificado nas Figuras 13 e 14 Desse modo, os cidadãos devem compreender, mesmo que informalmente, os procedimentos matemáticos utilizados para gerar um dado estatístico, como também as noções matemáticas que estão relacionadas com a estatística na apresentação dos dados em gráficos ou tabelas.

Em continuidade, o *Conhecimento de Contexto* também é necessário para a construção do Letramento Estatístico, tendo em vista que toda informação estatística está relacionada ao um determinado contexto. Assim, quando o indivíduo o conhece, é possível dar significado e compreender tal informação. Deste modo, esse conhecimento é o principal determinante para a familiaridade do leitor com as fontes que veiculam informações e dados estatísticos.

Se um indivíduo não estiver familiarizado com o contexto no qual as informações e dados estatísticos foram coletados, possivelmente ele encontrará dificuldades para realizar análises e levantar conclusões, bem como adotar uma postura crítica para verificar a confiabilidade das estatísticas. Maciel (2023) acrescenta que o conhecimento de contexto atua como um representante da fonte segura da informação. Logo, conhecer o contexto do qual a informação estatística emergiu, resguarda o cidadão das interpretações equivocadas e, principalmente, das conclusões tendenciosas.

Gal (2002, 2019) enfatiza que ao ensinar Estatística, em sala de aula, o contexto deve ser a pedra angular de todas as atividades, pois o objetivo do processo de ensino e aprendizagem deve ser a formação de cidadãos que estejam dispostos a se envolver eficazmente com dados estatísticos e, principalmente, a dar sentido às estatísticas que fluem para eles como consumidores de informações em seus respectivos contextos. Para isso, o autor chama a atenção de que não basta apenas a abordagem de dados reais, em sala de aula, mas sim, relacionar esses dados a contextos significativos e importantes. Isso se faz necessário porque o mais relevante não são os dados brutos em si, mas as conclusões que podemos obter através deles. Logo, o contexto é o motivador das problematizações que podem ser formuladas no decorrer do processo de ensino e aprendizagem, como também, das respostas que são geradas com base nos dados fornecidos.

O autor ainda aponta que para um contexto ser significativo e importante, há duas condições explícitas. A primeira exige que o contexto deve ser autêntico, ou seja, deve ocorrer naturalmente do mundo exterior, não envolvendo dados artificiais ou fictícios. Assim, os professores devem levar para sala de aula contextos que mobilizem o interesse nos estudantes a pensar estatisticamente. A outra condição é que o contexto seja capaz emergir uma genuína “necessidade de saber”, ou seja, desperte nos estudantes a motivação para responder a questões problemas que podem ser pessoais, empresariais ou governamentais e que geram o processo de análise estatística. Gal (2019) finaliza enfatizando que a combinação dessas duas condições pode potencializar as ações do processo de ensino e aprendizagem, aumentando o interesse dos estudantes por uma estatística que faça sentido para eles.

Como último elemento de conhecimento, o autor aponta que os *Questionamentos Críticos* compreendem a capacidade do cidadão em analisar de

forma crítica as informações estatísticas veiculadas por diversos meios, avaliando a origem, construção e formulação, assim como, o propósito e a veracidade dessas informações. O autor destaca que as informações estatísticas podem surgir de diferentes meios de comunicação e serem veiculadas por distintos agentes. Isso implica na possibilidade dessas informações serem modeladas para os objetivos de cada agente. Assim, os cidadãos devem estar atentos e avaliar razoabilidade das declarações veiculadas pelos meios de comunicação. Devem questionar a autenticidade das mensagens, a origem e a confiabilidade das evidências que embasam as informações ou conclusões expostas e considerar outras possíveis interpretações das conclusões que lhes são comunicadas.

Esse tipo de conhecimento envolve perguntas sobre a origem dos dados, o uso de amostras, o tamanho e a representatividade da amostra, e se as informações veiculadas são verdadeiras. Para exemplificar esse conhecimento, retomamos a Figura 10 apresentada anteriormente, na qual a apresentadora, para explicar a diferença entre a projeção e o número de casos confirmados da COVID-19, recorreu a técnica visual ao utilizar o palmo da mão para argumentar sobre o distanciamento entre as duas linhas do gráfico. Diante disso, o telespectador com base nesse conhecimento poderia questionar a autenticidade da mensagem e a confiabilidade dos argumentos apresentados. Logo, uma postura questionadora e a ativação do senso crítico mediante são fundamentais mediante a qualquer tipo de informação estatística.

Embora essas cinco bases de conhecimentos tenham sido apresentadas de forma separada para fins didáticos, Gal (2002) destaca que elas se sobrepõem e não atuam de forma independente. Como exemplo, podemos citar a análise de uma pesquisa amostral por parte do cidadão. Ele deve mobilizar as *habilidades de letramento* para compreender as informações estatísticas sobre as conclusões da pesquisa, o *conhecimento estatístico* e *matemático* também se faz necessário para a compreensão de conceitos, técnicas e procedimentos empregados no processo da pesquisa. Além disso, deve reconhecer o *contexto* onde os dados foram coletados e as implicações das conclusões para o mesmo. Por fim, deve tecer *questionamentos críticos* para avaliar a razoabilidade das informações divulgadas e dos procedimentos adotados ao longo da pesquisa.

O segundo componente do Letramento Estatístico proposto por Gal (2002) são os *Elementos de Disposição*, os quais referem-se à postura crítica, disposição e concepções dos indivíduos frente à determinada informação estatística. Deste modo,

compreende-se que enquanto os Elementos Cognitivos estão relativos mais a aspectos técnicos da aprendizagem estatística, os Elementos de Disposição contemplam um caráter mais emocional e afetivo. Através desses elementos, espera-se que os indivíduos adotem uma postura ativa ao lidarem com situações que envolvem a Estatística.

Os Elementos de Disposição englobam três elementos: Crenças, atitudes e postura crítica. Nesse sentido, as crenças estão relacionadas com as concepções, ideias ou opiniões sobre diversos temas ou sobre si mesmo. Gal (2002, 2019) aponta que elas podem ser positivas “*a Curva Normal é um excelente modelo para a análise de dados*”, “*a Estatística é essencial para compreender a nossa realidade*” como também negativas, “*a Estatística Inferencial é muito difícil de ser compreendida*” ou “*as pesquisas amostrais de intenção do voto são sempre falsas*”. Assim, as crenças podem facilitar ou dificultar a disposição dos cidadãos para pensar sobre a Estatística. Gal (2002, 2019) destaca que elas são desenvolvidas ao longo do tempo, são pouco flexíveis e fortemente desenvolvidas pelo contexto em que o indivíduo está inserido.

No que diz respeito as Atitudes, elas são consideradas como as materializações das crenças dos indivíduos e refletem os seus aspectos emocionais. Por conta disso, estão intrinsecamente ligadas a tomada de decisões frente à informação estatística. Gal (2002) menciona que as atitudes são emoções relativamente duradouras e intensas que se formam gradualmente pela assimilação repetida de reações emocionais favoráveis ou desfavoráveis ao longo do tempo. Elas se manifestam ao longo de um continuo positivo-negativo (gosto/não gosto de pesquisa estatística, a estatística inferencial é agradável/ a estatística inferencial é desagradável) e podem ser alteradas através de experiências educativas apropriadas que incentivem a participação e o interesse dos cidadãos pela Estatística.

Já a Postura Crítica refere-se à capacidade de se adotar um comportamento questionador ao lidar com situações estatísticas, como por exemplo, verificar criticamente se as informações veiculadas na mídia podem ser enganosas, tendenciosas ou incompletas. Essa postura exige, do indivíduo, raciocinar estatisticamente e ter disposição para se comunicar e pensar sobre os conceitos da Estatística. Nesse viés, os cidadãos devem desenvolver uma visão positiva de si mesmos, reconhecendo-se como indivíduos capazes de raciocinar sobre estatística e probabilidade, além de demonstrar disposição e interesse em aplicar essa capacidade em contextos pertinentes. Isso implica na compreensão do valor dos métodos

estatísticos e estejam abertos à ideia de que estudos bem planejados têm o potencial de gerar conclusões mais sólidas e confiáveis do que aquelas baseadas em relatos pessoais ou experiências isoladas.

Diante do exposto, é pertinente ainda destacar que, de acordo com Gal (2002, 2019), os elementos de conhecimento e de disposição não devem ser concebidos como algo fixo ou independente, mas sim como um conjunto, que na prática, estão altamente relacionados nas ações e comportamento dos indivíduos em contextos que contemplam a Estatística. Espera-se que o indivíduo letrado estatisticamente tenha a capacidade de analisar criticamente informações relativas ao campo estatístico, ao mesmo tempo que demonstra vontade e interesse em “pensar estatisticamente” em diferentes situações.

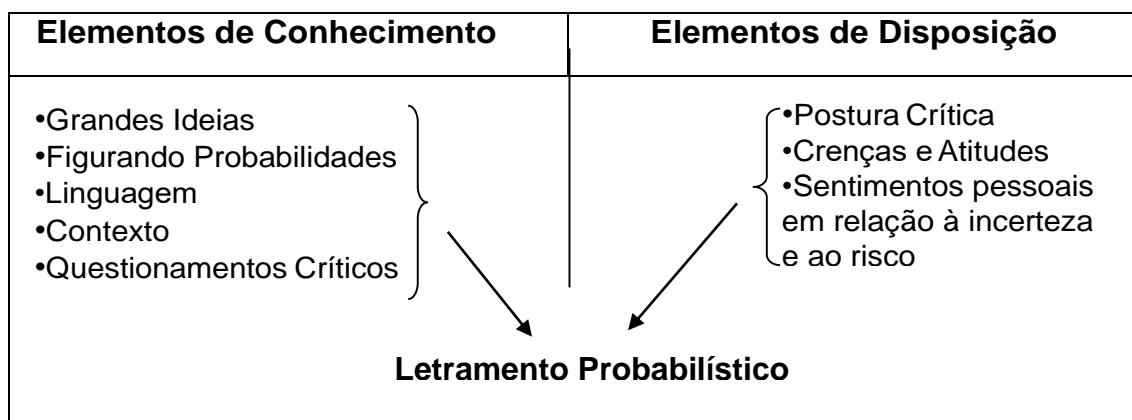
Além disso, Gal (2002, 2019) argumenta que o Letramento Estatístico deve ser considerado como é um processo dinâmico e contínuo de aprendizagem e está relacionado com as experiências e interesses de cada indivíduo. Nesse contexto, a escola, enquanto instituição formadora, deve sistematizar possibilidades para o desenvolvimento do letramento estatístico dos estudantes. Nessa direção, as orientações veiculadas pela BNCC (Brasil, 2018) vão ao encontro dos pressupostos do Letramento Estatístico de Gal (2002) ao indicar que os estudantes devem ter oportunidades de desenvolver habilidades de coleta, organização, análise e interpretação de dados, além de se envolverem com abordagens práticas e contextualizadas que impliquem na aplicação de conceitos estatísticos em situações do cotidiano e em diferentes áreas do conhecimento. Além disso, destaca a necessidade de promover o pensamento crítico e a tomada de decisão baseada em evidências, capacitando os estudantes a avaliar criticamente informações estatísticas apresentadas na mídia e em outras fontes.

Além do Letramento Estatístico, o referido autor também apresenta, de modo análogo, o modelo para o Letramento Probabilístico. Segundo Gal (2005) o indivíduo é considerado letrado em Probabilidade quando detém um conjunto de habilidades básicas que o possibilita ler e interpretar informações e dados probabilísticos presentes no seu cotidiano e na tomada de decisões.

Deste modo, o modelo de Letramento Probabilístico, sistematizado no Quadro 3, contempla dois componentes: O primeiro trata dos *Elementos de conhecimento* que contempla cinco elementos mais específicos: *Grandes Ideias, Figurando Probabilidades, Linguagens, Contexto e Questionamentos Críticos*. Já os *Elementos*

*da Disposição*, por sua vez, estão relacionados com *Postura Crítica e Crenças e Atitudes e os Sentimentos pessoais em relação à incerteza e ao risco* envolvendo os conceitos probabilísticos.

Quadro 3 – Modelo do Letramento Probabilístico



Fonte: Adaptado de Gal (2005).

Inicialmente, Gal (2005) destaca a necessidade do desenvolvimento desse letramento tendo em vista que a maioria dos cidadãos se portam como consumidores de dados e, consequentemente, estão inseridos em contextos que são permeados por situações que envolvem fenômenos aleatórios. O autor destaca que noções de probabilidade, incerteza e risco estão presentes em várias informações divulgadas pela mídia, relacionadas a áreas da saúde, econômica, ambiental, marketing, dentre outras. Além disso, tanto os cidadãos comuns quanto os profissionais necessitam calcular estimativas da probabilidade de determinados eventos mesmo que de modo informal. Diante disso, os cidadãos necessitam interpretar e compreender situações que envolvem conceitos probabilísticos para realizarem a tomada de decisões.

Porém, Gal (2005) argumenta que o pensamento e o comportamento dos indivíduos em situações probabilísticas são afetados por múltiplas bases de conhecimento e disposições. Nesse sentido, o primeiro elemento de conhecimento do seu modelo, *Grandes Ideias*, está relacionado com o domínio de termos e conceitos fundamentais para a compreensão da probabilidade, do quais autor destaca a aleatoriedade, variação e a incerteza. Em segundo plano, cita a derivação, representação, interpretação e implicações de declarações probabilísticas.

Sobre a Aleatoriedade, o autor aponta que se trata de um conceito de complexa assimilação e que ainda tem sido alvo de debate por não ter uma única definição precisa. Assim, Gal (2005) baseado nas ideias de Bennett (1998) indica esse termo

como uma propriedade de um resultado para um evento que ocorreu de forma desordenada, sem um roteiro prévio. Outra visão, define a aleatoriedade como uma ideia de que vários fenômenos ocorrem ao acaso, sem um projeto determinístico subjacente que permita a previsibilidade do seu resultado (Beltrami, 1999). Já a independência é uma ideia basilar da probabilidade porque, naturalmente, os eventos não estão totalmente conectados e a ocorrência de um evento não pode ser prevista a partir de outro. Já a ideia da variação é pertinente porque notadamente os eventos e fenômenos apresentam variações em seus comportamentos e resultados. O teórico menciona que a compreensão desses três termos é base para o domínio de outras ideias probabilísticas complementares, como a previsibilidade, a incerteza e noções relacionadas a risco e confiança.

A partir disso, Gal (2005) ressalta que essas ideias podem ser representadas de forma verbal, por símbolos matemáticos ou termos estatísticos e se faz necessário os cidadãos compreenderem de modo intuitivo a essência e significado desses termos em informações probabilísticas que permeiam inúmeras situações do cotidiano. Como exemplo, destacamos uma notícia divulgada pela mídia que menciona os termos incerteza e risco relacionados ao mercado financeiro.

Figura 15 – Notícia do mercado financeiro com menção a incerteza e risco

## Incerteza sobre meta fiscal elevou risco de investimentos, diz Copom em ata

Na última quarta (1º), BC cortou a taxa básica de juros em 0,5 ponto, a 12,25% ao ano

Fonte: Folha de São Paulo (07/11/2023).

Com base no elemento grande ideias, essa situação relaciona os termos incerteza e risco ao conceito de probabilidade. No mercado financeiro, a incerteza e o risco desempenham papéis fundamentais na tomada de decisões de investimento. Assim, a incerteza tem relação com a falta de conhecimento ou previsibilidade sobre eventos futuros, enquanto o risco é a medida da incerteza associada a um evento específico. Esses conceitos estão intrinsecamente ligados ao entendimento da probabilidade de diferentes resultados econômicos, a qual fundamenta as estratégias de investimento e as tomadas de decisão.

No que diz respeito ao elemento *Figurando Probabilidades*, Gal (2005) argumenta que os indivíduos devem compreender as diferentes formas para o cálculo de probabilidade de eventos, para que a partir dessa compreensão possam entender e comunicar sobre as informações e dados probabilísticos. O autor destaca que os cidadãos devem reconhecer que existem diferentes maneiras de chegar a resultados probabilísticos. Nessa direção, Batanero (2005) apresenta cinco tipos de significados ou abordagens para a Probabilidade: intuitiva, clássica, frequentista, subjetiva e axiomática.

O significado Intuitivo refere-se à ideia de que a probabilidade é uma medida da plausibilidade de um evento ocorrer, baseada em evidências disponíveis ou na compreensão do processo. Esse significado refere-se à nossa intuição ou senso de probabilidades, muitas vezes baseado em suposições informais. Silva e Carvalho (2022) apontam que essa ideia teve início em contextos de jogos de azar, nos quais eram formuladas visões pessoais sobre as chances de ganho ou perda. Nesse sentido, essa perspectiva intuitiva está relacionada ao pensamento probabilístico e à comunicação inseridos em contextos de probabilidade. Por exemplo, expressões como “é provável”, “é pouco provável”, “é possível” e “é impossível” esboçam uma intuição sobre elementos probabilísticos e fazem com que se tenha uma reação perante fenômenos que acontecem aleatoriamente.

O significado clássico, também chamado de laplaciano, apresenta a ideia da probabilidade vinculada ao cálculo de uma razão entre o número de elementos do evento, comumente também chamado de casos favoráveis, pelo número de casos do espaço amostral, paralelamente mencionado também como todos os casos possíveis. Assim, se o evento A é subconjunto de um espaço amostral E, sua probabilidade de ocorrência é dada por  $p(A) = \frac{n(A)}{n(E)}$ , no qual n corresponde a quantidade de elementos em cada conjunto. Gal (2005) e Lima e Borba (2019) relatam que esse significado é o preferencialmente abordado nas aulas de matemática e nos livros didáticos no âmbito da Educação Básica.

O significado Frequentista, por sua vez, é uma abordagem que define a probabilidade de um evento com base na frequência relativa com que esse evento ocorre em um grande número de repetições do mesmo experimento. Em outras palavras, a probabilidade de um evento é vista como a proporção de vezes que o evento ocorre em relação ao número total de repetições do experimento, quando esse

número tende ao infinito. Por exemplo, se lançarmos uma moeda justa, o significado frequentista de probabilidade diria que a probabilidade de obter cara é de 0,5 ou 50% pois, em um grande número de lançamentos da moeda, espera-se que aproximadamente metade dos resultados seja cara.

Já o significado subjetivo define a probabilidade como uma medida da incerteza ou crença pessoal de um indivíduo sobre a ocorrência de um evento. Segundo essa interpretação, a probabilidade reflete a confiança ou grau de certeza que um indivíduo atribui a um determinado evento, com base em suas próprias percepções, conhecimentos e crenças. Portanto, a probabilidade é vista como uma expressão subjetiva da incerteza, variando de pessoa para pessoa e podendo mudar ao longo do tempo, à medida que novas informações são adquiridas ou as circunstâncias mudam.

Essa abordagem subjetiva é frequentemente utilizada em situações onde as probabilidades não podem ser determinadas objetivamente com base em dados empíricos ou onde diferentes indivíduos podem ter diferentes interpretações dos mesmos eventos. Por exemplo, ao avaliar a probabilidade de sucesso de um projeto de negócios, um investidor pode levar em consideração sua experiência no setor, análise de mercado e intuição pessoal para fazer uma estimativa subjetiva da probabilidade de retorno positivo do investimento.

Na esfera da probabilidade subjetiva, Batanero (2005) destaca a probabilidade bayesiana, a qual se trata de uma abordagem estatística que se baseia no teorema de Bayes para atualizar a probabilidade de uma hipótese com base em novas evidências observadas. No contexto bayesiano, a probabilidade é interpretada como uma medida de incerteza ou grau de crença sobre a verdade de uma determinada hipótese, dada a informação disponível. O teorema de Bayes descreve como essa crença é atualizada à medida que novas evidências são obtidas. Assim, a inferência bayesiana permite incorporar informações prévias (*a priori*) sobre o problema em análise, bem como atualizar essas informações à medida que novas evidências são observadas. Isso a torna uma ferramenta poderosa para tomar decisões em face de incerteza e para modelar uma ampla variedade de problemas em diferentes áreas, incluindo ciência, medicina, engenharia, entre outros. Diante disso, essas duas perspectivas apresentam relação. Enquanto a probabilidade subjetiva é uma abordagem que reconhece e incorpora as crenças pessoais de um indivíduo, a probabilidade bayesiana é uma técnica que usa o teorema de Bayes para atualizar

probabilidades com base em evidências observadas, permitindo uma combinação formal de crenças subjetivas e dados observados.

Por fim, o significado axiomático é uma abordagem que define a probabilidade de eventos por meio de um conjunto de axiomas ou princípios matemáticos fundamentais. A partir desses axiomas, é possível derivar uma série de propriedades importantes e teoremas da probabilidade. O significado axiomático de probabilidade fornece uma base matemática sólida para a teoria das probabilidades. Como por exemplo, o axioma da complementaridade, o qual fundamenta que a probabilidade do evento complementar de um evento A é igual a 1 menos a probabilidade de A. De forma prática, se em um contexto a probabilidade de chover é 60%, consequentemente, a de não chover é de 40%.

Dando continuidade ao modelo de Letramento Probabilístico proposto por Gal (2005) o terceiro elemento de conhecimento é a *Linguagem*. Sobre isso, o autor menciona que os cidadãos devem compreender as diversas formas utilizadas para representar e comunicar sobre probabilidade, pois em nossa sociedade precisamos frequentemente representar, comunicar e dialogar sobre conceitos e resultados probabilísticos. Isso inclui a compreensão abstrata sobre termos do campo probabilístico, tais como: variabilidade, aleatoriedade, independência, previsibilidade, imprevisibilidade certeza, incerteza chance, probabilidade e risco. Como também as probabilidades reais de eventos ocorrerem que podem ser representadas quantitativamente, por uma escala de 0 a 1, por uma fração ou porcentagem.

Notadamente, compreender essa linguagem requer também domínio sobre o conhecimento matemático subjacente. Além disso, o autor cita a necessidade de ter familiaridade com a linguagem verbal que pode ser utilizada para indicar noções de probabilidades, incluindo frases ou termos que remetem a incerteza ou a certeza. Como exemplo desse conhecimento, apresentamos uma notícia veiculada por um portal que utiliza o termo “chance” para apresentar a informação relacionada a testes rápidos no contexto da COVID-19.

Figura 16 – Notícia em contexto de saúde com menção à chance

## **Testes rápidos podem ter até 75% de chance de erro**

Ministério da Saúde vê "limitações importantes" em testes rápidos e pede cautela para gestores do SUS

Fonte: Portal terra (01/04/2020).

Diante dessa informação, o cidadão letrado probabilisticamente, compreenderá que a chance mencionada no texto está sendo usada para descrever a probabilidade de um evento ocorrer, especificamente, a probabilidade de um teste rápido produzir um resultado incorreto. Logo, a palavra “chance” é uma forma de se referir à probabilidade de um evento acontecer, indicando a possibilidade ou a incerteza associada a esse evento.

Em continuidade, o quarto elemento de conhecimento, *Contexto*, está relacionado com a compreensão do papel e dos significados das informações e dados probabilísticos em diferentes contextos, isto é, nas situações onde as noções de probabilidade podem aparecer no dia a dia de um indivíduo. O autor aponta que paralelamente à compreensão dos conceitos probabilísticos, torna-se necessário a compreensão do contexto em que este conceito está inserido, tornando-o significativo, pois “há uma necessidade de aprender sobre probabilidade ou incerteza em diferentes circunstâncias da vida” (Gal, 2005, p. 52).

O autor enfatiza a importância de considerar o ambiente, as informações disponíveis e a finalidade da probabilidade ao interpretar e comunicar informações probabilísticas em diferentes situações. Isso ajuda a garantir uma compreensão mais completa e precisa das probabilidades em seus contextos específicos de aplicação. As Figuras 15 e 16 apresentadas anteriormente exemplificam a ideia de contexto, ao envolverem conceitos probabilísticos em cenários, respectivamente, da economia e da saúde.

O último elemento de conhecimento do Letramento Probabilístico, *Questionamentos Críticos*, está relacionado com a capacidade dos indivíduos desenvolverem um pensamento crítico sobre as informações e dados probabilísticos. Segundo o autor, este elemento é necessário porque os questionamentos críticos também podem estar relacionados a julgamento equivocados relativos às estimativas

de probabilidades, pois muitos indivíduos “tendem a estimar as probabilidades de forma imprecisa ou pensar sobre aleatoriedade, independência, variação ou riscos desvinculados dos aspectos formais” (Gal, 2005 p. 55).

Por meio desse conhecimento, o autor indica que os cidadãos devem tecer questões críticas quanto ao contexto da informação estatística, discriminando sua natureza, como também, a fonte da informação, o processo realizado para estimar probabilidades, as interpretações e significados das informações probabilísticas. Nesse sentido, os questionamentos críticos desempenham um papel importante porque ajudam a promover uma reflexão mais profunda e uma compreensão mais abrangente das informações probabilísticas. Esses questionamentos são fundamentais para o desenvolvimento do letramento probabilístico que envolve não apenas a compreensão básica de conceitos probabilísticos, mas também a capacidade de interpretar e avaliar criticamente as informações probabilísticas encontradas em diversas situações do cotidiano.

Além disso, como citado anteriormente, o modelo de Letramento Probabilístico também contempla os *Elementos da Disposição*. O primeiro elemento, *Postura Crítica*, refere-se à atitude ou disposição mental que os indivíduos desenvolvem em relação às informações probabilísticas. Essa postura crítica implica uma abordagem reflexiva e questionadora ao lidar com informações que envolvem incerteza e probabilidades. Por exemplo, ao se deparar com o resultado de uma pesquisa amostral, os cidadãos devem ter uma postura crítica para analisar e julgar todo o procedimento de pesquisa adotado, incluindo o método de Amostragem empregado, tamanho da amostra, nível de confiança e margem de erro.

Além do mais, é importante levar em consideração, na construção desse letramento, as *Crenças e Atitudes* pessoais envolvendo o pensamento probabilístico, isto é, as concepções e o sentimento de cada indivíduo, bem como a vontade e o interesse em pensar sobre a probabilidade e seus significados. As crenças estão relacionadas com as convicções de um indivíduo sobre as probabilidades, contemplando a confiança na precisão das informações probabilísticas, convicções sobre a própria capacidade de entender e lidar com probabilidades, entre outras. Já as atitudes referem-se às inclinações emocionais ou predisposições que um indivíduo deve ter em relação a probabilidades, riscos e incertezas. Isso pode incluir sentimentos de confiança, medo, desconfiança, otimismo ou pessimismo.

Por fim, os *Sentimentos pessoais em relação a incerteza e risco* referem-se a fatores afetivos e disposições emocionais que uma pessoa experimenta ao lidar com informações probabilísticas e situações de incerteza. Esses sentimentos podem desempenhar um papel crucial na forma como uma pessoa interpreta e responde a informações probabilísticas. O medo pode surgir quando a incerteza é percebida como uma ameaça aos objetivos pessoais. Por outro lado, a confiança pode impulsionar a disposição de alguém em enfrentar a incerteza, especialmente se houver uma crença sólida em suas habilidades de análise e tomada de decisões.

Como exemplo desses componentes de Disposição, podemos citar a possibilidade de um cidadão estar direcionado a investir no mercado financeiro. Inicialmente, ele deve fazer várias análises junto as instituições de mercado que fornecem probabilidades sobre o desempenho futuro de diferentes empresas. Paralelamente, também deve mobilizar suas crenças, as quais podem denotar os riscos quanto à imprevisibilidade desse mercado. Como resultado, mesmo que as análises forneçam informações promissoras sobre o potencial de crescimento de suas aplicações, esse cidadão pode ser cético em confiar nessas probabilidades. Suas crenças sobre a imprevisibilidade do mercado e sua própria capacidade de compreendê-lo afetam suas atitudes em relação às informações probabilísticas. Ele pode optar por não investir ou investir em uma estratégia mais conservadora, baseada em suas crenças e atitudes em relação ao risco e a incerteza relativos ao mercado financeiro.

Diante do exposto, Gal (2005) salienta que os elementos da disposição possuem uma relevância equivalente aos elementos de conhecimento no contexto do letramento probabilístico. Isso ocorre porque ambos estão intrinsecamente interligados, e qualquer um desses fatores pode exercer uma influência positiva ou negativa sobre as decisões dos indivíduos.

Apresentados os modelos do Letramento Estatístico e Probabilístico proposto por Gal (2002, 2005), cabe ainda destacarmos que apesar de abordarmos esses elementos separadamente, eles se interligam e interagem entre si de forma complexa. Disto, devemos compreender que o ensino de Estatística e Probabilidade deve oferecer ferramentas para o desenvolvimento de tais elementos, em sua totalidade, a fim de propiciar a construção do Letramento Estatístico e Probabilístico dos estudantes.

O autor ainda concebe que os tais Letramentos devem começar a serem construídos pelos indivíduos desde o início de sua escolarização até a vida adulta, tendo em vista que os estudantes já se deparam com situações envolvendo conceitos estatísticos e probabilísticos. Assim, ao abordar tais conceitos, em sala de aula, é necessário levar em conta os diferentes significados que eles podem assumir, a linguagem, o contexto em que estão inseridos, como também os questionamentos críticos envolvendo seus significados e a disposição dos estudantes e professores para pensar e racionar sobre esses conceitos.

### 3.2 O CICLO INVESTIGATIVO NO ENSINO DE ESTATÍSTICA E PROBABILIDADE

No capítulo anterior, discutimos que a Estatística, enquanto ciência, se ocupa da coleta, organização, análise e tratamento de dados e informações, além de lidar com situações probabilísticas não-determinísticas, ou seja, que envolvem a incerteza (Bayer *et al*, 2005).

Nesse contexto, no âmbito do Letramento Estatístico e Probabilístico apresentados anteriormente, Gal (2002, 2005) observa que os indivíduos, em suas interações sociais, comumente assumem o papel de consumidores de dados ao se apropriarem de informações veiculadas por diversas fontes, as quais são resultados de investigações. Ademais, também destaca a presença dos produtores de dados, ou seja, os indivíduos envolvidos em contextos de pesquisa estatística.

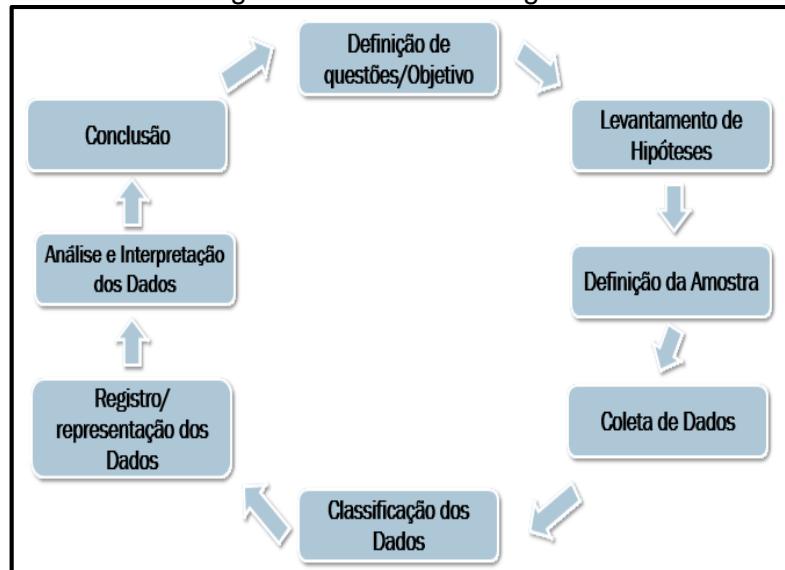
Em se tratando do processo de pesquisa estatística, Guimarães e Gitirana (2013) ressaltam a importância da abordagem e execução do ciclo investigativo de pesquisa estatística na Educação Básica. Elas argumentam que os estudantes não apenas precisam compreender os conceitos, técnicas e representações estatísticas e probabilísticas, mas também devem desenvolver competências e habilidades que os capacitem a conduzir processos de investigação.

Nessa perspectiva, as autoras defendem que a pesquisa deve ser vista como o eixo estruturador da formação estatística dos estudantes em todos os níveis de escolaridade. Através do envolvimento no processo de investigação, os estudantes têm a oportunidade de aprofundar o conhecimento sobre os conceitos, técnicas e representações do campo estatístico.

Guimarães e Gitirana (2013) formularam um modelo para o ciclo investigativo de uma pesquisa, representado na Figura 17, o qual envolve as seguintes fases, por

ordem de realização: *Definição da questão/objetivo*, *Levantamento de hipóteses*, *Definição da amostra*, *Coleta de dados*, *Organização/Classificação dos dados*, *Registro/Representação dos dados*, *Análise de dados* e *Conclusão*.

Figura 17 – Ciclo Investigativo



Fonte: Guimarães e Gitirana (2013)

Discorrendo sobre cada uma dessas fases, as autoras apontam que o primeiro passo de uma pesquisa é a definição da questão/objetivo. Nela se formula a indagação motivadora da investigação, a qual o pesquisador pretende responder. Logo, essa fase é crucial para garantir que a pesquisa seja direcionada e estruturada, resultando em possíveis descobertas significativas.

Em seguida, na fase de *Levantamento de Hipóteses* são realizadas reflexões sobre as variáveis da pesquisa e suas relações, juntamente com a formulação de hipóteses como resposta à questão da pesquisa, tendo por base, justificativas que ao final da investigação podem ser confirmadas ou refutadas. Assim, nesta etapa, os pesquisadores formulam suposições ou conjecturas sobre as relações entre variáveis ou fenômenos que serão investigados.

A terceira fase é destinada para a *definição da amostra* da pesquisa. No capítulo 1 deste estudo vimos que, no âmbito da Estatística, a amostra é classificada como um subconjunto representativo de uma população e esta, por sua vez, é caracterizada como o universo de elementos que possuem pelo menos uma característica em comum. Nesse sentido, em pesquisas estatísticas, especialmente quando lidamos com populações de grande tamanho, é mais viável e prático trabalhar

com uma amostra, a qual permite que se obtenham conclusões confiáveis, capazes de caracterizar e representar as respectivas populações de forma precisa. Uma amostra bem definida e representativa é a base sobre a qual se constrói todo a estrutura da análise estatística e das conclusões da pesquisa.

Em continuidade, a quarta fase é responsável pela coleta de dados. Nela, o pesquisador deve, a partir de instrumentos adequados, coletar os dados que o permitirá, em fases posteriores, analisá-los e interpretá-los com vistas a responder à questão da pesquisa. Essa fase permite que os pesquisadores selezionem os métodos e instrumentos mais apropriados para obter as informações desejadas. Isso pode incluir a realização de entrevistas, aplicação de questionários, observação direta, entre outros métodos, dependendo da natureza da pesquisa e das características da amostra/população estudada. Dessa forma, os dados coletados são a matéria-prima da análise estatística, a qual permite que as conclusões sejam robustas e representativas da realidade (Guimarães e Gitirana, 2013).

Após a coleta, a quinta fase do ciclo é direcionada para a classificação dos dados. O procedimento de classificar possibilita verificar, em um conjunto de elementos, aqueles que têm a mesma característica a partir de um critério determinado. Isso permite, ao pesquisador, a identificação de padrões, tendências e relações significativas entre as variáveis estudadas. Essa fase possibilita uma melhor organização dos dados com vistas à análise e interpretação de forma mais objetiva.

Na sequência, a sexta fase é reservada para o *Registro/representação de dados*. Nessa etapa, o pesquisador deve utilizar ferramentas adequadas de representação, como por exemplo, gráficos ou tabelas, que lhe propicie, de maneira sucinta, a compreensão do perfil dos dados e a relações estabelecidas entre as variáveis. O registro adequado dos dados é essencial para preservar a integridade das informações coletadas. Isso envolve a organização dos dados de forma sistemática, garantindo que todas as observações sejam registradas de maneira precisa e completa. Um registro cuidadoso dos dados ajuda a minimizar erros durante a análise estatística, garantindo a confiabilidade dos resultados.

Além disso, a representação dos dados em formatos visualmente acessíveis e informativos é importante para facilitar a análise e a interpretação. Gráficos, tabelas e outras formas de representação visual podem ajudar os pesquisadores a identificar padrões, tendências e relações entre as variáveis de forma direta e intuitiva. Isso

permite que os resultados da pesquisa sejam comunicados de maneira mais eficaz para diferentes públicos, tornando as informações mais acessíveis e compreensíveis.

Posteriormente ao registro, a sétima etapa é destinada a *Análise dos dados*. Nela, o pesquisador deve se ocupar do tratamento e análise dos dados, utilizando de conceitos estatísticos e probabilísticos, que lhe permitam fazer inferências estatísticas sobre esses dados. A análise dos dados possibilita que os pesquisadores testem suas hipóteses e respondam às questões de pesquisa formuladas no início do estudo. Por meio de técnicas e modelos estatísticos e probabilísticos apropriados, é possível explorar as relações entre as variáveis e determinar se existem associações estatisticamente significativas.

Por fim, a última etapa envolve a *Conclusão*, na qual o pesquisador deve formular conclusões para a pesquisa de acordo com os dados analisados. Assim, nessa etapa são estabelecidas conexões entre a questão da pesquisa e os resultados, confrontando-os com as hipóteses iniciais, o que pode levar a surgir novas questões, e, assim, retomar todo o ciclo investigativo. A *Conclusão* propicia a discussão das implicações práticas e teóricas dos resultados da pesquisa. Os pesquisadores podem explorar como seus achados podem influenciar práticas, políticas ou intervenções no mundo real, bem como identificar lacunas no conhecimento que sugerem direções para a realização de outras pesquisas.

Diante disso, Guimarães e Gitirana (2013) defendem que a escola, enquanto instituição formadora, deve oportunizar aos estudantes a realização do ciclo investigativo, tendo em vista que ele favorece a interação desses estudantes com as práticas sociais e com a natureza; além de incentivar a linguagem oral, o contato com diferentes representações de dados. Dessa forma, o ciclo investigativo proporciona aos estudantes uma oportunidade de aplicar os conceitos estatísticos em contextos do mundo real, o que pode tornar o aprendizado mais significativo. Ao engajarem-se em processos de pesquisa que abordam questões do seu próprio interesse ou relacionadas ao seu ambiente, os estudantes podem ver a importância prática da estatística e como ela pode ser aplicada em diversas áreas de conhecimento.

Corroborando com essas ideias, Barberino e Magalhães (2016) apontam a importância do desenvolvimento de projetos de pesquisa estatística na escolarização básica por acreditarem que os mesmos contribuem para a formação de estudantes críticos. Os autores defendem que a pesquisa estatística deve partir de situações relacionadas com o contexto social dos estudantes, o que pode facilitar a construção

de conceitos de modo significativo, não se prendendo apenas ao tratamento de técnicas operacionais.

O estudo conduzido por Gonçalves e Lima (2017) visou desenvolver uma abordagem educacional, pautada na realização de uma pesquisa estatística por parte dos alunos do Ensino Médio. Os autores observaram que a pesquisa foi eficaz para promover uma postura ativa dos estudantes enquanto pesquisadores, além de utilizar a matemática como ferramenta para resolver problemas do cotidiano. Verificou-se que os estudantes aprofundaram seu entendimento em conceitos fundamentais, como população, amostra, variável e medidas estatísticas. Conclui-se que o estudo de conceitos estatísticos, com foco na pesquisa, é crucial para desenvolver sujeitos críticos diante de problemas cotidianos, destacando a Estatística como uma ferramenta para compreender e transformar a realidade.

Muller e Nunes (2016) enfatizam que a proposta de se realizar uma pesquisa estatística no decorrer do processo de ensino e aprendizagem propicia a abordagem de conceitos estatísticos presentes na Inferência Estatística e a construção e ampliação de conhecimentos sobre tais conceitos. Além disso, os autores apontam que a realização de pesquisa estatística pelos estudantes se constitui como um importante recurso para o desenvolvimento da autonomia, o trabalho colaborativo, senso de cidadania, bem como a visão crítica sobre dados relativos aos seus contextos sociais.

Indo ao encontro dessa temática, a BNCC (Brasil, 2018) orienta que no ensino de Probabilidade e Estatística, nos anos finais do Ensino Fundamental e no Ensino Médio da Educação Básica, seja vivenciado o planejamento e execução de uma pesquisa amostral sobre questões relevantes do cotidiano dos estudantes. Esse documento enfatiza a importância de abordar a pesquisa estatística de forma interdisciplinar, integrando-a a diferentes áreas do conhecimento. Esse procedimento possibilita que os estudantes reconheçam a aplicabilidade da estatística em diferentes contextos e desenvolvam uma compreensão mais ampla de sua importância na sociedade. Além disso, o referido documento destaca a necessidade de promover o pensamento crítico e a tomada de decisão informada por meio da pesquisa estatística.

Nesse contexto, a presente pesquisa desenvolveu, na etapa do estudo experimental, a realização de um ciclo investigativo por parte de estudantes do 3º ano do Ensino Médio. O foco recaiu sobre as etapas de Definição da Amostra, Registro e Representação dos Dados, e Análise dos Dados. A intervenção pedagógica realizada

visou promover a aprendizagem e, por conseguinte, o desenvolvimento do letramento estatístico e probabilístico (Gal, 2002, 2005) dos estudantes em relação a conceitos como Amostragem e Curva Normal, bem como, suas relações. Esta abordagem contou com o suporte do Geogebra e a utilização de lápis e papel, enquanto recursos didáticos.

## 4 ORIENTAÇÕES CURRICULARES E REVISÃO DE LITERATURA

No que diz respeito ao processo de ensino e aprendizagem da Amostragem e Curva Normal no Ensino Médio, abordaremos, neste capítulo, as recomendações dos documentos oficiais sobre o ensino da Estatística e da Probabilidade com foco nesses conceitos. Além disso, apresentamos estudos anteriores que versam sobre a compreensão de estudantes e professores sobre esse tema, bem como, possibilidades para o seu ensino e aprendizagem. Essas pesquisas foram consultadas nas principais fontes acadêmicas, como o Catálogo de Teses e Dissertações da Capes, a Biblioteca Digital Brasileira de Teses e Dissertações, e em eventos e revistas na área da Educação Matemática e Estatística de nível nacional e internacional.

### 4.1 RECOMENDAÇÕES DOS DOCUMENTOS OFICIAIS SOBRE O ENSINO DA ESTATÍSTICA E DA PROBABILIDADE

No âmbito do processo de ensino e aprendizagem, as orientações curriculares exercem um importante papel ao estabelecerem referências técnicas que subsidiam as ações de todo corpo escolar, assim como, para a elaboração de recursos didáticos. Deste modo, é pertinente compreendermos o que tem sido proposto para o processo de ensino e aprendizagem de Amostragem e da Curva Normal no Ensino Médio.

Apresentaremos nessa seção, as recomendações dos documentos oficiais para Educação Básica, veiculadas pelos Parâmetros Curriculares Nacionais - (PCN's), o currículo para o Ensino Médio do Estado de Pernambuco (PCPE) e a Base Nacional Curricular Comum (BNCC) que buscam propiciar a formação educacional dos indivíduos. Deste modo, analisaremos, especificamente, as expectativas de aprendizagem e as recomendações para ensino de Estatística e Probabilidade para o Ensino Médio, com foco nos conteúdos curriculares propostos para o ensino de Amostragem e da Curva Normal.

De início, focando nos Parâmetros Curriculares Nacionais para o Ensino Médio - PCNEM (Brasil, 2000,2002), cabe frisar que os mesmos não são mais considerados como o principal documento norteador para a Educação Básica no Brasil, mas os abordaremos por entender que é pertinente compreendermos a evolução dos documentos curriculares. Isso posto, os PCNEM (Brasil, 2000,2002) apresentam, de

forma sistematizada, os blocos de conteúdos e as competências e habilidades que devem ser desenvolvidos e contemplados em cada ciclo da Educação Básica.

De modo geral, esse documento define o conhecimento matemático como um saber que permeia diversas áreas do conhecimento, tornando-se uma necessidade humana para a construção de uma visão de mundo e o desenvolvimento de capacidades e habilidades que propiciem, aos estudantes, a leitura, interpretação e compreensão de nossa realidade. Deste modo, o referido documento indica três competências gerais a serem alcançadas no Ensino Médio com o ensino de Matemática:

- Representação e comunicação, que envolvem a leitura, a interpretação e a produção de textos nas diversas linguagens e formas textuais características dessa área do conhecimento;
- investigação e compreensão, competência marcada pela capacidade de enfrentamento e resolução de situações-problema, utilização dos conceitos e procedimentos peculiares do fazer e pensar das ciências;
- contextualização das ciências no âmbito sócio-cultural, na forma de análise crítica das ideias e dos recursos da área e das questões do mundo que podem ser respondidas ou transformadas por meio do pensar e do conhecimento científico. (Brasil, 2002, p. 113).

No que diz respeito ao ensino de Estatística e Probabilidade, além da apropriação e ampliação do conhecimento sobre os conceitos abordados no Ensino Fundamental, orientam para uma abordagem nas aulas de Matemática voltada para a descrição de dados, representações gráficas, análise de dados, contemplando os conceitos de médias, moda e mediana, variância e desvio padrão, bem como, a ideia de possibilidades e o cálculo de probabilidades. Objetiva-se que o estudante do Ensino Médio analise com criticidade os significados de informações e dados, de forma a contemplar toda a investigação sobre os mesmos e propiciar a tomada de decisões.

Estatística: descrição de dados; representações gráficas; análise de dados: médias, moda e mediana, variância e desvio padrão.

- Identificar formas adequadas para descrever e representar dados numéricos e informações de natureza social, econômica, política, científico-tecnológica ou abstrata.
- Ler e interpretar dados e informações de caráter estatístico apresentados em diferentes linguagens e representações, na mídia ou em outros textos e meios de comunicação.
- Obter médias e avaliar desvios de conjuntos de dados ou informações de diferentes naturezas.

- Compreender e emitir juízos sobre informações estatísticas de natureza social, econômica, política ou científica apresentadas em textos, notícias, propagandas, censos, pesquisas e outros meios.  
• Probabilidade: possibilidades; cálculo de probabilidades.
  - Reconhecer o caráter aleatório de fenômenos e eventos naturais, científico, tecnológicos ou sociais, compreendendo o significado e a importância da probabilidade como meio de prever resultados.
  - Quantificar e fazer previsões em situações aplicadas a diferentes áreas do conhecimento e da vida cotidiana que envolvam o pensamento probabilístico.
  - Identificar em diferentes áreas científicas e outras atividades práticas modelos e problemas que fazem uso de estatísticas e probabilidades.
- (Brasil, 2002, p. 127)

Diante disso, os PNCEM (Brasil, 2000,2002) não apresentam de forma explícita orientações para o ensino de Amostragem. Apesar disso, podemos considerar que as orientações para esse ensino estão pautadas de forma implícita, sobretudo através de alguns tópicos de recomendações relacionados com a divulgação e informações estatísticas e a realização de pesquisas, como em “Compreender e emitir juízos sobre informações estatísticas de natureza social, econômica, política ou científica apresentadas em textos, notícias, propagandas, censos, pesquisas e outros meios” (Brasil, 2000, p.127). Com relação a Curva Normal, também não há recomendação explícita, mas de modo análogo à Amostragem, entendemos que há recomendação que contempla esse conceito de modo implícito, a exemplo de “Identificar em diferentes áreas científicas e outras atividades práticas modelos e problemas que fazem uso de estatísticas e probabilidades” (Brasil, 2000, p.128), que abarca o conceito da Curva Normal, tendo em vista que a mesma se trata de um modelo presente na inferência estatística, fazendo uso de estatísticas e probabilidades e que está presente em diferentes áreas científicas. Diante do exposto, podemos concluir que cabe ao professor, a partir desse documento, reconhecer e compreender esses dois conceitos tratados de forma implícita para que seja garantido suas abordagens em sala de aula, e, por conseguinte, favorecer a aprendizagem dos estudantes.

A partir das propostas veiculadas pelos PCN's (Brasil, 1998,2000,2002), o estado de Pernambuco desenvolveu os “Parâmetros Curriculares para a Educação Básica do Estado de Pernambuco” - PCPE (Pernambuco, 2013). O referido documento contempla os blocos de conteúdos curriculares para cada disciplina, as expectativas de aprendizagem dos estudantes, ano a ano, do Ensino Fundamental ao

Ensino Médio, bem como orientações voltadas para o processo de ensino e aprendizagem e as práticas pedagógicas a serem desenvolvidas em sala de aula.

Nesse contexto, no que diz respeito à disciplina de Matemática, o PCPE (Pernambuco, 2013) propõe a divisão de conceitos matemáticos em cinco blocos de conteúdos, a saber: Geometria, Estatística e Probabilidade, Álgebra e Funções, Grandezas e Medidas e Números e Operações. Desta forma, cada referido bloco é subdividido por tópicos de conteúdos, expectativas de aprendizagem e orientações para o ensino. Tal organização permeia todas as etapas da escolarização básica, isto é, do 1º ano dos anos iniciais do Ensino Fundamental ao 3º ano do Ensino Médio, de forma que conceitos matemáticos sejam retomados e ampliados, progressivamente, a cada ano escolar. Desta forma, “*o objetivo é o de que, independentemente da etapa em que o professor lecione, ele conheça as orientações dos outros anos e das outras etapas, para se apropriar da lógica interna de construção dos conceitos matemáticos*” (Pernambuco, 2013, p. 15).

Nessa perspectiva, explorando especificamente o bloco de conteúdos “Estatística e Probabilidade” destaca-se, a seguir, a divisão de conteúdos em tópicos gerais que devem ser abordados ao longo das etapas da Educação Básica, retomando e aprofundando os mesmos, a cada ano escolar:

- Coleta e organização de dados: elaboração de problemas e planos de pesquisas; organização e categorização de dados; população e amostra; análise de dados coletados; tipos de variáveis; frequências absoluta, relativa e acumulada.
- Representação de dados: construção de tabelas e gráficos; identificação e interpretação de informações apresentadas em tabelas e gráficos; comparação e conversão de diferentes representações de dados; elementos constitutivos de gráficos; agrupamento de dados em classes.
- Medidas estatísticas: média, moda, mediana e quartil; comparação de conjuntos de dados por meio de medidas de tendência central; amplitude, desvio médio, variância e desvio padrão.
- Probabilidade: eventos determinísticos e aleatórios; resultados possíveis de um experimento; cálculo de probabilidades. (Pernambuco, 2013, p.18 e 19).

No que diz respeito ao ensino de Estatística e Probabilidade no Ensino Médio, foco do nosso estudo, os PCPE (Pernambuco, 2013) apresentam as expectativas de aprendizagem a serem alcançadas e desenvolvidas pelos estudantes nas três séries que compõem esse nível de escolarização, as quais destacamos a seguir:

- Realizar uma pesquisa considerando todas as suas etapas (planejamento, seleção de amostras, elaboração e aplicação de instrumentos de coleta, organização e representação dos dados, interpretação, análise crítica e divulgação dos resultados).
- Selecionar uma amostra adequada para uma determinada pesquisa
- Construir tabelas e gráficos de diferentes tipos (barras, colunas, setores e gráficos de linha, histograma), preferencialmente utilizando recursos tecnológicos.
- Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extração, interpolação e interpretação.
- Utilizar informações expressas em gráficos ou tabelas para fazer inferências.
- Resolver problema com dados apresentados em tabelas ou gráficos.
- Analisa informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.
- Determinar frequências relativas, acumuladas e acumuladas relativas de dados agrupados.
- Calcular e interpretar medidas de tendência central (média, moda e mediana) para um conjunto de dados numéricos não agrupados.
- Resolver e elaborar problema que envolva a interpretação de tabelas e gráficos de diferentes tipos.
- Calcular e interpretar medidas de dispersão (amplitude, desvio médio, variância e desvio padrão) para um conjunto de dados numéricos não agrupados.
- Determinar a probabilidade de ocorrência de um evento, explorando representações diversas.
- Determinar a probabilidade da união de dois eventos, explorando representações diversas.
- Organizar tabelas com dados numéricos agrupados ou não agrupados.
- Determinar a probabilidade da união e da intersecção de eventos.
- Determinar a probabilidade condicional (Pernambuco, 2013, p.18 e 19).

Deste modo, com relação à Amostragem esse documento cita e orienta explicitamente o ensino deste tema ao propor a realização de pesquisas estatísticas que envolvem a adequada seleção de amostras. Além disso, nas orientações metodológicas e de ensino, esse documento recomenda que “conceito de amostra deve ser ampliado e os diferentes tipos de Amostragem, por exemplo, aleatória simples, proporcional, estratificada etc., devem ser igualmente discutidos com o estudante, em função de sua adequabilidade à pesquisa”. (Pernambuco, 2013, p.79). Assim, espera-se que o estudante ao vivenciar o ciclo investigativo de uma pesquisa tenha a oportunidade de discutir conceitos que estão subjacentes à Amostragem.

Já em relação à Curva Normal, não há uma referência explícita nas expectativas de aprendizagem citadas anteriormente. No entanto, o documento apresenta uma orientação metodológica direcionada para o ensino da Curva Normal. Assim, o tópico de conteúdo denominado “Representação de dados” traz que:

A ideia de que os fenômenos na natureza seguem um determinado padrão (assemelhando-se à curva normal, em formato de sino), em que os valores extremos têm menor frequência e os valores mais próximos da média ocorrem em maior número de vezes (maior frequência) deve ser destacada e discutida (Pernambuco, 2013, p.45).

Deste modo, espera-se que o professor de Matemática aborde o conceito da Curva Normal, em sala de aula, contemplando os diversos fenômenos do nosso cotidiano em que esse conceito pode ser aplicado, enfatizando os conceitos estatísticos, como a frequência de dados de forma simétrica em torno da média. Mas em contrapartida, essa orientação não faz referência ao conceito e ao cálculo de probabilidade que podem ser abordados através da Curva Normal. Logo, aqui também cabe ao professor ter domínio necessário sobre o tema para também abordá-los, e, assim, articular a Probabilidade com a Estatística na abordagem do modelo da Curva Normal.

Em linhas gerais, podemos notar que há um avanço do PCPE (Pernambuco, 2013) em relação ao PCNEM (Brasil, 2000,2002) no que diz respeito ao ensino de Estatística e Probabilidade e, mais especificamente a abordagem da Amostragem e da Curva Normal, inerentes a realização de pesquisas e ao tratamento e estudos de fenômenos.

Em continuidade, o Governo Federal, através do Ministério da Educação, elaborou, mais recentemente, a Base Nacional Comum Curricular – BNCC (Brasil, 2018) para a Educação Básica com o objetivo de reformular o currículo e apresentar as competências (gerais e específicas), habilidades e aprendizagens essenciais que todos os estudantes devem desenvolver durante cada etapa da Educação Básica. A BNCC é considerada atualmente, o principal documento norteador e que vem servindo de base para a reformulação e desenvolvimento dos currículos estaduais e dos livros didáticos.

Neste documento, a área de conhecimento denominada “A matemática e suas tecnologias” propõe a estrutura curricular da disciplina de Matemática no Ensino

Médio, em três blocos de conteúdos, a saber: Números e Álgebra, Geometria, e Medidas, e Probabilidade e Estatística. Nesta última etapa da escolarização básica, a BNCC (Brasil, 2018) propõe que o conhecimento matemático, construído no Ensino Fundamental, seja consolidado, ampliado e aprimorado, buscando possibilitar aos estudantes a ampliação do conhecimento, a capacidade de resolver situações-problemas com mais autonomia e segurança e a construção de uma visão de que a Matemática é essencialmente aplicada à realidade em diferentes contextos. Deste modo, o referido documento destaca as competências específicas de Matemática e suas tecnologias a serem desenvolvidas no Ensino Médio:

1. Utilizar estratégias, conceitos e procedimentos matemáticos para interpretar situações em diversos contextos, sejam atividades cotidianas, sejam fatos das Ciências da Natureza e Humanas, das questões socioeconômicas ou tecnológicas, divulgados por diferentes meios, de modo a contribuir para uma formação geral.
2. Propor ou participar de ações para investigar desafios do mundo contemporâneo e tomar decisões éticas e socialmente responsáveis, com base na análise de problemas sociais, como os voltados a situações de saúde, sustentabilidade, das implicações da tecnologia no mundo do trabalho, entre outros, mobilizando e articulando conceitos, procedimentos e linguagens próprios da Matemática.
3. Utilizar estratégias, conceitos, definições e procedimentos matemáticos para interpretar, construir modelos e resolver problemas em diversos contextos, analisando a plausibilidade dos resultados e a adequação das soluções propostas, de modo a construir argumentação consistente.
4. Compreender e utilizar, com flexibilidade e precisão, diferentes registros de representação matemáticos (algébrico, geométrico, estatístico, computacional etc.), na busca de solução e comunicação de resultados de problemas.
5. Investigar e estabelecer conjecturas a respeito de diferentes conceitos e propriedades matemáticas, empregando estratégias e recursos, como observação de padrões, experimentações e diferentes tecnologias, identificando a necessidade, ou não, de uma demonstração cada vez mais formal na validação das referidas conjecturas (Brasil, 2018. p. 533).

Dando prosseguimento à nossa análise, no que diz respeito ao ensino de Estatística e Probabilidade no Ensino Médio, a BNCC (Brasil, 2018) orienta a ampliação, consolidação e o aprofundamento dos conhecimentos e conceitos estatísticos e probabilísticos vistos ao longo dos anos escolares no Ensino Fundamental. Logo, espera-se que o estudante do Ensino Médio desenvolva, no âmbito da Estatística e Probabilidade, as seguintes competências e Habilidades:

(EM13MAT102) Analisar tabelas, gráficos e amostras de pesquisas estatísticas apresentadas em relatórios divulgados por diferentes meios de comunicação, identificando, quando for o caso, inadequações que possam induzir a erros de interpretação, como escalas e amostras não apropriadas.

(EM13MAT202) Planejar e executar pesquisa amostral sobre questões relevantes, usando dados coletados diretamente ou em diferentes fontes, e comunicar os resultados por meio de relatório contendo gráficos e interpretação das medidas de tendência central e das medidas de dispersão (amplitude e desvio padrão), utilizando ou não recursos tecnológicos.

(EM13MAT310) Resolver e elaborar problemas de contagem envolvendo agrupamentos ordenáveis ou não de elementos, por meio dos princípios multiplicativo e aditivo, recorrendo a estratégias diversas, como o diagrama de árvore.

(EM13MAT311) Identificar e descrever o espaço amostral de eventos aleatórios, realizando contagem das possibilidades, para resolver e elaborar problemas que envolvem o cálculo da probabilidade.

(EM13MAT106) Identificar situações da vida cotidiana nas quais seja necessário fazer escolhas levando-se em conta os riscos probabilísticos (usar este ou aquele método contraceptivo, optar por um tratamento médico em detrimento de outro etc.).

(EM13MAT312) Resolver e elaborar problemas que envolvem o cálculo de probabilidade de eventos em experimentos aleatórios sucessivos.

(EM13MAT316) Resolver e elaborar problemas, em diferentes contextos, que envolvem cálculo e interpretação das medidas de tendência central (média, moda, mediana) e das medidas de dispersão (amplitude, variância e desvio padrão).

(EM13MAT406) Construir e interpretar tabelas e gráficos de frequências com base em dados obtidos em pesquisas por amostras estatísticas, incluindo ou não o uso de softwares que inter-relacionem estatística, geometria e álgebra.

(EM13MAT407) Interpretar e comparar conjuntos de dados estatísticos por meio de diferentes diagramas e gráficos (histograma, de caixa (box-plot), de ramos e folhas, entre outros), reconhecendo os mais eficientes para sua análise.

(EM13MAT511) Reconhecer a existência de diferentes tipos de espaços amostrais, discretos ou não, e de eventos, equiprováveis ou não, e investigar implicações no cálculo de probabilidades (Brasil, 2018, p. 546).

Diante disso, percebe-se que com relação aos PNCEM (Brasil, 2000), houve um aumento nas competências e habilidades a serem alcançadas pelos estudantes do Ensino Médio. No que se refere à Amostragem as habilidades “EM13MAT102” e “EM13MAT202” contemplam, respectivamente, a análise crítica de amostras de pesquisas que são divulgadas por diversos meios de comunicação e o planejamento e execução de pesquisas amostrais sobre fatos relevantes do nosso cotidiano abarcando o estudo das medidas de centralidade e dispersão.

Já no que diz respeito à Curva Normal, podemos observar que, de modo semelhante aos PNCEM (Brasil, 2000), esse documento não traz explicitamente alguma orientação para o ensino e abordagem desse tópico. Logo, também cabe ao professor, tendo esse documento como meta, compreender essa orientação de modo implícito em algumas habilidades, como as “EM13MAT202”, “EM13MAT316” e “EM13MAT511” que, em síntese, respectivamente recomendam o trabalho com pesquisas amostrais, comunicando os resultados das pesquisas por meio de gráficos e a interpretação das medidas de centralidade e dispersão; a resolução e elaboração de problemas, em diferentes contextos, que envolvem cálculo e interpretação dessas medidas e o reconhecimento da existência de diferentes tipos de espaços amostrais para o cálculo de probabilidades.

Por último, baseado nas normativas difundidas pela a BNCC (Brasil,2018), o estado de Pernambuco elaborou um novo e mais atual currículo para a Educação Básica. Focando na etapa do Ensino Médio, o currículo (Pernambuco, 2021), traz as mesmas competências específicas de Matemática para essa etapa que são recomendadas pela BNCC, apresentadas anteriormente. Além disso, observamos que as habilidades voltadas para a Estatística e a Probabilidade são bastantes semelhantes as recomendadas na base nacional, a saber:

(EM13MAT202PE15) Realizar pesquisa amostral, utilizando a coleta de dados, de acordo com a realidade da sua região, comunicando os resultados por meio de relatórios, contendo gráficos e interpretação das medidas de tendência central e das medidas de dispersão (amplitude e desvio padrão) com e/ou sem apoio de tecnologias digitais.

(EM13MAT316PE32) Resolver e elaborar situações-problema, em contextos diversos, que envolvam o cálculo e a interpretação das medidas de tendência central (média, moda, mediana) e das medidas de dispersão (amplitude, variância e desvio padrão), com e/ou sem apoio de tecnologias digitais.

(EM13MAT406PE38) Construir e interpretar tabelas e gráficos de frequências com base em dados de pesquisas estatísticas relacionadas ao cotidiano (gravidez na adolescência, sexualidade, entre outros), com e/ou sem apoio de tecnologias digitais que inter-relacionem estatística, geometria e álgebra.

(EM13MAT102PE03) Interpretar e utilizar tabelas e gráficos a partir dos dados neles contidos, construindo argumentos e/ou inferências e identificando possíveis inadequações que induzem ao erro de interpretação, como escalas e amostras não apropriadas. Tabelas e Gráficos: argumentos e/ou inferências, inadequações.

(EM13MAT102PE04) Analisar, criticamente, amostras de pesquisas estatísticas apresentadas em relatórios divulgados por diferentes meios de comunicação.

- (EM13MAT106PE12) Identificar e interpretar situações do cotidiano, envolvendo riscos probabilísticos em que é necessário fazer escolhas como, por exemplo, usar este ou aquele método contraceptivo, optar por um tratamento médico em detrimento de outro, como nos demais campos de conhecimento
- (EM13MAT311PE27) Identificar e descrever o espaço amostral de eventos aleatórios, realizando contagem das possibilidades para resolver e elaborar situações-problema que envolvam o cálculo da probabilidade
- (EM13MAT312PE28) Resolver e elaborar situações-problema que envolvam o cálculo de probabilidade (simples, da união, da interseção, condicional) de eventos em experimentos aleatórios sucessivos.
- (EM13MAT511PE50) Reconhecer a existência de diferentes tipos de espaços amostrais, discretos ou não, e de eventos, equiprováveis ou não, e investigar implicações no cálculo de probabilidades
- (EM13MAT407PE39) Interpretar e comparar conjuntos de dados estatísticos por meio de diferentes diagramas, tabelas e gráficos, como por exemplo, histogramas de caixa (box-plot), de ramos e folhas, entre outros, identificando os mais eficientes para a análise de uma determinada situação problema, com e/ou sem apoio de tecnologias digitais (Pernambuco, 2021, p. 200).

Diante disso, com relação à Amostragem podemos observar nas habilidades “EM13MAT202PE15” e “EM13MAT102PE04” que os estudantes devem, respectivamente, realizar pesquisas amostrais e analisar criticamente amostras de pesquisas estatísticas, propostas que, como já dito, são semelhantes às recomendadas pela BNCC (Brasil, 2018).

No tocante a Curva Normal, de modo também semelhante à BNCC (Brasil, 2018), o currículo de Pernambuco não apresenta explicitamente alguma orientação para o ensino e abordagem desse conceito. No entanto, comprehende-se essa orientação de modo implícito em algumas habilidades, como as “EM13MAT202PE15”, “EM13MAT316PE32” e “EM13MAT511PE50” que, em síntese, orientam para a realização de pesquisas amostrais, com a divulgação dos resultados das pesquisas por meio de gráficos e interpretação das medidas de centralidade e dispersão, a resolução e elaboração de problemas, em diferentes contextos, que envolvem cálculo e interpretação das referidas medidas e o reconhecimento da existência de diferentes tipos de espaços amostrais para o cálculo de probabilidades.

Em síntese, ao observarmos as diretrizes e recomendações propostas pelos quatro documentos oficiais abordados e analisados anteriormente, podemos concluir que os mesmos se constituem como importantes materiais de apoio ao processo de ensino e aprendizagem da Matemática e mais especificamente da Amostragem e da Curva Normal. No entanto, ao analisar tais documentos, cabe destacarmos que os

mesmos apresentam recomendações um tanto gerais, que exigem do professor o reconhecimento dos conceitos que são abarcados pela Amostragem e pela Curva Normal e são tratados de forma implícita por algumas das recomendações, o que pode não garantir a sua abordagem em sala de aula, e, por conseguinte, não favorecer à aprendizagem dos estudantes sobre esse tópico conceitual.

Em relação à Amostragem, por exemplo, há ausência de competências e habilidades que citem explicitamente a representatividade e variabilidade amostral, o tamanho da amostra, margem de erro, técnicas de Amostragem e discussão de possíveis efeitos de vieses na seleção de amostras. Assim, em sala de aula, é importante a abordagem da Amostragem contemplando os conceitos que estão subjacentes a mesma.

Além disso, sobre o conceito da Curva Normal, como vimos, há uma recomendação para o seu ensino unicamente nos PCPE (Pernambuco, 2013). No entanto, essa recomendação só faz referência à distribuição dos dados, a simetria e o conceito de média, presentes na Curva Normal, não enfatizando o conceito e o cálculo de probabilidades que podem ser realizados através desse modelo. Em relação aos PNCEM (Brasil, 2000), e a BNCC (Brasil, 2018), e o PCPE (Pernambuco, 2021) como também destacado anteriormente, neles há orientações que contemplam de forma implícita a abordagem da Curva Normal, o que gera a dependência do domínio conceitual dos professores para reconhecer e abordar esse tema em sala de aula.

Neste cenário, compreendemos que os referidos documentos deixam lacunas sobre temas relativos à Amostragem e à Curva Normal, o que pode acarretar, por parte dos estudantes, na não compreensão da importância e aprendizagem da inter-relação entre a Estatística e Probabilidade por meio da Curva Normal, tendo em vista que os mesmos podem ser observados, comumente, em diversas áreas de conhecimento e fenômenos do nosso cotidiano.

#### 4.2 ALGUNS ESTUDOS ANTECEDENTES SOBRE ENSINO E APRENDIZAGEM DE AMOSTRAGEM

O estudo de Begué, Gea, Batanero e Béltran (2017) teve o objetivo de avaliar a compreensão intuitiva sobre Amostragem de 302 estudantes, sendo 157 do 2º ano (12-13 anos) e 145 do 4º ano (15-16 anos) de duas escolas públicas do Ensino Médio

da Espanha, os quais responderam a uma atividade que os solicitava fornecer quatro valores prováveis para o número de caras ao se jogar uma moeda justa 100 vezes. Com essa atividade, os autores focaram na compreensão intuitiva da relação entre a proporção de uma característica da população e a proporção correspondente nas amostras retiradas, bem como na percepção da variabilidade da proporção nas amostras da mesma população e o efeito do tamanho da amostra em tal variabilidade.

No experimento da atividade, o número de moedas que caem com a face voltada para cima segue uma distribuição binomial, na qual 100 é o tamanho da amostra e 50 é o número esperado dessas referidas faces. No âmbito desse estudo, as autoras consideraram os participantes como tendo um bom entendimento intuitivo da proporção da amostra, quando o valor médio do número esperado de moedas com a face para cima nas quatro amostras fornecidas foi próximo a 50 (valor teórico). As respostas foram consideradas próximas, quando a referida média situou no intervalo [45-55]. Essa faixa contém 68% das observações na distribuição normal, que se aproxima da distribuição binomial considerada, e é formulada pela determinação da média teórica e um desvio padrão para mais e para menos.

A análise estatística das respostas dos estudantes aponta um bom entendimento da relação entre a população e as proporções da amostra, dada a proximidade entre a proporção teórica na população (0,5) e o valor médio da distribuição de todas as respostas dos alunos (0,512). Além disso, observou-se que a variabilidade da Amostragem foi calculada em uma quantidade além do esperado por 47% dos alunos da 2<sup>a</sup> série e 37% dos alunos da 4<sup>a</sup> série. Observou-se também que 21 alunos forneceram quatro valores idênticos, o que denota uma concepção determinística de Amostragem. Evidenciou-se ainda que os resultados nos alunos do 4º ano foram melhores tanto na percepção do valor esperado quanto na percepção da variabilidade. Assim, a maioria dos alunos, principalmente os do quarto ano, atingiu o nível de raciocínio proporcional sobre a amostra. No entanto, menos de um quarto dos alunos conseguiu perceber adequadamente o efeito do tamanho da amostra na variabilidade da mesma e, portanto, não atingiram o nível de raciocínio distribucional esperado sobre a Amostragem. Os autores concluem que esses resultados sugerem a necessidade de um maior aprofundamento no ensino da Amostragem através de uma abordagem com base em materiais de simulações que favoreça alguma experiência sobre a variabilidade aleatória em diferentes amostras da mesma

população, e, consequentemente, o entendimento dos estudantes sobre as propriedades de Amostragem.

Melo (2017), por sua vez, realizou um estudo que objetivou investigar as contribuições de uma sequência didática para o ensino e aprendizagem de Amostragem em 5 turmas do Ensino Médio, sendo duas do 3º ano, duas do 2º ano e uma do 1º ano de uma escola pública do estado de São Paulo. Na primeira etapa, os estudantes realizaram um estudo introdutório sobre alguns conceitos estatísticos básicos, como população, amostra, tipos de variáveis, frequências, medidas de tendência central, medidas de dispersão e tipos de gráficos. Em seguida, através de textos, o autor discutiu com os estudantes sobre a necessidade e o papel da amostra em pesquisas estatísticas, nas quais a população de interesse é muito grande, além do tamanho da amostra e sua influência na realização e representatividade de uma pesquisa amostral. Na última etapa, os estudantes realizaram uma pesquisa estatística sobre temas sociais, proporcionou uma reflexão sobre os resultados observados e uma revisão dos conceitos da Amostragem a medida que o ciclo da pesquisa foi desenvolvido. O autor ressalta a importância de se realizar atividades práticas para o ensino e aprendizagem de Estatística, como o exemplo de pesquisas, porque possibilitam os estudantes refletirem sobre aspectos do seu próprio contexto, e, consequente, melhoram suas percepções do papel da Estatística em nosso cotidiano.

A pesquisa de Begué, Batanero, Gea e Béltran (2020) objetivou analisar o conhecimento sobre compreensão da representatividade e variabilidade amostral de estudantes do ensino secundário e médio espanhol e se esse conhecimento melhora na medida em que os mesmos avançam de ano escolar. Participaram do estudo 536 estudantes, dos quais 157 eram do 2º ano secundário (13–14 anos), 45 do 4º ano secundário (15–16 anos) e 234 do 2º ano do ensino médio (17–18 anos). Esses estudantes responderam a uma atividade na qual é abordado um experimento de lançamento de 100 tachinhas. Os autores constataram que uma parcela dos estudantes avançou com os passar dos anos no entendimento sobre a representatividade amostral. Porém, no que diz respeito a posição final da tachinha ao cair no chão, alguns estudantes apontaram para uma média equiprovável, o que não condiz com a realidade tendo em vista que a tachinha não é simétrica e, portanto, tem mais chance de cair com a ponta para cima do que para baixo. Em acréscimo, os autores apontam como um resultado alarmante, o fato da maioria dos estudantes

apresentarem dificuldades sobre a variabilidade amostral e que esse conhecimento não evoluiu com o avançar dos anos escolares.

Esses resultados evidenciam, na visão dos autores, uma desconexão entre o que é proposto no currículo educacional espanhol com o conhecimento dos estudantes em situações práticas. Com base nisso, Begué *et al* (2020) sugerem que o processo de ensino e aprendizagem de Amostragem seja iniciado desde os primeiros anos do nível secundário espanhol, de modo a propiciar de forma gradual a aprendizagem dos conteúdos relativos a esse tema, contemplando situações que permitam os estudantes realizarem experimentos práticos.

Innabi (2007) conduziu uma investigação envolvendo 360 estudantes do Ensino Médio dos Emirados Árabes Unidos. Os participantes foram solicitados a responder a uma questão que indagava sobre a frequência semanal com que os estudantes da Universidade dos Emirados Árabes iam à biblioteca durante o semestre letivo. Para isso, foram apresentadas cinco opções de resposta, variando em dois fatores: o tamanho da amostra (6 ou 600) e o método de seleção (Amostragem probabilística e não probabilística). Os estudantes, então, deveriam escolher uma opção que considerassem válida seguida de uma justificativa. Os resultados da pesquisa indicaram que apenas 20% dos estudantes levaram em consideração de maneira apropriada ambos os fatores para fundamentar suas decisões. Além disso, muitos tendiam a utilizar sua própria experiência e expectativa como critérios de validação da pesquisa. Se a conclusão estivesse alinhada com suas expectativas, era considerada válida; caso contrário, era julgada como não válida. Diante disso, a autora aponta que no processo de ensino e aprendizagem da Amostragem na Educação Básica, uma técnica que poderia ser usada para mudar os equívocos dos estudantes é confrontar suas concepções com exemplos e situações que os levem a reconhecer o erro em suas crenças e, assim, possam aprofundar seus conhecimentos.

Nesse contexto, o estudo de Jacob e Doerr (2013) teve o objetivo de investigar o conhecimento de 14 estudantes americanos do Ensino Médio sobre amostra e distribuição amostral. Esses destacam que, de modo geral, os estudantes do Ensino Médio investigados demonstraram ter conhecimento sobre a distribuição amostral, incluindo o reconhecimento da distribuição normal dos dados e decréscimo na variabilidade à medida que o tamanho da amostra aumenta, bem como, as probabilidades associadas aos intervalos mais próximos ou distantes da média da distribuição. Porém, a maioria dos alunos apresentaram dificuldades para relacionar

os dados de uma única amostra com a distribuição amostral associada, com base na variabilidade. Diante disso, os autores enfatizam que a realização de atividades que explorem a incerteza, com a oportunidade de os alunos coletarem uma amostra e compará-la com a distribuição amostral, podem ajudá-los a entender como uma única amostra pode ser usada para estimar características de uma população.

O estudo de Reyes (2019) teve o objetivo de investigar as concepções sobre Amostragem de 1241 alunos chilenos, abrangendo turmas do 8º ano (13 - 14 anos) do ensino fundamental e dos 2º e 4º anos do ensino médio (15 - 16 anos; 17 - 18 anos, respectivamente). Para isso, foi aplicado aos estudantes um teste contendo 11 questões que abordavam diversos conceitos relacionados à Amostragem. A pesquisadora observou que os estudantes demonstram compreensão do conceito de amostra e conseguem atribui-lo significado em diferentes contextos. Como também, aplicaram corretamente o raciocínio proporcional, ao determinarem a proporção da população dada a proporção da amostra. Porém, a maioria dos estudantes, de todos os níveis, enfrentou dificuldades ao selecionar uma amostra representativa da população, além de não conseguir identificar vieses em amostras devido às suas crenças, relacionadas as experiências pessoais com a Amostragem. Essas crenças geralmente estão centradas em argumentos que destacam o tamanho da amostra sem considerar a aleatoriedade.

No tocante ao processo de ensino e aprendizagem da Amostragem na Educação Básica, Batanero, Begué e Gea (2018) apresentam uma reflexão sobre os algumas crenças errôneas que devem ser superadas pelos estudantes desde a Educação Básica, a fim de possibilitar o desenvolvimento gradual do conhecimento sobre essa temática, a saber: *Insensibilidade ao tamanho da amostra*, a qual consiste em conceber que uma amostra, por menor que seja, e independentemente de como ela é selecionada, sempre representará a respectiva população. A *falácia do jogador*, por sua vez, está relacionada ao fato de se acreditar que o resultado de um experimento aleatório afetará a probabilidade de eventos futuros. Já o *Viés de equiprobabilidade* consiste em acreditar que os resultados de qualquer fenômeno aleatório, em particular, qualquer valor de uma estatística amostral, são igualmente prováveis. Em continuidade, a *Falácia de composição* está ligada à ideia da transferência de uma propriedade/inferência que é observada em um experimento simples para os experimentos composto, com mais de um evento. Como exemplo, a probabilidade de sair cara no lançamento de uma moeda ser a mesma de sair duas

caras no lançamento de duas moedas. Por fim, na *Ilusão de controle*, o indivíduo acredita, equivocadamente, ser capaz de controlar os resultados de um processo aleatório, ou de uma amostra, sendo comumente visto em jogadores compulsivos que sempre acreditam na concretização das suas previsões.

Diante disso, as autoras enfatizam que no ensino e aprendizagem da Amostragem, se faz necessário a abordagem de dados reais, em sala de aula, levando os estudantes a refletir e a interpretar a realidade. Destacam também que a variabilidade aleatória desempenha um papel importante na estatística inferencial e que os estudantes precisam aprender e identificar os tipos de variabilidade: variabilidade de resultados em um experimento aleatório; variabilidade nos dados; variabilidade em uma variável aleatória e variabilidade nas amostras ou distribuição amostral. Ressaltam ainda que a Distribuição é um conceito estritamente estatístico e que alguns estudantes têm dificuldade com o mesmo como, por exemplo, diferenciar os três tipos de distribuição que são trabalhados na inferência: a distribuição teórica de probabilidade, que modela uma variável aleatória retirada de uma população e que depende de um parâmetro que geralmente é desconhecido; a distribuição do conjunto de dados que constituem uma amostra aleatória simples, da qual diferentes inferências estatísticas são usadas para estimar parâmetros populacionais; a distribuição de uma estatística (como a média amostral) em todas as possíveis amostras do mesmo tamanho e condições (distribuição amostral).

Em acréscimo, enfatizam que a Amostragem implica na realização de métodos de seleção de amostra, com a finalidade de se estimar um parâmetro populacional. Assim, os estudantes devem compreender que a ideia básica na Amostragem é a formulação de estimativas que possam ser generalizadas para a respectiva população. Além disso, a apontam que Probabilidade é fundamental em toda a estatística e torna-se necessário compreender a relação da mesma com a Amostragem através da probabilidade de um evento em amostras diferentes, das ideias de representatividade da amostra, a variabilidade e a maior precisão da estimativa de probabilidades a partir dos tamanhos das amostras. Como última ideia necessária para a compreensão da Amostragem, as autoras destacam o contexto. Nesse sentido, os estudantes além de realizarem cálculos estatísticos devem reconhecer o contexto dos dados e dos resultados obtidos, o que implica em melhor relacionar as características das amostras com as da população.

Ben-Zvi, Bakker e Makar (2015) também discutem sobre a importância dos estudantes da Educação Básica raciocinarem estatisticamente a partir da Amostragem. Os autores justificam que a seleção de amostras representativas e as inferências realizadas por meio delas estão no centro da Estatística. Logo, torna-se necessário que os estudantes compreendam os conceitos que estão subjacentes à Amostragem. A compreensão da variabilidade amostral, ou seja, como as amostras retiradas de uma mesma população variam é importante para se fazer estimativas e decisões. Além disso, essa variabilidade, juntamente com a representatividade devem ser cuidadosamente compreendidas e equilibradas para que os estudantes aprendam sobre a inferência estatística. Os autores enfatizam que o excesso de confiança na representatividade amostral pode levar os estudantes a pensar que uma amostra reflete tudo de sua respectiva população, ao mesmo tempo que o excesso de confiança na variabilidade amostral pode levar os estudantes a pensar que uma amostra não reflete nada sobre a população. Além disso, é necessário um olhar para as distribuições de médias amostrais para muitas amostras extraídas de uma única população, pois permitem o entendimento de como uma amostra se compara às demais, levando-os a determinar se uma amostra reflete muito ou pouco aspectos de sua população. As amostras e variabilidade amostral também se baseiam em ideias básicas de aleatoriedade e chance, ou o estudo da probabilidade, A ideia do centro dos dados amostrais, ou seja, a compreensão da média da distribuição amostral assim como a ideia de modelo (Distribuição Normal como um modelo que se ajusta a distribuições amostrais sob certas condições) também estão subjacentes à Amostragem.

Diante disso, os autores defendem que o processo de ensino e aprendizagem de Amostragem, na Educação Básica, conte com situações que propiciem aos estudantes a seleção de amostras para a realização de pesquisas estatísticas e experimentos que os levem a compreender a representatividade amostral, como também, de simulações que permitam o estudo sobre a distribuição amostral e, consequentemente, a variabilidade das amostras. Os autores finalizam pontuando que tanto materiais concretos, assim como softwares podem ser úteis nesse processo, desde que os professores conheçam seus potenciais e os utilizem em momentos adequados.

A pesquisa de Reyes, Ruz, Portillo e Contreras (2019) analisou o conhecimento sobre o conceito de amostra de 148 estudantes do quarto ano do Ensino Médio

Chileno (último ano da Educação Básica no Chile). Observaram que, embora a maioria apresente domínio sobre o significado da palavra amostra, bem como, a pertinência de se realizar pesquisa estatística amostral, uma parcela considerável deles ainda demonstra não refletir sobre as implicações do tamanho amostral, o que pode levá-los a conceber que independentemente do tamanho, a amostra representa sempre a população a que pertence e, consequente, acarreta em erros nas interpretações em análises estatísticas. Logo, evidencia-se que nessa etapa de escolarização, se faz necessário propiciar o entendimento ao estudante para além do significado da amostra e da pesquisa amostral, contemplando também aspectos sobre o tamanho e variabilidade amostral.

Já a pesquisa de Souza, Silva e Pires (2022) analisou como 12 professores de Matemática raciocinam sobre conceitos da Amostragem a partir de uma mensagem divulgada em redes sociais durante as eleições municipais da cidade de São Paulo. A mensagem abordava os percentuais de votos dos candidatos em diferentes momentos da apuração e o autor embasou-se na Amostragem para questionar o fato de que a proporcionalidade dos votos contabilizados dos candidatos teve uma variação pequena, ou seja, a medida que apuração avançava e os votos válidos aumentavam, os percentuais para cada candidato manteve um padrão, isto é, com índices sempre próximos ao divulgado na primeira parcial da apuração. Estatisticamente, esse fato se justifica pela Lei dos grandes números ao afirmar que a regularidade dos dados amostrais, quando verificados em uma grande quantidade de casos similares, tende a se estabilizar cada vez mais, à medida que aumenta a quantidade de casos observados. A partir da vivência desse fato, tal mensagem foi utilizada de forma inadequada com a intenção de se questionar a legalidade do sistema eleitoral brasileiro.

Os autores observaram que boa parte dos professores raciocinaram de forma equivocada sobre aspectos da Amostragem, aleatoriedade e lei dos grandes números. Eles sustentaram a crença de que a mensagem era falsa e que o conteúdo foi criado pelo próprio autor porque seria muita coincidência ou algo difícil o fato de os candidatos manterem uma estabilidade em seus percentuais de votos à medida que a apuração avançada. Esses resultados evidenciam que os professores participantes, muito provavelmente, não abordam tais aspectos em suas aulas. Por fim, os autores defendem que é necessário um maior investimento na formação inicial dos professores de Matemática, com ênfase para uma Educação Estatística que

contemple processos comunicativos nas redes sociais, com vistas a proporcionar, aos estudantes, um olhar mais crítico e reflexivo para os dados e informações que são divulgados.

Diante do exposto, os estudos aqui apresentados apontam, em linhas gerais, que há lacunas no processo de ensino e aprendizagem de Amostragem na etapa do Ensino Médio, o que indica a necessidade de um maior aprofundamento em pesquisas na área. Em paralelo, comprehende-se que há poucas investigações que propõem um estudo de intervenção. Diante desse cenário, a presente pesquisa visa contribuir com esse processo, ao desenvolver um estudo intervencivo que contempla a análise da contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada.

#### 4.3 ALGUNS ESTUDOS ANTECEDENTES SOBRE O ENSINO E APRENDIZAGEM DA CURVA NORMAL

Tauber (2001) realizou um estudo sobre a Curva Normal com 57 graduandos de diferentes áreas, como Matemática, Ciências Econômicas e Empresariais e Ciências Sociais recém ingressos na Universidade de Granada, na Espanha. Nesse estudo, inicialmente, a autora realizou um diagnóstico sobre o tema com os estudantes e percebeu que os mesmos, apesar serem recém concluintes da escola secundária, equivalente ao Ensino Médio no Brasil, apresentaram inconsistências no conhecimento sobre a Estatística Inferencial, contemplando a Amostragem e a Curva Normal. Na segunda etapa da pesquisa, a autora desenvolveu uma sequência didática na qual, abordou, aos estudantes, os fundamentos teóricos da Curva Normal e utilizou o software Statgraphics para realização de três atividades que envolveram perguntas sobre leitura direta de dados estatísticos, definições, reconhecimento das propriedades, a representação gráfica da Curva Normal e o cálculo de probabilidades.

A partir das respostas fornecidas pelos estudantes às atividades e ao questionário avaliativo, aplicado após a sequência didática, a autora concluiu que eles apresentaram um avanço na compreensão do conceito da Curva Normal e a sua representação, bem como, no cálculo de probabilidades associado à Curva. Além disso, verificou que a utilização do computador como recurso didático serviu de elemento motivador para a compreensão do significado da Curva Normal, uma vez que foi possível enxergar de forma mais dinâmica e exploratória os elementos da

Curva Normal, como as medidas de centralidade e dispersão, e o gerenciamento das distribuições dos dados para a resolução dos problemas relacionados com a Curva.

Diante disso, a autora destaca a importância do ensino da Curva Normal, tendo em vista que muitos fenômenos físicos, biológicos e sociológicos do nosso cotidiano podem ser modelados através da mesma. Por fim, a Tauber (2001) pontua que, para esse ensino, é importante a realização de sequências didáticas que envolvam a utilização lúdica de recursos computacionais que possibilitem o envolvimento dos estudantes com a coleta, organização e análise exploratória e significativa de dados reais, tendo em vista que há uma complexidade do significado e entendimento do conceito da Curva Normal. Assim, o ensino desse conceito não pode ser reduzido apenas à sua definição, mas inclui um sistema interligado de conceitos estatísticos e probabilísticos que os estudantes devem reconhecer e relacioná-los para resolução de problemas envolvendo esse tipo de distribuição.

Por sua vez, Lima (2009) realizou um estudo com 11 alunos da cidade de São Paulo e egressos do terceiro ano do Ensino Médio, no qual desenvolveu uma sequência didática para o ensino da Curva Normal baseada na utilização do software Excel e na resolução de atividades interativas envolvendo análises exploratórias de dados reais. O autor concluiu que, mediante a sequência didática desenvolvida, os estudantes conseguiram compreender e reconhecer a representação gráfica, as propriedades e as características da Curva Normal e foi possível relacionar conteúdos de Estatística com os de Probabilidade a partir da Curva Normal. Por fim, Lima (2009) aponta que o Excel se configurou com um recurso didático que facilitou o estudo sobre a Curva Normal, promovendo uma melhor interação dos alunos com esse objeto matemático.

A investigação conduzida por Filho e Costa (2021) teve o objetivo de analisar as potencialidades de uma sequência didática para a introdução do conceito de Curva Normal em uma turma do terceiro ano do Ensino Médio do Instituto Federal do Estado de Tocantins, localizado na região norte do Brasil. A sequência envolveu a realização de atividades com lançamento de dados, análise de gráficos o estudo do significado das medidas de centralidade e de dispersão presentes na Curva, bem como, os conceitos de chance e probabilidade associados aos valores dos lançamentos e a área sob a Curva. Os autores concluíram que a sequência possibilitou a compreensão do conceito da Curva Normal, como também, dos conceitos estatísticos e probabilísticos que estão presentes nela. Destacam ainda que a utilização de

atividades práticas pode se constituir com um material didático com potencial para que tornar o estudo da Estatística e da Probabilidade mais atrativo ao articular o teórico com processos práticos que possibilitem as reflexões dos conceitos abordados.

Monroy e Herrera (2019), por sua vez, realizaram um estudo com 53 estudantes mexicanos, divididos em dois grupos, que cursaram o Bachillerato, equivalente aos dois últimos anos do Ensino Médio no Brasil, e destinado para quem pretende ingressar na Universidade. Nesse estudo, os autores buscaram investigar o conhecimento desses estudantes a partir da resolução com lápis e papel de um questionário contendo três problemas sobre a Curva Normal e que exigiam o cálculo de probabilidades a partir do método da Curva Normal Padronizada. Essa pesquisa foi realizada após ambos os grupos terem estudado sobre a Curva Normal no âmbito da disciplina Estatística e Probabilidade, porém, com abordagens de ensino distintas. O professor que ensinou o primeiro grupo utilizou como recurso didático apenas o software Fathom. Já o segundo grupo, teve o ensino pautado pelo professor apenas com a utilização de atividades com lápis e papel. Os autores observaram que a maioria dos estudantes soube aplicar corretamente, a nível processual, a fórmula e a tabela da Curva Normal Padronizada para o cálculo da área e consequentemente a probabilidade de intervalos sob a Curva, mas em termos conceituais, os estudantes apresentaram dificuldades em reconhecer o significado e algumas etapas dos procedimentos e a importância da padronização para análises estatísticas em uma distribuição de dados. Além disso, os autores verificaram que, de modo geral, os estudantes do segundo grupo tiveram um melhor desempenho do que os do primeiro. Esse fato, segundo os autores, decorreu porque o professor que ensinou ao primeiro grupo abordou a Curva Normal unicamente a partir do software e não teve a oportunidade de trabalhar a resolução de problemas por meio de atividades impressas para que os estudantes resolvessem problemas aplicando o método da padronização da Curva Normal. Diante disso, os autores concluem que é pertinente a abordagem do conceito da Curva Normal na Educação Básica e que esse ensino tenha um equilíbrio e integração entre a utilização de recursos computacionais e de atividades por escrito que propiciem aos estudantes a compreensão da representação gráfica e dos significados que podem ser observados em situações com a Curva Normal e a capacidade de resolver problemas envolvendo o cálculo de probabilidades associado a intervalos sob a Curva.

O estudo de Herrera, Monroy e Hernandez (2018) investigou os efeitos da aplicação de uma sequência didática, em duas turmas do Ensino Médio do México para o ensino e aprendizagem da Distribuição Normal. Foram utilizados um vídeo sobre probabilidade, o contexto histórico da Curva Normal, o software Fathom e atividades com lápis e papel, com situações problemas para a abordagem das características e propriedades da Curva Normal. Concluíram o processo de ensino e aprendizagem foi proveitoso e que os recursos, tanto o software, como as atividades com lápis e papel, tem suas potencialidades e desafios.

González, Ojeda e Palacios (2018) realizaram uma pesquisa com dois professores de matemática mexicanos com o objetivo de investigar a compreensão sobre a Distribuição Normal desses professores, sendo um atuante no Ensino Médio e o outro no Nível Superior. Ambos docentes tinham experiência com o ensino de Estatística e Probabilidade e, mais especificamente, com o ensino da Curva Normal em seus respectivos níveis de atuação. Os autores constataram que os professores tiveram dificuldades de interpretar a situação problema abordada e, consequentemente, não aplicaram corretamente o método da Curva Normal Padronizada para solucionar a questão. Além disso, ficou evidenciado que as mesmas dificuldades apresentadas pelos professores foram observadas em seus respectivos alunos, na resolução de um problema semelhante. Diante disso, os autores pontuam que esses alunos repetiam e o colocavam em prática aquilo que é ensinado. Ademais, foi constatado que os professores, em sala de aula, replicavam o discurso apresentado pelos livros que eles adotavam como referência e que o ensino da Curva Normal tinha como foco, apenas a operacionalização e tratamento das fórmulas e cálculos. Deste modo, os autores concluem que o ensino da Curva Normal, desde a Educação Básica, também deve incluir a compreensão dos significados desse conceito e dos conceitos estatísticos e probabilísticos abarcados nesse modelo em diversos contextos, para que seja possível a correta interpretação dos fenômenos e problemas que fazem uso da Curva Normal.

Gonçalves (2014), por sua vez, realizou um estudo bibliográfico onde destaca a importância do ensino da Curva Normal, pois ela é essencial para a inferência estatística, mas que ainda é pouco abordada nas aulas de Matemática no Ensino Médio no Brasil. Diante disso, em seu estudo, é apresentada uma proposta de ensino da Curva Normal a ser abordada nessa etapa de ensino, através da resolução de uma situação problema que indagava a massa média de uma população de 40 kg de

laranjas e a estimativa da quantidade de laranjas dessa população. O processo para a resolução desse problema inclui a seleção aleatória de uma amostra de laranjas, a construção da tabela de frequência com a massa das laranjas da amostra, a construção do histograma dessas massas, posteriormente, a verificação de padrões, o cálculo de medidas estatísticas, a aplicação do Modelo da Curva Normal e a construção de intervalos de confiança. Após a seleção da mostra de laranjas, todas as etapas seguintes contaram com a utilização do programa Geogebra para a análise e descrição dos dados.

Por fim, o autor destaca que através de propostas de ensino como essa, o professor tem a oportunidade de abordar, em sala de aula, o conceito da Curva Normal e outros conceitos pertencentes a Estatística Inferencial, tendo em vista que no Brasil, ainda é comum o ensino de Estatística ser restrito apenas a área descritiva. Dessa forma, a proposta de estudo apresentada, com a utilização prática e lúdica do Geogebra, busca propiciar aos alunos do Ensino Médio, o conhecimento relativo ao campo da Inferência Estatística, abarcando a Curva Normal e sua aplicabilidade em situações do nosso cotidiano, abordando os conceitos estatísticos e probabilísticos presentes nesse tipo de distribuição.

A pesquisa de Araújo (2020) investigou uma proposta de ensino para a abordagem articulada entre a Estatística e a Probabilidade por meio da Curva Normal com 12 professores de Matemática do Ensino Médio que atuam na rede pública do ensino do estado de Pernambuco. Para tanto, esse estudo foi dividido em duas etapas. Na primeira, realizou-se uma análise, através de um estudo diagnóstico, dos conhecimentos didáticos/matemáticos iniciais de professores participantes sobre a inter-relação entre a Estatística e a Probabilidade por meio da Curva Normal. Os resultados dessa etapa indicaram que os professores apresentavam lacunas no conhecimento didático-matemático sobre o tema pautado, tanto no que diz respeito ao domínio conceitual, quanto para o seu ensino.

Com base nesse diagnóstico, o autor realizou a segunda etapa, na qual foi desenvolvida uma proposta de ensino voltada para a abordagem articulada entre a Estatística e a Probabilidade por meio da Curva Normal, através de um encontro formativo. Esse encontro foi pautado na realização de uma sistematização teórica e de três atividades sobre o referido tema. A partir das análises de dados das discussões e das respostas da atividades realizadas nessa segunda etapa, os professores conseguiram avançar na construção, ressignificação e ampliação de seus

conhecimentos sobre articulação entre a Estatística e a Probabilidade por meio da Curva Normal, tanto na perspectiva Matemática, como também, na perspectiva didática, Araújo (2020) concluiu que é pertinente o investimento nas formações acadêmicas e em formações continuadas que possibilitem a apropriação e ampliação de conhecimentos didático-matemáticos, por parte dos professores de Matemática do Ensino Médio, principalmente sobre a Inferência Estatística e o modelo da Curva Normal, possibilitando aos estudantes o entendimento da relação entre a Estatística e a Probabilidade, e dos fenômenos e situações do nosso cotidiano que comumente explicitam essa relação.

De modo análogo ao apresentado na seção anterior, esses estudos demonstram, em linhas gerais, que há lacunas no processo de ensino e aprendizagem da Curva Normal na etapa do Ensino Médio, como também, há poucas investigações que propõem um estudo de intervenção que relate a Amostragem com a Curva Normal. Diante disso, acreditamos que o nosso estudo irá contribuir com esse processo ao promover a aprendizagem dessa temática a alunos do 3º ano do Ensino Médio a partir diferentes recursos (Lápis e papel e o software Geogebra).

## 5 MÉTODO

Este capítulo é dedicado à abordagem dos procedimentos metodológicos adotados para o desenvolvimento do nosso estudo. Para tanto, apresentamos os objetivos, o tipo de metodologia para cada etapa da pesquisa e os procedimentos para coleta e análise dos dados.

### 5.1 OBJETIVOS

#### Objetivo Geral

- Investigar o que estudantes do 3º ano do Ensino Médio sabem e pode aprender sobre Amostragem, Curva Normal e suas relações.

#### Objetivos Específicos

- Analisar o que tem sido proposto sobre Amostragem, Curva Normal e suas relações nos materiais para o ensino de Matemática (Livro de Projetos e Livros de Conhecimento) do PNLD 2021.
- Investigar o que foi proposto nas questões da prova de Matemática e suas Tecnologias do ENEM de 2022 sobre Amostragem, Curva Normal e suas relações.
- Identificar os conhecimentos iniciais de estudantes do 3º ano do Ensino Médio sobre Amostragem e Curva Normal e suas relações.
- Analisar a contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada a partir de um processo interventivo.

### 5.2 PROCEDIMENTOS METODOLÓGICOS

#### 5.2.1 Metodologia do estudo 1

No estudo 1, analisamos o que tem sido proposto nos materiais para o ensino de Matemática do Ensino Médio, aprovados pelo Programa Nacional do Livro e do

Material Didático - PNLD (Brasil, 2021), com foco para os conceitos da Amostragem, Curva Normal e suas relações. Esse tipo de estudo se caracteriza, de acordo com Gil (2008), como uma análise documental, a qual se utiliza de fontes primárias, isto é, materiais que ainda não receberam um tratamento analítico e que possibilitam o diagnóstico de dados qualitativos e quantitativos com categorias de análise que podem ser construídas a priori ou a posteriori. No presente estudo, os documentos analisados foram os livros didáticos, nos tipos: Livros de Projetos (integradores e de Vida) e os Livros de Conhecimento de Matemática, classificados, respectivamente, como os objetos 1 e 2 do referido PNLD. As categorias de análise foram formuladas a posteriori, a partir do levantamento das abordagens dos conceitos da Amostragem e da Curva Normal nos livros.

Para alcançar o objetivo supracitado, realizamos uma análise em todos os Livros didáticos aprovados pelo PNLD 2021 com relação ao ensino de Matemática no Ensino Médio, sendo 14 livros de projetos Integradores, 24 livros de projeto de Vida e 20 livros de conhecimento de Matemática, os quais são apresentados em volume único, ou seja, podem ser utilizados em todo o Ensino Médio.

Os livros de projetos Integradores contemplam projetos de trabalho que permitem a integração de diferentes áreas do conhecimento por meio de processos de ensino e de aprendizagem contextualizados que possibilitam os estudantes enxergarem essa integração em suas realidades. Os projetos de vida, por sua vez, incluem projetos voltados para reflexões sobre a perspectiva pessoal e social do estudante, a partir de temáticas sobre o autoconhecimento, lugar no mundo e planejamento para o futuro profissional.

Já os livros de conhecimento analisados são divididos em dois grupos temáticos: obras da área de conhecimento de Matemática e suas tecnologias (compostos por 10 volumes que abordam a temática de Estatística e Probabilidade) e a obra específica de Ciência Humanas sociais aplicadas em diálogo com a Matemática (composto de 10 volumes). Dessa forma, analisamos um total de 58 volumes de livros didáticos.

Realizamos um levantamento, página a página, para identificar as situações que envolviam o conceito de Amostragem e da Curva Normal e classificamos essas situações nos tipos “Explicação”, quando o autor aborda e explica o conteúdo ao leitor e “Resolução”, quando é proposto algum tipo de atividade. Cabe destacar que nas situações de resolução com mais de um item, como por exemplo, questão X com

alternativas a, b, c, essas foram contabilizadas como três situações de resolução, pois poderiam apresentar diferentes classificações entre si.

Por fim, classificamos as situações de Explicação e Resolução em categorias de análise que envolviam o conceito de Amostragem e da Curva Normal.

a) Amostragem:

- Conceito da Amostra: Explícito e Implícito
- Contexto: Dados reais ou fictícios
- Amostra/população estudada: refere-se a pessoas ou não pessoas
- Habilidades exploradas:
  - Seleção e representatividade da amostra,
  - Análise e Reflexão sobre dados de pesquisas amostrais,
  - Realização de Pesquisa Estatística
  - Técnicas de Amostragem,
  - Identificação da Margem de erro,
  - Comparação entre amostras,
  - Cálculo de medidas estatísticas dos dados amostrais,
  - Identificação o tamanho da Amostra,
  - Classificação da Pesquisa como censitária ou amostral,
  - Classificação do tipo de variável
  - Construção de gráfico para a representação dos dados amostrais,
  - Cálculo de probabilidades a partir dos dados amostrais
  - Contagem a partir de dados amostrais.

b) Curva Normal

- Conceito da Curva Normal: Explícito ou Implícito
- Contexto: Dados reais ou fictícios
- Representação: Gráfica, numérica ou algébrica
- Habilidades exploradas:
  - Análise da representação Gráfica,
  - Cálculo ou identificação do Desvio-Padrão,
  - Cálculo ou identificação da Média, Moda e Mediana,
  - Cálculo ou identificação da média e do desvio-padrão,
  - Cálculo da Porcentagem associada a intervalos da Distribuição
  - Identificação da Frequência dos dados

### 5.2.2 Metodologia do Estudo 2

No estudo 2, também do tipo documental (Gil, 2008), analisamos o que foi proposto nas questões da prova de Matemática e suas Tecnologias do Exame Nacional do Ensino Médio, em sua edição de 2022, sobre Amostragem, Curva Normal e suas relações. Para isso, realizamos nosso estudo em duas etapas. Inicialmente, fizemos uma análise na Matriz de Referência do ENEM, a qual foi elaborada em 2009 e reúne os eixos cognitivos, competências e habilidades exigidas aos candidatos que participarem do exame. Nesta etapa, focando na prova de Matemática e suas tecnologias, elencamos quais dessas normativas tem relação com o nosso tema de estudo.

Na sequência, analisamos todas as questões da prova de Matemática e suas tecnologias, destacando as temáticas das questões que pertencem a unidade de Estatística e Probabilidade e, em seguida, as possíveis questões que tem relação com a Amostragem e a Curva Normal, bem como, a forma de abordagem.

### 5.2.3 Metodologia do Estudo 3

No estudo 3, identificamos os conhecimentos iniciais de estudantes do 3º ano do Ensino Médio sobre a temática supracitada. Esse tipo de estudo é definido por como sendo uma pesquisa de Levantamento, a qual de acordo com Shaughnessy, Zechmeister e Zechmeister (2012, p. 151, 152) “envolve selecionar uma amostra (ou amostras) e usar um conjunto predeterminado de questões” com o intuito de “avaliar os pensamentos, opiniões e sentimentos das pessoas”.

Para tanto, participaram do estudo três turmas do 3º ano do Ensino Médio, selecionadas por conveniência de, respectivamente, três escolas públicas da cidade de Nazaré da Mata - PE, totalizando 108 estudantes. Nossa escolha por estudantes desse nível de escolaridade está pautada nas orientações curriculares apresentadas no capítulo 3 deste estudo. Assim, acreditados que esses estudantes devem, ao concluir a Educação Básica, demonstrar competências e habilidades voltadas para a Estatística e a Probabilidade, e mais especificamente, para os fundamentos da Amostragem e da Curva Normal.

Esse estudo foi direcionado para a realização de uma diagnose com vistas a revelar o nível de aprendizagem de turmas do 3º ano do Ensino Médio sobre a referida

temática, como também, validar nosso instrumento de análise. Logo, cabe destacar que os estudantes que participaram desse estudo não foram os mesmos que participaram do processo intervencivo (Estudo 4), no qual, realizaremos com uma outra amostra de estudantes, um pré-teste, a intervenção de ensino e um pós-teste.

Para identificar esse conhecimento, utilizamos, como instrumento de coleta de dados, um questionário composto por 6 questões que juntas contemplaram 16 itens. Os estudantes tiveram o tempo de uma aula, o equivalente a 50 minutos para respondê-lo. Mediante as respostas apresentadas pelos participantes, as categorizamos nos tipos: Resposta Adequada, Resposta Parcialmente Adequada ou Resposta Inadequada de acordo com a habilidade a ser demonstrada em cada questão.

No que diz respeito à atribuição de pontuação aos desempenhos apresentados, nas questões 1º, 2º e 6º “A” foram observados apenas respostas inadequadas e adequadas. Assim, para essas questões, atribuiu-se 0 (zero) ponto quando a resposta se evidenciou inadequada, ou seja, o estudante não demonstrou ter domínio da habilidade e 1 (um) ponto quando a resposta se caracterizou adequada, isto é, o estudante demonstrou ter domínio da habilidade explorada. Já nas questões 3º, 4º, 5º e 6º “B” evidenciou-se os três tipos de respostas destacadas anteriormente. Logo, atribuiu-se 0 (zero) ponto quando a resposta se evidenciou inadequada, 1 (um) ponto quando a resposta se caracterizou parcialmente adequada, isto é, quando o estudante demonstrou ter domínio parcial da habilidade em questão e 2 (dois) pontos quando a resposta se mostrou adequada. Nesse sentido, os 16 itens do questionário totalizaram 23 pontos. A seguir apresentamos as questões e na sequência (Quadro 4), destacamos os critérios de pontuação juntamente com as categorias de análises emergentes das respostas dos estudantes.

## **Questão 1**

1º Para cada uma das pesquisas, a seguir, identifique a população e a amostra correspondente:

Pesquisa A	Qual é a população?	Qual é a amostra?
A empresa de serviço de proteção ao crédito (SPC-Brasil) fez uma pesquisa		

com 600 brasileiros para analisar o perfil do consumidor consciente.		
<b>Pesquisa B</b>	<b>Qual é a população?</b>	<b>Qual é a amostra?</b>
Uma pesquisa realizada pelo Sesc analisou 2024 blogs brasileiros com o intuito de verificar as linguagens artísticas utilizadas para a apresentação dos conteúdos.		
<b>Pesquisa C</b>	<b>Qual é a população?</b>	<b>Qual é a amostra?</b>
14 macacos foram submetidos a uma pesquisa que tinha por finalidade testar a eficácia de uma vacina.		

**Habilidade:** Identificar a população e amostra relativa a cada pesquisa.

## Questão 2

2º A direção de uma escola quer saber os hábitos da utilização de celular dos estudantes com a finalidade de discutir com eles o uso consciente. Para isso, ela fará uma pesquisa estatística e decidiu que não irá entrevistar todos, mas, sim, uma amostra de estudantes da escola.

- a) Quais características dos estudantes a direção poderia utilizar para selecionar uma amostra de modo que os resultados representem toda a escola? Justifique a escolha.
- b) Como a direção poderia selecionar a amostra de estudantes para representar toda a escola?

**Habilidade:** a) Levantar critérios para a representatividade amostral, considerando a variabilidade natural (características dos estudantes). b) Levantar critérios para a seleção de uma amostra representativa dos estudantes.

## Questão 3

3º Os donos de uma escola de idiomas (inglês e espanhol) desejam ampliar os serviços oferecidos. Para isso, querem saber o percentual de seus estudantes que tem interesse em se matricular em aulas de francês. Diante disso, quatro pesquisadores entrevistaram 80 estudantes de um total de 800 da escola de idiomas. Porém, cada pesquisador selecionou uma amostra diferente.

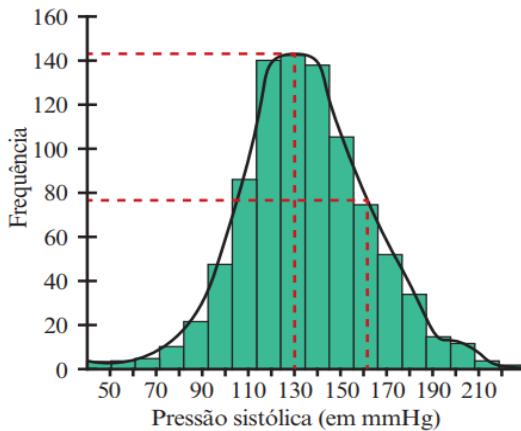
Julgue cada uma das amostras, a seguir, em adequada ou inadequada para representar o resultado da pesquisa, em questão, para toda a escola. Além disso, não esqueça de justificar sua resposta.

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Eduardo entrevistou 80 estudantes que estavam passando no refeitório.		
<b>Amostra 2:</b> Sabendo que a escola tinha o mesmo número de homens e mulheres, Luana entrevistou 80 estudantes, sendo 40 homens e 40 mulheres, escolhendo-os, aleatoriamente, de cada curso de idiomas que a escola possui atualmente (inglês e espanhol).		
<b>Amostra 3:</b> Natália tinha o nome de todos os 800 estudantes, colocou-os em um chapéu e então tirou, aleatoriamente, 80 deles para realizar a entrevista.		
<b>Amostra 4:</b> Renato enviou um questionário para todos os estudantes e depois usou os 80 primeiros que foram devolvidos a ele.		

**Habilidade:** Julgar e justificar a adequação ou não dos métodos de Amostragem.

#### Questão 4

4º) O gráfico, a seguir, representa a distribuição das medidas da pressão arterial de 900 pessoas. Uma pessoa apresentou pressão arterial 80 mmHg. De acordo com o gráfico, essa pessoa apresenta o valor médio da pressão arterial desse grupo? Justifique sua resposta.



**Habilidade:** Reconhecer o gráfico como uma distribuição normal, identificar a média como o centro da Curva e julgar se a medida apresentada é igual à média.

### Questão 5

Durante alguns dias foi medida a temperatura (em graus Celsius) em duas cidades brasileiras às 6h da manhã. Os resultados estão apresentados nas Figuras 1 e 2, a seguir. Sabendo-se que a média de temperatura das duas cidades é 5, qual cidade apresenta uma maior variabilidade das temperaturas e o que ele significa nessa situação? Justifique sua resposta.

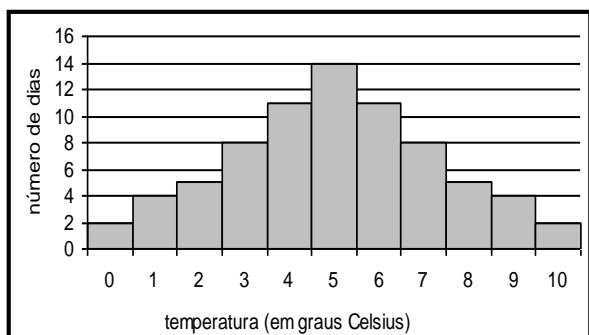


Figura 1 - Medidas de temperatura da cidade A no período de 46 dias.

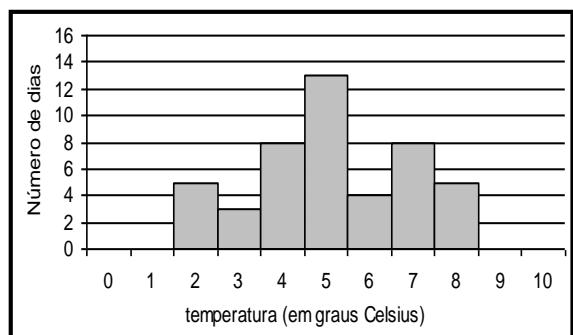
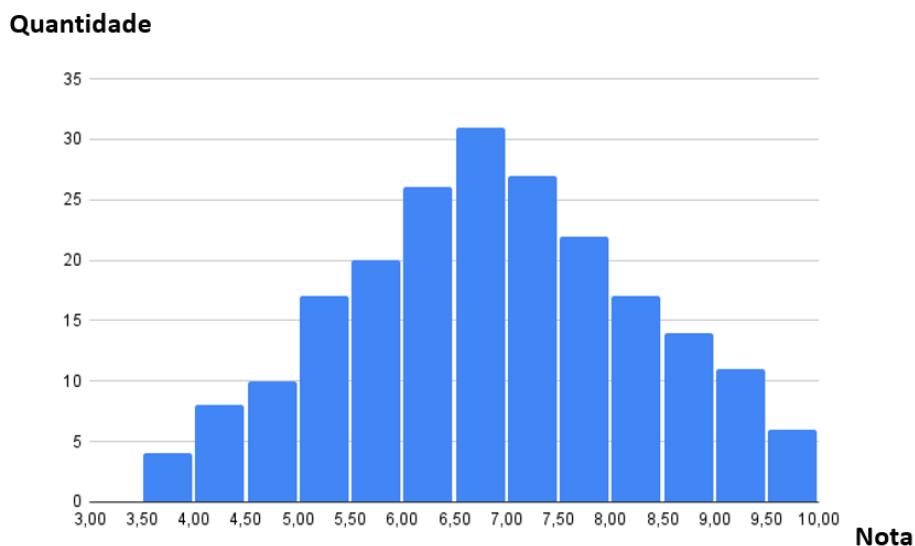


Figura 2 - Medidas de temperatura da cidade B no período de 74 dias

**Habilidade:** Reconhecer através do gráfico, qual distribuição apresenta maior variabilidade.

## Questão 6

O gráfico abaixo apresenta a distribuição das notas que os estudantes do Ensino Médio de uma escola pública de Pernambuco obtiveram na disciplina de Matemática no segundo bimestre de 2022. Essa distribuição é aproximadamente normal e a sua média é 6,7.



- a) Qual a probabilidade aproximada de selecionarmos um estudante dessa escola, ao acaso, e ele possuir a nota de matemática maior ou igual a média?
  
- b) De acordo com o gráfico, se o diretor dessa escola selecionar uma amostra, ao acaso, de 30 estudantes do Ensino Médio, a probabilidade de que a média amostral seja igual a 4,5 é baixa ou alta? Justifique sua resposta.

### Habilidades:

- a) Reconhecer que metade da Curva (nota maior ou igual que a média) tem probabilidade de 50%.
- b) Reconhecer que a média amostral tende a se aproximar da média populacional, e, consequentemente, compreender que a probabilidade solicitada para a nota é baixa por se distanciar da média.

Quadro 4 – Critérios de correção por item do teste diagnóstico

<b>Questão</b>	<b>Habilidades</b>	<b>Pontuações</b>
1 – Itens A <sub>1</sub> , A <sub>2</sub> , B <sub>1</sub> , B <sub>2</sub> , C <sub>1</sub> e C <sub>2</sub>	Identificar a população e amostra relativa a cada pesquisa.	0 pontos: resposta inadequada 1 ponto: resposta adequada
2 – Item A	Levantar critérios para a representatividade amostral, considerando a variabilidade natural (características dos estudantes)	0 pontos: resposta inadequada 1 ponto: adequada
2 – Item B	Levantar critérios para a seleção de uma amostra representativa dos estudantes.	0 pontos: resposta inadequada 1 ponto: adequada
3 – Itens A, B, C, D	Julgar e justificar a adequação ou não dos métodos de Amostragem	0 ponto: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada
4	Reconhecer o gráfico como uma distribuição normal, identificar a média como o centro da Curva e julgar se a pessoa tem pressão igual à média.	0 ponto: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada
5	Reconhecer através do gráfico, qual distribuição apresenta maior variabilidade	0 pontos: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada
6 – Item A	Reconhecer que metade da Curva (nota maior ou igual que a média) tem probabilidade de 50%	0 pontos: resposta inadequada 1 ponto: resposta adequada
6 – Item B	Reconhecer que a média amostral tende a se aproximar da média populacional, e, consequentemente, compreender que a probabilidade solicitada para a nota é baixa por se distanciar da média.	0 ponto: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada

Fonte: O autor (2022).

#### 5.2.4 Metodologia do Estudo 4

O estudo 4 foi desenvolvido com base nos resultados encontrados nos três estudos anteriores. Nele, tivemos o objetivo de analisar a contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada a partir de um processo interventivo. Esse estudo foi realizado em duas turmas de estudantes do 3º ano do Ensino Médio da rede pública de ensino da cidade de Nazaré da Mata. Assim como relatado no estudo 3, nossa

escolha por estudantes desse nível de escolaridade está pautada nas respectivas orientações curriculares apresentadas e discutidas nesta pesquisa.

Metodologicamente, este estudo se classifica como experimental. De acordo com Shaughnessy, Zechmeister e Zechmeister (2012), os experimentos permitem que os pesquisadores verifiquem se um tratamento ou programa altera o comportamento efetivamente dos participantes. Para isso, um experimento deve envolver a manipulação de um ou mais fatores e a medição dos efeitos dessa manipulação sobre o comportamento dos indivíduos. Em nosso estudo, buscamos verificar se a intervenção de ensino realizada mediante aos recursos utilizados (Lápis e Papel e o Software Geogebra) provocou alguma mudança no comportamento, ou seja, na aprendizagem dos estudantes participantes sobre a Amostragem e a Curva Normal.

Nas duas turmas, o estudo 4 foi desenvolvido em três etapas: pré-teste, a intervenção de ensino e o pós-teste. Nessa direção, destacamos que o pré-teste teve como objetivo investigar os conhecimentos prévios dos estudantes sobre o tema do nosso estudo e pós-teste visou analisar as contribuições da proposta de intervenção de ensino. O questionário de ambos os testes foi semelhante ao instrumento aplicado no estudo 3. No entanto, foi necessário realizarmos ajustes em algumas questões a partir das respostas fornecidas pelos estudantes nesse estudo. Mais especificamente, para o pré-teste, e, por conseguinte, o pós-teste, fizemos alterações na terceira e quinta questão do teste diagnóstico do terceiro estudo. Desse modo, esses testes foram compostos por 6 questões que contemplaram 17 itens e totalizaram 25 pontos. Em cada um, os estudantes tiveram o tempo de uma aula, o equivalente a 50 minutos para respondê-lo.

A intervenção de ensino, por sua vez, foi desenvolvida em quatro momentos distintos, um por dia, ao longo do mês de setembro de 2023. Cada encontro teve a duração de uma hora e quarenta minutos, ou seja, duas aulas por dia, totalizando oito aulas em cada turma. Para a realização das atividades durante a intervenção, uma turma teve como recurso o lápis e papel e a outra o software Geogebra. No nosso modelo experimental, uma turma desempenhou o papel de controle da outra, tendo em vista que nosso olhar foi direcionado para as contribuições dos dois recursos empregados. A turma que utilizou o lápis e papel teve o total de 34 estudantes participantes, já a turma que recorreu ao Geogebra envolveu a participação de 38 estudantes. É relevante ressaltar que tanto o pré-teste quanto o pós-teste contaram com a presença de todos os estudantes de ambas as turmas. Já em alguns momentos da intervenção ocorreu ausência pontual de um ou outro estudante nas duas turmas,

no entanto, essa irregularidade na frequência não comprometeu o andamento da nossa intervenção de ensino.

Durante esse processo, os estudantes responderam a atividades em forma de questionários e impressas, as quais foram seguidas de discussões coletivas envolvendo toda a turma e o professor pesquisador sobre os conceitos abordados. Diante disso, frisamos que todos os momentos da intervenção foram conduzidos sob a mediação do professor pesquisador, as atividades impressas foram recolhidas, e as discussões e socializações foram audiogravadas, sendo descritas e analisadas no presente texto.

Para melhor delineamento das etapas do nosso estudo experimental, abordamos nas próximas seções o pré-teste, a descrição dos momentos desenvolvidos na intervenção de ensino e o pós-teste. Cabe salientar que o primeiro momento da intervenção, no qual introduzimos a temática do estudo aos estudantes, foi o mesmo para ambas as turmas. A partir do segundo momento, as atividades foram adaptadas para o tipo de recurso de cada turma, porém, contemplaram os mesmos objetivos de aprendizagem.

#### **5.2.4.1 Pré-Teste**

##### **Questão 1**

1º Para cada uma das pesquisas, a seguir, identifique a população e a amostra correspondente:

<b>Pesquisa A</b>	<b>Qual é a população?</b>	<b>Qual é a amostra?</b>
A empresa de serviço de proteção ao crédito (SPC-Brasil) fez uma pesquisa com 600 brasileiros para analisar o perfil do consumidor consciente.		
<b>Pesquisa B</b>	<b>Qual é a população?</b>	<b>Qual é a amostra?</b>
Uma pesquisa realizada pelo Sesc analisou 2024 blogs brasileiros com o intuito de verificar as linguagens artísticas utilizadas para a apresentação dos conteúdos.		

Pesquisa C	Qual é a população?	Qual é a amostra?
14 macacos foram submetidos a uma pesquisa que tinha por finalidade testar a eficácia de uma vacina.		

**Habilidade:** Identificar a população e amostra relativa a cada pesquisa.

### Questão 2

2º A direção de uma escola quer saber os hábitos da utilização de celular dos estudantes com a finalidade de discutir com eles o uso consciente. Para isso, ela fará uma pesquisa estatística e decidiu que não irá entrevistar todos, mas, sim, uma amostra de estudantes da escola.

- a) Quais características dos estudantes a direção poderia utilizar para selecionar uma amostra de modo que os resultados representem toda a escola? Justifique a escolha.
- b) Como a direção poderia selecionar a amostra de estudantes para representar toda a escola?

**Habilidade:** a) Levantar critérios para a representatividade amostral, considerando a variabilidade natural (características dos estudantes). b) Levantar critérios para a seleção de uma amostra representativa dos estudantes.

### Questão 3

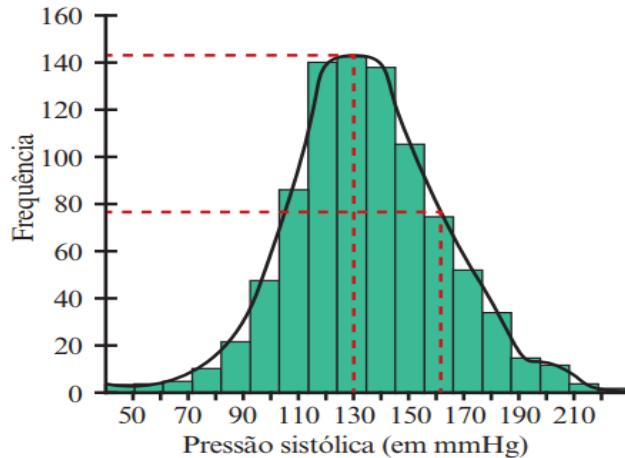
3º Os donos de uma escola de idiomas (inglês e espanhol) desejam ampliar os serviços oferecidos. Para isso, querem saber o percentual de seus estudantes que tem interesse em se matricular em aulas de francês. Diante disso, cinco pesquisadores entrevistaram 80 estudantes de um total de 800 da escola de idiomas. Porém, cada pesquisador selecionou uma amostra diferente. Julgue cada uma das amostras, a seguir, em adequada ou inadequada para representar o resultado da pesquisa em questão. Além disso, não esqueça de justificar sua resposta.

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Eduardo entrevistou 80 estudantes que estavam passando no refeitório.		
<b>Amostra 2:</b> Sabendo que a escola tinha o mesmo número de homens e mulheres, Luana entrevistou 80 estudantes, sendo 40 homens e 40 mulheres, escolhendo-os, aleatoriamente, de cada curso de idiomas que a escola possui atualmente (inglês e espanhol).		
<b>Amostra 3:</b> Natália tinha o nome de todos os 800 estudantes, colocou-os em um chapéu e então tirou, aleatoriamente, 80 deles para realizar a entrevista.		
<b>Amostra 4:</b> Renato enviou um questionário para todos os estudantes e depois usou os 80 primeiros que foram devolvidos a ele.		
<b>Amostra 5:</b> Arthur enumerou e ordenou todos os 800 alunos da escola em uma lista. Em seguida, a cada intervalo de 10 alunos, ele selecionou o sexto, totalizando 80 entrevistados.		

**Habilidade:** Julgar justificar a adequação ou não dos métodos de Amostragem.

#### Questão 4

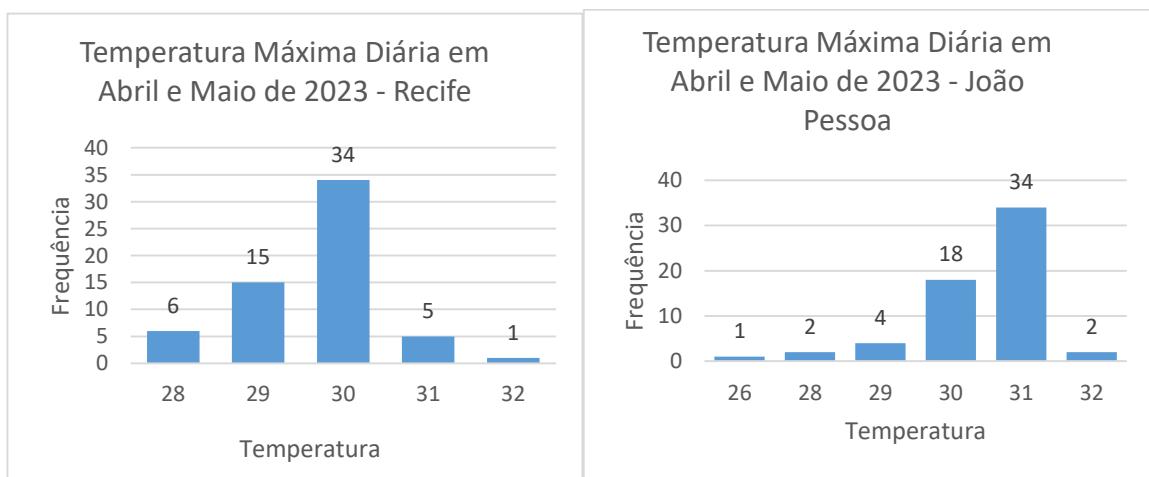
4º) O gráfico, a seguir, representa a distribuição das medidas da pressão arterial de 900 pessoas. Uma dessas pessoas apresentou a pressão arterial 80 mmHg. De acordo com o gráfico, essa pessoa apresenta o valor médio da pressão arterial desse grupo? Justifique sua resposta.



**Habilidade:** Reconhecer o gráfico como uma distribuição normal, identificar a média como o centro da Curva e julgar se a pessoa tem pressão igual à média.

### Questão 5

5º Durante os meses de abril e maio de 2023, foi medida a temperatura máxima diária (em graus Celsius) nas capitais Recife e João Pessoa. Os resultados estão apresentados nos gráficos, a seguir. Sabendo-se que a média das temperaturas registradas nas duas capitais é aproximadamente 30 graus, qual delas apresenta maior variabilidade nas medidas das temperaturas? Justifique sua resposta!

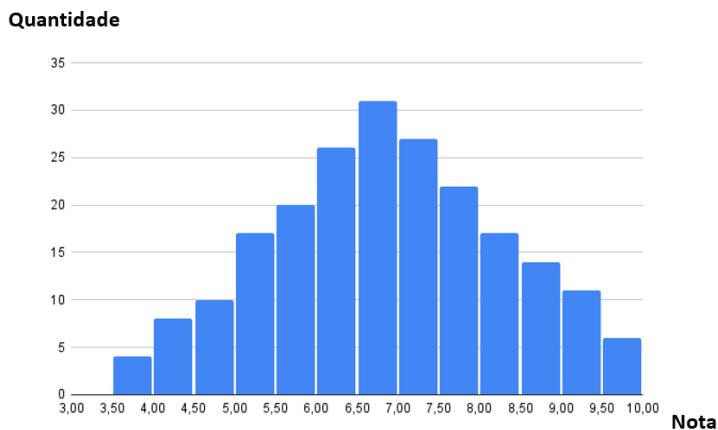


Fonte: AccuWeather.com, 2023.

**Habilidade:** Reconhecer através do gráfico, qual distribuição apresenta maior variabilidade das temperaturas.

## Questão 6

6º O gráfico abaixo apresenta a distribuição das notas que os estudantes do Ensino Médio de uma escola pública de Pernambuco obtiveram na disciplina de Matemática no segundo bimestre de 2022. Essa distribuição é aproximadamente normal e a sua média é 6,7.



- a) Qual a probabilidade aproximada de selecionarmos um estudante dessa escola, ao acaso, e ele possuir a nota de matemática maior ou igual a média?
- b) De acordo com o gráfico, se o diretor dessa escola selecionar uma amostra, ao acaso, de 30 estudantes do Ensino Médio, a probabilidade de que a média amostral seja igual a 4,5 é baixa ou alta? Justifique sua resposta

### Habilidades:

- a) Reconhecer que metade da Curva (nota maior ou igual que a média) tem probabilidade de 50%.
- b) Reconhecer que a média amostral tende a se aproximar da média populacional, e, consequentemente, compreender que a probabilidade solicitada para a nota é baixa por se distanciar da média.

### 5.2.4.2 Intervenção de Ensino

#### 1º Momento: Reflexões iniciais sobre a Amostragem e a Curva Normal

O primeiro momento, como já mencionado anteriormente, foi comum para ambas as turmas e teve o objetivo de propiciar reflexões iniciais sobre os conceitos da Amostragem e da Curva Normal. Nesse momento, inicialmente, o professor

pesquisador promoveu uma breve reflexão sobre a temática explorada no pré-teste e, posteriormente, foi solicitado que os estudantes, em dupla, resolvessem uma atividade para, na sequência, os conceitos serem discutidos. A atividade contemplou a Amostragem a partir de pesquisas em contextos reais que possibilitaram discussões sobre os conceitos de população e amostra, a representatividade da pesquisa amostral relacionada ao tipo de método probabilístico, o tamanho da amostra e a variabilidade dos dados. Além disso, foi abordado o modelo da Curva Normal com ênfase para a sua representação gráfica e as medidas de centralidade e dispersão que estão presentes nesse modelo.

### **Atividade 1**

1º) Para cada item abaixo, identifique a população e amostra relativa a cada pesquisa:

- a) O instituto Ibope Conecta fez uma pesquisa com 2.000 internautas das classes A, B, C e D, de todas as regiões do Brasil para saber o que mais incomoda os internautas brasileiros na internet. (Fonte: De Leonardo, Livro de Conhecimento de Matemática – Editora Moderna, p. 35, 2021).
- b) O Procon do Rio de Janeiro fez uma pesquisa em 344 postos da capital fluminense para verificar os preços dos combustíveis. Fonte: <https://vejario.abril.com.br/cidade/gasolina-fiscalizacao-postos-procon-reducao/>

**Habilidade:** Identificar a população e amostra relativa a cada pesquisa.

Através da primeira questão, procedemos com uma reflexão sobre o conceito de População e Amostra na Estatística. Logo, discutimos que a População remete ao conjunto de todos os elementos que possuem pelo menos uma característica em comum (Triola, 2008) e a amostra se trata de um subconjunto da mesma. Desta forma, ressaltamos a importância de, no campo estatístico, conceber que a população não engloba só pessoas, mas também pode ser formada por animais, objetos, estabelecimentos e etc.

2º O dipirona é um medicamento indicado para dor e febre. Um estudo clínico com 628 crianças verificou a eficácia desse medicamento. Os resultados indicam que, após o uso do dipirona, 82% das crianças tiveram melhora significativa (Wong *et al*, 2001).

- Qual a amostra utilizada nesse estudo?
- Qual a importância do processo de amostragem para a fabricação de medicamentos?
- A partir dos resultados do estudo relatado, você compraria esse medicamento para uma criança com quadro clínico de dor e febre? Por que?

**Habilidade:** Identificar a amostra e refletir sobre a importância de pesquisas estatísticas para a área da saúde.

A segunda questão, por sua vez, abordou uma situação envolvendo processos de pesquisa estatística na área da saúde, e mais especificamente, a fabricação de medicamentos, algo que é comum em nosso cotidiano. Através dessa situação, procedemos com a reflexão sobre a amostra utilizada na pesquisa e a importância da estatística inferencial para as pesquisas na área de saúde, com ênfase para a contribuição da Estatística para a comprovação científica da eficácia de medicamentos.

3º) O instituto Nacional de Ciência e tecnologia em Comunicação (INCT-CPCT) realizou uma pesquisa com 2206 jovens brasileiros para saber o que eles pensam sobre ciência e tecnologia. A pesquisa foi realizada com jovens do sexo masculino e feminino, de 15 a 24 anos e de todas as regiões do Brasil. O nível de confiança é de 95% e a margem de erro de 2%. A Figura, a seguir, apresenta os principais resultados da pesquisa.



Fonte: Souza, Livro de Projetos Integradores – Editora FTD (2021, p. 83).

- a) Qual o tamanho da amostra?
- b) Na sua opinião, por que o pesquisador entrevistou uma amostra e não a população toda?
- c) Na sua opinião, por que as características como o sexo e a região geográfica foram importantes para a seleção da amostra?
- d) O que significa a margem de erro de 2% nessa pesquisa?

**Habilidade:** Identificar o tamanho da amostra, refletir sobre a importância da pesquisa amostral e critérios para a representação amostral, compreender a variabilidade da pesquisa amostral associada ao tamanho da amostra e ao método de Amostragem.

Por meio da terceira questão, aprofundamos nossas reflexões relativas à Amostragem, a partir de discussões sobre a importância e a necessidade da realização de uma pesquisa amostral, em comparação com a pesquisa censitária quando se tem uma população muito grande. Como também, discutimos os critérios de seleção da amostra para que a mesma seja representatividade, como foco para os métodos probabilísticos, a variabilidade, nos tipos natural, amostral e de medida (Reyes, 2019), o nível de confiança, intervalo de confiança e o tamanho da amostra. Nesse momento, o professor pesquisador também utilizou uma apresentação em slide para a abordagem de uma sistematização teórica sobre esses conceitos, tendo em vista que no pré-teste os estudantes apresentaram lacunas significativas sobre esse tema.

4º) Um site de notícia brasileiro lançou uma enquete, apresentada na imagem abaixo, para saber o perfil do consumidor brasileiro. Você considera que esse tipo de pesquisa pode ser representativo (generalizado) para toda população do Brasil? Justifique sua resposta.

A imagem mostra uma interface de enquete com um fundo cinza. No topo, uma caixa de diálogo com fundo branco pergunta: "Você já comprou alguma coisa por impulso?". Abaixo, há dois botões de resposta: "Sim" e "Não", cada um com uma caixinha para marcar. Ambos os botões contêm uma única opção de resposta, com o cursor sobre a opção "Sim".

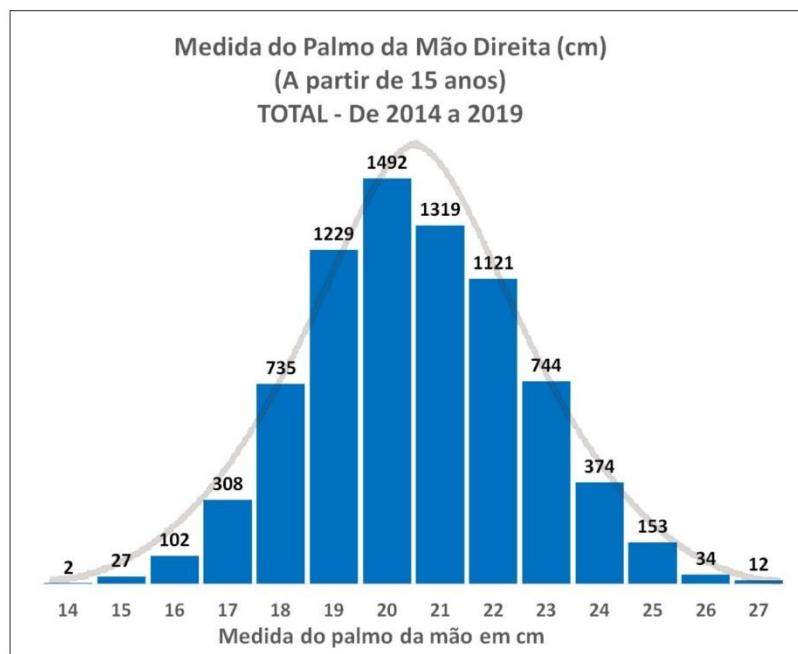
Sim	Não
<input type="radio"/>	<input type="radio"/>

Fonte: Portal G1 (2013).

**Habilidade:** Reconhecer vieses de não representatividade amostral.

A quarta questão, por sua vez, propiciou discussões sobre os métodos de Amostragem não probabilísticos e como eles não refletem os resultados para toda a população. Assim, com o apoio de uma apresentação em slide, procederemos com a abordagem dos métodos por Conveniência, Resposta Voluntária, Julgamento e por Quotas (Triola, 2008), destacando sua importância para determinados tipos de pesquisa de exploração e o caráter não representativo.

5º) A imagem, a seguir, apresenta a distribuição das medidas do palmo da mão direita (cm) de 7652 pessoas, entre homens e mulheres, a partir de 15 anos, que participaram de um estudo. Com base nesses dados responda aos itens dessa questão.



Fonte: CONRE-3, 2020.

- Na sua opinião, essa distribuição de dados, estabelece, aproximadamente, algum tipo de simetria?
- Uma pessoa que tem a medida da palma da mão igual a 22 cm corresponde ao valor da média da palma da mão de todos os participantes dessa pesquisa? Justifique sua resposta.
- Uma fábrica de luvas de borracha analisou esses dados e decidiu fabricar a mesma quantidade de luvas para cada tamanho. A escolha foi correta? Justifique sua resposta.
- Duas pessoas desse grupo foram escolhidas ao acaso, a primeira possui a medida do palmo da mão igual a 17 cm. Já a segunda tem essa medida igual a 22 cm. Qual

das duas possui a medida com maior desvio-padrão em relação à média de todo o grupo?

**Habilidade:** Reconhecer as características do gráfico da distribuição normal, identificar a média como o centro da Curva e o desvio padrão como indicador do distanciamento para a média.

Já com a última questão, introduzimos o estudo sobre modelo da Curva Normal com ênfase para a sua representação gráfica e as medidas de centralidade e dispersão que estão presentes nesse modelo. Considerando que no pré-teste, a maioria dos estudantes de ambas as turmas demonstrou não conhecer esse modelo, recorremos a mesma estratégia na abordagem dos métodos de Amostragem, ou seja, antes da resolução da questão procedemos com uma sistematização teórica sobre Curva Normal, utilizando o gráfico da quarta questão do pré-teste, o qual abordou a distribuição da pressão arterial de um conjunto de pessoas. Dessa forma, buscarmos ressaltar a importância da Curva Normal para análise de dados contínuos na Inferência Estatística pelo fato da mesma modelar de uma gama de fenômenos do nosso cotidiano. Nessa direção, discutimos a ideia de que determinados fenômenos na natureza seguem um determinado padrão, assemelhando-se à Curva Normal, em que os valores extremos têm menor frequência e os valores mais próximos da média ocorrem em maior número de vezes.

## **2º Momento: Etapas do Ciclo Investigativo e Variabilidade Amostral**

No segundo momento da nossa intervenção, vivenciamos algumas etapas do ciclo investigativo (amostra, coleta de dados, registro e representação dos dados e análise e interpretação dos dados) e discutimos aspectos da variabilidade na Amostragem nos tipos natural, amostral e de medida (Reyes, 2019). Para isso, a atividade iniciou com a elaboração de um banco de dados com as alturas de todos os estudantes das turmas, totalizando em uma turma 34 dados e na outra, 38. Cabe destacar que consideramos cada turma como uma população de interesse, a qual reúne todos os estudantes do 3º ano participantes do nosso estudo e que utilizam o lápis e papel ou o Geogebra como recurso, ou seja, esses dados correspondem à quantidade de estudantes de cada turma, como já mencionado neste capítulo. Além disso, a escolha da altura (cm) como variável a ser trabalhada se deu pelo fato de ser

uma variável contínua e a Curva Normal descrever comportamento de variáveis com essa característica. Para a realização dessa atividade, os estudantes foram organizados em duplas e, como já pontuado, as questões foram adaptadas para o tipo de recurso utilizado por cada turma.

## Atividade 2

### Turma Lápis e Papel

1º) Caros estudantes, para darmos início a nossa atividade, realize o seguinte procedimento:

- a) Faça uma pesquisa com seus colegas da classe e colete a altura (em cm) de cada um deles.
- b) Organize os dados em uma lista com o número de identificação dos alunos e suas respectivas alturas.

**Habilidade:** Vivenciar a etapa de coleta de dados de uma pesquisa estatística.

2º) A média da altura dos estudantes de sua classe é igual a \_\_\_\_\_. O desvio-padrão é igual a \_\_\_\_\_. O que essas medidas significam?

**Habilidade:** Refletir sobre o significado da média e do desvio-padrão, enquanto medida de centralidade e dispersão, respectivamente.

Após a realização da coleta de dados, para a resolução da segunda questão o professor pesquisador forneceu a média e o desvio-padrão das medidas das alturas dos alunos participantes da pesquisa. Para tanto, utilizou como ferramenta o software Excel e explicou todo o procedimento aos estudantes com o auxílio do aparelho Datashow para a projeção da tela do notebook. Tal procedimento foi adotado para dar viabilidade ao andamento da intervenção de ensino, tendo em vista a quantidade de dados e que o intuito da questão esteve direcionado para o significado das medidas e não apenas o cálculo em si. Ademais, possibilitou aos estudantes o conhecimento de um recurso que permite o cálculo dessas medidas.

**Habilidade:** Refletir sobre o significado da média e do desvio-padrão, enquanto medida de centralidade e dispersão, respectivamente.

3º) A partir da lista com os dados das alturas de todos os estudantes, selecione, aleatoriamente, uma amostra de estudantes para cada tamanho destacado a seguir, e na sequência, calcule a média das alturas de cada amostra.

Tamanho da Amostra	Média da altura (cm)
Amostra 1: 4 estudantes	
Amostra 2: 8 estudantes	
Amostra 3: 12 estudantes	
Amostra 4: 16 estudantes	

- a) Compare as medidas das suas médias das amostras com as de outra dupla de colegas. Os resultados são iguais ou diferentes?
- b) De acordo com o tamanho da amostra, qual delas tem a média mais próxima da média de populacional (todos os alunos da turma)?
- c) Na sua opinião, para uma amostra representativa há alguma relação entre o tamanho da amostra e a margem de erro?
- d) Se essa pesquisa fosse realizada com toda a escola, quais características da população deveriam ser consideradas para a seleção de uma amostra representativa?

**Habilidade:** Compreender a variabilidade em pesquisas amostrais. Refletir sobre a relação entre o tamanho da amostra e a variabilidade para a representatividade amostral.

Por meio dessa terceira questão e apoiados em Reyes (2019) discutimos os três principais tipos de variabilidade amostral: a variabilidade natural, associada à natureza de cada elemento amostral, a variabilidade amostral, relacionada à variação estatísticas amostrais, e a variabilidade de medição, a qual representa a diferença entre o resultado amostral e o verdadeiro resultado da população, conhecida também como margem de erro. Dessa forma, refletiremos o quanto a variabilidade relacionada também com o tamanho da amostra impacta na representatividade amostral.

## Turma Geogebra

O segundo momento com a turma que teve como recurso o Geogebra foi semelhante ao realizado com a turma anterior. Inicialmente, os estudantes realizaram uma pesquisa e elaboraram um banco de dados com as alturas (cm) de todos os estudantes dessa turma, totalizando 38 dados. Nessa essa etapa, cada dupla de estudantes possuiu um notebook ou tablet com o software Geogebra disponível para utilização. Os dados das alturas coletados foram plotados na ferramenta “Planilha” do Geogebra, a qual permite a organização e listagem de dados em linhas e colunas (Figura 18). A partir disso, as questões da atividade do segundo momento, nessa turma, foram adaptadas para a utilização do Geogebra, a quais foram resolvidas com a orientação e apoio do professor pesquisador.

Figura 18 – Ferramenta Planilha no Geogebra

Fonte: Geogebra (2023).

- 1º) Caros estudantes, para darmos início a nossa atividade, realize o seguinte procedimento:
  - a) Faça uma pesquisa com seus colegas da classe e colete a altura (em cm) de cada um deles.
  - b) Insira os dados na planilha do Geogebra com o número de identificação dos alunos e suas respectivas alturas.

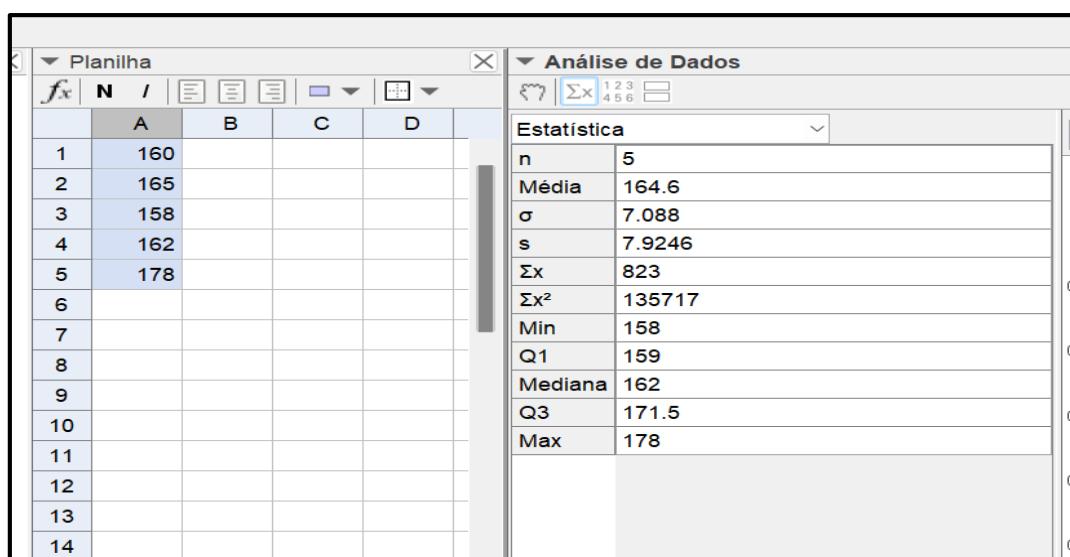
**Habilidade:** Vivenciar a etapa de coleta de dados de uma pesquisa estatística.

2º) Utilizando a ferramenta análise univariada, determine a média, e o desvio-padrão das medidas das alturas dos estudantes da sua turma. O que cada uma dessas medidas significa?

**Habilidade:** Refletir sobre o significado da média e do desvio-padrão, enquanto medida de centralidade e dispersão, respectivamente.

A ferramenta Análise Univariada utilizada nessa questão permitiu a análise estatística dos dados inseridos nas células da planilha. A seguir (Figura 19) exemplifica essa ferramenta com a análise e cálculo de medidas estatísticas de 5 dados.

Figura 19 – Ferramenta de Análise Univariada no Geogebra



Fonte: Elaborado pelo autor a partir do Geogebra (2023).

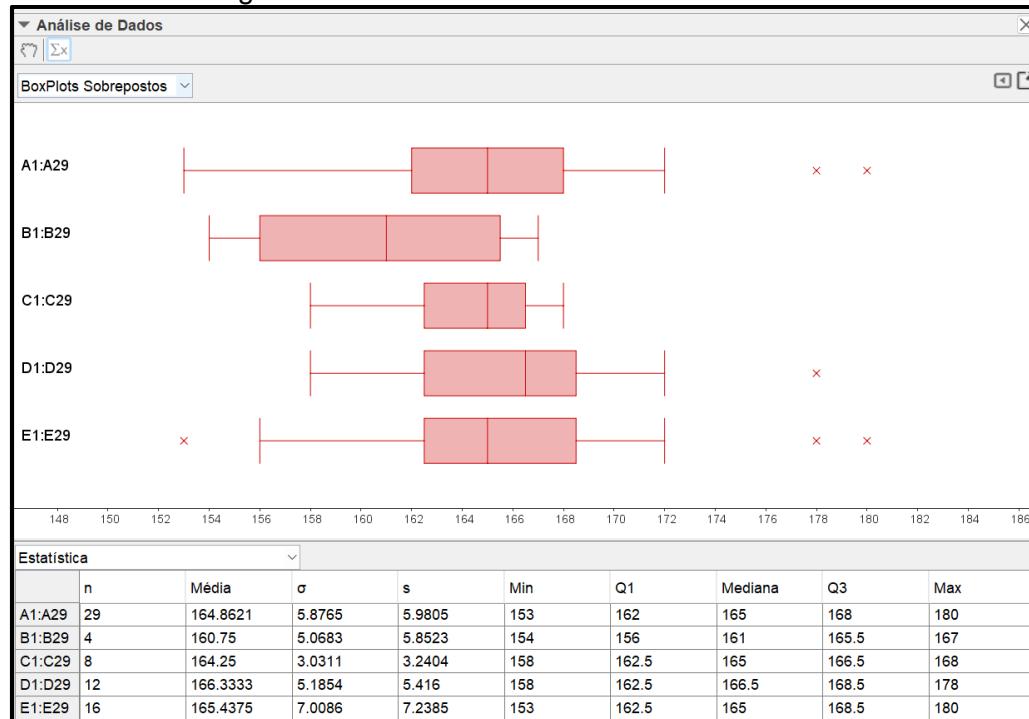
3º) A partir da lista com os dados das alturas de todos os estudantes, selecione, aleatoriamente, no Geogebra, uma amostra de estudantes para cada tamanho destacados a seguir. Na sequência, utilizando a ferramenta análise Multivariada no Geogebra, calcule a média das alturas de cada amostra.

Tamanho da Amostra	Média da altura (cm)
Amostra 1: 4 estudantes	
Amostra 2: 8 estudantes	
Amostra 3: 12 estudantes	
Amostra 4: 16 estudantes	

- a) Compare as medidas das suas médias das amostras com as de outra dupla de colegas. Os resultados são iguais ou diferentes?
- b) De acordo com o tamanho da amostra, qual delas tem a média mais próxima da média de populacional (todos os alunos da turma)?
- c) Na sua opinião, para uma amostra representativa há alguma relação entre o tamanho da amostra a margem de erro?
- d) Se essa pesquisa fosse realizada com toda a escola, quais características da população deveriam ser consideradas para a seleção de uma amostra representativa?

Para a realização da terceira questão, a partir da lista populacional já inserida na planilha do Geogebra, os estudantes selecionaram, aleatoriamente, os elementos das amostras de acordo com cada tamanho e inseriram os dados amostrais em outras colunas na mesma planilha. Em seguida, utilizando a ferramenta de análise Multivariada, os estudantes procederam com a análise das estatísticas das amostras selecionadas de forma simultânea, comparando as estatísticas entre elas e com relação à população. Graficamente, a ferramenta ainda disponibiliza a representação de Boxplots simultâneos para a realização das análises (Figura 20).

Figura 20 – Ferramenta de Análise Multivariada

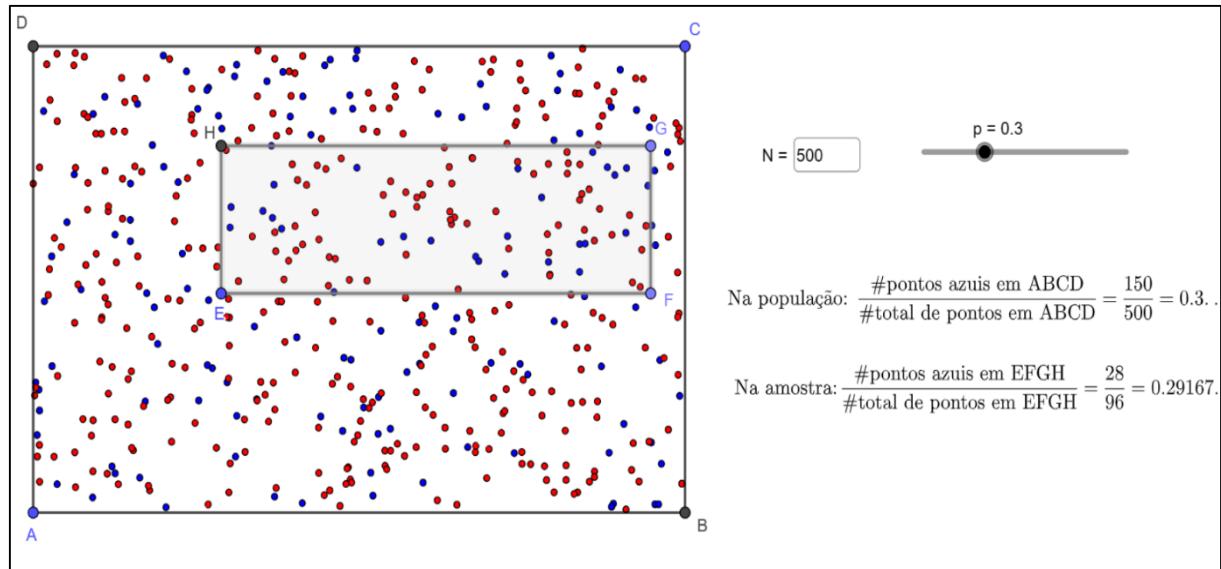


Fonte: Elaborado pelo autor a partir do Geogebra (2023).

Após a realização dessa atividade, os estudantes ainda manipularam, no Geogebra, outra ferramenta didática, desenvolvida por Bortolossi (2020), que ilustra o

efeito de tamanho da amostra para a determinação da proporção de uma característica da população e a proporção correspondente nas amostras retiradas, bem como na percepção da variabilidade da proporção nas amostras da mesma população e o efeito do tamanho da amostra em tal variabilidade. A Figura 21 apresenta um recorte dessa ferramenta.

Figura 21 – O efeito do tamanho da amostra para a determinação da proporção de elementos amostrais



Fonte: Bortolossi (2020). Link: <https://www.geogebra.org/m/mruvtfq>

Nessa ferramenta, há um retângulo ABCD de área maior que representa uma população, cujo os elementos são os pontos azuis e vermelhos. O retângulo EFGH, por sua vez, representa uma amostra dessa população. Ao lado, o controle “N” indica o tamanho populacional e o controle “p” determina a proporção entre os elementos azuis e o total de elementos da população. Consequentemente, é mostrado abaixo o quantitativo de elementos na população que determina o valor de  $p$  e também a quantidade de elementos da amostra que faz a proporção amostral se aproximar do valor de  $p$ , caracterizando a variabilidade. Além disso, é possível manipular o tamanho da área do retângulo EFGH e deslocá-lo no interior do retângulo maior. A partir disso, dentre algumas conclusões, é possível observar que quanto maior a área do retângulo EFGH, ou seja, o tamanho da amostra, mais haverá uma tendência da proporção de pontos azuis na amostra se aproximar da proporção verificada na população.

### **3º Momento: Caracterização da Curva Normal e o cálculo de probabilidades**

O terceiro momento foi reservado para o aprofundamento da representação gráfica do modelo da Curva Normal vista no primeiro momento, como também, para o estudo das medidas estatísticas e o cálculo de probabilidades no referido modelo, através dos dados das alturas coletados na questão anterior. Nessa etapa, foi solicitado que os estudantes, novamente em duplas, resolvessem a terceira atividade.

#### **Atividade 3**

##### **Turma Lápis e Papel**

1º) A partir dos dados coletados na atividade anterior, realize os seguintes procedimentos.

- a) Determine a mediana e a moda das medidas das alturas coletadas.
- b) Esboce um gráfico de histogramas para as medidas das alturas (em cm) coletadas. Localize, no gráfico, a média, moda e mediana das medidas das alturas. Você identifica alguma relação entre elas?

**Habilidade:** Esboçar graficamente o modelo da Curva Normal, identificar as medidas de centralidade como centro da distribuição.

Essa primeira questão teve o objetivo de refletir sobre a relação entre a média, mediana e moda, as quais, no modelo normal, apresentam valores iguais, ou seja, estão no centro da distribuição. Inicialmente, cada dupla procedeu com o cálculo da moda e mediana e junto com a média já fornecida na atividade anterior, elaboraram o esboço do gráfico de barras com as medidas das alturas. Na sequência, refletimos sobre a localização dessas medidas e as características do gráfico da Curva Normal.

2º) Com base no esboço do gráfico realizado na atividade anterior, responda:

- a) Se escolhermos uma pessoa dentre todos os estudantes dessa turma, ao acaso, qual a probabilidade aproximada de que ela possua a altura maior que a média do grupo?

- b) Se o professor escolher um aluno dessa turma, ao acaso, qual a probabilidade aproximada de que o aluno escolhido tenha a altura entre 168 cm e 178 cm?
- c) Qual a probabilidade aproximada de um aluno dessa turma ter a altura entre 158 cm e 188 cm?
- d) Com base nas propriedades teóricas do Modelo Normal, há alguma relação entre as respostas dos itens “b” e “c”?

**Habilidade:** Realizar cálculos probabilísticos associados a intervalos da Curva Normal.

Através do primeiro item, discutimos a probabilidade associada à metade da área da Curva, tendo em vista que a média está localizada no centro da mesma. Assim, refletimos, enquanto propriedade da Curva Normal, que a probabilidade desejada é de 50%. Para realização dos cálculos probabilísticos dos itens b) e c) utilizamos a média e o desvio-padrão que foram fornecidos na atividade anterior. Antes da resolução, o professor pesquisador apresentou uma sistematização teórica aos estudantes sobre o método da distribuição padronizada que está descrita no capítulo 1 deste estudo, o qual possibilitou o cálculo das probabilidades solicitadas nesses dois itens. Em seguida, na alternativa “d”, de forma mais intuitiva, os estudantes procederam com a observação de que os intervalos de valores dos itens b e c são simétricos, em relação à média. Consequentemente, as probabilidades solicitadas são iguais. Diante disso, abordamos, mais uma vez a propriedade da simetria da Curva Normal, associando o cálculo de probabilidades a área de intervalo sob a Curva Normal.

### Turma Geogebra

O terceiro momento, assim como o realizado com a outra turma, destinou-se ao aprofundamento da representação gráfica do modelo da Curva Normal, abarcando o estudo das medidas estatísticas e o cálculo de probabilidades no referido modelo, com base nos dados das alturas coletados na questão anterior. Para essa turma, recorremos a ferramenta de análise univariada, a qual permite a construção da representação gráfica e a ferramenta Calculadora de Probabilidades, que, por sua vez, possibilita o cálculo de probabilidades associado a intervalos sob a Curva Normal

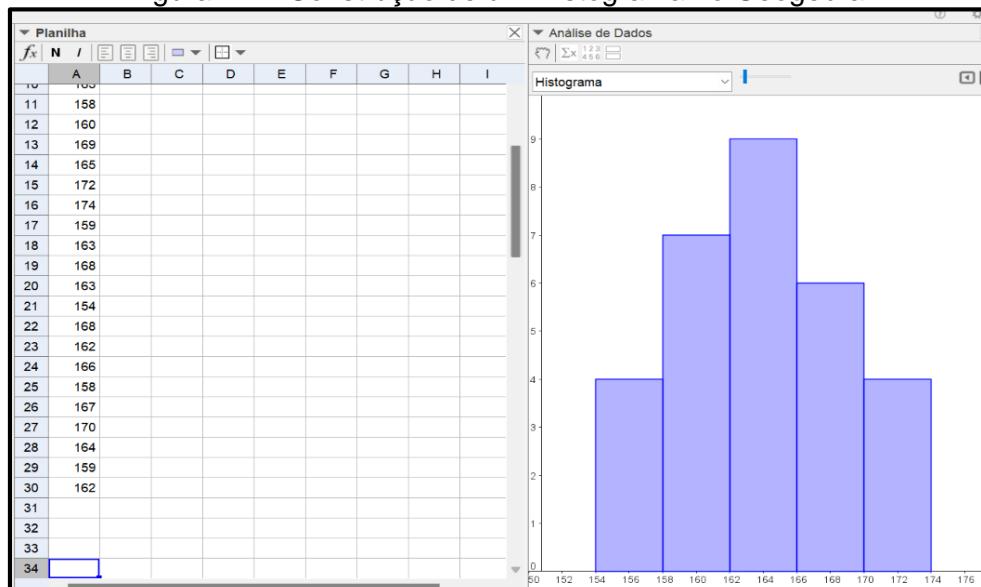
1º) A partir dos dados coletados na atividade anterior, realize os seguintes procedimentos:

- Com a ferramenta análise univariada e o comando “moda”, determine, respectivamente, a mediana e a moda das medidas das alturas coletadas.
- Utilizando a ferramenta de análise univariada no Geogebra, faça a construção de um histograma das medidas das alturas (em cm) coletadas. Localize, no gráfico, a média, moda e mediana das medidas das alturas. Você identifica alguma relação entre elas?

**Habilidade:** Esboçar graficamente o modelo da Curva Normal, identificar as medidas de centralidade como centro da distribuição.

Como já visto, essa primeira atividade tem o objetivo de refletir sobre a relação entre a média, mediana e moda, as quais, no modelo normal, apresentam valores iguais, ou seja, estão no centro da distribuição. Para isso, inicialmente, cada dupla procedeu com o cálculo mediana e moda utilizando ferramentas do Geogebra. Posteriormente, realizaram a construção, no Geogebra, do gráfico de histograma com as medidas das alturas. Na sequência, refletimos sobre a localização dessas medidas no gráfico, en quanto propriedade da Curva Normal. A seguir, exemplificamos (Figura 22) a construção de  $n = 30$ , na qual a representação gráfica se aproxima do modelo normal.

Figura 22 – Construção de um Histograma no Geogebra



Fonte: Elaborado pelo autor a partir do Geogebra (2023).

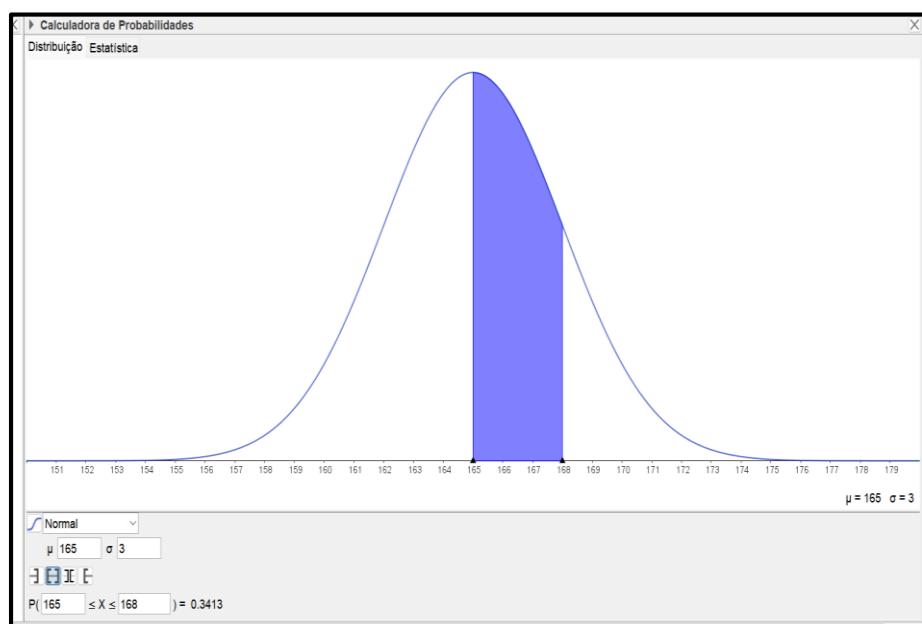
2º Com base na construção do gráfico realizado na atividade anterior, responda:

- Se escolhermos uma pessoa dentre todos os estudantes dessa turma, ao acaso, qual a probabilidade de que ela possua a altura maior que a média do grupo?
- Se o professor escolher um aluno dessa turma, ao acaso, qual a probabilidade que o aluno escolhido tenha a altura entre 169 cm e 180 cm
- Qual a probabilidade de um aluno dessa turma ter a altura entre 158 cm e 169 cm?
- Com base nas propriedades teóricas do Modelo Normal, há alguma relação entre as respostas dos itens “b” e “c”?

**Habilidade:** Realizar cálculos probabilísticos associados a intervalos da Curva Normal

Para a realização desses cálculos de probabilidades, recorremos a ferramenta “Calculadora de Probabilidades” do Geogebra, a qual possibilita, através dos parâmetros média e desvio-padrão dos dados, calcular a área de intervalos da Curva Normal, o que, como já vimos, corresponde a probabilidade de a variável em questão assumir valor em um determinado intervalo de interesse. Assim como pontuado para a turma com o recurso de lápis e papel, a realização desses cálculos se deu a partir da média e o desvio-padrão que foram fornecidos na atividade anterior. A Figura 23 retrata um exemplo do cálculo de probabilidades entre intervalos a partir da ferramenta “calculadora de probabilidades” no Geogebra.

Figura 23 – Cálculo de Probabilidade na Curva Normal



Fonte: Elaborado pelo autor a partir do Geogebra (2023).

## **4º Momento: Teorema Central do Limite: Relacionando a Amostragem com a Curva Normal**

O quarto e último momento foi destinado para o estudo do Teorema Central do Limite, com ênfase para a relação existente nele, entre a Amostragem e o modelo da Curva Normal. Para tanto, a quarta atividade, em ambas as turmas, foi desenvolvida em 4 etapas.

## Atividade 4

## Turma Lápis e Papel

1º Para a realização dessa atividade, cada trio deve seguir estes procedimentos:

- a) Etapa 1: Cada integrante do trio deve fornecer a sua altura. O trio formado se constitui uma população. Em seguida, deve-se calcular a média populacional, a qual será dada pela média aritmética simples entre os três valores da altura.
  
  - b) Etapa 2: Extração das Amostras e Obtenção das Médias Amostrais.

amostras possíveis de tamanho 2 da população. Em seguida, realizem o cálculo das médias de cada amostra, de modo que a quantidade de médias amostrais seja igual à quantidade de amostras selecionadas. O Quadro a seguir deve ser preenchido com os elementos de cada amostra e suas respectivas médias.

c) Etapa 3: Cálculo da Média das Médias amostrais.

Nessa etapa, o trio deve realizar cálculo da média das médias amostrais, as quais foram obtidas no item anterior, sendo determinada pela média aritmética simples entre as nove médias amostrais. Há alguma relação entre a média encontrada nessa etapa 3 com a média populacional obtida na etapa 1?

d) Etapa 4: Construção de um histograma para as nove médias encontradas na Etapa 3.

Construa um gráfico de histograma para as nove médias encontradas na Etapa 3. Há alguma relação entre o gráfico construído e o modelo da Curva Normal?

**Habilidade:** Reconhecer a relação entre a Amostragem e o modelo da Curva Normal através do Teorema Central do Limite.

Na primeira etapa, os estudantes foram divididos em trios, nos quais cada integrante forneceu a sua altura, em cm. Nesse momento, cada trio formou uma população de interesse. Por exemplo, poderemos ter um trio com os números [155, 160 e 165]. Em seguida, cada trio procedeu com o cálculo da média populacional, a qual é dada pela média aritmética simples entre os três valores da altura. No nosso exemplo, a média populacional corresponde a:  $155 + 160 + 165/3 = 160$ .

Dando prosseguimento, na segunda etapa, os estudantes selecionaram, com reposicionamento de elementos, todas as amostras possíveis de tamanho ( $n=2$ ) da população. Seguindo o nosso exemplo, os pares possíveis são: (155,155); (155,160); (155, 165); (160,160); (160,155); (160,165); (165,165); (165,155) e (165,160). Na sequência, os estudantes procederam com o cálculo das médias de cada amostra, de modo que a quantidade de médias amostrais foi igual à quantidade de amostras selecionadas. O Quadro 5 apresenta a média para cada uma dessas amostras.

Quadro 5 – Média das nove amostras

Amostra	Média
(155,155)	155
(155,160)	157,5
(155, 165)	160
(160,160)	160

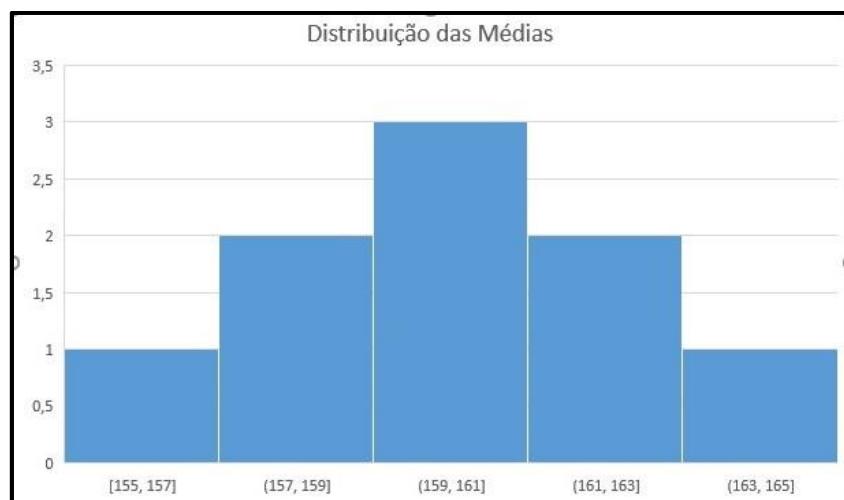
(160,155)	157,5
(160,165)	162,5
(165,165)	165
(165,155)	160
(165,160)	162,5

Fonte: O autor (2023).

Já na terceira etapa, os estudantes realizaram o cálculo da média das médias amostrais, as quais foram obtidas no item anterior. Com o nosso exemplo, a média desejada é igual a  $\frac{155+157,5+160+160+157,5+162,5+165+160+162,5}{9} = 160$ . Em seguida, direcionados a comparar o valor obtido com o resultado da média populacional, obtida na Etapa 1.

Como podemos observar, a média calculada nesta etapa tem valor igual a média da população, ou seja, 160. Nesse momento, o professor pesquisador discutiu com os estudantes uma das conclusões mais importantes do referido teorema que é a essência da estatística inferencial: a distribuição amostral das médias tem média igual à da população. Além disso, como abordado no capítulo 1 do nosso estudo, esse teorema também conclui que a distribuição das médias amostrais apresentará uma normalidade. Para verificar isso, os estudantes procederam com a realização da quarta etapa, na qual construíram um histograma para representar os diferentes valores assumidos pela média amostral calculadas na segunda etapa. Como exemplo, destacamos as médias amostrais que abordamos no exemplo

Figura 24 – Gráfico da distribuição das médias Amostrais



Fonte: O autor (2023).

Após a construção do gráfico, foi discutido que a distribuição das médias amostrais também obedece a uma Curva Normal. Logo, há uma concentração de valores próximos à média amostral que, conforme visto na etapa anterior, é igual à média populacional. Prosseguimos discutindo que, de acordo com o Teorema Central do Limite, independentemente da distribuição de probabilidade associada a um fenômeno de interesse, a média das amostras possui uma distribuição aproximadamente normal, com média igual à média populacional ( $\mu$ ). Isso significa que, para tamanhos de amostra adequados, a distribuição das médias amostrais se aproxima de uma distribuição normal, o que nos permite utilizar as propriedades da distribuição normal na inferência estatística. Podemos, portanto, calcular probabilidades relacionadas aos valores assumidos pela média amostral em uma determinada amostra usando a área sob a Curva Normal. Finalizamos o debate ressaltando a importância dessas conclusões, uma vez que fornecem uma base sólida para a inferência sobre a população com base em amostras, permitindo a construção de intervalos de confiança com precisão estatística.

### **Turma Geogebra**

De modo semelhante a outra turma, o quarto e último momento foi destinado para o estudo do Teorema Central do Limite, destacando a relação, entre a Amostragem e o modelo da Curva Normal. A atividade desenvolvida seguiu as mesmas etapas da outra turma, tendo como recurso o Geogebra.

1º) Para a realização dessa atividade, cada trio deve seguir estes procedimentos:

a) Etapa 1: Cada integrante do trio deve fornecer a sua altura e inseri-la na tabela do Geogebra. Em seguida, utilizando a ferramenta de análise univariada, deve-se calcular a média da altura desse trio, o qual constitui uma população.

b) Etapa 2: Extração das Amostras e Obtenção das Médias Amostrais.

Nessa etapa, o trio deve selecionar, com reposicionamento de elementos, todas as amostras possíveis de tamanho 2 da população e colocar os dados de cada amostra na tabela do Geogebra. Em seguida, devem realizar o cálculo das médias de cada amostra, utilizando a ferramenta análise multivariada, de modo que a quantidade de médias amostrais seja igual à quantidade de amostras selecionadas.

c) Etapa 3: Cálculo da Média das Médias amostrais.

Nessa etapa, o trio deve colocar em uma coluna da tabela do Geogebra as nove médias obtidas, em seguida, deve-se realizar cálculo da média desses nove valores, utilizando a ferramenta análise univariada.

Há alguma relação entre a média encontrada nessa etapa 3 com a média populacional obtida na etapa 1?

d) Etapa 4: Construção de um histograma para as novas médias encontradas na Etapa 3

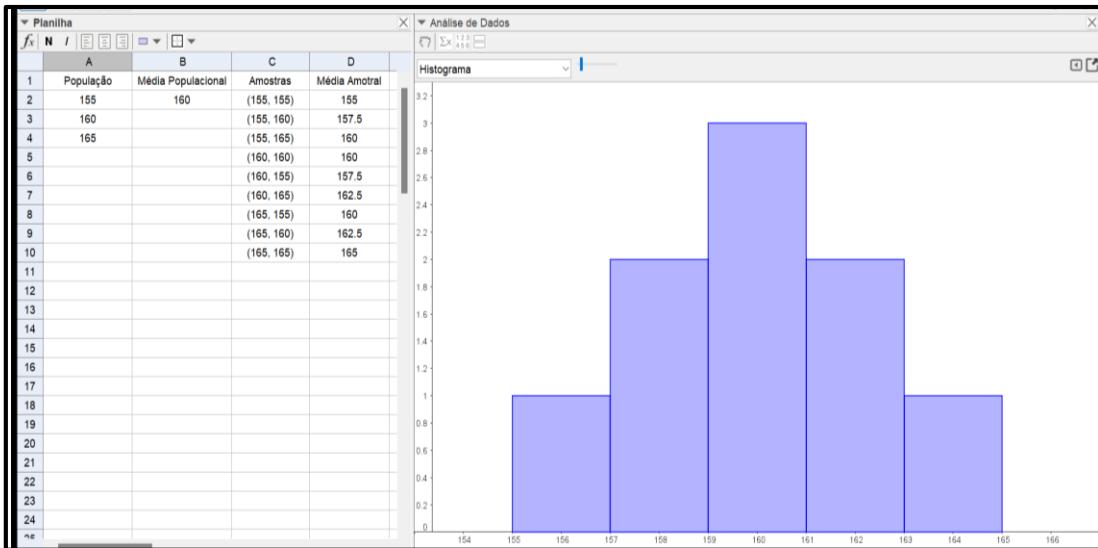
Utilizando a ferramenta análise univariada com os dados da etapa anterior, o trio deve selecionar o Histograma para a exibição gráfica das novas médias. Há alguma relação entre o gráfico construído e o modelo da Curva Normal?

No desenvolvimento da atividade na turma Geogebra, os estudantes, inicialmente, foram divididos em trios, nos quais cada integrante forneceu a sua altura, em cm. Nesse caso, cada trio também formou uma população de interesse. Em seguida, os estudantes inseriram na planilha do Geogebra os valores das unidades populacionais. Na sequência, calcularam a média populacional, utilizando a ferramenta de análise univariada no Geogebra.

A segunda etapa foi direcionada para a extração das amostras e obtenção das médias amostrais. Para tanto, os estudantes selecionaram, com reposicionamento de elementos, todas as amostras possíveis de tamanho ( $n = 2$ ) da população. Em seguida, inseriram essas amostras na planilha do Geogebra para a realização da terceira etapa, a qual se ocupou do cálculo da média das médias amostrais. Nessa direção, com a utilização da ferramenta de análise univariada, os estudantes calcularam a média das médias amostrais e compararam o seu valor com a média populacional. Assim, como foi realizado na outra turma, o professor pesquisador discutiu uma das conclusões mais importantes do referido teorema que é a essência da estatística inferencial: a distribuição amostral das médias tem média igual à da população.

A quarta e última etapa foi reservada para a elaboração do gráfico para as médias amostrais. Logo, os estudantes construíram um gráfico de histogramas para representar os diferentes valores assumidos pela média amostral calculadas no item anterior. Como exemplo, destacamos (Figura 25) todo esse processo adotando como uma população o trio [155, 160 e 165], o mesmo exemplo utilizado para a outra turma.

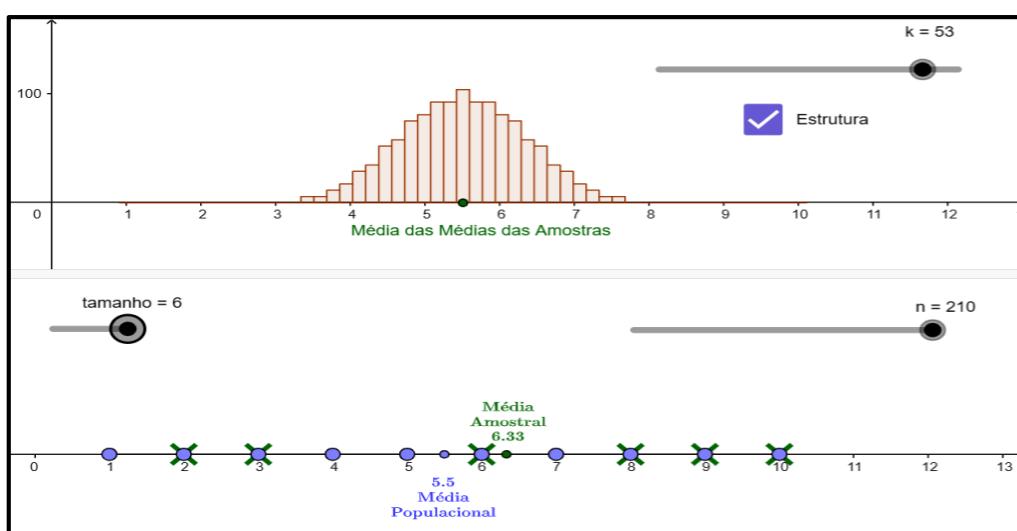
Figura 25 – Gráfico da distribuição das médias Amostrais no Geogebra



Fonte: Elaborado pelo autor a partir do Geogebra (2023).

Após a construção do gráfico, procedemos com a mesma discussão realizada com a outra turma, apresentada anteriormente. Assim, ressaltamos as conclusões do Teorema Central do Limite, com ênfase para a relação entre a distribuição das médias amostrais e a média populacional, e a caracterização do gráfico da Curva Normal. Por fim, os estudantes ainda manipularam, no Geogebra, outra ferramenta didática, desenvolvida por Bortolossi (2015) que ilustra as consequências teóricas desse teorema. A Figura 26 apresenta um recorte dessa ferramenta.

Figura 26 – Ferramenta didática no Geogebra sobre Amostragem e Curva Normal



Fonte: Bortolossi (2015). (<https://www.geogebra.org/m/PmxGbN4A>)

Como podemos observar, na parte inferior da construção, a população é constituída por 10 números no intervalo [0, 10] (círculos azuis). O valor da média populacional também é exibido. Além disso, o controle deslizante “tamanho” manipula o tamanho da amostra em 4, 5 ou 6. Já controle deslizante “n” possibilita selecionar as diferentes amostras possíveis com o tamanho determinado. Na parte superior, o histograma apresenta a média das médias amostrais. O controle “k”, por sua vez, manipula o número de classes no histograma. Através dessas manipulações, os estudantes puderam concluir que quanto maior o tamanho da amostra, menor a amplitude gráfica da distribuição das médias amostrais, ou seja, menor é a variabilidade de medida e, consequentemente, mais concentrada a média amostral está em torno da média de todas as médias amostrais, que, como já explorado anteriormente, coincide com a média populacional. Com base nisso e finalizando a nossa intervenção, refletimos que quanto maior o tamanho da amostra, menor é a variabilidade de medida, ou seja, a amostra selecionada terá uma probabilidade maior de que sua média amostral esteja mais próxima da média populacional. E além disso, o histograma construído, obedece, graficamente, ao modelo da Curva Normal.

#### **5.2.4.3 Pós-teste**

Apresentamos nessa seção as questões do pós-teste, o qual foi semelhante ao pré-teste e compreende a última etapa do nosso modelo experimental. O mesmo teve o intuito de avaliar a aprendizagem dos estudantes após a realização da intervenção de ensino.

##### **Questão 1**

1º Para cada uma das pesquisas, a seguir, identifique a população e a amostra correspondente:

<b>Pesquisa A</b>	<b>Qual é a população?</b>	<b>Qual é a amostra?</b>
Para descobrir o nível de satisfação dos usuários de transporte público do Recife, entrevistou-se 2.400 pessoas que utilizam esse meio de locomoção.		
<b>Pesquisa B</b>	<b>Qual é a população?</b>	<b>Qual é a amostra?</b>

**Habilidade:** Identificar a população e amostra relativa a cada pesquisa.

Foram testados 8 modelos de carros populares para avaliar a segurança desses veículos.		
<b>Pesquisa C</b>	<b>Qual é a população?</b>	<b>Qual é a amostra?</b>
77 gatos participaram de uma pesquisa que tinha por objetivo saber se eles reconheciam o próprio nome mesmo quando chamados por um estranho.		

### Questão 2

2º Uma professora de Artes quer conhecer os hábitos musicais dos estudantes da escola com a finalidade de discutir, com eles, a cultura popular. Para isso, ela fará uma pesquisa estatística. Decidiu que não irá entrevistar todos, mas uma amostra de estudantes da escola.

- a) Quais características dos estudantes a professora de Artes poderia utilizar para selecionar uma amostra de modo que os resultados representem toda a escola? Justifique a escolha.
- b) Como a professora de Artes poderia selecionar a amostra de estudantes para representar toda a escola?

#### Habilidades:

- a) Levantar critérios para a representatividade amostral, considerando a variabilidade natural (características dos estudantes)
- b) Levantar critérios para a seleção de uma amostra representativa dos estudantes

### Questão 3

3º Uma empresa quer saber a satisfação de seus funcionários quanto aos equipamentos disponíveis, as instalações, a carga horária de trabalho, os benefícios oferecidos e etc. Para isso, cinco pesquisadores entrevistaram 60 funcionários, de um total de 600. Porém, cada pesquisador selecionou uma amostra diferente. A seguir, julgue cada uma das amostras em adequada ou inadequada para representar o resultado da pesquisa em questão. Além disso, não esqueça de justificar sua resposta.

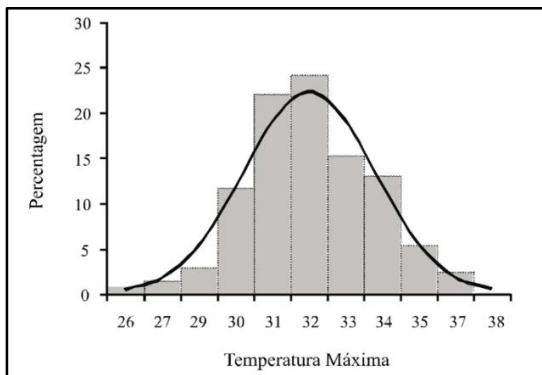
Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Alice entrevistou 60 funcionários que estavam passando em um dos corredores da empresa.		
<b>Amostra 2:</b> Sabendo que a empresa tinha o mesmo número de funcionários homens e mulheres, Gustavo entrevistou 30 homens e 30 mulheres, escolhendo-os aleatoriamente, de cada setor da empresa, totalizando 60 funcionários.		
<b>Amostra 3:</b> Fernanda tinha o nome de todos os 600 funcionários, colocou-os em um chapéu e então tirou, aleatoriamente, 60 deles para realizar a entrevista.		
<b>Amostra 4:</b> Luís enviou um questionário para todos os funcionários e depois usou os 60 primeiros que foram devolvidos a ele.		
<b>Amostra 5:</b> Carol enumerou e ordenou todos os 600 funcionários da empresa em uma lista. Em seguida, a cada intervalo de 10 funcionários, ela selecionou o oitavo, totalizando 60 entrevistados.		

**Habilidade:** Julgar e justificar a adequação ou não dos métodos de Amostragem.

#### Questão 4

4º O gráfico representa a distribuição das temperaturas máximas diárias, em graus Celsius, registradas em uma cidade do sertão nordestino durante o mês de fevereiro. Um meteorologista selecionou um dia desse mês, aleatoriamente, e observou que temperatura máxima nesse dia foi de 35º. De acordo com o gráfico, essa temperatura

observada por ele representa a média das temperaturas dessa distribuição? Justifique sua resposta.



**Habilidade:** Reconhecer o gráfico como uma distribuição normal, identificar a média como o centro da Curva e julgar se a medida apresentada é igual à média.

### Questão 5

5º Em uma pesquisa foi medido o comprimento (cm) da palma da mão direita de homens e mulheres, maiores de 15 anos. A Figura 1, a seguir, apresenta a distribuição das medidas do palmo da mão direita (cm) de homens cuja média é igual a 21,9 cm. Já a Figura 2 apresenta a distribuição das medidas do palmo da mão direita (cm) de Mulheres e sua média é igual a 19,4 cm. Com base nas informações das figuras, qual das duas distribuições possui maior variabilidade das medidas do palmo da mão direita (cm)? Justifique sua resposta!

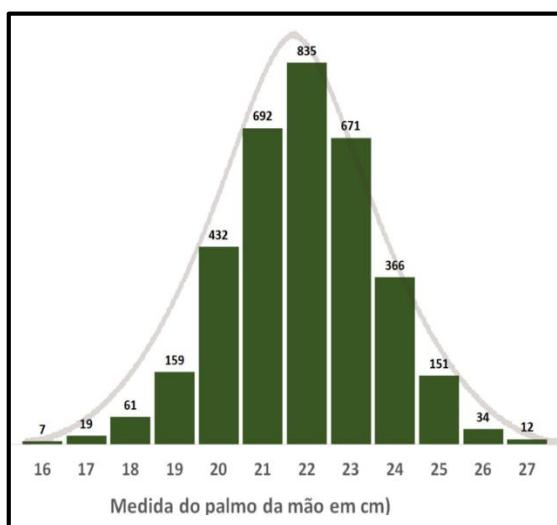


Figura 1 – Distribuição das medidas da palma da mão de homens. Média: 21,9 cm

**Habilidade:** Reconhecer através do gráfico, qual distribuição apresenta maior variabilidade.

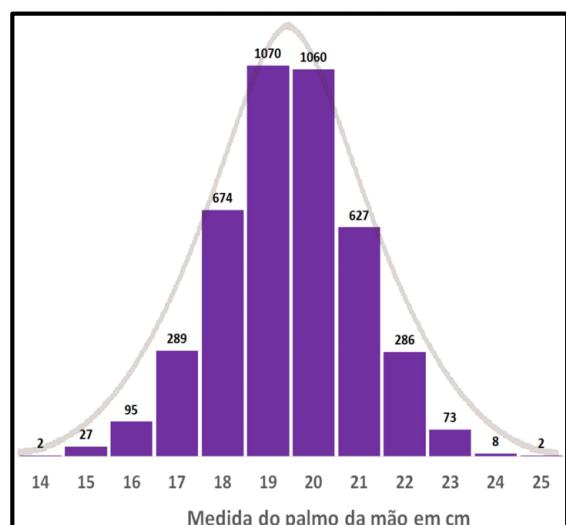
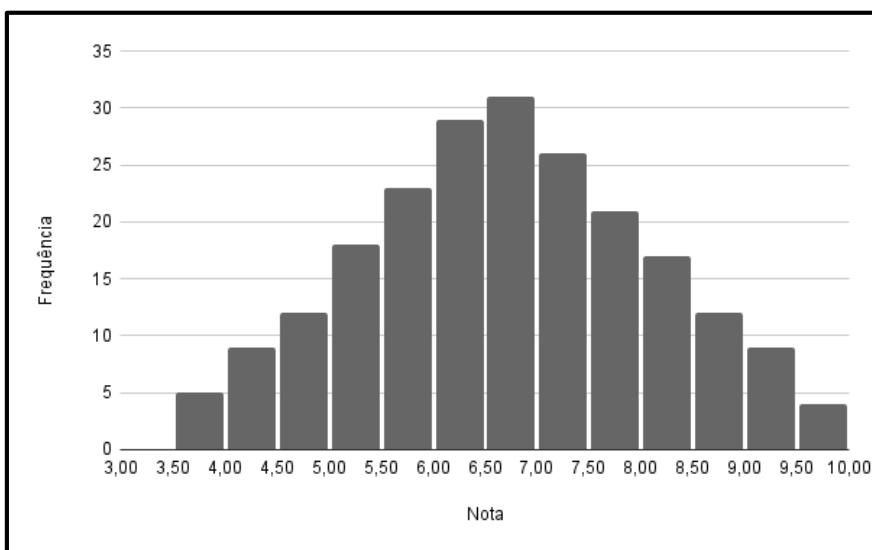


Figura 2 – Distribuição das medidas da palma da mão de Mulheres. Média: 19,4 cm

### Questão 6

6º O gráfico apresenta a distribuição das notas que os estudantes das turmas do 2º ano do Ensino Médio de uma escola pública de Pernambuco obtiveram em um teste de Matemática que valeu no máximo 10 pontos. Essa distribuição é aproximadamente normal e a sua média é igual a 6,4. Com base nas informações do gráfico, responda aos itens dessa questão.



- Qual a probabilidade aproximada de selecionarmos um desses estudantes, ao acaso, e ele possuir a nota no teste menor ou igual que a média?
- De acordo com o gráfico, se o professor de Matemática selecionar uma amostra, ao acaso, de 30 estudantes que responderam ao teste, a probabilidade de que a média amostral seja igual a 9,0 é baixa ou alta? Justifique sua resposta

#### Habilidades:

- Reconhecer que metade da Curva (nota maior ou igual que a média) tem probabilidade de 50%
- Reconhecer que a média amostral tende a se aproximar da média populacional, e, consequentemente, compreender que a probabilidade solicitada para a nota é baixa por se distanciar da média.

Por fim, destacamos que o pré-teste e o pós-teste envolveram 17 itens que totalizaram 25 pontos. No quadro 6 apresentamos esses itens com suas respectivas habilidades

e os critérios de pontuação juntamente com as categorias de análises emergentes das respostas dos estudantes.

Quadro 6 – Critérios de correção por item do pré-teste e pós-teste

<b>Questão</b>	<b>Habilidades</b>	<b>Pontuações</b>
1 – Itens A <sub>1</sub> , A <sub>2</sub> , B <sub>1</sub> , B <sub>2</sub> , C <sub>1</sub> e C <sub>2</sub>	Identificar a população e amostra relativa a cada pesquisa.	0 pontos: resposta inadequada 1 ponto: resposta adequada
2 – Item A	Levantar critérios para a representatividade amostral, considerando a variabilidade natural (características dos estudantes)	0 pontos: resposta inadequada 1 ponto: adequada
2 – Item B	Levantar critérios para a seleção de uma amostra representativa dos estudantes.	0 pontos: resposta inadequada 1 ponto: adequada
3 – Itens A, B, C, D e E	Julgar e justificar a adequação ou não dos métodos de Amostragem	0 ponto: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada
4	Reconhecer o gráfico como uma distribuição normal, identificar a média como o centro da Curva e julgar se a medida apresentada é igual à média.	0 ponto: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada
5	Reconhecer através do gráfico, qual distribuição apresenta maior variabilidade	0 pontos: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada
6 – Item A	Reconhecer que metade da Curva (nota maior ou igual que a média) tem probabilidade de 50%	0 pontos: resposta inadequada 1 ponto: resposta adequada
6 – Item B	Reconhecer que a média amostral tende a se aproximar da média populacional, e, consequentemente, compreender que a probabilidade solicitada para a nota é baixa por se distanciar da média.	0 ponto: resposta inadequada 1 ponto: resposta parcialmente adequada 2 pontos: resposta adequada

Fonte: O autor (2023).

## 6 RESULTADOS DO ESTUDO 1

Neste capítulo apresentamos os resultados do Estudo 1, o qual objetivou analisar o que tem sido proposto nos livros didáticos do Ensino Médio, aprovados pelo Programa Nacional do Livro e do Material Didático – PNLD (Brasil, 2021), para o ensino de Amostragem, Curva Normal e suas relações. Nesse sentido, essa análise foi desenvolvida em duas etapas: na primeira, analisamos os Livros de Projetos (integradores e de Vida), que constituem o objeto 1 desse PNLD e na segunda etapa, analisamos os Livros de Conhecimento, os quais são classificados como o objeto 2 do referido PNLD.

### 6.1 LIVROS DE PROJETOS INTEGRADORES

#### 6.1.1 Amostragem

Como indicado no capítulo do Método, selecionamos e analisamos os 14 livros de projetos integradores aprovados pelo PNLD 2021. De início, fizemos um levantamento, página a página, para identificar as situações que envolviam o conceito de Amostragem e da Curva Normal, nos tipos “Explicação”, e “Resolução”. Nessa direção, no que tange ao conceito de Amostragem, identificamos um total de 127 situações, sendo 21 de explicação e 106 de Resolução.

Inicialmente, como primeira categoria de análise, classificamos as situações quanto à forma que o conceito de amostra é apresentado. Assim, denotamos de amostra explícita às situações que notadamente têm por objetivo discutir conceitos relacionados à amostra. No entanto, também há situações que não tem essa finalidade, mas possibilita ao professor fazer uma conexão com o conceito de amostra e, consequentemente, discuti-lo com os estudantes e, assim, as classificamos como amostra implícita (Tabela 1).

Tabela 1 – Quantitativo das situações de Amostra por tipo de abordagem

<b>Conceito</b>	<b>Amostragem</b>	
	<b>Situação</b>	
	<b>Explicação (n = 21)</b>	<b>Resolução (N= 106)</b>
Explícito	15 (71%)	68 (64%)
Implícito	06 (29%)	38 (36%)

Fonte: Dados da Pesquisa (2022).

Diante dessas informações, podemos concluir que há uma predominância de situações que envolvem o conceito de amostra de forma explícita, tanto nas de explicação como as de resolução. Assim, acreditamos que situações que trabalhem os conceitos de forma explícita facilitam as ações dos estudantes e também do professor durante o processo de ensino e aprendizagem. Para exemplificar esses dois tipos de abordagem do conceito de amostra, destacamos os recortes, a seguir.

Figura 27 – Situação de explicação com abordagem explícita ao conceito de Amostra

**Pesquisa avalia percepção de jovens sobre ciência e tecnologia**

Temas de ciência e tecnologia despertam grande interesse entre os jovens brasileiros, superando assuntos relacionados a esportes e comparável aos de religião. A maioria porém, incluindo os jovens de curso superior, não consegue citar o nome de uma instituição nacional de pesquisa nem de algum cientista brasileiro. A constatação é da pesquisa **O que os jovens brasileiros pensam da ciência e da tecnologia?**, do Instituto Nacional de Ciência e Tecnologia em Comunicação Pública da Ciência e Tecnologia (INCT/CPCT) [...].

Realizado pela primeira vez no Brasil, o estudo teve abrangência nacional e emprego da técnica de *survey*, para aplicação de questionário estruturado, presencial, junto a amostra da população brasileira de jovens entre 15 e 24 anos. A pesquisa quantitativa ouviu 2 206 pessoas [...]. Ela envolveu também etapas cognitiva e qualitativa, para saber dos jovens suas opiniões e atitudes sobre ciência e tecnologia.

[...]

**Principais resultados**

Para seleção dos entrevistados, foi utilizada amostra probabilística até o penúltimo estágio, com aplicação de cotas amostrais de sexo, idade e escolaridade no último estágio. O intervalo de confiança é de 95 por cento. As entrevistas, realizadas por equipe treinada, foram feitas em domicílio entre os meses de março e abril de 2019.

Dentre os resultados, é possível destacar:

- A maioria dos jovens brasileiros manifesta grande interesse para temas de ciência e tecnologia, tanto as mulheres quanto os homens, e em quase todos os grupos sociais; o interesse por Ciência e Tecnologia, em geral, é maior que o por esportes, e comparável com o interesse por religião;
- Os jovens possuem, em geral, uma imagem positiva da figura do cientista e, em sua maioria, acreditam que homens e mulheres têm a mesma capacidade para ser cientista, e devem ter as mesmas oportunidades;

Fonte: Bueno, Livro de Projetos Integradores (2021, p. 87).

Figura 28 – Situação de resolução com abordagem implícita ao conceito de Amostra

**O que os moradores pensam sobre sua cidade?**



2. Em grupo, escolham um dos temas a seguir para elaborar uma entrevista ou uma pesquisa a ser realizada com os moradores do bairro.

Na elaboração de cada pergunta, estejam atentos para o objetivo, a linguagem clara e a ordem que vai ocupar na entrevista ou pesquisa. Depois de realizada a entrevista, apresentem as informações para a turma. Pode ser em forma de vídeo ou texto.

Fonte: Bueno, Livro de Projetos Integradores (2021, p. 163).

A primeira imagem retrata os resultados de uma pesquisa estatística sobre o tema ciência e tecnologia, através da qual o autor apresenta e discute esse tema, seus principais resultados e ainda destaca o tipo de Amostragem utilizado, o tamanho da amostra e o intervalo de confiança. Já a segunda imagem se configura como uma situação de resolução na qual há a abordagem de forma implícita do conceito de amostra, tendo em vista que há uma proposta de pesquisa sobre o que os moradores de um bairro pensam sobre a sua cidade. Essa situação trata o conceito de amostra

de modo implícito, o que exigirá, do professor, levantar questionamentos aos estudantes de qual será a amostra da pesquisa ou se a mesma será censitária e se os resultados observados no bairro podem ser generalizados para toda a população.

Na segunda categoria de análise (Tabela 2), classificamos as situações de acordo com o contexto, nos tipos real ou fictício. Nessa análise, consideramos uma situação fictícia aquela imaginada ou idealizada e que não representa uma situação real. Já as situações com contexto real são aquelas que utilizam fontes de dados reais, como exemplo as de órgão governamentais ou institutos de pesquisa. Nesse sentido, verificamos que nas 127 situações que envolvem o conceito de Amostragem, há a predominância de contextos reais.

Tabela 2 – Quantitativo das situações de Amostragem por tipo de contexto

<b>Amostragem</b>		
<b>Contexto dos dados</b>	<b>Situação</b>	
	<b>Explicação (n = 21)</b>	<b>Resolução (N= 106)</b>
Real	16 (76%)	78 (74%)
Fictício	05 (24%)	28 (26%)

Fonte: Dados da Pesquisa (2022).

Esses índices podem ser justificados pelo fato de que há uma maior ênfase de situações envolvendo a análise e realização de pesquisas estatísticas e, consequentemente, contempla mais contextos com dados reais, como pode ser observado nas Figuras apresentadas anteriormente. Diante disso, concordamos com Guimarães e Gitirana (2013) que os livros didáticos devem priorizar a abordagem de situações com dados reais, pois isso possibilita uma maior reflexão dos estudantes sobre o contexto social ao qual ele está inserido, levando a uma melhor compreensão de mundo e a tomada de decisões a partir do que verdadeiramente se sabe sobre os aspectos que os rodeiam. Ben-Zvi, Bakker e Makar (2015) complementam que a utilização de situações reais envolvendo à Amostragem pode facilitar o entendimento sobre a variabilidade e representatividade amostral.

Foram encontradas 47 (quarenta e sete) situações, sendo 02 de explicação e 45 de resolução que propunham a realização de pesquisas. Como afirmado anteriormente, esse resultado é bastante positivo, uma vez que acreditamos ser fundamental levar os alunos a construírem e realizarem pesquisas estatísticas para compreenderem seus contextos reais.

Em continuidade, verificamos o tipo de população/amostra de cada situação, o qual pode ser pessoas ou objetos. Nesse sentido, os dados, a seguir, apontam uma predominância de situações que envolvem o conceito de população/amostra se referindo apenas a pessoas.

Tabela 3 – Quantitativo das situações de Amostragem por tipo de população/amostra

Amostragem		
População/Amostra	Situação	
	Explicação (n = 21)	Resolução (N= 106)
Pessoas	20 (95%)	96 (91%)
Objetos	01 (5%)	10 (9%)

Fonte: Dados da pesquisa (2022).

Sobre esse tema, alguns estudos como Watson e Kelly (2005) e Gomes (2013), apontam que os estudantes apresentam dificuldades para conceber o conceito de população, na Estatística, como algo que não envolve apenas pessoas, principalmente, por entenderem esse termo da mesma forma que ele é compreendido no senso comum. Diante disso, julgamos que é necessário que os livros didáticos busquem um equilíbrio para abordagem o conceito de população por meio de situações que envolvam pessoas, como também, objetos, animais e etc. A seguir, apresentamos dois recortes de um dos livros analisados, a primeira abordada o conceito de população/amostra com pessoas, através da realização de uma pesquisa estatística e a segunda contempla esse conceito com objetos, ao apresentar e questionar sobre resultado de uma pesquisa realizada com uma amostra de blogs.

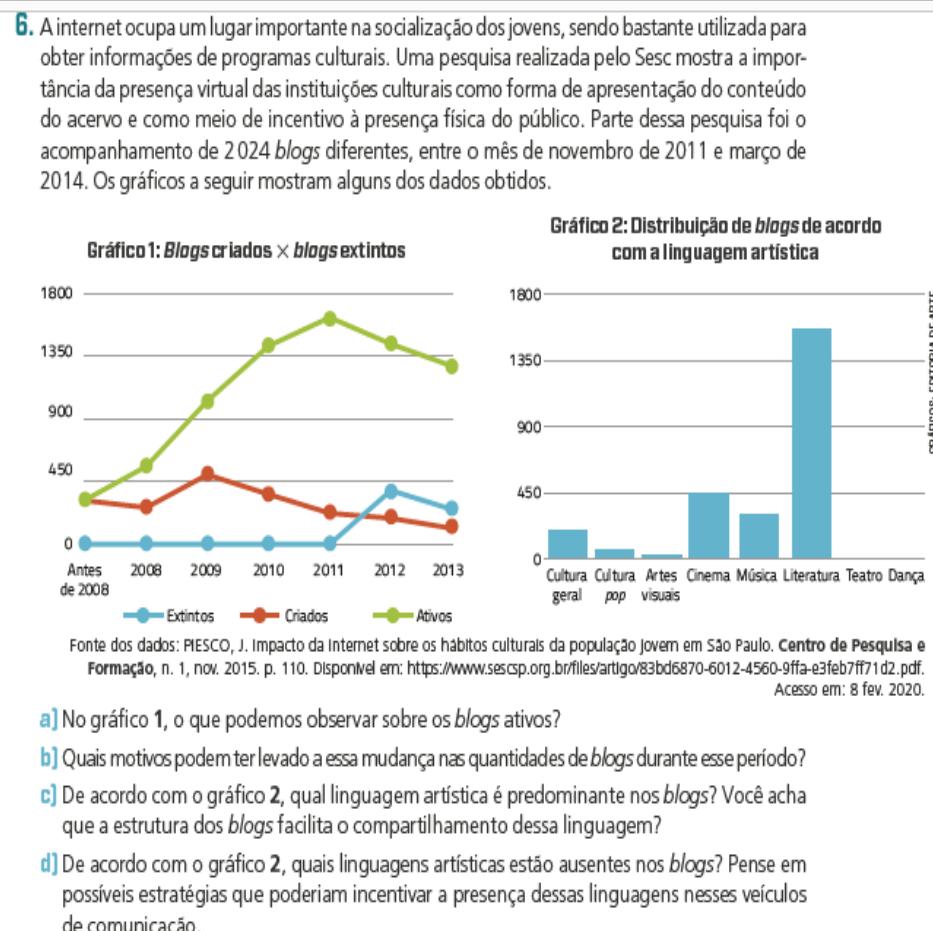
Figura 29 – Situação de resolução envolvendo o conceito de população/amostra com pessoas

3. Com o grupo de projeto, vocês vão planejar a pesquisa e coletar dados sobre mulheres e homens da região em que se encontra a escola. Esses dados serão relativos, entre outras coisas, ao tempo dedicado ao trabalho remunerado e não remunerado, quantidade de filhos e escolaridade.

Depois da coleta, vocês devem se reunir e organizar os dados coletados em tabelas e representá-los utilizando gráficos. É possível que os números que vocês encontrarem sejam diferentes dos que estudaram, uma vez que estão avaliando apenas um recorte do Brasil; além disso, os dados vão ajudá-los a perceber características pontuais do local em que a pesquisa for feita. Para a pesquisa, o grupo deve atentar aos seguintes tópicos:

Fonte: Furtado *et al*, Livro de Projetos Integradores (2021, p. 112).

Figura 30 – Situação de resolução envolvendo o conceito de população/amostra com objetos



Fonte: Souza, Livro de Projetos Integradores (2021, p. 191).

Por fim, a nossa quarta categoria de análise diz respeito às habilidades que são exploradas em cada situação (Tabela 4).

Tabela 4 – Quantitativo das situações de Amostragem por habilidades

<b>Amostragem</b>		
<b>Habilidades</b>	<b>Situação</b>	
	<b>Explicação (n = 21)</b>	<b>Resolução (N= 106)</b>
Seleção e Representatividade de Amostra	0	2 (2%)
Análise dos dados de pesquisa amostral	18 (86%)	14 (13%)
Realização de Pesquisa Estatística	2 (10%)	45 (42%)
Técnicas de Amostragem	1 (4%)	2 (2%)
Identificar a margem de erro	0	4 (4%)
Comparação entre Amostras	0	1 (1%)
Cálculo de medidas Estatísticas	0	31 (29%)

Identificar o tamanho da Amostra	0	7 (7%)
Pesquisa censitária ou amostral		
Classificação do tipo de Variável		
Construção de gráfico para a representação dos dados amostrais		
Cálculo de probabilidades a partir dos dados amostrais		
Contagem a partir de dados amostrais		

Fonte: Dados da pesquisa (2022).

Podemos observar que nas situações de explicação, há a predominância da habilidade de análise dos dados de pesquisas amostrais. Esse índice pode ser justificado pelo fato de que os autores para explicarem as temáticas dos projetos integradores, a serem desenvolvidas pelos estudantes, recorreram a pesquisas já realizadas sobre diversos temas inerentes a nossa sociedade, com o intuito de instigar a reflexão sobre a importância das temáticas e como elas estão presentes em nosso cotidiano. Para exemplificar, destacamos (Figura 31) os resultados de uma pesquisa TIC Domicílios, realizada, em 2019, pelo Centro Regional de Estudos para o Desenvolvimento da Sociedade da Informação (CETIC) sobre o acesso à internet no Brasil, no âmbito de um projeto integrador proposto que envolve o tema de mídias digitais.

Figura 31 – Situação de explicação - Análise dos dados de pesquisa amostral

**Veja outros destaques da pesquisa:**

- Nas regiões urbanas, a conexão é um pouco maior do que a média: 74% da população está ligada à internet.
- Pela primeira vez, metade da zona rural brasileira está conectada – 49% da população disse ter acesso à rede em 2018, acima dos 44% de 2017.
- Também pela primeira vez, metade da camada mais pobre do Brasil está oficialmente na internet: 48% da população nas classes D e E, acima de 42% em 2017.
- São 46,5 milhões de domicílios com acesso à internet, 67% do total.
- Entre os usuários da internet, 48% adquiriu ou usou algum tipo de serviço *on-line*, como aplicativos de carros, serviços de *streaming* de filmes e música, ou pedido de comida.

Fonte: De Leonardo, Livro - Projetos Integradores (2021a, p. 186).

Diante disso, destacamos que embora não haja uma homogeneização nas situações de Explicação das habilidades relativas ao conceito de Amostragem, tendo em vista que os livros de projetos não tem a atribuição de explicar detalhadamente os conceitos como ocorre nos livros de conhecimento, a exploração dos resultados de pesquisas estatísticas com dados reais possibilita uma maior compreensão de mundo e podem contribuir para o aprimoramento dos elementos de Conhecimento e de Disposição (Gal, 2002,2019) dos estudantes, acarretando em uma maior domínio e motivação para refletir sobre as suas crenças e atitudes frentes a esses resultados.

No que diz respeitos às situações de Resolução, observamos que há uma maior variedade de habilidades envolvidas, tendo uma maior frequência de situações voltadas para Realização de pesquisa Estatística, seguidas dos Cálculos de Medidas Estatística e a Análise de dados de pesquisa amostral. Nesse sentido, verificamos que as realizações de pesquisas estavam vinculadas às etapas do projeto, como forma dos estudantes irem a campo para pesquisarem dados em seus cotidianos que viabilizassem a execução dos projetos. Assim como pontuado anteriormente, acreditamos que a abordagem e execução de pesquisas estatísticas é de grande valia porque possibilita os estudantes refletirem sobre os dados reais que estão presentes em seus contextos sociais o que pode potencializar seu conhecimento de mundo e a tomadas de decisões. Barberino e Magalhães (2016) apontam a importância do desenvolvimento de projetos de pesquisa estatística na escolarização básica por acreditarem que esses projetos contribuem para a formação de estudantes críticos, ao partirem de situações relacionadas com o contexto social dos mesmos, além de facilitar a construção de conceitos, a medida em que o projeto vai se desenrolando, e não se prende apenas tratamento de técnicas operacionais.

No entanto, observamos que essas situações não exploravam todas fases do ciclo investigado (Guimarães e Gitirana, 2013), se concentrando apenas nas etapas do Problema de pesquisa, coletas de dados, representação dos dados, análise dos dados e conclusão. Logo, essas propostas podem não contribuir, por exemplo, para o desenvolvimento de habilidades, por parte dos estudantes, voltadas para o levantamento de hipóteses e a classificação dos dados, as quais são etapas essenciais de um ciclo de investigação (Guimarães e Gitirana, 2013). Como exemplo, exemplificamos (Figura 32) a proposta de uma pesquisa estatística voltada para o levantamento de informações sobre os estudantes que possuem algum tipo de deficiência, no âmbito de um projeto que trata sobre a acessibilidade.

Figura 32 – Situação de explicação – Realização de Pesquisa Estatística



- 5.** Para resolver esta atividade, junte-se a colegas e formem grupos com quatro integrantes. Vocês realizarão uma pesquisa estatística na escola em que estudam, ou em outra escola do município, com o objetivo de obter e analisar informações sobre estudantes com deficiência, tais como: quantidade, tipos de deficiência, desafios que eles enfrentam etc. Para isso, podem ser realizados os passos a seguir.

Fonte: Souza, Livro de Projetos Integradores (2021, p. 20).

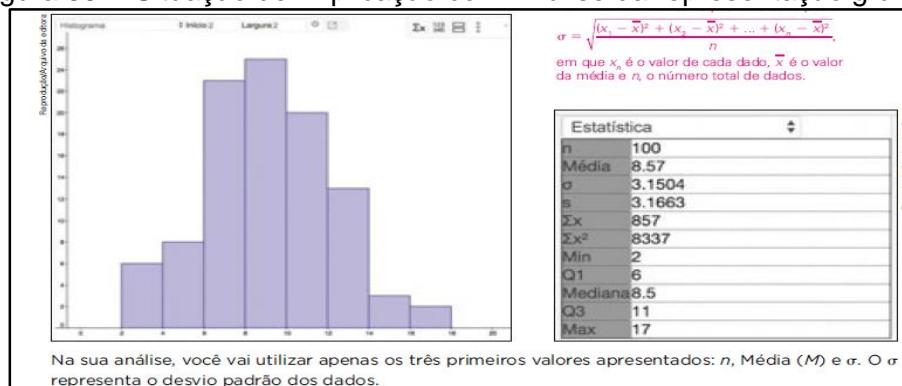
Além disso, evidenciou-se que a partir das pesquisas realizadas, as situações de resolução também eram voltadas com uma frequência significativa para o cálculo de medidas estatísticas, como por exemplo, as medidas de tendência central e dispersão, seguidas da análise dos dados amostrais, por meio de questionamento após o tratamento de dados. Diante disso, reconhecemos a importância da aplicação do algoritmo para o cálculo dessas medidas como forma de analisar e refletir sobre os dados, no entanto, as situações propostas, evidenciaram pouca abordagem sobre a seleção e representatividade da amostra, técnicas de Amostragem e margem de erro, isto é, tópicos relativos ao conceito de Amostragem que possibilitam compreender como se dá o processo de inferência estatística. Ben-zvi, Bakker e Makar (2015), Cazares (2019) chamam a atenção de que para a compreensão do conceito de Amostragem é necessário o domínio de diversos outros conceitos, como os citados anteriormente, e desta forma, o processo de ensino e aprendizagem da Amostragem deve possibilitar a exploração dos mesmos.

Diante disso, acreditamos que os projetos integradores propostos podem contribuir para a compreensão do contexto social em que os estudantes estão inseridos, por meio da abordagem e realização de pesquisas estatísticas que contemplam dados reais e promovem a reflexão sobre o referido contexto em diversas temáticas. Ao mesmo tempo, julgamos que os projetos poderiam explorar mais conceitos relativos ao campo da Amostragem, concatenando com a realização das pesquisas, como por exemplo, o processo de seleção e representatividade da amostra, através das técnicas de Amostragem, como também a reflexão sobre a margem de erro, variabilidade amostral, o tamanho da amostra e a comparação entre amostras, aspectos que influência na representatividade amostral e são importantes na Inferência Estatística (Beguè *et al*, 2017; Batanero *et al*, 2018).

#### 6.1.2 Curva Normal

Com relação ao conceito da Curva Normal, identificamos apenas 2 situações de Explicação que abordavam esse conceito de forma implícita, envolvendo a habilidade de análise da representação gráfica em um contexto fictício. A primeira situação (Figura 33), o autor aborda a distribuição dos dados relativos a variável tempo (min) e, para isso, apresenta a representação gráfica desses dados, envolvendo uma variável contínua e, além disso, é apresentado algumas medidas estatísticas relativas a esses dados, dando ênfase para a média e o desvio-padrão, os quais constituem os dois parâmetros que definem uma distribuição Normal.

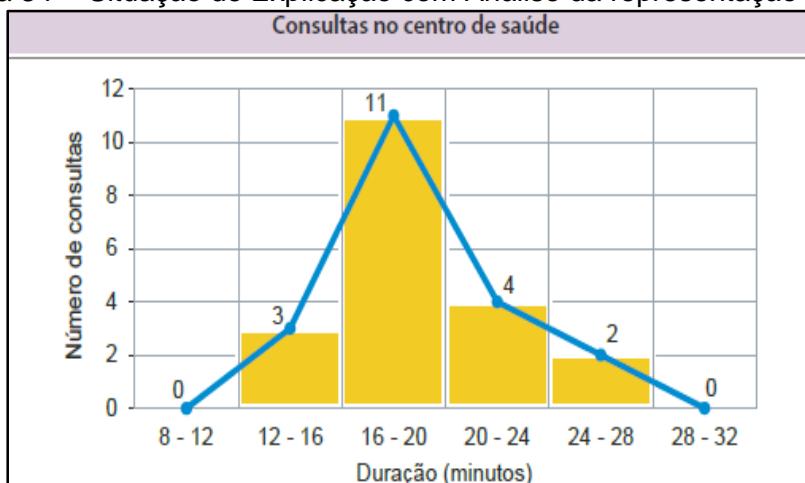
Figura 33 – Situação de Explicação com Análise da representação gráfica



Fonte: Santos, Livro de Projetos Integradores (2021, p. 20).

Já a segunda situação (Figura 34) também aborda a distribuição dos dados relativos à variável tempo (min). Essa retrata, de modo fictício, o tempo demandado para o atendimento a pacientes em um centro de saúde através de uma representação gráfica com uma variável contínua e, assim, contempla o conceito da Curva Normal de modo implícito

Figura 34 – Situação de Explicação com Análise da representação gráfica



Fonte: Gay, Livro de Projetos Integradores (2021, p. 47).

Através de situações como essas, o professor pode ressaltar que quanto maior a quantidade de dados contínuos em uma distribuição, mais haverá a tendência da mesma se configura como uma distribuição normal, caracterizando o formato de uma curva. Logo, a ideia de que determinados fenômenos na natureza seguem um determinado padrão (assemelhando-se à Curva Normal), em que os valores extremos têm menor frequência e os valores mais próximos da média ocorrem em maior número de vezes (maior frequência) pode ser explorada em sala de aula, favorecendo, aos estudantes, habilidades para analisarem de modo crítico uma variedade de fenômenos do nosso cotidiano que apresentam uma Distribuição Normal ou aproximadamente normal de dados.

Diante disso, acreditamos que a pouca abordagem desse conceito nos livros analisados se deve ao fato de que não há uma preocupação em se discutir, principalmente nas etapas de representação e análise de dados, sobre modelos probabilísticos que permitem a realização de inferências estatísticas. Verificamos que nesses livros há uma ênfase para a realização de pesquisas estatísticas, como etapas dos projetos, mas a partir delas é explorado predominantemente as análises descritivas em detrimento às inferências que podem ser realizadas, através dos referidos modelos, que possibilitem um maior poder de análise e conclusão da pesquisa, por parte dos estudantes.

Diante disso, tendo em vista que a Curva Normal é considerada como o principal modelo de análise de dados presente na área Inferencial (Batanero, Tauber e Sánchez, 2004), destacamos que e a sua abordagem deve ser iniciada já na Educação Básica, por acreditar que seu processo de ensino e aprendizagem possibilita a modelagem de fenômenos presentes em nosso cotidiano em diferentes áreas do conhecimento e também a inter-relação entre a Estatística e a Probabilidade, áreas que, comumente, são ensinadas de forma totalmente independente por muitos docentes, na escolarização básica (Monroy e Herrera, 2019; Araújo, 2020).

## 6.2 LIVROS DE PROJETOS DE VIDA

Como pontuado no capítulo do Método da nossa pesquisa, as obras de projetos de vida apresentam propostas de projetos que englobam as dimensões pessoal (encontro consigo mesmo), cidadã (encontro com o mundo), dimensão profissional (encontro com o futuro). Consequentemente, observamos que os projetos englobam com maior ênfase as habilidades voltas para as ciências humanas e sociais. Logo, ao

fazermos o levantamento das situações relativas ao tema na nossa pesquisa, encontramos um total de 49 situações envolvendo unicamente o conceito de Amostragem, sendo 28 de explicação e 21 de resolução, uma quantidade menor comparada com os livros de projetos integradores.

Inicialmente, como primeira categoria de análise, analisamos e classificamos as situações quanto à forma que o conceito de amostra é apresentado, seja explícito ou implícito. A tabela, a seguir, apresentam os resultados dessa categoria.

Tabela 5 – Quantitativo das situações de Amostra por tipo de abordagem

Amostragem		
Conceito	Situação	
	Explicação (n = 28)	Resolução (N= 21)
Explícito	22 (78%)	13 (62%)
Implícito	06 (22%)	08 (38%)

Fonte: Dados da Pesquisa (2022).

Através da primeira categoria de análise, verificamos uma predominância de situações que envolve o conceito de amostra de forma explícita, tanto nas de explicação como as de resolução. Logo, assim como pontuado sobre as obras de projetos integradores, acreditamos que situações que trabalhem esse conceito de forma explícita facilitam a compreensão dos estudantes, ao se depararem com uma variedade de situações envolvendo à Amostragem, além de dar mais subsídios para as ações do professor durante o processo de ensino e aprendizagem. A seguir, exemplificamos (Figura 35) uma situação de explicação, com o conceito de amostra explícito, na qual é apresentada os resultados de uma pesquisa nacional por amostras de domicílios (PNAD) sobre a participação dos negros no ensino superior público. Na sequência, apresentamos (Figura 36) uma situação de resolução na qual é apresentada os resultados de uma pesquisa amostral e, em seguida são feitos alguns questionamentos com base nesses resultados.

Figura 35 – Situação de Explicação com o conceito de amostra explícito

### Pela primeira vez, negros são maioria no ensino superior público

Segundo o IBGE, as matrículas de pretos e pardos somam 50,3%

A proporção de pessoas pretas ou pardas (que compõem a população negra) cursando o ensino superior em instituições públicas brasileiras chegou a 50,3% em 2018. Apesar desta parcela da população representar 55,8% dos brasileiros, é a primeira vez que os pretos e pardos ultrapassam a metade das matrículas em universidades e faculdades públicas.

Os dados estão no informativo Desigualdades Sociais por Cor ou Raça no Brasil, divulgado hoje (13), no Rio de Janeiro, pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A comparação foi feita com as informações do suplemento de educação da Pesquisa Nacional por Amostra de Domicílio — Contínua (Pnad Contínua), que começou a ser aplicado em 2016.

A pesquisa mostra que a população negra está melhorando seus índices educacionais, tanto de acesso como de permanência, apesar de ainda se manter bem atrás dos índices medidos entre as pessoas brancas.

A proporção de jovens de 18 a 24 anos pretos ou pardos no ensino superior passou de 50,5% em 2016 para 55,6% em 2018. Entre os brancos, a proporção é de 78,8%.

Na mesma faixa etária, o número de pretos e pardos com menos de 11 anos de estudo e que não estavam frequentando a escola caiu de 30,8% em 2016 para 28,8% em 2018, enquanto o indicador para a população branca é de 17,4%.

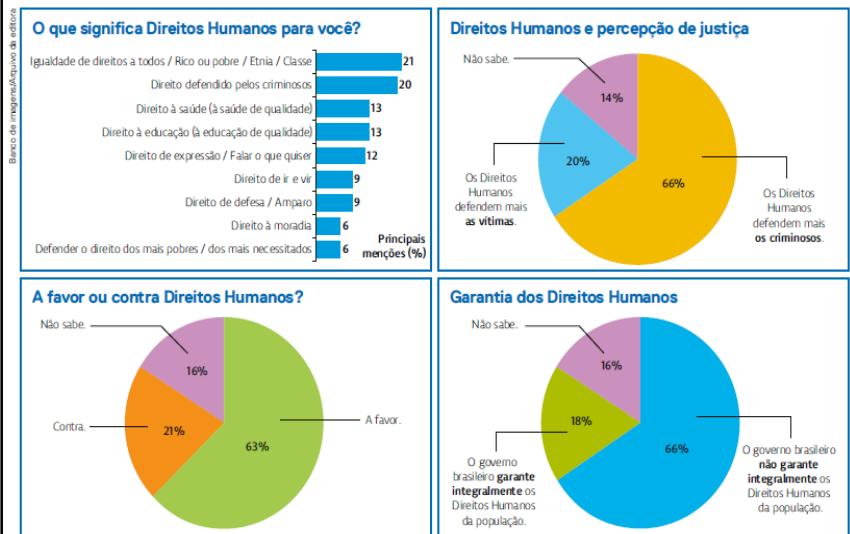
[...]



Fonte: Danza e Silva, Livro de Projetos de Vida (2021, p. 101).

Figura 36 – Situação de Resolução com o conceito de amostra explícito

Em 2018, o Instituto Ipsos realizou uma enquete para saber o que os brasileiros pensam sobre Direitos Humanos e qual a sua relação com o conceito de justiça. Chamado Pulso Brasil, o levantamento coletou a opinião de 1,2 mil pessoas em um total de 72 municípios brasileiros. Confira os resultados nos gráficos a seguir, que apresentam margem de erro de 3 pontos percentuais.



Fonte: Alchorne e Carvalho, Livro Projetos de Vida (2021, p. 94).

A segunda categoria de análise, por sua vez, se voltou para a classificação das situações de acordo com o contexto, nos tipos real ou fictício (Tabela 6). A partir desses dados, podemos concluir que há unicamente a presença de dados reais em todas situações observadas. Logo, isso constitui um fator positivo pois, como já pontuado na seção anterior, possibilita uma maior compreensão sobre o contexto social em que os estudantes estão inseridos, propiciando uma melhor compreensão de mundo e a tomada de decisões. A evidência de que os livros de projetos focam na abordagem dos resultados de pesquisas já realizadas (Figura 35) e (Figura 36), sem dúvida, colabora para que se explorem dados reais de diferentes contextos sociais.

Tabela 6 – Quantitativo das situações de Amostragem por tipo de contexto

<b>Amostragem</b>		
<b>Contexto dos dados</b>	<b>Situação</b>	
	<b>Explicação (n = 28)</b>	<b>Resolução (n = 21)</b>
Real	28 (100%)	21(100 %)
Fictício		

Fonte: Dados da pesquisa (2022).

Através da terceira categoria de análise, verificamos que todas as situações, seja de explicação ou de resolução, contemplavam o tipo de população/amostra apenas como pessoas. Os dados, a seguir, sintetizam essa informação.

Tabela 7 – Quantitativo das situações de Amostragem por tipo de população/amostra

<b>Amostragem</b>		
<b>População/Amostra</b>	<b>Situação</b>	
	<b>Explicação (n = 28)</b>	<b>Resolução (N= 21)</b>
Pessoas	28 (100%)	21 (100%)
Objetos		

Fonte: Dados da pesquisa (2022).

As Figuras 35 e 36, apresentadas anteriormente, que abordam a análise de pesquisas estatísticas, exemplificam essa conclusão. Diante disso, assim como pontuamos nas análises realizadas com os livros de projetos integradores, o fato de se abordar a população/amostra, na estatística, apenas como pessoas pode levar os estudantes a não conceberem que esse conceito também pode se referir a um conjunto de objetos ou animais. Logo, reiteramos que os livros didáticos necessitam

buscar um equilíbrio para abordagem o conceito de população por meio de situações que envolvam pessoas, como também, outros tipos de elementos.

Por fim, verificamos ainda as habilidades que são exploradas nas situações de explicação e resolução. Os dados, a seguir, evidenciam os resultados dessa análise.

Tabela 8 – Quantitativo das situações de Amostragem por habilidades

<b>Habilidades</b>	<b>Amostragem</b>	
	<b>Situação</b>	
	<b>Explicação (n = 28)</b>	<b>Resolução (n = 21)</b>
Seleção e Representatividade de Amostra	0	0
Análise dos dados de pesquisa amostral	28 (100%)	5 (24%)
Realização de Pesquisa Estatística	0	14 (67%)
Cálculo de medidas Estatísticas	0	2 (9%)
Técnicas de Amostragem		
Identificação da margem de erro		
Comparação entre Amostras		
Identificação do tamanho da Amostra		
Pesquisa censitária ou amostral		
Classificação do tipo de variável		
Construção de gráfico para a representação dos dados amostrais		
Cálculo de probabilidades a partir dos dados amostrais		
Contagem a partir de dados amostrais		

Fonte: Dados da Pesquisa (2022).

Diante disso, podemos observar que nas situações de explicação, há unicamente a exploração da habilidade de análise dos dados de pesquisas amostrais (Figura 35). Assim, de modo semelhante as obras de projetos integradores, esse índice se justifica pelo fato de que, os autores recorreram a pesquisas já realizadas, sobre diversos temas inerentes à nossa sociedade, para explicarem as temáticas dos projetos, a serem desenvolvidas pelos estudantes, com o intuito que os mesmos reflitam sobre a importância das temáticas e como elas estão presentes em nosso dia a dia.

Já nas situações de resolução, há a predominância de situações envolvendo a realização de pesquisas estatísticas. Essas atividades, assim como nos livros de projetos integradores, estão ligadas a algumas etapas dos projetos, como forma dos

estudantes pesquisarem dados relativos aos temas estudados. No entanto, também verificamos que essas propostas de pesquisas não contemplam todas as fases do ciclo investigativo (Guimarães e Gitirana, 2013), ficando restritas principalmente, ao problema de pesquisa, coleta de dados, representação dos dados, análise e conclusão. Diante disso, mais uma vez reiteramos sobre a importância do trabalho com as fases de uma investigação como forma dos estudantes refletirem sobre as ações em cada etapa, ao mesmo tempo que aprofundam à aprendizagem de conceitos estatísticos. Como exemplo para uma situação com essa habilidade, destacamos (Figura 37) uma proposta para a realização de uma pesquisa com idosos para que os estudantes conheçam suas histórias de vida.

Figura 37 – Situação de Resolução com a realização de pesquisa estatística

Faça, com a turma e o professor, uma entrevista com idosos que moram ou frequentem os arredores da escola ou que sejam familiares de alguém da turma, a fim de conhecê-los e ouvir suas histórias de vida. Primeiramente, prepare-se criando um roteiro. Veja uma sugestão a seguir.

Lembre-se: você já realizou uma entrevista no primeiro bloco, **Quem é você?**. Use o que você aprendeu naquele momento para fazer essa nova entrevista. [Junto com a turma, desenvolva um roteiro para realizar as entrevistas.](#)

- Comece com perguntas sobre os dados pessoais, como nome, local e data de nascimento. Depois, faça perguntas que ajudem a construir a narrativa de forma cronológica, como: Quem são seus pais e avós? Como e onde foi sua infância?
- Pergunte sobre experiências marcantes de sua trajetória de vida.
- Ajude o entrevistado a relacionar sua história de vida com o tempo presente, avaliando quais aspectos da sua vida contribuíram para que se tornasse quem é hoje.
- Para finalizar, pergunte ao entrevistado: Que conselho de vida você daria para os mais jovens?

Fonte: Danza e Silva, Livro de Projetos de Vida (2021, p. 41).

### 6.3 LIVROS DE CONHECIMENTO

A partir dos procedimentos adotados na nossa metodologia, analisamos os 20 livros de conhecimento aprovados no PNLD 2021, sendo 10 volumes dos livros de Matemática e suas tecnologias, que abordam a Estatística e a Probabilidade e os outros 10 referentes a obra específica de Ciência Humanas e Matemática. A seguir, apresentamos os resultados de nossa análise iniciando com as situações que envolvem o conceito nos livros de Matemática e suas Tecnologias.

#### 6.3.1 Amostragem

Com relação ao conceito de Amostragem, verificamos que os 10 livros de Matemática e suas tecnologias que abordam a temática de Estatística e Probabilidade contemplam 436 situações, sendo 126 de explicação e 310 de resolução. Como primeira categoria de análise, classificamos essas situações de acordo com forma de

abordagem do conceito de amostra, nos tipos explícito e implícito. A tabela, abaixo, apresenta os resultados dessa categorização.

Tabela 9 – Quantitativo das situações de Amostra por tipo de abordagem

Conceito	Amostragem	
	Situação	
	Explicação (n = 126)	Resolução (N= 310)
Explícito	124 (71%)	300 (64%)
Implícito	03 (29%)	10 (36%)

Fonte: Dados da Pesquisa (2022).

Como podemos observar, há uma grande predominância de situações nos tipos explicação e resolução com a abordagem do conceito de amostra de forma explícita, o que é bem importante. Esse fato, assim como já observado e pontuado nos livros de projetos, possibilita uma melhor exploração desse conceito durante o processo de ensino e aprendizagem da Matemática, possibilitando o professor abordá-lo em diferentes contextos e atividades, o que pode facilitar a aprendizagem dos estudantes. Para exemplificarmos essa conclusão, a abordamos (Figura 38) uma situação de explicação, na qual o autor discute a importância de uma pesquisa amostral para o nosso cotidiano em ocasiões em que a população é muito grande, a exemplo de pesquisas eleitorais. Em seguida, apresentamos (Figura 39) uma situação de resolução na qual é posto um contexto de uma pesquisa amostral e se questiona alguns aspectos como o tamanho da população e da amostra e possíveis variáveis a serem pesquisadas.

Figura 38 – Situação de explicação com o conceito de amostra explícito

### População e amostra estatística

Quando uma pesquisa estatística é planejada, diversos elementos relacionados a ela devem ser definidos, como as variáveis que constarão no questionário. Outro fator importante no planejamento é estabelecer quem será consultado nessa pesquisa. Para obter a altura dos alunos de uma turma do Ensino Médio, por exemplo, é possível consultar todos os alunos da turma, que constituem a população ou universo estatístico dessa pesquisa.

Em outros casos, é impossível ou inviável consultar todos os indivíduos da pesquisa. Durante as campanhas eleitorais para presidente da República, por exemplo, é inviável questionar todos os eleitores do país sobre sua intenção de voto, uma vez que isso demandaria muito tempo e recursos financeiros. Nesses casos, seleciona-se um subconjunto da população, chamado amostra, que possa representar essa população da melhor maneira possível. Para que a amostra seja representativa, ela deve ter as principais características da população à qual pertence. No exemplo apresentado, os eleitores selecionados para responder a sua intenção de voto para a eleição presidencial correspondem à amostra da pesquisa.

Fonte: Teixeira, Livro Matemática e suas Tecnologias (2021, p. 93).

Figura 39 – Situação de resolução com o conceito de amostra explícito

1. Uma empresa de confecções planeja abrir uma filial em uma cidade com 125.000 habitantes. Para fazer uma pesquisa sobre a viabilidade do projeto, contratou uma agência de publicidade. A agência informou que consultará 800 pessoas da região.
  - a) Qual é o tamanho da população da pesquisa? **125.000**
  - b) Qual é o tamanho da amostra? **800**  
resposta pessoal
  - c) Cite quatro possíveis variáveis que podem ser objetos dessa pesquisa.

Fonte: De Leonardo, Livro Matemática e suas Tecnologias (2021, p. 20).

Em continuidade, através da segunda categoria de análise, verificamos que tanto nas situações de explicação como nas de resolução, há uma maior frequência de dados fictícios. Os dados dessa categoria estão sistematizados na tabela a seguir

Tabela 10 – Quantitativo das situações de Amostragem por tipo de contexto

Amostragem		
Contexto dos dados	Situação	
	Explicação (n = 126)	Resolução (n = 310)
Real	52 (41%)	120 (39%)
Fictício	74 (59%)	190 (61%)

Fonte: Dados da pesquisa (2022).

Diante disso, podemos concluir que os Livros de conhecimento de Matemática e suas tecnologias, nos volumes de Estatística e Probabilidade, vão na contramão dos livros de projetos ao apresentarem mais situações com dados fictícios. Guimarães e Gitirana (2013) apontam a necessidade dos livros didáticos priorizarem situações com dados reais tendo em vista que as situações que de fato refletem a realidade podem proporcionar uma maior reflexão por parte dos estudantes e colaborar para o desenvolvimento do seu Letramento Estatístico (Gal, 2022), como por exemplo, suas atitudes, crenças, levantamentos de hipóteses frente a essas situações reais. Em complemento, Ben-Zvi, Bakker e Makar (2015) enfatizam que a utilização de situações reais que contemplam o conceito de Amostragem pode facilitar o entendimento sobre a variabilidade e representatividade amostral. A seguir, destacamos duas Figuras dos livros que representam situações com dados fictícios.

Figura 40 – Situação de Explicação com dados fictícios

**2.** Uma empresa pretende fazer uma pesquisa com funcionários de diversos setores, garantindo que tenha representantes de forma proporcional de cada setor, pois não será possível realizar a pesquisa com todos os funcionários. Observe a seguir dados referentes a esses funcionários e responda às questões.

- Determine quantos funcionários de cada setor precisam ser entrevistados para que 20% da empresa responda à pesquisa.
- Que tipo de amostra é mais apropriada para esse estudo?

**Resolução**

- Calculando 20% do total de funcionários, temos:

$$1600 \cdot 0,20 = 320$$

Portanto, 320 funcionários vão responder à pesquisa.

Para escolher de forma proporcional os funcionários de cada setor, fazemos:

$$\text{Setor A: } 250 \cdot 0,20 = 50$$

$$\text{Setor C: } 450 \cdot 0,20 = 90$$

$$\text{Setor E: } 410 \cdot 0,20 = 82$$

$$\text{Setor B: } 135 \cdot 0,20 = 27$$

$$\text{Setor D: } 355 \cdot 0,20 = 71$$

Portanto, deverão ser entrevistados 50, 27, 90, 71 e 82 funcionários, respectivamente, dos setores A, B, C, D e E.

- A amostra estratificada é a recomendada para essa pesquisa.

> Funcionários por setor	
Setor	Quantidade de funcionários
A	250
B	135
C	450
D	355
E	410
Total	1600

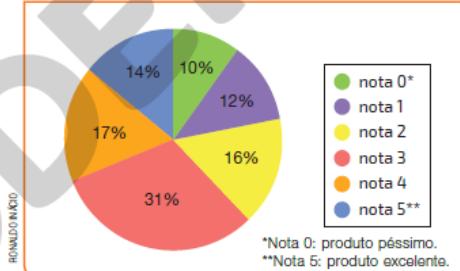
Fonte: Dados fictícios.

Fonte: Bonjorno et al., Livro Matemática e suas Tecnologias (2021, p. 65).

Figura 41 – Situação de Resolução com dados fictícios

**12** Uma empresa fez uma pesquisa com 3 000 consumidores de um de seus produtos. Na pesquisa, cada consumidor atribuía uma nota de 0 até 5. Os resultados estão apresentados no gráfico a seguir. As informações apresentadas no gráfico são fictícias.

Avaliação dos consumidores do produto em 2020



Fonte de pesquisa: Registros da empresa.

- Calcule a média, a moda e a mediana das notas atribuídas a esse produto na pesquisa e dê o significado de cada uma delas.
- Conforme as medidas calculadas no item a, escreva um pequeno texto com algumas conclusões sobre as notas atribuídas a esse produto na pesquisa.

*Respostas no final do livro.*

*Resposta pessoal. Espera-se que, nesse texto, os alunos*

Fonte: Teixeira, Livro Matemática e suas Tecnologias (2021, p. 119).

Na Figura 40 o autor explica a resolução da questão a partir de um contexto fictício, envolvendo uma pesquisa com os funcionários de uma empresa. Já na Figura 41 há também um contexto fictício que aborda uma pesquisa realizada com os consumidores de uma empresa. Além disso, destaca-se que na fonte do gráfico é adotado um termo fictício de acordo com o tema explorado na questão.

Em sequência, na categoria de análise referente ao conceito de Amostra, verificamos as situações quanto ao tipo de população/amostra (Tabela 11).

Tabela 11 – Quantitativo das situações de Amostragem por tipo de população/amostra

<b>Amostragem</b>		
<b>População/Amostra</b>	<b>Situação</b>	
	<b>Explicação (n = 126)</b>	<b>Resolução (N= 310)</b>
Pessoas	103 (82%)	249 (80%)
Objetos	23 (18%)	61 (20%)

Fonte: Dados da pesquisa (2022).

Diante disso, podemos concluir que tanto nas situações de explicação como nas de resolução há uma frequência maior com o tipo de população/amostra se referindo a pessoas (Figura 40 e 41). Assim como já discorremos na seção sobre os livros de projetos, é importante que as situações relativas ao conceito de Amostragem busquem um equilíbrio na abordagem do tipo de população/amostra, se referindo a pessoas e também a objetos ou animais (Watson e Kelly, 2005; Gomes, 2013).

Por fim, a última categoria de análise se voltou para verificar nos livros em questão, as habilidades que são exploradas em cada situação (Tabela 12).

Tabela 12 – Quantitativo das situações de Amostragem por habilidades

<b>Amostragem</b>		
<b>Habilidades</b>	<b>Situação</b>	
	<b>Explicação (n = 126)</b>	<b>Resolução (n = 310)</b>
Seleção e Representatividade de Amostra	13 (11%)	21 (7%)
Análise dos dados de pesquisa amostral	30 (24%)	27 (9%)
Realização de Pesquisa Estatística	12 (9%)	42 (14%)
Definir a pesquisa como Censitária ou Amostral	03 (2%)	19 (6%)
Cálculo de medidas Estatísticas	12 (9%)	91 (30%)
Técnicas de Amostragem	38 (30%)	34 (11%)
Identificar a margem de erro	04 (3%)	11 (3%)
Comparação entre Amostras	0	02 (1%)
Identificar o tamanho da Amostra	01 (1%)	19 (6%)
Classificar o tipo de variável	01 (1%)	11 (3%)

Construção de gráficos para representação dos dados amostrais	07 (6%)	11 (3%)
Cálculo de probabilidades a partir dos dados amostrais	04 (3%)	18 (6%)
Realizar um tipo de contagem a partir dos dados amostrais	01(1%)	04 (1%)

Fonte: Dados da Pesquisa (2022).

Diante desses dados, podemos observar que nas situações de explicação há uma maior frequência para a abordagem sobre as técnicas de Amostragem, o que possibilita ao estudante entender os diferentes processos para a formulação de uma amostra. No entanto, verificamos que a maioria dos livros abordam apenas as técnicas de Amostragem probabilísticas dos tipos Aleatória Simples, Estratificada e Sistemática. Poucos livros abordam os tipos de Amostragem não-probabilísticas, a exemplo das por conveniência e por julgamento. Sobre esse tema, a BNCC (Brasil, 2018) e o Currículo de Pernambuco (Pernambuco, 2021) orientam que no Ensino Médio deve ser realizado pesquisas amostrais com dados coletados em diferentes fontes e realidades, como também, a análise crítica de pesquisas amostrais divulgadas em diferentes meios de comunicação. Diante disso, acreditamos que os Livros Didáticos poderiam abordar os tipos de Amostragem em sua totalidade, o que inclui as não-probabilísticas que também estão presentes em diversas situações do nosso cotidiano.

Além disso, verificamos que nas situações de explicação há pouca abordagem que permita o estudante refletir sobre outros conceitos que permeiam a Amostragem, como a variabilidade, a influência do tamanho da amostra, distribuição dos dados amostrais, margem de erro, dentre outros. Begué *et al* (2017) e Batanero *et al* (2018) argumentam que na educação básica ainda há a necessidade de um maior aprofundamento no ensino da Amostragem e de fornecer aos alunos uma reflexão com base em materiais de simulações que favoreçam alguma experiência sobre a variabilidade aleatória em diferentes amostras da mesma população, e, consequentemente, o entendimento dos estudantes sobre as propriedades de Amostragem. A seguir, destacamos um exemplo da abordagem sobre os tipos de Amostragem.

Figura 42 – Situação de explicação sobre técnicas de Amostragem

A seleção de uma amostra pode ser feita de diversas maneiras. Veja a seguir alguns tipos de amostra.

- **Amostra aleatória simples:** quando os elementos da amostra são escolhidos aleatoriamente, ou seja, ao acaso. Todos os elementos têm a mesma probabilidade de ser escolhidos.
  - **Amostra sistemática:** quando os elementos da população estão ordenados, de modo que possam ser identificados por sua posição. O primeiro elemento da amostra é escolhido de modo aleatório, e os demais são retirados periodicamente. Por exemplo, de 5 em 5 ou de 8 em 8.
  - **Amostra estratificada:** quando a população é dividida em subgrupos, chamados de estratos, heterogêneos, e os indivíduos são escolhidos aleatoriamente de cada estrato.

Fonte: Dante e Viana, Livro Matemática e suas Tecnologias (2021, p. 14)

Já nas situações de resolução, observamos uma maior ênfase para a aplicação dos cálculos de medidas estatísticas, com o intuito de que o aluno domine o processo algorítmico dessas medidas. Cabe salientar que compreendemos que uma das principais funções do ensino de matemática na Educação Básica é levar o estudante à compreensão e ao tratamento das operações matemáticas através dos processos algorítmicos. No entanto, esses dados evidenciam uma abordagem voltada para o processo mecânico desses cálculos e para as análises descritivas em detrimento das análises inferenciais que frequentemente são realizadas no nosso cotidiano a partir dos dados amostrais. Oliveira e Rosa (2020) enfatizam que a relação entre os conteúdos estatísticos presentes no currículo e a realidade dos estudantes ainda é pouco explorada pelos professores. Na escolarização básica ainda é dada uma maior ênfase a metodologia tradicional, muitas vezes utilizando a aplicação de exercícios baseados em fórmulas prontas e totalmente desvinculados do contexto dos estudantes. Logo, acreditamos que essa abordagem dificulta a construção de conhecimento de forma significativa e torna-se necessário a efetivação de um ensino contextualizado para os estudantes, que os levem a compreender e interpretar os dados e informações estatísticas reais, tanto de modo descritivo como inferencial, inseridos em seus contextos. Abaixo, destacamos uma Figura que retrata essa habilidade mais presente nos livros nas situações de resolução.

Figura 43 – Situação de Resolução com cálculos de medidas estatísticas

#### **16. (FEMM-MG) Dentre as amostras:**

<b>Amostra 1</b>	0	1	2
<b>Amostra 2</b>	26	27	28
<b>Amostra 3</b>	1.002	1.005	1.003
<b>Amostra 4</b>	-5	10	25

qual é a amostra que possui maior desvio padrão?



Fonte: De Leonardo, Livro Matemática e suas Tecnologias, (2021, p. 87)

### 6.3.2 Curva Normal

No que diz respeito ao conceito da Curva Normal, verificamos um total de 52 situações que abordavam esse conceito, sendo 15 de explicação e 37 de resolução. Destaca-se, inicialmente, que há uma disparidade entre a quantidade de situações sobre Amostragem e a da Curva Normal. Isso pode ser explicado pelo fato dos livros didáticos analisados explorarem com uma maior ênfase as análises descritivas dos dados amostrais e pouco abordam situações em que se pode realizar inferências estatísticas através de modelos de distribuição de dados, como o modelo da Curva Normal.

Jacob e Doeer (2014) destacam que os estudantes do Ensino Médio, são capazes de construir conhecimentos sobre a distribuição amostral, incluindo o reconhecimento da distribuição normal dos dados e decréscimo na variabilidade à medida que o tamanho da amostra aumenta e às probabilidades associadas aos intervalos mais próximos ou distantes da média da distribuição. Diante disso, os autores enfatizam a importância da realização de atividades que explorem a incerteza, com a oportunidade dos alunos coletarem uma amostra e compará-la com a distribuição amostral, explorando modelos de distribuição de probabilidades, a exemplo do modelo da Curva Normal para que realizem inferências.

Em continuidade, a nossa primeira categoria de análise classificou as situações observadas de acordo com o tipo de abordagem do conceito da Curva Normal, isto é, explícito ou implícito. Os dados dessa categorização estão apresentados na tabela, a seguir.

Tabela 13 – Quantitativo das situações de Curva Normal por tipo de abordagem

<b>Curva Normal</b>		
<b>Conceito</b>	<b>Situação</b>	
	<b>Explicação (n = 15)</b>	<b>Resolução (N= 37)</b>
Explícito	10 (67%)	20 (54%)
Implícito	05 (33%)	17 (46%)

Fonte: Dados da Pesquisa (2022).

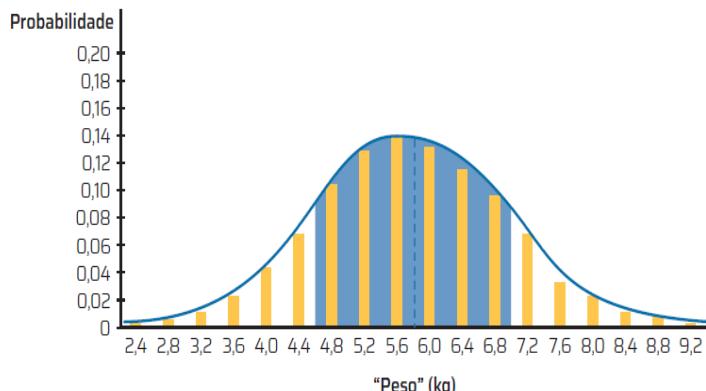
Diante disso, concluímos que as situações que contemplam o conceito da Curva Normal, são, em sua maioria, abordadas de forma explícita. Assim como pontuamos na seção sobre a Amostragem, acreditamos também que a abordagem explícita desse conceito pode facilitar as ações docentes e discentes no âmbito de seu

processo de ensino e aprendizagem. A seguir, exemplificamos Figuras que retratam essa conclusão.

**Figura 44 – Situação de explicações com abordagem explícita**

Em uma curva de distribuição, a frequência ou a probabilidade de certo intervalo é igual à área limitada pelo eixo horizontal, pelas retas verticais que passam nos extremos do intervalo e pela curva de distribuição.

No exemplo que estamos analisando, a área destacada abaixo corresponde à zona de normalidade. Ela mostra onde está concentrada a maioria dos indivíduos pesquisados.



Fonte: Smole e Diniz, Livro Matemática e suas Tecnologias (2021, p. 147).

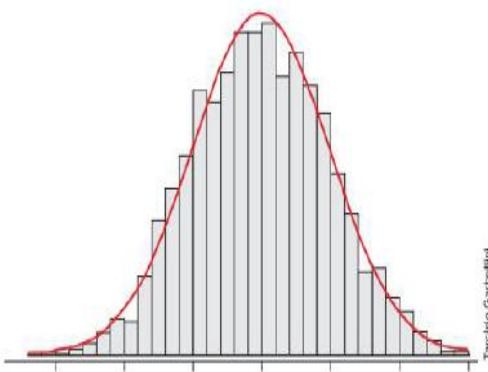
**Figura 45 – Situação de Resolução com abordagem explícita**

**Junte-se a mais três colegas para fazer esta atividade.**

**1ª parte**

Leiam e interpretem as informações a seguir.

Considerem que vocês estão diante de um histograma que representa a distribuição de um conjunto de valores, como representado abaixo. Da forma que esse histograma está, podemos traçar uma curva como a que está indicada a seguir. Em Estatística, ela é chamada de **curva normal**.



Fonte: Freitas et al, Livro Matemática e suas Tecnologias (2021, p. 33).

A Figura 44 apresenta uma situação de explicação na qual o autor aborda uma distribuição da massa (kg) de recém-nascidos, explorando a representação da Curva da distribuição dos dados e algumas propriedades do modelo da Curva Normal. Já na Figura 45 há um retrato de uma situação de resolução na qual os alunos devem

pesquisar exemplos de distribuição simétricas, a exemplo da Curva Normal, na qual as medidas de centralidade representam o centro e o eixo de simetria da distribuição.

Em prosseguimento a nossa análise, verificamos por meio da segunda categoria que as situações, nos tipos explicação e resolução, referentes ao conceito da Curva Normal contemplam mais dados fictícios (Tabela 14).

Tabela 14 – Quantitativo das situações de Curva Normal por tipo de contexto

<b>Curva Normal</b>		
<b>Contexto dos dados</b>	<b>Situação</b>	
	<b>Explicação (n = 15)</b>	<b>Resolução (n = 37)</b>
Real	4 (27%)	4 (11%)
Fictício	11 (73%)	33 (89%)

Fonte: Dados da pesquisa (2022).

Diante disso, assim como já discutido em seções anteriores, reforçamos que a maior frequência de situações com dados fictícios não é algo positivo porque não possibilita ao estudante e também ao professor, no processo de ensino e aprendizagem, a abordagem de situações reais ligadas aos seus contextos, o que pode instigar a uma maior reflexão sobre a suas realidades, como também refinar os elementos de conhecimento e de disposição relativos ao letramento estatístico.

A terceira categoria de análise, por sua vez, foi direcionada para a verificação do tipo de representação dos dados da Curva Normal, podendo ser gráfica, algébrica ou numérica. Nesse sentido, ao analisarmos as 52 situações, observamos uma maior frequência para a representação gráfica (Tabela 15).

Tabela 15 – Quantitativo das situações de Curva Normal por tipo de representação

<b>Curva Normal</b>		
<b>Contexto dos dados</b>	<b>Situação</b>	
	<b>Explicação (n = 15)</b>	<b>Resolução (n = 37)</b>
Gráfica	14 (93%)	29 (78%)
Numérica	01 (7%)	08 (22%)
Algébrica	0	0

Fonte: Dados da pesquisa (2022).

Nesse entendimento, acreditamos que a representação gráfica da Curva Normal é a forma ideal para o processo de ensino e aprendizagem desse conceito,

tendo em vista que através dessa representação torna-se mais fácil visualizar a noção da Curva, bem como as medidas abarcadas por esse modelo, principalmente a média, que corresponde ao ponto central e o desvio-padrão que indica a dispersão do conjunto, isso implica que ao variá-los, o gráfico pode apresentar movimentos de translação e achatamento. No entanto, também acreditados que os professores podem abordar em sala de aula outros tipos de representação, como a numérica, em que os dados são apresentados em sequência e a algébrica, haja vista que a distribuição normal corresponde a uma função de densidade. Lima (2009) aponta que os alunos do Ensino Médio são capazes de compreender as propriedades e as características da Curva Normal em diferentes contextos. Logo concluímos que para o seu ensino é importante o professor abordar os diferentes tipos de representação para que o aluno conceba uma distribuição normal além de sua representação gráfica.

Por fim, através da última categoria de análise, analisamos as habilidades que são exploradas nas situações envolvendo o conceito da Curva Normal (Tabela 16).

Tabela 16 – Quantitativo das situações de Curva Normal por habilidades

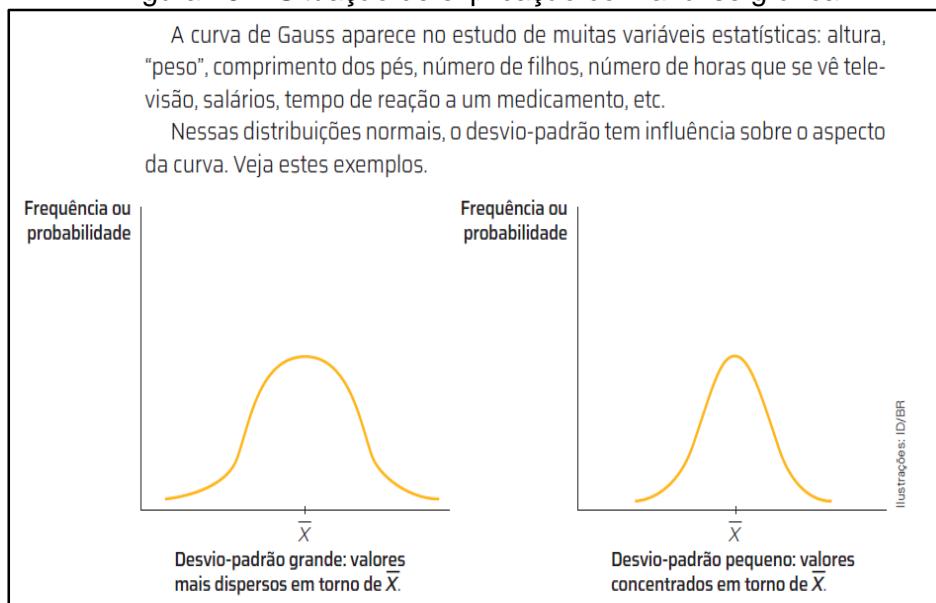
<b>Curva Normal</b>		
<b>Habilidades</b>	<b>Situação</b>	
	<b>Explicação (n = 15)</b>	<b>Resolução (n = 37)</b>
Análise Gráfica	3 (20%)	3 (8%)
Calculo ou identificação das medidas de centralidade	3 (20%)	12 (32%)
Calculo ou identificação das medidas de Dispersão	1 (7%)	7(19%)
Calculo ou identificação das medidas de Centralidade e Dispersão	1 (7%)	1(3%)
Cálculo de Probabilidades	3 (20%)	6(16%)
Cálculo da porcentagem associada a intervalos da Curva	2 (13%)	7 (19%)
Identificação da Frequênciados Dados	2 (13%)	1(3%)

Fonte: Dados da Pesquisa (2022).

Diante dessas informações, podemos observar que há uma igualdade na quantidade situações de explicação que exploram as habilidades de Análise gráfica, Cálculo ou identificação das medidas de centralidade e o Cálculo de probabilidades. Logo, assim como apontam González *et al* (2018) acreditamos que o ensino da Curva Normal, desde a Educação Básica, deve incluir a compreensão dos significados desse

conceito e dos conceitos estatísticos e probabilísticos abarcados nesse modelo em diversos contextos, para que seja possível a correta interpretação dos fenômenos e problemas que fazem uso da Curva Normal. Assim, através dessas situações o professor pode explorar a representação gráfica desse modelo (Figura 46), abordando a ideia da Curva de normalidade (Figura 47) e os cálculos das medidas de centralidade e dispersão e de probabilidades que são abarcadas pela Curva Normal (Figura 48).

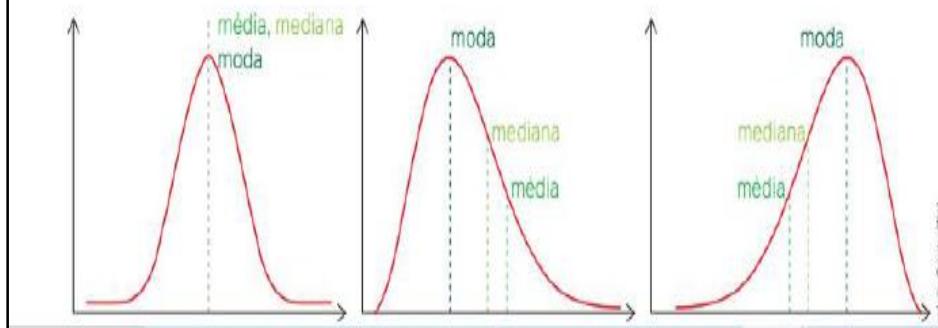
Figura 46 – Situação de explicação com análise gráfica



Fonte: Smole e Diniz, Livro Matemática e suas Tecnologias (2021, p. 148)

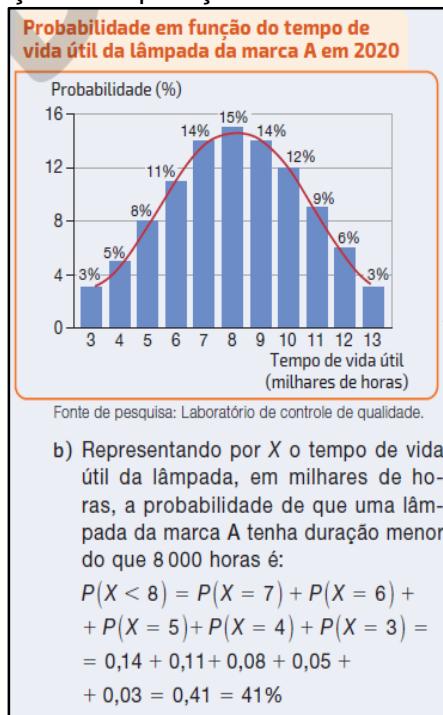
Figura 47 – Situação de explicação com o cálculo ou identificação das medidas de dispersão

Além de representarem medidas de tendência central, a média aritmética, a mediana e a moda são chamadas de medidas de posição central. Entretanto, dependendo da posição que elas ocupam uma em relação às outras, podemos classificar essa distribuição de valores (no eixo vertical está representada a frequência de valores) em três casos principais:



Fonte: Freitas et al, Livro Matemática e suas tecnologias (2021, p. 33).

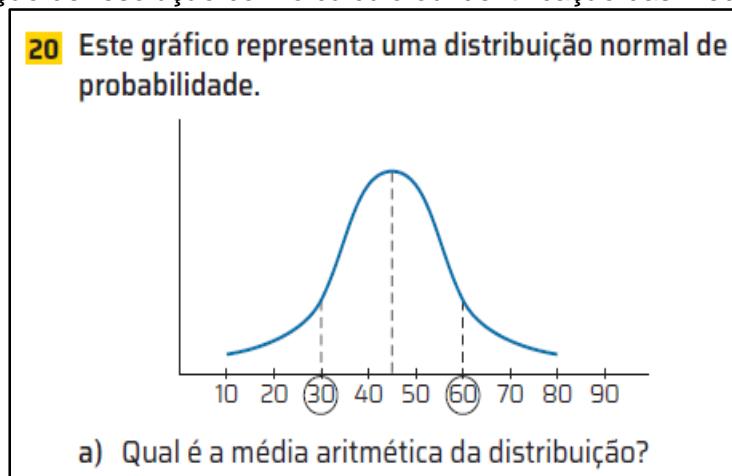
Figura 48 – Situação de explicação com o cálculo de probabilidades



Fonte: Freitas et al, Livro Matemática e suas tecnologias (2021, p. 149).

Já nas situações de resolução há uma frequência maior para o cálculo ou identificação das medidas de centralidade, que estão presentes no centro da Curva Normal e estabelecem o eixo de simetria da distribuição. Logo, a exploração de atividades que possibilitem a realização de inferências estatísticas e mais especificamente abordem as medidas de centralidade da Curva Normal, vão além da simples aplicação dos cálculos e propriedades dessas medidas e propiciam aos professores os estudantes a leitura das mesmas em diferentes fenômenos que seguem uma distribuição normal (Figura 49).

Figura 49 – Situação de resolução com o cálculo ou identificação das medidas de dispersão



Fonte: Smole e Diniz, Livro Matemática e suas tecnologias (2021, p. 151).

## 6.4 LIVROS DE CIÊNCIAS HUMANAS E MATEMÁTICA

Como já pontuado na metodologia, o segundo grupo dos livros de conhecimento é composto por 10 livros de Ciências Humanas e Matemática, os quais apresentam temas que possibilitam a interdisciplinaridade entre esses. Nessa direção, a nossa análise verificou, inicialmente, que esses livros contemplam situações referentes apenas ao conceito de Amostragem. Percebemos que, assim como nos demais livros, há uma ênfase voltada para as análises descritivas dos dados amostrais. Logo, a ausência de situações que contemplem o conceito da Curva Normal chamou nossa atenção porque frequentemente são realizadas pesquisas amostrais que dizem respeitos a diversos fenômenos sociais e que podem ser analisados pelo modelo da Curva Normal, tendo em vista que o mesmo é o principal modelo de análise de dados na Inferência Estatística.

Observamos um total de 89 situações envolvendo o conceito de Amostragem, sendo 48 de explicação e 41 de resolução. Procedemos com a nossa análise a partir das mesmas categorias adotadas nos outros livros. A seguir, apresentamos os resultados da primeira categoria, ou seja, quanto ao tipo de abordagem do conceito de amostra.

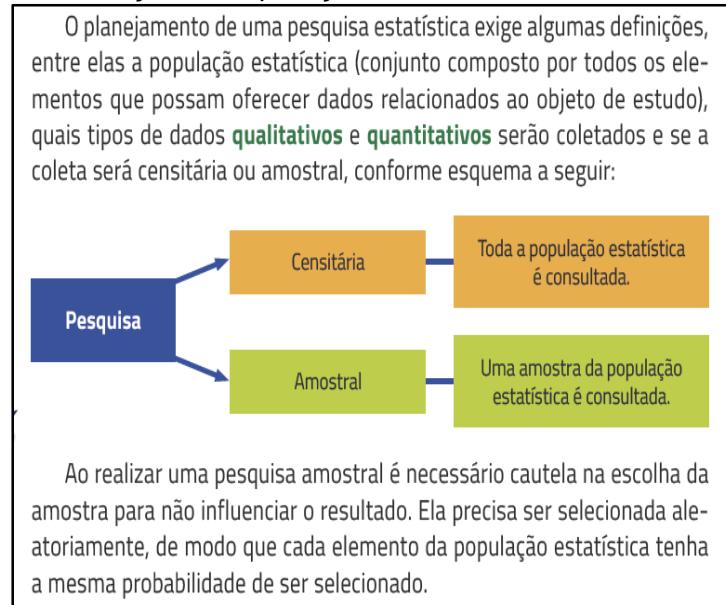
Tabela 17 – Quantitativo das situações de Amostra por tipo de abordagem

<b>Amostragem</b>		
<b>Conceito</b>	<b>Situação</b>	
	<b>Explicação (n = 48)</b>	<b>Resolução (N= 41)</b>
Explícito	44 (92%)	35 (85%)
Implícito	04 (8%)	06 (15%)

Fonte: Dados da Pesquisa (2022).

Para exemplificar uma situação de explicação com o conceito de amostra explícito (Figura 50) é abordado os processos de planejamento de uma pesquisa estatística e a diferenciação entre a pesquisa censitária e a amostral.

Figura 50 – Situação de explicação com o conceito de Amostra explícito



Fonte: Selke et al, Livro de Ciências Humanas e Matemática (2021, p. 58).

Já a Figura abaixo apresenta a situação de resolução com o conceito de amostra explícito (Figura 51) é questionado a relação que existe entre o universo e a amostra estatística e o ato de provar uma colher de um prato para avaliar a qualidade geral.

Figura 51 – Situação de Resolução com o conceito de Amostra explícito

- Explique, com suas palavras, quais as analogias possíveis entre: o tema e o objeto de uma pesquisa, o universo e a amostra estatísticos e o ato de provar uma colher de um prato para avaliar a qualidade geral.

Fonte: Salomão, Livro de Ciências Humanas e Matemática (2021, p. 82)

Diante disso, reafirmamos que uma quantidade maior de situações que abordam com o conceito de amostra explícita é um fato positivo porque permite o professor e os estudantes estabelecerem diferentes formas para a abordagem dessa temática no decorrer do livro, facilitando o seu processo de ensino e aprendizagem.

Em prosseguimento à análise, a terceira categoria se voltou para a verificação do tipo de contexto que abordagem as situações, nos tipos real ou fictício. Nessa direção, observamos que nas situações de explicação e de resolução há uma maior frequência de dados reais (Tabela 18).

Tabela 18 - Quantitativo das situações de Amostragem por tipo de contexto

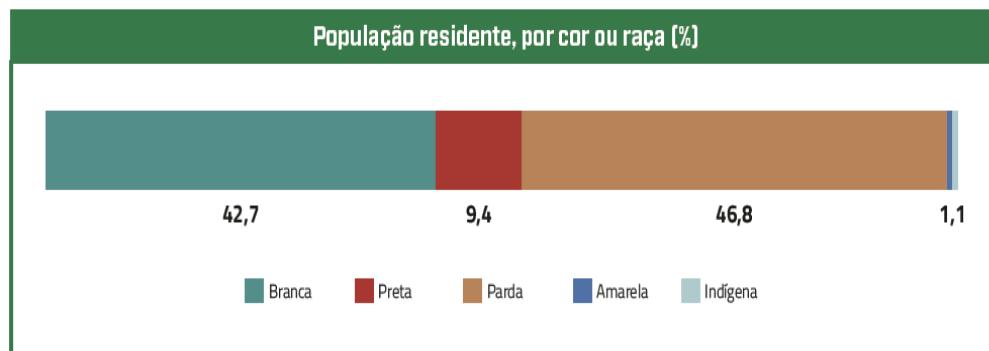
Contexto dos dados	Curva Normal	
	Situação	
	Explicação (n = 48)	Resolução (n = 37)
Real	31 (27%)	34 (11%)
Fictício	17 (73%)	03 (89%)

Fonte: Dados da pesquisa (2022).

De modo semelhante aos livros de projetos, reiteramos que a maior frequência de situações com dados reais possibilita aos estudantes refletirem com mais propriedade sobre os temas sociais que são abordados nos livros (Figura 52) que aborda o resultado de uma pesquisa sobre a cor e a raça dos brasileiros e uma proposta de realização de pesquisa estatística sobre a mobilidade urbana. Situações como essas, contribuem para um maior aprendizado sobre o contexto social no qual os estudantes estão inseridos. (Guimarães e Gitirana, 2013).

Figura 52 - Situação de explicação com Dados Reais

Em 2019, a Pesquisa Nacional por Amostra de Domicílios (PNAD) apresentou a seguinte composição populacional, fundamentada na declaração da população sobre a cor ou a raça à qual se identifica. Observe o gráfico:



EDITORIA DE ARTE

Fonte: IBGE EDUCA. Cor ou raça, Rio de Janeiro, c2020. Disponível em: <https://educa.ibge.gov.br/jovens/conheca-o-brasil/populacao/18319-cor-ou-raca.html>. Acesso em: 20 ago. 2020.

A diversidade brasileira vai muito além de traços e cores. Ela envolve também a história e as lutas de cada povo, que fazem parte de uma construção cultural que evidencia ainda a heterogeneidade dos modos e das condições de vida no país.

Fonte: Selke *et al*, Livro de Ciências Humanas e Matemática (2021, p. 58).

Em sequência, verificamos, por meio da terceira categoria de análise, que há uma predominância de situações que envolvem o conceito de população/amostra se referindo apenas a pessoas (Tabela 19).

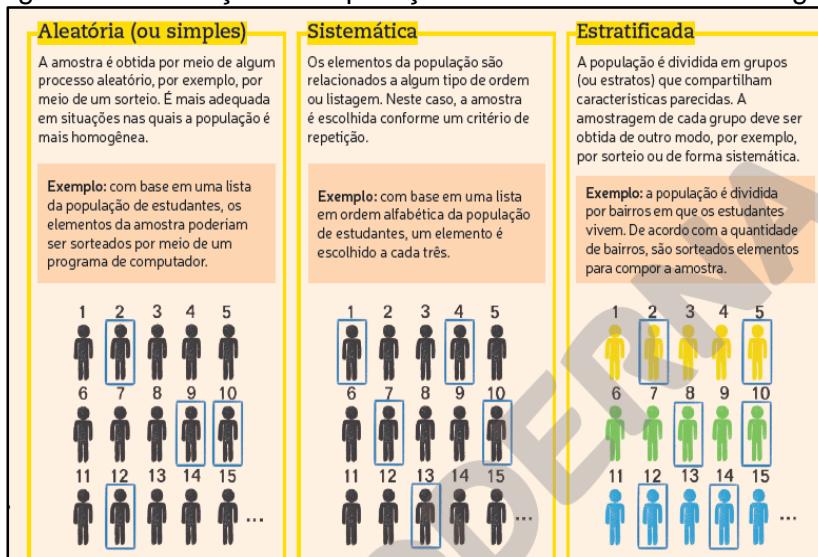
Tabela 19 – Quantitativo das situações de Amostragem por tipo de população/amostra

População/Amostra	Amostragem	
	Explicação (n = 48)	Resolução (N= 41)
Pessoas	42 (82%)	35 (80%)
Objetos	06 (18%)	06 (20%)

Fonte: Dados da pesquisa (2022).

A Figura 53 retrata uma situação de explicação contemplando o tipo de população/amostra como pessoas. Nela, o autor explica os tipos de Amostragem aleatória simples, estratificada e sistemática em um contexto que envolve estudantes como elementos da amostra.

Figura 53 – Situação de explicação com técnicas de Amostragem



Fonte: Zammataro e Neto, Livro de Ciências Humanas e Matemática (2021, p. 69)

Já a Figura 54 exemplifica uma situação de resolução na qual o tipo de população/amostra também se refere a pessoas. Nela é proposta uma pesquisa sobre possíveis melhorias no espaço escolar e, com isso, considera a população/amostra os estudantes da escola

Figura 54 – Situação de explicação com a realização de pesquisa amostral

6. Para validar a escolha das questões a ser resolvidas, vocês deverão também considerar a opinião dos outros estudantes da escola. Assim, vocês vão elaborar um questionário de uma página e aplicá-lo a uma amostra de alunos, considerando como população os estudantes de toda a escola.

Fonte: Oliveira, Livro de Ciências Humanas e Matemática (2021, p. 122)

Diante desses dados, reiteramos mais uma vez, a necessidade dos livros abordarem o conceito de população/amostra para além de pessoas, possibilitando ao estudante o entendimento que no campo estatístico, esses termos podem se referir a objetos, animais, etc. e assim, contribuir para o desenvolvimento do Letramento Estatístico.

Por fim, a última categoria de análise foi direcionada para a verificação das habilidades que são exploradas nas situações de explicação e resolução que contemplam o conceito de Amostragem (Tabela 20).

Tabela 20 - Quantitativo das situações de Amostragem por habilidades

Amostragem		
Habilidades	Situação	
	Explicação (n = 48)	Resolução (n = 41)
Seleção e Representatividade de Amostra	08 (17%)	05 (12%)
Análise dos dados de pesquisa amostral	21(44%)	09 (22%)
Realização de Pesquisa Estatística	3 (6%)	20 (49%)
Técnicas de Amostragem	13 (27%)	0
Identificar a margem de erro	02 (4%)	07 (17%)
Classificar a pesquisa em Censitária ou Amostral	01 (2%)	0
Cálculo de Medidas Estatísticas		
Comparação entre Amostras		
Identificar o tamanho da Amostra		
Classificação do tipo de Variável		
Construção de gráfico para a representação dos dados amostrais		
Cálculo de probabilidades a partir dos dados amostrais		
Contagem a partir de dados amostrais		

Fonte: Dados da Pesquisa (2022).

Podemos observar que nas situações de explicação há a predominância da habilidade voltada para a análise de dados de pesquisa amostral. Nesse sentido, percebemos que esses livros se assemelham aos livros de projetos, nos quais os

autores abordam os resultados de pesquisas estatísticas para lançar e discutir o tema com o estudante (Figura 55).

Figura 55 – Situação de explicação com a análise dos dados de pesquisa amostral

<b>Quase 4 em cada 10 jovens de 19 anos não concluíram o Ensino Médio, aponta levantamento</b> <i>Entre eles, 62% já estão fora da escola e 55% pararam de estudar ainda no Ensino Fundamental. Dados fazem parte do monitoramento do Todos pela Educação, com base na Pnad Contínua, do IBGE.</i>  <i>Quase quatro (36,5%) em cada dez brasileiros de 19 anos não concluíram o Ensino Médio em 2018, idade considerada ideal para esta etapa de ensino. Entre eles, 62% não frequentam mais a escola e 55% pararam de estudar ainda no Ensino Fundamental.</i>  <i>[...]</i> <i>Se no Ensino Médio a taxa de conclusão ainda é baixa, no Fundamental o cenário não é muito diferente: 24,2% dos jovens de 16 anos ainda não concluíram esta etapa de estudo. Entre eles, 23% não estão mais na escola.</i>
---

Fonte: Souza e Vaz, Livro de Ciências Humanas e Matemática (2021, p. 91).

Já nas situações de resolução há a predominância da habilidade voltada para a realização de pesquisas estatísticas. De modo semelhante aos livros de projetos, os autores apresentam propostas de pesquisas para que os estudantes se apropriem, com uma maior profundidade, sobre os temas sociais estudados. Assim, concordamos com Gonçalves e Lima (2017) que a realização de pesquisas estatísticas é um importante método para possibilitar a formação de sujeitos críticos frente a problemas relacionados a seus contextos, apresentando a matemática como uma ferramenta para compreensão e transformação da realidade. Para exemplificar uma situação com essa habilidade, destacamos a Figura 56 que retrata uma proposta de pesquisa para investigar aspectos relacionados à mobilidade urbana.

Figura 56 – Situação de resolução com a realização de pesquisa amostral

 <b>1</b> Formem grupos de acordo com as orientações do professor a fim de pesquisar sobre mobilidade em relação à escola, ao bairro ou ao município onde ela se localiza. Sigam as orientações.  a) Definam qual será a população a ser investigada e suas características. Além disso, é preciso estimar a quantidade de pessoas dessa população, ou seja, definir sua amostragem. b) Sigam os passos apresentados nas páginas 68 a 70 para definir uma amostra representativa da população a ser investigada e organizem-se para viabilizar a pesquisa. c) Decidam qual(is) pergunta(s) vão fazer aos entrevistados, a fim de identificar o problema de mobilidade que pretendem investigar. d) Após a coleta dos dados, organizem-nos em uma tabela e esboce um ou mais gráficos, conforme a necessidade. Em seguida, produzam um relatório que apresente os principais resultados, por meio de um texto e de representações gráficas, além das conclusões a que vocês chegaram.
--

Fonte: Zammataro e Neto, Livro de Ciências Humanas e Matemática (2021, p. 70).

Dessa forma, podemos concluir que os livros de projetos integradores e de vida e os livros de conhecimento apresentam interessantes propostas para o trabalho com a Matemática no Ensino Médio, contemplando tópicos da Estatística e da Probabilidade, principalmente na abordagem dos resultados de pesquisas e realização das mesmas.

No entanto, foi possível verificar que enquanto os livros de projetos e os livros de ciências humanas contextualizam temáticas através de exploração de pesquisas estatísticas com dados reais, os livros de Matemática e suas tecnologias tem um direcionamento para o tratamento dos cálculos de medidas estatísticas e de probabilidade com ênfase para a abordagem de dados fictícios. Além disso, todos esses livros também dão uma maior ênfase às análises descritivas dos dados amostrais e pouco exploram a relação entre a Amostragem e o modelo da Curva Normal, através da Inferência Estatística.

Acreditamos que é possível, em sala de aula e nos livros didáticos, um maior aprofundamento na Estatística Inferencial que contemple aspectos da Amostragem relacionando com o modelo da Curva Normal, através de situações que envolvam, por exemplo, a incerteza, variabilidade amostral, propriedades da Amostragem, e o modelo da Curva Normal para a distribuição de probabilidades dos dados, propiciando aos estudantes um leque maior de possibilidades para a análise de dados, a qual se constitui de uma das etapas do Ciclo Investigativo (Guimarães e Gitirana, 2013).

Destacamos que o processo de ensino e aprendizagem da Amostragem concatenado com a Curva Normal pode possibilitar, aos estudantes, o desenvolvimento do Letramento Estatístico (Gal,2002), a partir das habilidades para a análise e interpretação de uma variedade de fenômenos que estão presentes no nosso cotidiano em diversas áreas do conhecimento, além de uma articulação entre a Estatística e a probabilidade, áreas que comumente ainda são ensinadas, na Educação Básica, de forma totalmente distanciadas.

## 7 RESULTADOS DO ESTUDO 2

Neste capítulo, apresentamos os resultados do Estudo 2, o qual objetivou analisar o que foi proposto nas questões da prova de Matemática e suas Tecnologias do Exame Nacional do Ensino Médio (ENEM) de 2022 sobre Amostragem, Curva Normal e suas relações. Nesse sentido, nossa análise focou nas questões da área de Matemática e suas tecnologias que foram propostas na edição do ENEM no ano de 2022.

### 7.1 UM BREVE HISTÓRICO DO ENEM

Inicialmente, é pertinente destacarmos que o ENEM, enquanto política pública, foi criado em 1998, pelo Governo Federal do Brasil, sendo gerido pelo Instituto Nacional de Estudos e Pesquisa Educacionais (INEP), autarquia vinculada ao Ministério da Educação (MEC). O modelo inicial do ENEM tinha como objetivos:

I - conferir ao cidadão parâmetro para autoavaliação, com vistas à continuidade de sua formação e à sua inserção no mercado de trabalho; II - criar referência nacional para os egressos de qualquer das modalidades do Ensino Médio; III - fornecer subsídios às diferentes modalidades de acesso à Educação Superior; IV - constituir-se em modalidade de acesso a cursos profissionalizantes pós-médio. (Brasil, 1998, artigo 1º).

Nesse contexto, Oliveira (2016), Morrone (2011) e Soares, Soares e Santos (2021) apontam que o ENEM se consolidou, em seu início, principalmente como uma ferramenta de avaliação anual da aprendizagem dos alunos do Ensino Médio em todo o Brasil, fornecendo dados para o MEC traçar políticas públicas visando o aprimoramento da Educação Básica, em consonância com os preceitos veiculados pela Lei de Diretrizes e Bases (Brasil, 1996) e os Parâmetros Curriculares Nacionais (Brasil, 1998). O modelo de prova contemplava 63 questões interdisciplinares de múltipla escolha e a prova de redação, sendo realizado em um único dia.

Com o passar dos anos, outras políticas públicas foram atreladas ao ENEM. Em 2004, o MEC instituiu o Programa Universidade para Todos (ProUni), um tipo de programa social que passou a utilizar a nota do ENEM como critério para a concessão de Bolsas de estudos, integrais e parciais, para estudantes em universidades privadas (Soares *et al*, 2021).

Em 2009, o ENEM passou uma grande reformulação em sua estrutura. Essa mudança buscou institucionalizar uma forma de seleção nos processos seletivos das universidades públicas, democratizando as oportunidades de acesso ao ensino superior, através do sistema de seleção unificado (SISU), que consiste em uma plataforma digital pela qual as instituições públicas de Ensino Superior oferecem vagas aos candidatos que realizaram o exame. Além disso, esse “Novo ENEM” buscou pavimentar a reestruturação dos currículos do Ensino Médio e possibilitou uma grande mobilidade acadêmica, tendo em vista que através do SISU, o estudante pode concorrer a qualquer instituição vinculada ao programa e não necessariamente a de sua localidade (Pinto, 2014).

Nesse novo modelo, o ENEM passou a ser composto pela prova de redação e por quatro provas objetivas divididas por área de conhecimento: Ciências Humanas e suas Tecnologias, Ciências da Natureza e suas Tecnologias; Linguagens, Códigos e suas Tecnologias e Matemática e suas Tecnologias. Cada prova é composta por 45 questões de múltipla escolha e, assim, o ENEM passou a ser realizado em dois dias. Atualmente, além do PROUNI e SiSU, destacados anteriormente, ele também é atrelado aos seguintes programas: Fundo de Financiamento Estudantil (FIES), que oportuniza o financiamento de cursos a juros reduzidos, ENEM Portugal, que possibilita a participação em processos seletivos de instituições de educação portuguesas e o Treineiros, que oportuniza aos estudantes que ainda não concluíram o Ensino Médio conhecer o exame e avaliar seus conhecimentos (Soares et al, 2021).

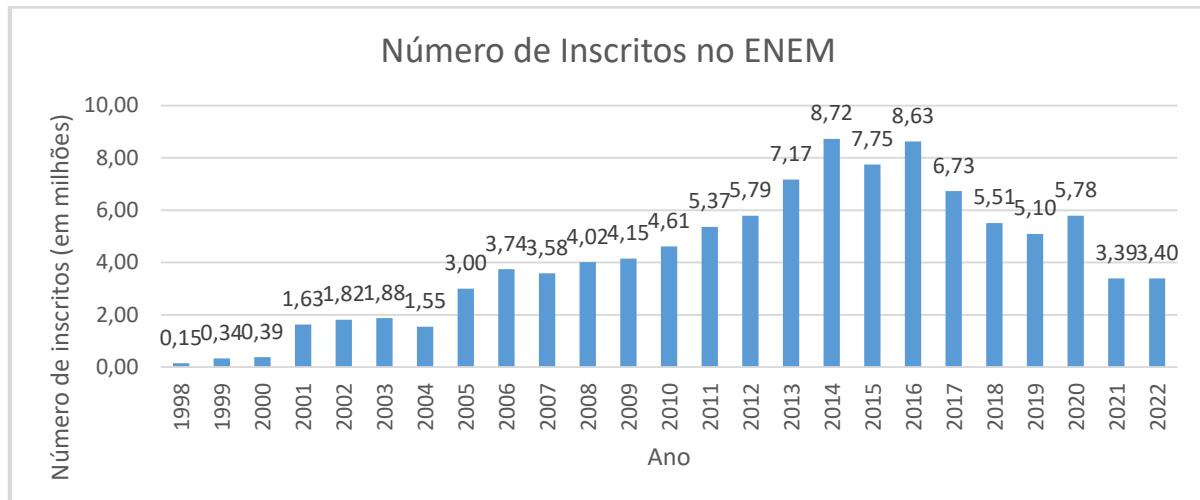
Em acréscimo, Silveira, Barbosa e Silva (2015) destacam a evolução do ENEM, em termos de principal finalidade desde a sua formulação.

Durante mais de dez anos este exame foi usado única e exclusivamente para avaliar as habilidades e competências de concluintes do Ensino Médio, sem o objetivo de selecionar para o ensino superior. Os exames de seleção, o concurso vestibular ao ensino superior, eram formulados por equipes locais país afora e formatos diferentes ocorriam nas diversas universidades. Da heterogeneidade entre os distintos concursos decorria certa diversidade cultural e de formação dos ingressantes no ensino superior. A partir de 2009 medidas governamentais estimularam o uso do ENEM não apenas como um processo de avaliação do Ensino Médio, mas como forma de acesso ao ensino superior no Brasil. O Sistema de Seleção Unificada (Sisu) passou a operar em larga escala no processo de alocação dos candidatos às vagas (Silveira; Barbosa; Silva, 2015, p 02.).

A partir da reformulação implementada em 2009, o ENEM ganhou uma considerável importância no contexto educacional brasileiro, principalmente porque se

tornou o maior vestibular realizado no Brasil e atrelado ao SISU, se consolidou como o principal instrumento de acesso ao Ensino Superior. Esse grau de importância pode ser verificado na quantidade de inscritos, ano a ano, desde a sua criação (Figura 57).

Figura 57 – Número de inscritos no ENEM em todas as edições até 2022



Fonte: INEP (2022).

Desses dados, destacamos o crescimento após a reestruturação ocorrida em 2009. Consequentemente, por ter se tornado um exame de referência, apesar de um decrescimento no número de inscrições nos últimos anos, o ENEM, desde a sua reestruturação, implicou em mudanças nas ações didático-pedagógicas nos ambientes escolares.

Andrade, Batista, Santos e Mocrosky (2020) apontam que a partir da implantação do Novo ENEM, as escolas que ofertam o Ensino Médio passaram a adotar metodologias para alcançar as especificidades da prova. Os professores, em sala de aula, têm dado mais ênfase aos conteúdos que historicamente são cobrados no exame, principalmente no 3º ano. Nesse contexto, Melão e Soares (2012) relatam que em boa parte das escolas, é comum a gestão escolar orientar os professores a incentivar os estudantes a se preparem e participarem do ENEM, bem como, abordar, em sala de aula, questões desse exame para os alunos aprofundarem seus conhecimentos.

Otobelli (2018), por sua vez, enfatiza que comumente, durante as aulas de Matemática no Ensino Médio, os professores trabalham com questões de edições anteriores do ENEM para abordar os conteúdos propostos no planejamento e, ao mesmo tempo, preparar os estudantes para se habituarem com o perfil de exigência da prova.

Diante do exposto, acreditamos que é pertinente analisar as questões de Matemática do ENEM com foco para a temática de nossa pesquisa, sobretudo, porque concordamos que no Ensino Médio e principalmente no 3º ano, muito do que é trabalhado em sala de aula tem como planejamento e meta os conteúdos abordados no referido exame.

## 7.2 MATRIZ DE REFERÊNCIA E A ANÁLISE DAS QUESTÕES

Antes de apresentarmos a nossa análise, cabe destacarmos que as questões que são propostas nas provas objetivas do Novo ENEM (Ciências Humanas e suas Tecnologias, Ciências da Natureza e suas Tecnologias; Linguagens, Códigos e suas Tecnologias e Matemática e suas Tecnologias) são fundamentadas em uma Matriz de Referência (Brasil, 2009), a qual reúne, os eixos cognitivos, competências e habilidades exigidas aos candidatos que participarem do exame.

Os eixos cognitivos são comuns a todas as áreas de conhecimento: dominar linguagens, compreender fenômenos, enfrentar situações-problema, construir argumentação e elaborar propostas (Brasil, 2009). Já as competências e as habilidades são específicas de cada área e englobam um conjunto de conhecimentos que são utilizados como norteadores para a avaliação dos estudantes. A seguir, apresentamos apenas as competências e habilidades de Matemática e suas tecnologias por se tratar da área que está relacionada diretamente como a temática de nossa pesquisa.

Competência de área 1 - Construir significados para os números naturais, inteiros, racionais e reais.

H1 - Reconhecer, no contexto social, diferentes significados e representações dos números e operações - naturais, inteiros, racionais ou reais.

H2 - Identificar padrões numéricos ou princípios de contagem.

H3 - Resolver situação-problema envolvendo conhecimentos numéricos.

H4 - Avaliar a razoabilidade de um resultado numérico na construção de argumentos sobre afirmações quantitativas.

H5 - Avaliar propostas de intervenção na realidade utilizando conhecimentos numéricos.

Competência de área 2 - Utilizar o conhecimento geométrico para realizar a leitura e a representação da realidade e agir sobre ela.

H6 - Interpretar a localização e a movimentação de pessoas/objetos no espaço tridimensional e sua representação no espaço bidimensional.

H7 - Identificar características de Figuras planas ou espaciais.

H8 - Resolver situação-problema que envolva conhecimentos geométricos de espaço e forma.

H9 - Utilizar conhecimentos geométricos de espaço e forma na seleção de argumentos propostos como solução de problemas do cotidiano.  
Competência de área 3 - Construir noções de grandezas e medidas para a compreensão da realidade e a solução de problemas do cotidiano.

H10 - Identificar relações entre grandezas e unidades de medida.

H11 - Utilizar a noção de escalas na leitura de representação de situação do cotidiano.

H12 - Resolver situação-problema que envolva medidas de grandezas.

H13 - Avaliar o resultado de uma medição na construção de um argumento consistente.

H14 - Avaliar proposta de intervenção na realidade utilizando conhecimentos geométricos relacionados a grandezas e medidas.

Competência de área 4 - Construir noções de variação de grandezas para a compreensão da realidade e a solução de problemas do cotidiano.

H15 - Identificar a relação de dependência entre grandezas.

H16 - Resolver situação-problema envolvendo a variação de grandezas, direta ou inversamente proporcionais.

H17 - Analisar informações envolvendo a variação de grandezas como recurso para a construção de argumentação.

H18 - Avaliar propostas de intervenção na realidade envolvendo variação de grandezas.

Competência de área 5 - Modelar e resolver problemas que envolvem variáveis socioeconômicas ou técnico-científicas, usando representações algébricas.

H19 - Identificar representações algébricas que expressem a relação entre grandezas.

H20 - Interpretar gráfico cartesiano que represente relações entre grandezas.

H21 - Resolver situação-problema cuja modelagem envolva conhecimentos algébricos.

H22 - Utilizar conhecimentos algébricos/geométricos como recurso para a construção de argumentação.

H23 - Avaliar propostas de intervenção na realidade utilizando conhecimentos algébricos.

**Competência de área 6 - Interpretar informações de natureza científica e social obtidas da leitura de gráficos e tabelas, realizando previsão de tendência, extração, interpolação e interpretação.**

**H24 - Utilizar informações expressas em gráficos ou tabelas para fazer inferências.**

**H25 - Resolver problema com dados apresentados em tabelas ou gráficos.**

**H26 - Analisar informações expressas em gráficos ou tabelas como recurso para a construção de argumentos.**

**Competência de área 7 - Compreender o caráter aleatório e não-determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas, determinação de amostras e cálculos de probabilidade para interpretar informações de variáveis apresentadas em uma distribuição estatística.**

**H27 - Calcular medidas de tendência central ou de dispersão de um conjunto de dados expressos em uma tabela de frequências de dados agrupados (não em classes) ou em gráficos.**

**H28 - Resolver situação-problema que envolva conhecimentos de estatística e probabilidade.**

**H29 - Utilizar conhecimentos de estatística e probabilidade como recurso para a construção de argumentação.**

**H30 - Avaliar propostas de intervenção na realidade utilizando conhecimentos de estatística e probabilidade. (Brasil, 2009, p. 5, 6 e 7)**

Com relação ao ensino de Estatística e Probabilidade, áreas que fazem parte do escopo do nosso estudo, destacamos as competências 6 e 7 e suas respectivas habilidades. Percebe-se, em linhas gerais, que o ENEM exige dos estudantes a capacidade de interpretar informações de natureza científica e social, obtidas da leitura e análise de gráficos e tabelas, e compreender o caráter aleatório e não-determinístico dos fenômenos naturais e sociais, englobando os instrumentos para o cálculo das medidas de tendência central, determinação de amostras e o cálculo de probabilidade.

Logo, enfatizamos que com relação à Amostragem, nota-se que na competência 7, há a uma referência explícita ao mencionar que o estudante deve ter domínio sobre instrumentos adequados para a determinação de amostras no contexto dos fenômenos naturais e sociais. No que diz respeito a Curva Normal, não há uma referência explícita, mas compreendemos que essas referências são consideravelmente abrangentes e que contemplam, implicitamente, o conhecimento relativo à Curva Normal, principalmente na competência 7 que cita “distribuição estatística” e nas habilidades voltadas para análise e interpretação de gráfico (H24, H25 H 26), o cálculo de medidas de tendência central ou de dispersão (H27), na resolução de situações-problema que envolva conhecimentos de estatística e probabilidade (H28) e na utilização de conhecimentos de estatística e probabilidade como recurso para a construção de argumentação (H29).

A partir desse entendimento, analisamos as 45 questões de Matemática do caderno de prova amarelo da edição do ENEM em 2022. Inicialmente, verificamos que nenhuma questão aborda o conceito de Amostragem e nem da Curva Normal. Logo, observa-se que a prova não explora situações ou fenômenos do nosso cotidiano que envolvem fundamentos da Estatística Inferencial, e mais especificamente, esses dois conceitos. Consequentemente, essa ausência pode acarretar, em sala de aula, em uma pouca abordagem, por parte do professor, para esses tópicos.

Em seguida, constatamos que na referida prova há 6 questões envolvendo a Estatística, sendo 4 voltadas para o conceito de média aritmética e 2 para o conceito

de mediana. Assim, só as medidas de tendência central foram exploradas nessa prova. Além disso, verificamos que essas questões, em linhas gerais, exploram a aplicação dos cálculos das medidas, com ênfase para a descrição dos dados. No que diz respeito ao conceito probabilidade, contabilizamos 2 questões, sendo essas com a abordagem unicamente voltada para o significado de probabilidade enquanto razão clássica.

Para exemplificar, apresentamos duas questões de cada tipo de conceito destacamos anteriormente. Com relação à média aritmética, as Figuras 58 e 59 retratam a aplicação desse conceito em situações-problemas nas quais exigem do estudante a compreensão de seu significado e a aplicação do seu algoritmo.

Figura 58 – Questão do ENEM que aborda o conceito de Média aritmética



Fonte: INEP (2022).

Nessa situação-problema é apresentado os valores da esperança de vida, em anos, ao nascer no período de 2008 a 2013. A partir das informações contida na questão, o estudante deveria interpretar que o valor registrado em cada ano corresponde à média aritmética dos valores registrado no ano antecessor e sucessor

immediatos. Assim, o valor registrado em 2013, é igual à média dos valores dos anos de 2012 e 2014. Logo, desenvolvendo o algoritmo da média aritmética que consiste em somar todos os valores da variável e dividir pelo número de dados, encontra-se que o valor registrado no ano de 2014.

Além disso, destacamos que essa questão apresentou os dados tendo como fonte o IBGE e foi a única a apresentar dados reais de todas as questões que envolvem a Estatística. Desse modo, evidencia-se que a prova do ENEM no ano de 2022, priorizou dados fictícios para elaborar as situações-problemas envolvendo essas áreas de conhecimento. Diante disso, assim como pontuado no estudo 1 e considerando que o ENEM é um exame de referência para a Educação Básica no Brasil, ressaltamos que a ausência de dados reais não é satisfatória porque não proporciona o exercício de reflexão por parte dos estudantes sobre a nossa realidade, como também, não colabora para o desenvolvimento do Letramento Estatístico, como ênfase para as atitudes, crenças e hipóteses pessoais frente a essas situações-problema. (Guimarães e Gitirana, 2013).

Figura 59 – Questão do ENEM que aborda o conceito de Média aritmética

<b>QUESTÃO 166</b>				
Jogadores	Jogo 1	Jogo 2	Jogo 3	Jogo 4
I	12	25	20	20
II	12	12	27	20
III	14	14	17	26
IV	15	18	21	21
V	22	15	23	15

O quinto e último jogo será realizado para decidir a equipe campeã e qual o melhor jogador da temporada. O jogador que precisa fazer a menor quantidade de pontos no quinto jogo, para igualar a média de pontos do melhor jogador da temporada passada, é o

**A** I.  
**B** II.  
**C** III.  
**D** IV.  
**E** V.

Fonte: INEP (2022).

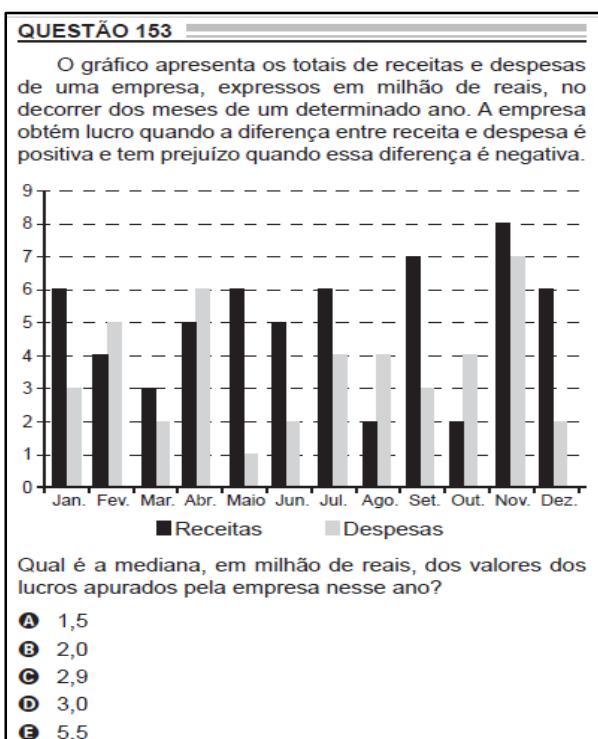
Na Figura 59 é exemplificado uma outra questão que envolve o conceito de média aritmética na qual é apresentado um quadro com dados dos pontos de 5 jogadores de basquete em 4 partidas. De modo semelhante a questão anterior, para

encontrar a resposta, o estudante deveria, inicialmente, compreender que o valor 18 corresponde à média dos pontos do melhor jogador nos últimos cinco jogos da temporada anterior, e assim, ao desenvolver o algorítmico da média aritmética, determina-se o jogador da atual temporada que precisa fazer a menor quantidade de pontos no último jogo para igualar com a referida média de pontos.

Como já pontuado anteriormente, essa questão apresentada dados fictícios para explorar o conceito de medida aritmética. Diante disso, concordamos com Pereira e Souza (2016) ao argumentarem que as questões do ENEM que envolvem a Estatística, embora sejam contextualizadas, prioritariamente exigem do estudante a aplicação dos algoritmos e processos mecânicos da Matemática, em detrimento às reflexões sobre os significados dos conceitos explorados nas situações. Essas informações, evidenciam que, tendo o ENEM como foco, o professor de Matemática pode acabar reforçando a abordagem tecnicista dos conceitos, em sala de aula, além de utilizar contextos fictícios das questões quando as exploram para lecionar os conteúdos.

No que diz respeito ao conceito de Mediana, destacamos as questões das Figuras 60 e 61. Suas abordagens contemplam contextos fictícios e ênfase para a aplicação dos cálculos a partir dos dados apresentados para que o estudante determine o valor da mediana solicitado.

Figura 60 – Questão do ENEM que aborda o conceito de Mediana



Fonte: INEP (2022).

Essa questão (Figura 60) apresenta em um gráfico de colunas, os dados de receitas e despesas de uma empresa. A partir disso, é solicitado ao estudante o valor da mediana nos meses em que a empresa teve lucro, totalizando 8 dados. Assim, o estudante deveria proceder com a colocação dos dados em ordem e, por se tratar de uma quantidade par, com a aplicação da propriedade operatória da Mediana que corresponde à média aritmética dos termos centrais.

Figura 61 – Questão do ENEM que aborda o conceito de Mediana

<b>QUESTÃO 173</b>	
<p>Uma das informações que pode auxiliar no dimensionamento do número de pediatras que devem atender em uma Unidade Básica de Saúde (UBS) é o número que representa a mediana da quantidade de crianças por família existente na região sob sua responsabilidade. O quadro mostra a distribuição das frequências do número de crianças por família na região de responsabilidade de uma UBS.</p>	
Número de crianças por família	Frequência
0	100
1	400
2	200
3	150
4	100
5	50

O número que representa a mediana da quantidade de crianças por família nessa região é

A 1,0.  
 B 1,5.  
 C 1,9.  
 D 2,1.  
 E 2,5.

Fonte: INEP (2022).

Essa segunda questão (Figura 61), por sua vez, apresenta um Quadro com a frequência do número de crianças por família e solicita ao estudante a mediana desses dados. Percebe-se que, por se tratar de um conjunto com mil dados, a questão exige do estudante a aplicação da propriedade da posição da mediana em uma sequência ordena, como também, a propriedade da Mediana para a quantidade para de dados, também exigida na questão anterior.

Diante do exposto, evidencia-se que a prova de Matemática do ENEM em 2022, no que se refere à Estatística, contempla apenas a área Descritiva, com ênfase para a aplicação de cálculos a partir, majoritariamente, de dados fictícios. Verifica-se que essas questões tem o intuito de verificar se estudante domina as habilidades voltadas as propriedades e o processo algoritmo relativos as medidas de centralidade. Logo, tendo em vista que o ENEM interfere no planejamento e prática docente em sala de aula, isso pode influenciar o professor de Matemática a priorizar a Estatística Descritiva, no Ensino Médio, com uma abordagem que se assemelha ao que é apresentado nas questões. Assim, acreditamos que o ENEM, enquanto exame de

referência para a Educação Básica, deve contemplar também em suas questões aspectos da Estatística Inferencial, reconhecendo que a Amostragem e o modelo da Curva Normal estão, em nosso cotidiano, frequentemente presentes em processos de investigação e em fenômenos de diversas áreas do conhecimento, o que torna essencial, para o Estudante em formação, refletir sobre tais conceitos e resolver situações-problema que abordem esse conhecimento. Em continuidade, destacamos, nas Figuras 62 e 63, duas questões do ENEM que contemplam a probabilidade.

Figura 62 – Questão do ENEM que aborda a Probabilidade

**QUESTÃO 152**

A *World Series* é a decisão do campeonato norte-americano de beisebol. Os dois times que chegam a essa fase jogam, entre si, até sete partidas. O primeiro desses times que completar quatro vitórias é declarado campeão.

Considere que, em todas as partidas, a probabilidade de qualquer um dos dois times vencer é sempre  $\frac{1}{2}$ . Qual é a probabilidade de o time campeão ser aquele que venceu a primeira partida da *World Series*?

(A)  $\frac{35}{64}$   
 (B)  $\frac{40}{64}$   
 (C)  $\frac{42}{64}$   
 (D)  $\frac{44}{64}$   
 (E)  $\frac{52}{64}$

Fonte: INEP (2022).

Figura 63 – Questão do ENEM que aborda a Probabilidade

**QUESTÃO 162**

Em um jogo de bingo, as cartelas contêm 16 quadrículas dispostas em linhas e colunas. Cada quadrícula tem impresso um número, dentre os inteiros de 1 a 50, sem repetição de número. Na primeira rodada, um número é sorteado, aleatoriamente, dentre os 50 possíveis. Em todas as rodadas, o número sorteado é descartado e não participa dos sorteios das rodadas seguintes. Caso o jogador tenha em sua cartela o número sorteado, ele o assinala na cartela. Ganhá o jogador que primeiro conseguir preencher quatro quadrículas que formam uma linha, uma coluna ou uma diagonal, conforme os tipos de situações ilustradas na Figura 1.

Preenchimento em linha	Preenchimento em coluna	Preenchimento em diagonal

**Figura 1**

O jogo inicia e, nas quatro primeiras rodadas, foram sorteados os seguintes números: 03, 27, 07 e 48. Ao final da quarta rodada, somente Pedro possuía uma cartela que continha esses quatro números sorteados, sendo que todos os demais jogadores conseguiram assinalar, no máximo, um desses números em suas cartelas. Observe na Figura 2 o cartão de Pedro após as quatro primeiras rodadas.

03	48	12	27
49	11	22	05
29	50	19	45
33	23	38	07

**Figura 2**

A probabilidade de Pedro ganhar o jogo em uma das duas próximas rodadas é

(A)  $\frac{1}{46} + \frac{1}{45}$   
 (B)  $\frac{1}{46} + \frac{2}{46 \times 45}$   
 (C)  $\frac{1}{46} + \frac{8}{46 \times 45}$   
 (D)  $\frac{1}{46} + \frac{43}{46 \times 45}$   
 (E)  $\frac{1}{46} + \frac{49}{46 \times 45}$

Fonte: INEP (2022).

Como podemos observar, a primeira questão apresenta um contexto real que exige do estudante o cálculo da probabilidade da união dos eventos que determinam como campeão da World Series, o time que obrigatoriamente vencer a primeira partida. Já a segunda, apresenta um contexto fictício de um jogo de bingo que, por sua vez, exige do estudante o cálculo da probabilidade da união dos eventos que fazem Pedro ganhar o jogo na quinta ou sexta partida. Como caráter mais pertinente, destacamos que ambas as questões abordam o conceito de probabilidade em sua concepção clássica, ou seja, como uma razão entre o número de casos favoráveis e o número de casos possíveis em um espaço amostral.

Lima e Borba (2019) pontuam que essa concepção de probabilidade é a mais presente na Educação Básica, tanto no trabalho, em sala de aula, como na abordagem desse tema nos livros didáticos. Logo, evidencia-se que o ENEM, em sua edição mais recente, reforça esse entendimento, o que na nossa visão não é algo satisfatório, o que pode reduzir a probabilidade apenas a esse contexto no processo de ensino e aprendizagem, haja vista que é comum os professores de Matemática utilizarem as questões do ENEM para a abordagem dos conteúdos.

Observamos ainda que nenhuma dessas questões relaciona a Estatística e da Probabilidade, o que também pode reforçar o ensino dessas áreas de forma totalmente independente na Educação Básica (Monroy e Herrera ,2019; Araújo, 2020). Assim, torna-se necessário explorar situações problemas que relacionem a Estatística e a Probabilidade, através da Inferência Estatística, oportunizando aos estudantes a construção de conhecimentos e reflexões sobre uma variedade de fenômenos em nosso cotidiano que abarcam, por exemplo, as propriedades da Amostragem, representatividade amostral, margem de erro, a influência do tamanho da amostra, o modelo da Curva Normal para a distribuição de probabilidades dos dados, dentre outros. Acreditamos que esse conhecimento deve permear a Educação Básica com vistas a uma formação que propicie, ao estudante, um olhar mais apurado para o processo de análise de dados e informações que comumente são veiculados pela mídia.

Por fim, ressaltamos que a prova do ENEM tem potencial para contribuir com o desenvolvimento do letramento estatístico e probabilístico dos estudantes, desde que contemple situações-problema com contextos reais que se aproxime do cotidiano dos mesmos a partir de elementos críticos e problematizadores. Deste modo, este exame deve não focar meramente no processo de cálculos matemáticos para se obter a resposta, mas também, instigar o processo de interpretação e reflexão dos conceitos

explorados, possibilitando o aprimoramento dos Elementos de Conhecimentos e Dispositionais (Gal, 2002,2005) com vistas a uma melhor compreensão de mundo e a tomada de decisões.

## 8 RESULTADOS DO ESTUDO 3

Neste capítulo apresentamos os resultados do Estudo 3, o qual teve por objetivo identificar os conhecimentos iniciais de estudantes do 3º ano do Ensino Médio sobre Amostragem, Curva Normal e suas relações. Nesse sentido, apresentaremos a nossa análise seguindo a ordem das questões que compõem o instrumento de coleta de dados, destacando o desempenho dos estudantes mediante as habilidades de cada item e as categorias de análises emergidas.

### Questão 1

A primeira questão apresenta três situações de pesquisas amostrais e buscou analisar a habilidade dos estudantes em identificar a população e amostra relativas às pesquisas abordadas. A Tabela 21 apresenta o percentual de resposta adequada para cada item (população e amostra) mediante as respostas dos 108 estudantes.

Tabela 21 – Percentual de resposta adequada por item da Questão 1

	<b>População</b>	<b>Amostra</b>
Pesquisa A	78%	88%
Pesquisa B	77%	83%
Pesquisa C	87%	87%

Fonte: Dados da Pesquisa (2022).

Diante desses dados, evidencia-se um bom desempenho por parte dos estudantes no processo de identificação da população e da amostra relacionadas a cada uma das pesquisas. Para cada item, o percentual de acerto foi maior que 75%. Além disso, destaca-se que, embora próximo, o percentual de acerto foi maior para a identificação da amostra nas três pesquisas. Tais resultados corroboram com os achados do estudo de Reyes *et al* (2019), no qual estudantes chilenos do Ensino Médio demonstraram reconhecer a amostra em situações contextos de pesquisa estatísticas, semelhantes as abordadas nessa questão. A Figura 64 exemplifica a resposta de um estudante com acerto para todos os itens.

Figura 64 – Resposta adequada de um estudante para a 1º Questão

Pesquisa A	Qual é a população?	Qual é a amostra?
A empresa de serviço de proteção ao crédito (SPC-Brasil) fez uma pesquisa com 600 brasileiros para analisar o perfil do consumidor consciente.	brasileiros	600
Pesquisa B	Qual é a população?	Qual é a amostra?
Uma pesquisa realizada pelo Sesc analisou 2024 blogs brasileiros com o intuito de verificar as linguagens artísticas utilizadas para a apresentação dos conteúdos.	blogs brasileiros	2024
Pesquisa C	Qual é a população?	Qual é a amostra?
14 macacos foram submetidos a uma pesquisa que tinha por finalidade testar a eficácia de uma vacina.	macacos	14

Fonte: Dados da pesquisa (2022).

Destacamos ainda que o menor percentual de acerto foi observado na identificação da população na pesquisa B. Verificamos que 15% dos alunos indicaram que a população seria os brasileiros e não os blogs. Esse dado evidencia que alguns estudantes ainda podem conceber esse conceito remetendo-se ao significado dessa palavra visto no senso comum, ou seja, se referindo apenas a pessoas (Watson e Kelly, 2005; Gomes, 2013). Logo, torna-se pertinente que a abordagem desse conceito, na Educação Estatística, retrate situações em que o significado envolva o universo de elementos que possuem pelo menos uma característica em comum, podendo envolver pessoas, como também, objetos.

## Questão 2

A segunda questão apresenta um contexto de pesquisa estatística amostral, na qual a direção de uma escola deseja selecionar uma amostra de estudantes para investigar os hábitos da utilização de celular com a finalidade de discutir com eles o uso consciente. No primeiro item, os estudantes participantes foram questionados sobre quais as características dos estudantes da escola deveriam ser consideradas para que a amostra fosse representativa. No segundo item, questionou-se como a direção da escola poderia selecionar a amostra dos estudantes para que a mesma fosse representativa, ou seja, qual o método de Amostragem deveria ser adotado para que os resultados amostrais pudessem ser generalizados para toda a escola.

Ao verificarmos o desempenho para essa questão, contatamos que no primeiro item, apenas 18% dos estudantes responderam adequadamente. Da parcela que respondeu de modo inadequado, ou seja, 82% dos estudantes, o tipo de resposta

mais frequente foi a de que a direção deveria levar em consideração apenas uma característica dos estudantes, o que não garante a representatividade da amostra. Essa conclusão é evidenciada na resposta de um estudante, apresentada na Figura 65, o qual argumenta que a direção escolar deveria se basear apenas na idade dos estudantes para seleção da amostra.

Figura 65 – Resposta Inadequada de um estudante para a 2º (A) Questão

- a) Quais características dos estudantes a direção poderia utilizar para selecionar uma amostra de modo que os resultados representem toda a escola? Justifique a escolha.

*Basear-se na idade dos alunos, verificando o grau de influência da Tecnologia conforme o Tempo*

[Basear-se na idade dos alunos, verificando o grau de influência da tecnologia conforme o tempo]

Fonte: Dados da pesquisa (2022).

Para esse item, esperava-se que os estudantes pudessem mencionar algumas características importantes para a seleção dos elementos da amostra, como o sexo, ano escolar, idade, se possui celular, região onde mora, desempenho escolar, dentre outros. Ou seja, são características inerentes à variabilidade natural dos elementos amostrais. Apesar desse resultado, como visto, 18% dos estudantes apresentaram respostas adequadas por elencarem algumas dessas características, como podemos observar na Figura 66.

Figura 66 – Resposta adequada de um estudante para a 2º (A) Questão

- a) Quais características dos estudantes a direção poderia utilizar para selecionar uma amostra de modo que os resultados representem toda a escola? Justifique a escolha.

*Gênero (homens, mulheres e não binários se portam de maneira distintas); médias bimestrais de português e matemática (para entender se o celular influencia ou não); e renda (pessoas de classes sociais distintas agem de forma distinta).*

[Gênero (homens, mulheres e não binários se portam de maneira distintas); médias bimestrais de português e matemática ( para entender se o celular influência ou não); e renda ( pessoas de classes sociais distintas agem de forma distinta)]

Fonte: Dados da pesquisa (2022).

Já no segundo item, observamos que 70% dos estudantes responderam de modo inadequado. A maioria desses estudantes mencionou um modo não probabilístico para a seleção da amostra, isto é, que não garante uma

representatividade. A Figura 67 retrata a resposta de um estudante que indica a seleção da amostra pelo método de resposta voluntária.

Figura 67 – Resposta Inadequada de um estudante para a 2º (B) Questão

b) Como a direção poderia selecionar a amostra de estudantes para representar toda a escola?

*Realizando um questionário voluntário sobre essas características*

Realizando um questionário voluntário sobre essas características

Fonte: Dados da pesquisa (2022).

No entanto, os demais estudantes citaram em suas respostas métodos probabilísticos para a seleção da amostra, possibilitando, assim, a representatividade amostral. A Figura 68 destaca a resposta de um estudante que cita o método de Amostragem por Amostra aleatória simples, ao argumentar que de cada sala deveria ser escolhido alguns estudantes aleatoriamente. Assim, apresenta um método probabilístico que, teoricamente, garante a representatividade da amostra.

Figura 68 – Resposta adequada de um estudante para a 2º (B) Questão

b) Como a direção poderia selecionar a amostra de estudantes para representar toda a escola?

*[Entrevistar aleatoriamente alguns estudantes por sala, por ter maiores chances de pegar alunos com hábitos mais variados, de modo a reunir mais dados.]*

[Entrevistar aleatoriamente alguns estudantes por sala, por ter maiores chances de pegar alunos com hábitos mais variados, de modo a reunir mais dados]

Fonte: Dados da pesquisa (2022).

Os resultados observados nessa questão evidenciam que embora esses estudantes tenham domínio sobre os conceitos de população e amostra, possivelmente ainda não tiveram uma formação, em sala de aula, que oportunizasse o ensino e aprendizagem dos aspectos que garantem a representatividade amostral, bem como, os tipos de métodos de Amostragem e sua relação com a generalização dos dados. Sobre isso, concordamos com Batanero *et al* (2018), Gomes e Guimarães (2018) e Luna e Guimarães (2021) ao mencionarem que além de compreender o significado da palavra amostra e sua relação com a população, os estudantes do Ensino Médio devem avançar no entendimento sobre a necessidade de realização

das pesquisas amostrais, a representatividade da amostra e variabilidade amostral para que possam se aprofundar nos fundamentos teóricos da Inferência Estatística.

### Questão 3

A Questão 3, por sua vez, abordou uma situação em que se desejava realizar uma pesquisa amostral, na qual é apresentada, ao estudante, 4 tipos de seleção de amostras e questiona-se qual (ais) método(s) podem ser considerados adequados ou não para se garantir uma representatividade da amostra. Nessa questão, as amostras 1 e 4 são classificadas como inadequadas porque envolvem os métodos não probabilístico, respectivamente, por conveniência e por resposta voluntária, assim, não garantem uma representatividade, como pontuado no capítulo 1 deste estudo. Já as amostras 2 e 3 são classificadas como adequadas porque envolvem um processo de seleção por métodos probabilísticos, por Amostragem Estratificada e por Aleatória Simples, respectivamente. O Quadro 7 apresenta o percentual de respostas adequadas dos estudantes para cada tipo de amostra.

Quadro 7 – Percentual de resposta adequada por item da Questão 3

Amostras	Desempenho
1	26%
2	44%
3	14%
4	27%

Fonte: Dados da Pesquisa (2022).

Na amostra 1, nota-se que apenas 26% dos estudantes responderam adequadamente, pontuando que a amostra indicada é inadequada e justificaram, em suas respostas, elencando vícios para a não garantia da representatividade, como o fato de se pesquisar apenas os estudantes que estavam passando pelo refeitório. A Figura 69 evidencia esse tipo de resposta.

Figura 69 – Resposta adequada de um estudante para a 3º (A) Questão

Amostras	Adequada ou Inadequada?	Por quê?
Amostra 1: Eduardo entrevistou 80 estudantes que estavam passando no refeitório.	Inadequada	Pois a amostra foi apenas alunos que passavam ali, e não deveria ser assim.

[Inadequada/ Pois a amostra foi apenas alunos que passavam ali, e não deveria ser assim]

Fonte: Dados da pesquisa (2022).

Além disso, 37% estudantes responderam de modo parcialmente adequado. Desse grupo, a maioria dos estudantes indicou a inadequação da amostra, porém justificou de modo inadequado, sobretudo por mencionar que o local da pesquisa, ou seja, o refeitório, não seria adequado para fazer a investigação, e sim para finalidades alimentícias, sem pontuar o fator da não representatividade que o método impõe. Os demais estudantes, também 37% do total, responderam de modo inadequado, a maioria por pontuarem que a amostra seria adequada e argumentarem, principalmente, que esse método garante a aleatoriedade na seleção dos elementos amostrais. As Figuras 70 e 71 exemplificam, respectivamente, essas duas conclusões.

Figura 70 – Resposta Parcialmente Adequada de um estudante para a 3º (A) Questão

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Eduardo entrevistou 80 estudantes que estavam passando no refeitório.	Inadequada	Pois é muito improvável eles não se interessarem visto que o objetivo era comer no refeitório.

[Inadequada/ Pois é muito improvável eles não se interessarem visto que o objetivo era comer no refeitório]

Fonte: Dados da pesquisa (2022).

Figura 71 – Resposta Inadequada de um estudante para a 3º (A) Questão

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Eduardo entrevistou 80 estudantes que estavam passando no refeitório.	Adequada	Foi uma análise feita de forma aleatória

[Adequada/ Foi uma análise feita de forma aleatória]

Fonte: Dados da pesquisa (2022).

Em continuidade a nossa análise, as respostas para a amostra 2 foram a que apresentaram o melhor desempenho, isto é, 44% os estudantes responderam adequadamente ao indicarem se tratar de uma amostra adequada e justificarem que a seleção contempla alunos de sexos e cursos diferentes, embora não tenham citado que se trata do método estratificado. Esse índice pode estar relacionado ao fato de que os estudantes perceberam e refletiram que a amostra garante uma representatividade por contemplar a variabilidade dos elementos no que se refere ao sexo e ao curso, como podemos observar na Figura 72 que exemplifica a resposta de um estudante.

Figura 72 – Resposta adequada de um estudante para a 3º (B) Questão

<b>Amostra 2:</b> Sabendo que a escola tinha o mesmo número de homens e mulheres, Luana entrevistou 80 estudantes, sendo 40 homens e 40 mulheres, escolhendo-os, aleatoriamente, de cada curso de idiomas que a escola possui atualmente (inglês e espanhol).	<i>adequado</i>	<i>abrange ambos os sexos e pessoas de cada curso de idiomas que a escola possui</i>
---	-----------------	--

Adequada/ Abrange ambos os sexos e pessoas de cada curso de idiomas que a escola possui

Fonte: Dados da pesquisa (2022).

Em acréscimo, 42% dos estudantes responderam parcialmente adequado. A maioria dos estudantes desse grupo, apesar de enxergarem adequação na amostra, não justificaram com critérios que ressaltassem a representatividade da amostra. Por exemplo, a Figura 73 retrata a resposta de um estudante que indica a amostra como adequada, mas justifica apenas que dizendo que a mesma foi abrangente, sem pontuar o porquê.

Figura 73 – Resposta Parcialmente adequada de um estudante para a 3º (B) Questão

<b>Amostra 2:</b> Sabendo que a escola tinha o mesmo número de homens e mulheres, Luana entrevistou 80 estudantes, sendo 40 homens e 40 mulheres, escolhendo-os, aleatoriamente, de cada curso de idiomas que a escola possui atualmente (inglês e espanhol).	<i>Adequada</i>	<i>foi abrangente</i>
---	-----------------	-----------------------

Adequada/ Foi abrangente

Fonte: Dados da pesquisa (2022).

Ainda com relação a amostra 2, 15% dos estudantes responderam de modo inadequado por indicarem que essa amostra não seria adequada. A Figura 74 exemplifica a resposta de uma estudante que justifica indicando que o processo de seleção da amostra não deveria ter divisão dos elementos de acordo com gênero. Esse tipo de resposta evidência que alguns desses estudantes não conhecem os métodos de Amostragem, e este, em particular, o método estratificado.

Figura 74 – Resposta Inadequada de um estudante para a 3º (B) Questão

<b>Amostra 2:</b> Sabendo que a escola tinha o mesmo número de homens e mulheres, Luaná entrevistou 80 estudantes, sendo 40 homens e 40 mulheres, escolhendo-os, aleatoriamente, de cada curso de idiomas que a escola possui atualmente (inglês e espanhol).	Inadequada	Não deve-se escolher isso pelo gênero.
---	------------	--

Fonte: Dados da pesquisa (2022).

Já as respostas para a amostra 3 apresentaram o pior desempenho. Observamos que 14% dos estudantes responderam adequadamente e indicaram que a amostra seria adequada para se garantir a representatividade e justificaram, em linhas gerais, argumentando que, por ser um sorteio aleatório, esse método permite que todos os elementos tenham a mesma chance de compor amostra de forma justa e imparcial, como podemos observar uma exemplificação na Figura 75.

Figura 75 – Resposta adequada de um estudante para a 3º (C) Questão

<b>Amostra 3:</b> Natália tinha o nome de todos os 800 estudantes, colocou-os em um chapéu e então tirou, aleatoriamente, 80 deles para realizar a entrevista.	<i>Adequada</i>	O sorteio é uma forma parcialmente justa de analisar esses estudantes
--	-----------------	---

Fonte: Dados da pesquisa (2022).

Além disso, constatamos que 17% dos estudantes responderam de modo parcialmente adequado. Isso significa que apontaram a amostra 3 como adequada, mas ao justificarem, argumentaram com critérios alheios, não apresentando elementos que justificam a amostra aleatória simples, como é o caso da amostra 3, como representativa. Em caráter de exemplo, a Figura 76 retrata a resposta de um estudante que apenas justifica que o método é interessante.

**Figura 76 – Resposta Parcialmente adequada de um estudante para a 3º (C) Questão**

<b>Amostra 3:</b> Natália tinha o nome de todos os 800 estudantes, colocou-os em um chapéu e então tirou, aleatoriamente, 80 deles para realizar a entrevista.	Adequada	Porque é uma maneira interessante. Interessante assim com a amostra 4.
--	----------	--

Fonte: Dados da pesquisa (2022).

Por fim, 62% dos estudantes responderam de modo inadequado ao pontuarem que se tratava de uma amostra inadequada e justificaram, sobretudo, que o sorteio realizado não seria uma boa maneira para selecionar os elementos da amostra. Esse dado evidencia que essa parcela dos estudantes não tem domínio em torno do significado do método da Amostra Aleatória Simples, ou não reconhecem esse método como probabilístico, ou seja, que garanta uma representatividade. A Figura 77 exemplifica essa conclusão.

Figura 77 – Resposta Inadequada de um estudante para a 3º (C) Questão

<b>Amostra 3:</b> Natália tinha o nome de todos os 800 estudantes, colocou-os em um chapéu e então tirou, aleatoriamente, 80 deles para realizar a entrevista.	<i>Inadequado</i>	<i>Uma seleção baseada em sorte,</i>
Inadequada/ Uma seleção baseada em sorte]		

Fonte: Dados da pesquisa (2022).

No que diz respeito à amostra 4, verificamos que 27% dos estudantes responderam de modo adequado por indicarem que amostra é inadequada e justificarem, de modo geral, que esse método não garante a representatividade porque só quem compõem a amostra são os elementos que se voluntariaram a participar da pesquisa, assim, não reflete os resultados para toda a população, como podemos observar na Figura 78.

Figura 78 – Resposta adequada de um estudante para a 3º (D) Questão

<b>Amostra 4:</b> Renato enviou um questionário para todos os estudantes e depois usou os 80 primeiros que foram devolvidos a ele.	<i>Inadequada</i>	<i>Esta pesquisa pode ter deixado de fora muitas opiniões de alunos que não poderiam responder de imediato.</i>
Inadequada/ Essa pesquisa pode ter deixado de fora muitas opiniões de alunos que não poderiam responder de imediato]		

Fonte: Dados da pesquisa (2022).

Além do mais, 36% tiveram suas respostas caracterizadas como parcialmente adequadas por indicarem que a amostra 4 é inadequada, mas justificaram com critérios alheios ao fator não probabilístico. Como exemplo, um estudante (Figura 79) indica a não adequação da amostra pelo fato da mesma ser algo muito padrão e necessita de uma criatividade.

Figura 79 – Resposta parcialmente adequada de um estudante para a 3º (D) Questão

<b>Amostra 4:</b> Renato enviou um questionário para todos os estudantes e depois usou os 80 primeiros que foram devolvidos a ele.	<i>inadequado</i>	<i>muito padrão, poderia algo mais criativo</i>
[Inadequada/ Muito padrão, poderia algo mais criativo]		

Fonte: Dados da pesquisa (2022).

Já 37% dos estudantes responderam de modo inadequado por afirmarem que se tratava de uma amostra adequada, em termos de representatividade, e, consequentemente, justificaram inadequadamente. Na Figura 80 representa a resposta do estudante que indica adequação da amostra porque contou com a disponibilidade e boa vontade dos elementos da população para compô-la. Logo, evidencia-se que alguns estudantes não têm domínio sobre o caráter não representativo da Amostragem não probabilística, em particular, a amostra voluntária.

Figura 80 – Resposta inadequada de um estudante para a 3º (D) Questão

<b>Amostra 4:</b> Renato enviou um questionário para todos os estudantes e depois usou os 80 primeiros que foram devolvidos a ele.	<i>ADEQUADA</i>	<i>POIS FOI FEITA COM BASE NA DISPONIBILIDADE E BOA VONTADE DOS PARTICIPANTES</i>
Adequada/ Pois foi feita com base na disponibilidade e boa vontade dos participantes		

Fonte: Dados da pesquisa (2022).

Diante dos resultados apresentados, enfatizamos que os resultados nessa terceira questão vão ao encontro do que se concluiu mediante as respostas apresentadas na segunda questão. Evidencia-se que, embora a maioria dos estudantes, compreendam o que significa população e amostra, eles não apresentam domínio conceitual sobre os métodos de seleção de amostra, demonstrando não ter familiaridade com os significados e a diferença entre os métodos probabilísticos e não probabilísticos e o quanto eles interferem na variabilidade e, consequentemente, na representatividade amostral. Assim como pontuado anteriormente, esse cenário indica que, possivelmente, esses estudantes ainda não tiveram a oportunidade de estudarem sobre os métodos de seleção de amostras em pesquisas e suas implicações para a generalização dos resultados na população de interesse.

#### 4º questão

Na quarta questão é apresentada uma situação que aborda distribuição norma das medidas da pressão arterial de 900 pessoas. A partir do gráfico apresentado, questionou-se se uma pessoa que possui a pressão arterial igual a 80 mmHg representa o valor médio do grupo. Logo, os estudantes deveriam identificar a média como o centro da Curva e julgar se o valor da pressão discutido é igual à média.

Ao analisarmos as respostas dos estudantes, observamos que 23% de modo adequado, 44% de modo parcialmente adequado e 33% de modo inadequado. A parcela dos estudantes que respondeu adequadamente (Figura 81) indicou que o valor de 80 mmHg não é igual a média porque não está no centro da Curva e sim mais à esquerda da Distribuição. Logo, esses estudantes realizam uma análise adequada do gráfico, reconhecendo a distribuição como uma normalidade e julgando que o valor não corresponde ao centro da Curva. Assim, esses estudantes evidenciam o domínio sobre a relação entre a representação gráfica da Curva Normal e a posição das medidas de tendência central.

Figura 81 – Resposta adequada de um estudante para a 4º Questão

[Apresenta valor abaixo na média, a média é 130]

Fonte: Dados da pesquisa (2022).

Dos que responderam parcialmente adequado, o tipo de resposta mais mencionado, foi que o valor da pressão da pessoa não representa a média do grupo, no entanto, assim, como no grupo anterior, se baseando nos valores indicados no eixo vertical do gráfico, o qual indicava a frequência e não os valores dos dados das pressões, presentes no eixo horizontal. Logo, esse grupo de estudantes também apresentou dificuldades para realizar uma leitura adequada do gráfico da Curva Normal e reconhecer a média como o centro da Distribuição. Para exemplificar esse tipo de resposta, destacamos a Figura 82.

Figura 82 – Resposta Parcialmente Adequada de um estudante para a 4º Questão

mas, pois as pessoas com a pressão 80, é quase inexistente.

[Não, pois as pessoas com a pressão 80 é quase inexistente]

Fonte: Dados da pesquisa (2022).

Já na parcela que respondeu inadequadamente, o principal tipo de resposta mencionada foi que o valor da pressão arterial da pessoa mencionada no enunciado da questão representa a média do grupo e justificaram com base nos valores da frequência dos dados indicadas no eixo vertical do gráfico, no qual o valor 80 que indica a frequência da pressão 160 está na região intermediária desse eixo. (Figura 83). Esse resultado indica que essa parcela dos estudantes não tem domínio sobre as características da representação gráfica da Curva Normal e, consequentemente, não realizam a leitura adequada para identificar a média dos dados no eixo horizontal e no centro da Curva.

Figura 83 – Resposta inadequada de um estudante para a 4º Questão

Sim, pois a pressão dela ficou em "equilíbrio"  
se comparada as outras.

[ Sim, pois a pressão dela ficou em “equilíbrio” se comparada com os outros]

Fonte: Dados da pesquisa (2022).

## 5º questão

Nessa questão é abordada uma situação-problema envolvendo o conceito de dispersão e média através de duas representações gráficas, sendo uma delas uma aproximação e a outra uma Curva Normal dos dados da temperatura de duas cidades. Sabendo que a média é a mesma para as duas distribuições, os estudantes deveriam indicar, através dos gráficos apresentados, qual cidade apresenta maior desvio-padrão das temperaturas, analisando com relação ao centro, ou seja, a média da distribuição.

Por meio das respostas apresentadas, verificamos que 25% dos alunos responderam adequadamente, indicando maior desvio-padrão na cidade B e justificando com base na maior dispersão dos dados em relação ao centro da

distribuição. No entanto, 74% dos estudantes responderam de modo inadequado ao indicarem que a cidade A possuía maior desvio-padrão e 1% respondeu de modo parcialmente adequado por indicar a cidade B e não apresentar justificativa.

Dos estudantes que responderam adequadamente, observamos que a resposta mais mencionada foi a indicação da cidade B pelo fato dela possuir uma maior variação de temperaturas em relação à média, centro da Distribuição. Alguns estudantes também fizeram a mesma indicação e justificaram com base na maior amplitude, ou seja, maior diferença entre a maior e menor temperatura em comparação com a cidade A. Logo, esses estudantes analisaram adequadamente a relação entre a média da distribuição e o desvio-padrão dos dados, seja em relação ao centro da distribuição ou pela amplitude dos dados. A Figura 84 retrata o tipo de resposta mais mencionado para essa categoria.

Figura 84 – Resposta adequada de um estudante para a 5º Questão

<p><i>Por apresentar um maior afastamento da média, a cidade 2 acaba tendo uma escala de temperaturas ao longo dos dias maior</i></p> <p>[ Por apresentar um maior afastamento da média, a cidade 2 acabou tendo uma escala de temperatura ao longo dos dias maior]</p>
---

Fonte: Dados da pesquisa (2022).

Do grupo que respondeu de modo inadequado, a resposta mais mencionada foi a cidade A pelo fato da mesma apresentar uma maior oscilação na frequência dos dados, verificada no eixo vertical do gráfico. Assim, essa parcela de estudante não procedeu com a análise correta do gráfico, ao refletirem apenas sobre a frequência dos dados e não atentarem para o valor das temperaturas registradas no eixo horizontal, o qual implicitamente contempla a relação entre a média, centro da distribuição e o desvio-padrão dos dados. Esse resultado vai ao encontro das conclusões verificadas na questão 4, isto é, uma parcela dos estudantes apresentou lacunas na análise do gráfico, não assimilando de modo adequado o que os eixos determinam na representação gráfica e as conclusões que podem ser obtidas a partir dessa representação. A Figura 85 exemplifica esse tipo de resposta.

Figura 85 – Resposta Inadequada de um estudante para a 5º Questão

Cidade a. Significa que essa cidade tem uma grande oscilação de temperatura.

[ Cidade a. Significa que essa cidade tem uma grande oscilação de temperatura]

Fonte: Dados da pesquisa (2022).

Em linhas gerais, esses resultados corroboram com as conclusões evidenciadas na questão quatro. A maioria desses estudantes não apresenta domínio no que se refere ao gráfico da Curva Normal e, consequentemente, não realizam a leitura gráfica adequada para relacionar a média com o desvio-padrão dos dados. Isso também pode indicar que possivelmente, não foi oportunizado, a esses estudantes, no âmbito da Educação básica e mais precisamente no Ensino Médio, um processo de ensino e aprendizagem que contemplasse o conceito da Curva Normal, sua representação gráfica, bem como, as medidas de centralidade e dispersão que são abarcadas nesse modelo.

## 6º questão

Essa questão contempla um gráfico da distribuição das notas que os estudantes do Ensino Médio de uma escola pública de Pernambuco obtiveram na disciplina de Matemática no segundo bimestre de 2022. É informado que a distribuição é aproximadamente normal e a sua média é 6,7. No item A, perguntou-se aos estudantes qual a probabilidade aproximada de selecionarmos um estudante dessa escola, ao acaso, e ele possuir a nota de matemática maior ou igual a média. Para esse questionamento, os estudantes deveriam reconhecer que metade da Curva (nota maior ou igual que a média) tem probabilidade aproximada de 50%, pelo fato da distribuição está bem próxima da normalidade. Já no item B, considerando a possibilidade de o diretor escolar selecionar uma amostra, ao acaso, de 30 estudantes do Ensino Médio, indagou-se se a probabilidade de que a média amostral seja igual a 4,5 é baixa ou alta. Para essa pergunta, os estudantes deveriam reconhecer que a média amostral tende a se aproximar da média populacional, e, consequentemente, compreender que a probabilidade solicitada para a nota é baixa por se distanciar da média.

Ao analisarmos as respostas dos estudantes ao item A, constatamos, inicialmente, que apenas 18% respondeu adequadamente e os demais, ou seja, 82% responderam de modo inadequado. Para o primeiro grupo, constatamos que os estudantes indicaram que a probabilidade questionada era de 50%, isso indica que analisaram de modo adequado o gráfico da questão e apresentam domínio sobre uma das propriedades da Curva Normal, a que associa a área da Curva com a probabilidade da variável em questão assumir um valor dentro de um determinado intervalo. A Figura 86 apresenta a resposta de um estudante com a referida classificação.

Figura 86 – Resposta adequada de um estudante para a 6º (A) Questão

Probabilidade de 50% serem aprovados na média ou acima da média.

[ Probabilidade de 50% serem aprovados na média ou acima da média]

Fonte: Dados da pesquisa (2022).

Já aqueles que responderam de forma inadequada, a resposta mais mencionada foi valores para probabilidade acima de 50%, o que não corresponde à metade da Curva, ou seja, nota maior ou igual que a média. Essa conclusão, mais uma vez, evidencia o que apresentamos nas questões quatro e cinco, de que os estudantes participantes apresentam lacunas no que diz respeito a Curva Normal e suas propriedades. A Figura 87 exemplifica essa conclusão.

Figura 87 – Resposta Inadequada de um estudante para a 6º (A) Questão

80% dos alunos estão acima da média

[80% dos alunos estão acima da média]

Fonte: Dados da pesquisa (2022).

Com relação ao item B, constatamos que apenas 5% dos estudantes responderam de modo adequado. Além disso, 3% deles responderam de modo parcialmente adequado e 92% responderam de modo inadequado. A parcela que respondeu adequadamente justificou que a média da amostral tende a ser próxima à da população e por isso a probabilidade questionada é baixa porque a nota 4,5 está

distante do centro da Curva, ou seja, afastava da média populacional que é de 6,7. Logo, evidencia-se que essa parcela de estudantes analisou os dados da questão e refletiram de modo adequado sobre a relação entre a média amostral e a populacional. A Figura 88 exemplifica essa conclusão.

Figura 88 – Resposta Adequada de um estudante para a 6º (B) Questão

*baixa, já que a amostra costuma ser próxima a retratada na realidade.  
Sendo assim, deve oscilar entre 6,7*

[Baixa, já que a amostra costuma ser próxima a retratada na realidade. Sendo assim, deve oscilar entre 6,7]

Fonte: Dados da pesquisa (2022).

Os estudantes que responderam parcialmente adequado indicaram que a probabilidade era baixa, porém não apresentaram justificativa. Já no grupo que respondeu de modo inadequado, a resposta mais mencionada foi que a probabilidade é baixa porque a frequência da nota 4,5 é baixa (Figura 89). Aqui, destaca-se que esse tipo de resposta se caracterizou inadequada porque esses estudantes confundiram a distribuição da variável nota, que indica a frequência dos valores, com a distribuição amostral das médias, a qual obedece ao Teorema Central do Limite. Logo, esse resultado evidencia que a maioria dos estudantes participantes, possivelmente, não tiveram a oportunidade de refletir, mesmo que de modo informal, como se dá a distribuição das médias amostrais e sua relação com a média da população, o que permite, sala de aula, relacionar a Amostragem com o modelo da Curva Normal.

Figura 89 – Resposta Inadequada de um estudante para a 6º (B) Questão

*É baixa pois o número de alunos com 4,5 foi pouco.*

[É baixa pois o número de alunos com 4,5 foi pouco]

Fonte: Dados da pesquisa (2022).

Apresentados os resultados do nosso estudo diagnóstico, destacamos por fim, e em linhas gerais, que a maioria desses estudantes demonstraram domínio sobre o significado da palavra amostra e sua relação com a população, porém não

demonstram avançar para além da compreensão desses significados. Considerando que o questionário aplicado teve uma pontuação máxima de 23 pontos, o desempenho médio de acertos foi igual a 10,9. Logo, evidenciou-se lacunas sobre a compreensão da representatividade da amostra, abarcando a variabilidade amostral e os métodos de Amostragem, como também, no modelo da Curva Normal, contemplando sua representação gráfica e propriedades, a exemplo das medidas de centralidade e dispersão e o cálculo de probabilidades. Além disso, nota-se fragilidades na compreensão da relação entre a distribuição das médias amostrais com os parâmetros populacionais. Diante desse Quadro, reforçamos a importância de se oportunizar, aos alunos do Ensino Médio, um processo de ensino e aprendizagem que conte cole os fundamentos da Amostragem concatenados ao modelo da Curva Normal para possibilitar o entendimento desse conhecimento nos vários fenômenos do nosso cotidiano, além de favorecer uma maior leitura crítica dos dados com vistas a uma melhor tomada de decisões e o desenvolvimento do letramento Estatístico e Probabilístico.

## 9 RESULTADOS DO ESTUDO 4

Neste capítulo, apresentamos os resultados do Estudo 4, o qual objetivou analisar a contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada no 3º ano do Ensino Médio.

Relembramos ainda que este estudo contemplou três etapas: pré-teste, a intervenção de ensino e o pós-teste, sendo desenvolvido em duas turmas do 3º ano do Ensino Médio. Para a resolução das atividades na etapa da intervenção, uma turma teve como recurso o Lápis e Papel e a outra o software Geogebra.

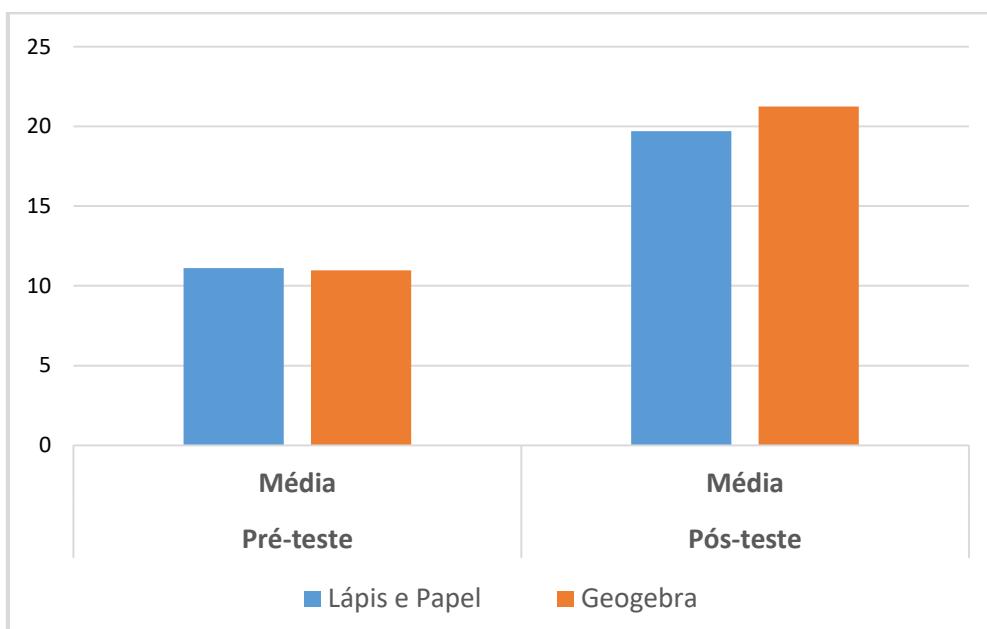
Iniciamos com a apresentação e análise do desempenho quantitativo geral das duas turmas no pré-teste e pós-teste, como também, o detalhamento do desempenho por cada questão e habilidade investigada nesses testes. Além de sistematizar esses dados quantitativos, também analisamos as respostas dos estudantes de forma qualitativa, buscando evidenciar como eles resolveram as questões e suas concepções.

Por último, descrevemos detalhadamente as ações desenvolvidas durante os momentos da intervenção de ensino, com ênfase para as atividades e as discussões coletivas realizadas, as quais evidenciam as contribuições dos recursos empregados para a aprendizagem dos estudantes participantes sobre os conceitos relacionados à Amostragem e à Curva Normal.

### 9.1 DESEMPENHO GERAL DAS TURMAS NO PRÉ-TESTE E NO PÓS-TESTE

Conforme detalhado em nossa metodologia, tanto o pré-teste quanto o pós-teste consistiram em 17 itens, abrangendo habilidades relacionadas aos conceitos de Amostragem e Curva Normal, com um total de 25 pontos. No pré-teste, observamos que o desempenho médio da turma Lápis e Papel foi de 11,12, enquanto a turma Geogebra alcançou uma média de 10,97. No entanto, no pós-teste, ambas as turmas apresentaram avanços significativos. A turma Lápis e Papel obteve uma média de 19,71, enquanto a turma Geogebra atingiu uma média de 21,24 (Gráfico 1).

Gráfico 1 – Desempenho das turmas no pré-teste e pós-teste



Fonte: Dados da pesquisa (2023).

Com base nesses dados, é perceptível que ambas as turmas apresentaram baixo desempenho no pré-teste, com médias próximas e abaixo de 50% do total de pontos. Além disso, o desvio-padrão das notas da turma Lápis e Papel foi 3,8 e da turma Geogebra foi igual a 3,6. Portanto, nesta fase inicial, as turmas tiveram desempenho semelhantes e ficaram evidentes lacunas no conhecimento sobre Amostragem e Curva Normal.

Além disso, o fato das médias das turmas no pré-teste serem similares vai ao encontro do que uma pesquisa experimental exige, isto é, as turmas participantes devem ter desempenho semelhantes no pré-teste. Consequentemente, isso garante que as turmas sejam equivalentes antes da intervenção de ensino e que os eventuais progressos verificados no pós-teste possam ser atribuídos ao processo interventivo e não meramente a fatores prévios (Shaughnessy, Zechmeister e Zechmeister, 2012).

Para garantir essa equivalência, realizamos uma análise estatística denominada teste-t para amostras independentes (Independent Samples Test), o qual, segundo Dancey e Reidy (2006), é utilizado quando há dois grupos independentes e deseja-se verificar se as diferenças entre as médias são discrepantes o suficiente para que se diga se elas apresentam diferenças significativas. O teste – t indicou que não há diferença significativa entre as duas turmas no pré-teste,  $t(70) = .164; p \leq .871$ . Logo, concluímos que as duas turmas apresentaram desempenho equivalentes.

Já no pós-teste, nota-se um aumento no desempenho médio, o qual, nas duas turmas, corresponde a mais de 75% do total de pontos. Embora que a turma Lápis e papel registrou um aumento no desvio-padrão das médias, indo a 4,6, ao passo que na turma Geogebra observou-se uma diminuição, com desvio-padrão igual a 3,1, indicando uma menor dispersão dos resultados. De modo geral, esse dado evidencia a contribuição da intervenção de ensino para o desenvolvimento da aprendizagem sobre a Amostragem e a Curva Normal. Além disso, averiguamos se as diferenças entre as médias de acertos obtidas por cada turma no pré-teste e pós-teste foram significativas, estatisticamente. Para tanto, realizamos o Teste-t para amostras emparelhadas ou relacionadas (*Paired Samples Test*). Essa análise se aplica quando se deseja avaliar se existe uma diferença significativa entre as amostras em condições relacionadas ou emparelhadas, ou seja, quando há dois eventos e os participantes tomam parte de ambos (Dancey e Reidy, 2006).

Nessa direção, verificamos que tanto na turma que utilizou o lápis e papel, test-t [ $t(33) = 7.607; p \leq .000$ ], como na turma que recorreu ao software Geogebra, test-t [ $t(37) = 13.544; p \leq .000$ ], houve uma diferença altamente significativa entre os desempenhos do pré-teste e do pós-teste. Logo, ratifica-se, com um olhar estatístico, que efetivamente a intervenção de ensino contribuiu para a construção e ampliação dos conhecimentos sobre a Amostragem e a Curva Normal, independentemente do recurso utilizado.

Além disso, também verificamos, estatisticamente, se houve diferença significativa entre as turmas exclusivamente entre os desempenhos no pós-teste. Para isso, recorremos mais uma vez ao teste-t para amostras independentes (*Independent Samples Test*). O resultado indicou que não há diferença significativa entre as duas turmas no pós-teste, test-t [ $t(70) = 1.650; p \leq .104$ ]. Diante disso, conclui-se que as duas turmas apresentam desempenhos equivalentes no pós-teste e ratifica-se que, independente do recurso utilizado, a intervenção de ensino contribuiu significativamente para a aprendizagem dos estudantes.

Apresentadas as análises gerais quantitativas do pré-teste e do pós-teste, passaremos a abordar, nas próximas seções, as análises quantitativas e qualitativas de cada questão desses dois testes.

### 9.1.1 Desempenho dos estudantes na questão 1 no pré-teste e pós-teste

A primeira questão do pré-teste e pós-teste teve o objetivo de investigar a habilidade dos estudantes em identificar a população e a respectiva amostra em três pesquisas estatísticas. O quadro 8 apresenta o percentual de resposta adequada para cada item (população e amostra) de cada pesquisa abordada, mediante as respostas dos estudantes das duas turmas.

Quadro 8 – Percentual de resposta adequada por item da questão 1

Item	Turma/Fase da Pesquisa			
	Lápis e papel		Geogebra	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Pesquisa A - População	82,5	94,0	71,0	89,5
Pesquisa A – Amostra	82,5	88,0	79,0	87,0
Pesquisa B - População	76,5	91,0	71,0	89,5
Pesquisa B – Amostra	82,5	85,5	82,0	87,0
Pesquisa C - População	79,5	91,0	66,0	89,5
Pesquisa C – Amostra	85,5	85,5	82,0	87,0

Fonte: Dados da Pesquisa (2023).

Esses resultados notabilizam que o desempenho relativo a identificação da população e da amostra que já era satisfatório no pré-teste evoluiu no pós-teste em cada item, com exceção na amostra da pesquisa C na turma do Lápis e papel, que se manteve igual. Destaca-se que o menor percentual de respostas adequadas no pré-teste na turma de Lápis e Papel, se verifica na identificação da população da pesquisa B, e na turma do Geogebra na identificação da população da pesquisa C. Nesses casos, a população envolveu elementos que não são pessoas, respectivamente, blogs e animais. Logo, uma parcela dos estudantes, de ambas as turmas, apresentava fragilidades em identificar o conceito de população estatística quando o mesmo estava inserido em contextos que não abarcavam pessoas, explicitando seu conhecimento estatístico e suas crenças (Gal, 2002). Estes relacionavam o conceito unicamente a um conjunto de pessoas, como é concebido no senso popular. Porém, nesses dois casos, verifica-se um aumento no percentual de respostas adequadas no pós-teste. Em linhas gerais, podemos concluir que após a intervenção de ensino, a maioria dos estudantes, aprimorou ainda mais o conhecimento sobre os conceitos de população e amostra. Tendo em vista que a maioria dos estudantes, das duas turmas, respondeu

adequadamente em ambos os testes, destacamos na a resposta de um estudante da turma de Lápis e Papel (Figura 90) no pré-teste e a resposta de um estudante da turma do Geogebra no pós-teste (Figura 91).

Figura 90 – Resposta adequada de um estudante para a 1ºquestão do pré-teste – turma Lápis e Papel

Pesquisa A	Qual é a população?	Qual é a amostra?
A empresa de serviço de proteção ao crédito (SPC-Brasil) fez uma pesquisa com 600 brasileiros para analisar o perfil do consumidor consciente.	OS CONSUMIDORES BRASILEIROS	600 BRASILEIROS
Pesquisa B	Qual é a população?	Qual é a amostra?
Uma pesquisa realizada pelo Sesc analisou 2024 blogs brasileiros com o intuito de verificar as linguagens artísticas utilizadas para a apresentação dos conteúdos.	BLOGS BRASILEIROS	2024 BLOGS BRASILEIROS
Pesquisa C	Qual é a população?	Qual é a amostra?
14 macacos foram submetidos a uma pesquisa que tinha por finalidade testar a eficácia de uma vacina.	MACACOS	14 MACACOS

Fonte: Dados da Pesquisa (2023).

Figura 91 – Resposta adequada de um estudante para a 1º questão do pós-teste – turma Geogebra

Pesquisa A	Qual é a população?	Qual é a amostra?
Para descobrir o nível de satisfação dos usuários de transporte público do Recife, entrevistou-se 2.400 pessoas que utilizam esse meio de locomoção.	Usuários de trans- porte público do Recife	2.400 usuários de transporte pú- blico do Recife
Pesquisa B	Qual é a população?	Qual é a amostra?
Foram testados 8 modelos de carros populares para avaliar a segurança desses veículos.	Carros populares	8 modelos de carro populares
Pesquisa C	Qual é a população?	Qual é a amostra?
77 gatos participaram de uma pesquisa que tinha por objetivo saber se eles reconheciam o próprio nome mesmo quando chamados por um estranho.	Gatos	77 gatos

Fonte: Dados da Pesquisa (2023).

### 9.1.2 Desempenho dos estudantes na questão 2 no pré-teste e pós-teste

A segunda questão dos testes envolveu um contexto de realização de pesquisa amostral e foi composta por dois itens. O primeiro item teve o objetivo de analisar o conhecimento dos estudantes sobre a variabilidade natural amostral (Reyes, 2019) e questionou-se sobre quais as características dos elementos amostrais deveriam ser consideradas para que a amostra fosse representativa. Já o segundo item teve o objetivo de analisar o conhecimento sobre os métodos de Amostragem e questionou-se como a amostra deveria ser selecionada de modo a se garantir a representatividade (Quadro 9).

Quadro 9 – Percentual de resposta adequada por item da questão 2

Item	Turma/Fase da Pesquisa			
	Lápis e papel		Geogebra	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
A	38,0	82,0	32,0	82,0
B	9,0	62,0	10,5	66,0

Fonte: Dados da Pesquisa (2023).

No que refere ao primeiro item do pré-teste, os resultados apontam que a maioria dos estudantes, de ambas as turmas, respondeu de modo inadequado, demonstrando não terem conhecimento sobre a variabilidade natural presente em amostras. Nessa questão, nos dois testes, foi abordado contextos de pesquisas amostrais com estudantes, o que contempla algumas características essenciais para se garantir a representatividade, a exemplo da série/turma, sexo, classe social, etc. Constatamos que a maioria daqueles que responderam de forma inadequada indicou apenas uma característica a ser considerada, o que para as pesquisas amostrais não garante a representatividade estatística. Begué *et al* (2020) enfatizam que a variabilidade é um fator importante para se compreender a representatividade amostral e que os estudantes da Educação Básica geralmente apresentam dificuldades na compreensão desse conceito. Como exemplo, na resposta de um estudante da turma de Lápis e Papel para esse item nota-se (Figura 92) que ele indica em sua resposta apenas a idade como característica necessária para a seleção de estudantes em uma pesquisa amostral sobre o uso consciente do celular. Logo, demonstrou, nessa fase do nosso estudo experimental, fragilidades no conhecimento

estatístico (Gal,2002) inerente a variabilidade amostral relacionada com a representação amostral.

Figura 92 – Resposta inadequada de um estudante para a 2º (A) questão do pré-teste – turma Lápis e Papel

Escolher estudantes com idades variadas, pois assim representarão vários.

Fonte: Dados da Pesquisa (2023).

Já na turma Geogebra a resposta (Figura 93) de um estudante exemplifica a mesma conclusão para esse item. Ele menciona que a direção escolar deveria selecionar os estudantes para a amostra com base apenas na responsabilidade/compromisso com os estudos, na ideia de que esses alunos seriam um modelo de comportamento para o uso do celular, o que não garantia a representatividade amostral.

Figura 93 – Resposta inadequada de um estudante para a 2º (A) questão do pré-teste – turma Geogebra

A direção poderia escolher os alunos mais responsáveis, pois através deles poderiam saber em quais aulas seriam necessárias o uso dos mesmos e também saber administrar o tempo (tempo para estudar e tempo para uso do telefone) nos intervalos, agrando, dessa forma a todos.

A direção poderia escolher alunos mais responsáveis, pois através deles poderiam saber em quais aulas seriam necessárias o uso dos mesmos e também saber administrar o tempo (tempo para estudar e tempo para o uso do telefone, nos intervalos, agrando, dessa forma a todos).

Fonte: Dados da Pesquisa (2023).

No entanto, percebe-se que no pós-teste o percentual de respostas adequadas no primeiro item deu um grande salto, alcançando 82% nas duas turmas. Esse resultado indica que a intervenção de ensino proporcionou, aos estudantes de ambas as turmas, a compreensão de que a variabilidade natural está inerente ao processo de pesquisa inferencial, sendo necessário o pesquisador conhecê-la e considerá-la, a partir do contexto envolvido, para se garantir a representatividade. A resposta adequada (Figura 94) do mesmo estudante da turma Lápis e Papel abordado anteriormente demonstra no pós-teste a construção do conhecimento estatístico

(Gal,2002) sobre algumas características necessárias, ou seja, a variabilidade natural da amostra relatada na questão. Em sua resposta, ele menciona o sexo, a idade, a turma e a cidade onde mora como características que devem ser consideradas em uma pesquisa sobre os hábitos musicais dos estudantes relacionados com a cultura popular.

Figura 94 – Resposta adequada de um estudante para a 2º (A) questão do pós-teste – turma Lápis e Papel

<p><i>Sexo, idade, turma e cidade onde mora. A seleção baseada nessas características vai garantir uma pesquisa que represente a escola.</i></p> <p>Sexo, idade, turma e a cidade onde mora. A seleção baseada nessas características vai garantir uma pesquisa que represente a escola.</p>
--

Fonte: Dados da Pesquisa (2023).

Na turma Geogebra a conclusão é semelhante. O mesmo estudante discutido anteriormente no pré-teste, respondeu adequadamente no pós-teste (Figura 95), elencando algumas características dos elementos amostrais que compõem a variabilidade natural para compor a amostra da referida pesquisa, ou seja, o sexo, idade e a turma. A sua resposta também menciona termos como “representatividade” e “variabilidade”, demonstrando ter avançado tanto no conhecimento estatístico, como nas habilidades de letramento (Gal,2002) para se comunicar sobre a Estatística.

Figura 95 – Resposta adequada de um estudante para a 2º (A) questão do pós-teste – turma Geogebra

<p><i>Sexo; idade; turma → Dessa forma a escolha será mais representativa (utilizando esses critérios e diferenciando eles, para poder ter uma maior variabilidade)</i></p> <p>Sexo, idade, turma. Dessa forma a escolha será mais representativa (utilizando esses critérios e diferenciando eles, para poder ter uma maior variabilidade)</p>
---

Fonte: Dados da Pesquisa (2023).

Já no item B, a maioria dos estudantes de ambas as turmas respondeu de modo inadequado no pré-teste, sobretudo, por indicarem que deveria ser formado uma amostra, mas não apresentaram elementos da Amostragem em suas respostas que indicasse um método de Amostragem a ser utilizado. Isso evidencia que esses estudantes não dominavam o conhecimento estatístico (Gal,2002) relativo aos métodos de seleção de amostra. Percebemos que, em linhas gerais, os estudantes

de ambas as turmas se preocuparam em indicar uma quantidade de elementos amostrais a serem selecionados, trazendo uma ideia de uma parte do todo dos elementos, a partir de um grupo/amostra. As respostas a seguir (Figuras 96 e 97) exemplificam, respectivamente, a concepção de um estudante da turma de Lápis e papel e um estudante da turma Geogebra para esse item.

Figura 96 – Resposta inadequada de um estudante para a 2º (B) questão do pré-teste – turma Lápis e Papel

*Toma certa quantidade de cada turma*

Uma certa quantidade de cada turma

Fonte: Dados da Pesquisa (2023).

Figura 97 – Resposta inadequada de um estudante para a 2º (B) questão do pré-teste – turma Geogebra

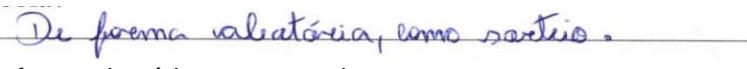
*Pegando pequena parte dos alunos de cada sala, assim, toda a escola seria representada por esses alunos*

Pegando pequena parte dos alunos de cada sala, assim, toda a escola seria representada por esses alunos.

Fonte: Dados da Pesquisa (2023).

Porém, a partir dos dados do pós-teste é perceptível um avanço em ambas as turmas sobre o conhecimento relacionado aos métodos de Amostragem, incluindo os probabilísticos que tornam as amostras representativas. A maioria dos estudantes, das duas turmas, respondeu adequadamente por, diferentemente do pré-teste, trazerem em suas respostas elementos que caracterizam uma seleção amostral por um método probabilístico. Isso notabiliza que a intervenção de ensino proporcionou as estudantes, independente do recurso utilizado, um aprofundamento sobre os métodos de Amostragem, contribuindo para o desenvolvimento dos elementos de conhecimento e de disposição do Letramento Estatístico (Gal, 2002). A resposta (Figura 98) do mesmo estudante da turma Lápis e Papel abordado anteriormente no pré-teste evidencia essa conclusão.

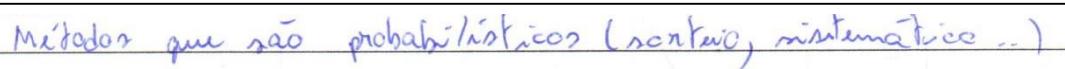
Figura 98 – Resposta adequada de um estudante para a 2º (B) questão do pré-teste – turma Lápis e Papel

 De forma aleatória, como sorteio.
---

Fonte: Dados da Pesquisa (2023).

Nota-se que no pós-teste o estudante responde adequadamente ao indicar que a seleção deveria ser aleatória, para garantir a representatividade, ou seja, indica que o método deveria ser probabilístico. Além disso, citada o sorteio que remete ao método de Amostragem Aleatória Simples, como um exemplo. De modo semelhante, a resposta (Figura 99) do mesmo estudante da turma Geogebra apresentado anteriormente para essa questão no pré-teste notabiliza a mesma conclusão. Percebe-se que ele menciona que os métodos de seleção deveriam ser probabilísticos para garantir a representatividade e indica os exemplos do sorteio (aleatório simples) e o sistemático. Diante disso, evidencia-se que a intervenção de ensino possibilitou a construção do conhecimento estatístico sobre os métodos de Amostragem, e consequentemente, o desenvolvimento do letramento estatístico (Gal,2002) dos estudantes.

Figura 99 – Resposta inadequada de um estudante para a 2º (B) questão do pré-teste – turma Geogebra

 Métodos que são probabilísticos (sorteio, sistemático...)
---

Fonte: Dados da Pesquisa (2023).

### 9.1.3 Desempenho dos estudantes na questão 3 no pré-teste e pós-teste

O objetivo da terceira questão de ambos os testes foi analisar como os estudantes julgavam a adequação de métodos de Amostragem para se garantir a representatividade em contextos de pesquisa amostral. Nela, são apresentadas 5 amostras, das quais, as amostras 1 e 4 são consideradas como inadequadas para a representatividade porque envolvem métodos não probabilísticos, respectivamente, por conveniência e por resposta voluntária. Por outro lado, as amostras 2, 3 e 5 são classificadas como adequadas para se garantir a representatividade da pesquisa amostral porque envolvem processos de seleção amostral por métodos

probabilísticos, respectivamente, Amostragem Estratificada, Amostragem Aleatória Simples e Amostragem Sistemática. A seguir, apresentamos (Quadro 10) o percentual de resposta adequada das duas turmas para cada item dessa questão nos dois testes.

Quadro 10 – Percentual de resposta adequada por item da questão 3

Item	Turma/Fase da Pesquisa			
	Lápis e papel		Geogebra	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Amostra 1	23,5	68,0	18,5	82,0
Amostra 2	44,0	91,0	50,0	95,0
Amostra 3	21,0	62,0	27,0	82,0
Amostra 4	18,0	62,0	21,0	84,0
Amostra 5	12,0	53,0	21,0	87,0

Fonte: Dados da Pesquisa (2023).

Esses dados nos mostram que para todas as amostras houve uma evolução muito significativa no percentual de respostas adequadas do pré-teste para o pós-teste nas duas turmas. Isso demonstra que a intervenção de ensino contribuiu para o aprimoramento do Letramento Estatístico (Gal,2002) dos estudantes, seja nos elementos de conhecimento ou de disposição, haja vista que tanto para realizar uma pesquisa amostral ou analisar criticamente os resultados da mesma, é necessário ter o domínio sobre os tipos de métodos de Amostragem e suas características do conhecimento dos estudantes sobre os tipos e métodos de seleção amostral. Destaca-se que o maior percentual de acerto se verifica na amostra estratificada nos dois testes e nas duas turmas. Esse tipo de Amostragem por envolver a divisão da população em subgrupos, ou seja, extratos e posteriormente a seleção aleatória dos elementos de cada grupo, pode ser um facilitador do entendimento de que o mesmo é probabilístico e que garante a representatividade.

No que diz respeito a turma Lápis e Papel, verifica-se um desempenho insatisfatório no pré-teste na análise das amostras. Isso reforça as conclusões obtidas na questão anterior, de que a maioria desses estudantes desconheciam os métodos de seleção amostra seja probabilístico ou não probabilístico. No pós-teste se verifica uma evolução significativa no percentual das cinco amostras. Porém, considerando que na amostra 5, o resultado no pós-teste é um pouco maior que a metade, uma parcela de estudantes dessa turma demonstra ainda necessitar de um contato formal maior sobre as características conceituais do método sistemático para avançar ainda

mais na compreensão. A resposta (Figura 100) de um estudante da turma Lápis e Papel no pré-teste notabiliza essa conclusão.

Figura 100 – Resposta parcialmente adequada de um estudante para a 3º questão do pré-teste – turma Lápis e Papel

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Eduardo entrevistou 80 estudantes que estavam passando no refeitório.	Inadequada	Não possuía requisitos para entrevistar
<b>Amostra 2:</b> Sabendo que a escola tinha o mesmo número de homens e mulheres, Luana entrevistou 80 estudantes, sendo 40 homens e 40 mulheres, escolhendo-os, aleatoriamente, de cada curso de idiomas que a escola possui atualmente (inglês e espanhol).	Inadequada	Poderia ser feito de melhor forma
<b>Amostra 3:</b> Natália tinha o nome de todos os 800 estudantes, colocou-os em um chapéu e então tirou, aleatoriamente, 80 deles para realizar a entrevista.	Inadequada	As entrevistas feitas foram aleatórias
<b>Amostra 4:</b> Renato enviou um questionário para todos os estudantes e depois usou os 80 primeiros que foram devolvidos a ele.	Adequada	Foi feito um filtro das pessoas a serem entrevistadas
<b>Amostra 5:</b> Arthur enumerou e ordenou todos os 800 alunos da escola em uma lista. Em seguida, a cada intervalo de 10 alunos, ele selecionou o sexto, totalizando 80 entrevistados.	Inadequada	Não selecionou de uma forma específica a serem entrevistados

Amostra 1: Inadequada/ Não possuía requisitos para entrevistar  
 Amostra 2: Inadequada/ poderia ser feito de melhor forma  
 Amostra 3: Inadequada/ As entrevistas feitas foram aleatórias  
 Amostra 4: Adequada/ Foi feito um filtro das pessoas a serem entrevistadas  
 Amostra 5: Inadequada/ Não selecionou de uma forma específica a serem entrevistados

Fonte: Dados da pesquisa (2023).

Percebe-se que o estudante responde adequadamente apenas a amostra 1, indicando, ao seu modo, que a mesma é inadequada para garantir a representatividade e justifica que esse método (conveniência) não possuía requisitos, ou seja, a seleção dos entrevistados era sem critério. Já as demais amostras ele responde de forma inadequada, ao mencionar que as amostras 2, 3 e 5 que são probabilísticas não garantiriam a representatividade e amostra 4, que não é probabilística traria uma representação. Destaca-se ainda que o estudante justifica que a amostra 3, aleatória simples, é inadequada porque tem o caráter aleatório. Constatamos que a maioria dos estudantes das duas turmas indicaram que essa amostra não traria a representação porque o sorteio não era um método adequado ou justo. Isso demonstra que na etapa do pré-teste a maioria dos estudantes não

conheciam as características dos métodos de Amostragem e explicitaram ideias conforme a suas crenças e experiências (Gal, 2002; Innabi, 2007).

Já no pós-teste (Figura 101) esse estudante da turma Lápis e Papel demonstra ter avançado e aprofundado seu conhecimento estatístico sobre os métodos de Amostragem.

Figura 101 – Resposta adequada de um estudante para a 3º questão do pós-teste – turma Lápis e Papel

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Alice entrevistou 60 funcionários que estavam passando em um dos corredores da empresa.	Inadequada	Pois priorizou funcionários passando nos corredores
<b>Amostra 2:</b> Sabendo que a empresa tinha o mesmo número de funcionários homens e mulheres, Gustavo entrevistou 30 homens e 30 mulheres, escolhendo-os aleatoriamente, de cada setor da empresa, totalizando 60 funcionários.	Adequada	Pois foi feito de forma aleatória
<b>Amostra 3:</b> Fernanda tinha o nome de todos os 600 funcionários, colocou-os em um chapéu e então tirou, aleatoriamente, 60 deles para realizar a entrevista.	Adequada	Pois foi realizado de forma aleatória
<b>Amostra 4:</b> Luís enviou um questionário para todos os funcionários e depois usou os 60 primeiros que foram devolvidos a ele.	Inadequado	Pois priorizou quem respondeu o formulário
<b>Amostra 5:</b> Carol enumerou e ordenou todos os 600 funcionários da empresa em uma lista. Em seguida, a cada intervalo de 10 funcionários, ela selecionou o oitavo, totalizando 60 entrevistados.	Adequado	Pois foi aleatório não dando privilégio a nenhum funcionário
Amostra 1: Inadequada/ pois priorizou funcionários passando nos corredores		
Amostra 2: Adequada/ pois foi feito de forma aleatória		
Amostra 3: Adequada/ pois foi realizado de forma aleatória		
Amostra 4: Inadequado/ pois priorizou quem respondeu o formulário		
Amostra 5: Adequado/pois foi aleatório não dando privilégio a nenhum funcionário		

Fonte: Dados da pesquisa (2023).

A partir dessa resposta, nota-se que o estudante passa conceber a forma aleatória de seleção dos elementos amostrais como algo adequado para garantir a representação. Assim, ele responde que as amostras 2, 3 e 5 são adequadas para a representatividade, como também, indica os vícios nas amostras 1 e 4 que as tornam

inadequadas para o mesmo propósito. Logo, a intervenção de ensino proporcionou aos estudantes a construção do conhecimento estatístico, bem como, uma postura crítica (Gal,2002) para julgar, a partir de contextos de pesquisas, a adequação de alguns métodos visando a representação amostral.

Já na turma Geogebra, também se verifica um desempenho insatisfatório no pré-teste, evidenciando que a maioria dos estudantes dessa turma também não conheciam as características dos métodos de Amostragem abordados, como pode ser observado na resposta (Figura 102) de um estudante dessa turma para a terceira questão.

Figura 102 – Resposta parcialmente adequada de um estudante para a 3º questão do pré-teste – Geogebra

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Eduardo entrevistou 80 estudantes que estavam passando no refeitório.	Inadequada	O ambiente onde é feito conta muito, no refeitório aos alunos provavelmente não iam dar muita atenção
<b>Amostra 2:</b> Sabendo que a escola tinha o mesmo número de homens e mulheres, Luana entrevistou 80 estudantes, sendo 40 homens e 40 mulheres, escolhendo-os, aleatoriamente, de cada curso de idiomas que a escola possui atualmente (inglês e espanhol).	Adequada	Ela foi nos alunos que já estavam aprendendo outros idiomas e incentivou eles a aprender mais, no caso o Francês agora.
<b>Amostra 3:</b> Natália tinha o nome de todos os 800 estudantes, colocou-os em um chapéu e então tirou, aleatoriamente, 80 deles para realizar a entrevista.	Inadequada	Esse forma de "sorteio" acredito que não seja a melhor opção
<b>Amostra 4:</b> Renato enviou um questionário para todos os estudantes e depois usou os 80 primeiros que foram devolvidos a ele.	Adequada	Esses 80 foram os que apresentavam ter mais interesse por darem importância ao questionário
<b>Amostra 5:</b> Arthur enumerou e ordenou todos os 800 alunos da escola em uma lista. Em seguida, a cada intervalo de 10 alunos, ele selecionou o sexto, totalizando 80 entrevistados.	Inadequada	Novamente não acredito que essa forma de selecionar não seja a melhor na ocasião

Amostra 1: Inadequada/ O ambiente onde é feito conta muito, no refeitório aos alunos provavelmente não iam dar muita atenção  
 Amostra 2: Adequada/ Ela foi nos alunos que já estavam aprendendo outros idiomas e incentivou eles a aprender mais, no caso o Francês agora.  
 Amostra 3: Inadequada/ Essa forma de "sorteio" acredito que não seja a melhor opção  
 Amostra 4: Adequada/ Esses 80 foram os que apresentavam ter mais interesse por darem importância ao questionário.  
 Amostra 5: Inadequada/ Novamente não acredito que essa forma de selecionar não seja a melhor na ocasião

Fonte: Dados da pesquisa (2023).

Percebe-se que o estudante responde adequadamente apenas na amostra 1 (por conveniência), justificando, ao seu modo, os vícios nesse processo de seleção. Chamamos a atenção para a justificativa da amostra 3, ou seja, a aleatória simples. Assim como pontuamos na resposta do estudante da outra turma, verifica-se que esse estudante também concebia o sorteio como algo que não era adequado para trazer a representação. Esse tipo de resposta também foi mencionado por outros estudantes dessa turma. Isso evidencia as crenças e atitudes relacionadas com o conhecimento estatístico (Gal,2002) que esses estudantes possuíam na etapa do pré-teste.

Porém, constata-se que no pós-teste dessa turma, o percentual de resposta adequada aumentou significativamente, com índice superior a 80% para todas as amostras. Logo, conclui-se que intervenção de ensino possibilitou o aprofundamento de seus conhecimentos sobre os métodos de Amostragem e suas relações com a representatividade amostral. A resposta (Figura 103) do mesmo estudante da turma Geogebra abordado no pré-teste anteriormente.

Figura 103 – Resposta adequada de um estudante para a 3º questão do pós-teste – turma Geogebra

Amostras	Adequada ou Inadequada?	Por quê?
<b>Amostra 1:</b> Alice entrevistou 60 funcionários que estavam passando em um dos corredores da empresa.	Inadequada	Não é adequada pois ela usou um método por conveniência dela de uma forma que não dá para representar um grupo grande de pessoas.
<b>Amostra 2:</b> Sabendo que a empresa tinha o mesmo número de funcionários homens e mulheres, Gustavo entrevistou 30 homens e 30 mulheres, escolhendo-os aleatoriamente, de cada setor da empresa, totalizando 60 funcionários.	Adequada	É um dos métodos adequados (probabilísticos)
<b>Amostra 3:</b> Fernanda tinha o nome de todos os 600 funcionários, colocou-os em um chapéu e então tirou, aleatoriamente, 60 deles para realizar a entrevista.	Adequada	Sorteio é um método justo.
<b>Amostra 4:</b> Luís enviou um questionário para todos os funcionários e depois usou os 60 primeiros que foram devolvidos a ele.	Inadequada	Dessa forma não é justo pois nem todos têm acesso e acaba se tornando um método para um grupo mais fechado.
<b>Amostra 5:</b> Carol enumerou e ordenou todos os 600 funcionários da empresa em uma lista. Em seguida, a cada intervalo de 10 funcionários, ela selecionou o oitavo, totalizando 60 entrevistados.	Adequada	Sistemático também pode representar uma população por existir uma grande variedade nas pessoas escolhidas.

Amostra 1: Inadequada/ Não é adequada pois ela usou um método por conveniência dela, de uma forma que não dá para representar um grupo grande de pessoas

Amostra 2: Adequada/ É um dos métodos adequados (probabilísticos)

Amostra 3: Adequada/ Sorteio é um método justo

Amostra 4: Inadequado/ Dessa forma não é justo pois nem todos têm acesso e acaba se tornando um método para um grupo mais fechado

Amostra 5: Adequado/Sistemático também pode representar uma população por existir uma grande variedade nas pessoas escolhidas

Diante disso, evidencia-se a evolução no conhecimento deste estudante ao responder adequadamente todas as amostras e apresentar suas respectivas as justificativas. A citação dos termos “conveniência” “probabilísticos” e “sistêmático” denotam o aprimoramento da habilidade de letramento para se comunicar sobre a estatística relacionada com o conhecimento estatístico relativo aos métodos de Amostragem (Gal,2002). Além disso, nota-se que o estudante passa a conceber o sorteio como algo justo e adequado para garantir, através do método aleatório simples, a representação amostral. Destaca-se ainda que nas amostras 1 e 4, não probabilísticas, o estudante aponta seus vieses no processo de seleção. Diante desses resultados da terceira questão, ratifica-se que a intervenção de ensino, independente do recurso utilizado nas duas turmas, proporcionou o desenvolvimento do letramento estatístico relativo aos métodos de Amostragem.

#### 9.1.4 Desempenho dos estudantes na questão 4 no pré-teste e pós-teste

A quarta questão nos dois testes teve o objetivo de analisar, a partir das situações apresentadas, se os estudantes reconheciam o gráfico abordado como uma Curva Normal, identificariam a média como o valor que está no centro da Curva e julgariam se medida anunciada na questão corresponde à média da distribuição. O quadro 11 apresenta o percentual por tipo de resposta das duas turmas para essa questão nos testes aplicados.

Quadro 11– Percentual por tipo de resposta da Questão 4

Categoria	Turma/Fase da Pesquisa			
	Lápis e papel		Geogebra	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Resposta Inadequada	29,0		32,0	5,0
Resposta Parcialmente Adequada	56,0	3,0	50,0	8,0
Resposta Adequada	15,0	97,0	18,0	87,0

Fonte: Dados da pesquisa (2023).

Comparando os resultados dessa questão nos dois testes, nota-se uma evolução no percentual de respostas adequadas. Com relação ao pré-teste, nas duas turmas, o tipo de resposta mais frequente foi a parcialmente adequada. Constatamos

que a maioria dos estudantes mencionaram que a medida apresentada na questão não representava de fato a média da distribuição, mas justificaram com critérios alheios que incluíram a indicação de que a média seria um valor maior, sem defini-lo; como também, de que não era possível, sem especificar todos os valores e suas respectivas frequências, calcular a média. Isso evidencia que esses estudantes concebiam a média como uma medida restrita ao seu processo do algoritmo, o qual consiste em somar os valores da variável e dividir pela quantidade de dados. Além disso, verificamos que alguns estudantes apresentaram critérios relacionados à frequência dos dados no eixo vertical do gráfico, não se atentando que a média e os demais valores da variável constavam no eixo horizontal, caracterizando assim, um equívoco na análise e interpretação do gráfico.

Logo, a maioria dos estudantes, de ambas as turmas, apresentavam fragilidades no conhecimento estatístico e matemático (Gal,2002) relacionados com a análise e interpretação de gráficos. Esse resultado nos mostrou que a maioria dos estudantes, das duas turmas, não compreendiam as características principais da Curva Normal, como o formato do gráfico e as medidas de centralidade e dispersão. Essa conclusão pode ser verificada (Figuras 104 e 105) na resposta parcialmente adequada de um estudante da turma Lápis e Papel e de um estudante da turma Geogebra, respectivamente, para essa questão no pré-teste.

Figura 104 – Resposta parcialmente adequada de um estudante para a 4º questão do pré-teste – turma Lápis e Papel

*Não, pois, no gráfico, 80 é a frequência de pessoas com determinada pressão sistólica.*

Não, pois no gráfico, 80 é a frequência de pessoas com determinada pressão arterial.

Fonte: Dados da Pesquisa (2023).

A partir dessa resposta, comprehende-se que o estudante indicou que o valor de 80mmHg da pressão arterial não correspondia à média da distribuição, porém justifica que esse valor seria o da frequência, presente no eixo vertical do gráfico abordado na questão. Logo, esse estudante não fez a leitura adequada do gráfico e não analisou adequadamente os eixos.

Figura 105 – Resposta parcialmente adequada de um estudante para a 4º questão do pré-teste – turma Geogebra

*Não, pois seu único valor de pressão arterial não é o suficiente para declarar valor médio, visto que no gráfico não é apresentado quantas pessoas tem tal pressão arterial*

Não, pois seu único valor da pressão arterial não é o suficiente pra declarar o valor médio, visto que no gráfico não é apresenado quantas pessoas tem tal pressão.

Fonte: Dados da Pesquisa (2023).

A resposta do estudante da turma Geogebra também evidencia fragilidades na compreensão das características do gráfico da Curva Normal. Nota-se que o estudante afirma não ser possível determinar o valor da distribuição porque não se tem o indicativo da frequência do mesmo. Assim, como destacamos anteriormente, esse estudante, possivelmente, concebia a média remetendo-a unicamente a aplicação de seu algoritmo, algo que na Educação Básica é fortemente abordado (Oliveira e Rosa, 2020), evidenciando sua crença paralelamente ao conhecimento estatístico e matemático (Gal, 2002). Ou seja, não compreendia a média como valor central da Curva Normal.

Porém, o resultado do pós-teste indicou que independente do recurso didático utilizado nas turmas, a intervenção de ensino contribuiu para a construção e aprimoramento do conhecimento dos estudantes sobre o conceito da Curva Normal e suas características. Observa-se que o percentual de respostas adequadas foi superior a 85% nas duas turmas A resposta adequada (Figura 106) do mesmo estudante da turma de Lápis e papel apresentado anteriormente evidencia essa conclusão.

Figura 106 – Resposta adequada de um estudante para a 4º questão do pré-teste – turma Lápis e Papel

*Não, pois a média tende a estar localizada no centro do gráfico, nesse caso, a média das temperaturas seria 32°C e não 35°C.*

Não, pois a média tenda a estar localizada no centro do gráfico, nsse caso, a m'dia das temperaturas seria 32°C não 35°C

Fonte: Dados da Pesquisa (2023).

A partir dessa resposta, constata-se o conhecimento do estudante sobre a Curva Normal e sua média, ao mencionar adequadamente que o valor apresentado na questão não corresponde à média da distribuição das temperaturas porque a mesma consiste ao valor central da curva, ou seja, 32°C. De modo análogo, na

resposta (Figura 107) do estudante da turma Geogebra também abordado anteriormente no pré-teste, podemos notar seu avanço na compreensão do modelo da Curva Normal ao responder adequadamente a indagação da questão. Nota-se que o estudante também indica que o valor da média é 32°C e não 35°C como mencionou a questão. Além disso, aponta que o valor da média é também os mais frequentes, remetendo a uma propriedade da Curva Normal, na qual a média, moda e mediana coincidem e correspondem ao valor central da Curva. Diante disso, destacamos que tanto a turma Lápis e Papel como a turma Geogebra avançou no conhecimento estatístico e matemático sobre a Curva Normal, evidenciando o aprimoramento do letramento estatístico inerente a esses conceitos.

Figura 107 – Resposta adequada de um estudante para a 4º questão do pré-teste – turma Lápis e Papel

*Não representa, pois a média de temperaturas, ou seja, a que mais ocorreu foi 32°, não 35°.*

Não representa, pois a média das temperaturas, ou seja, a que mais ocorreu foi 32°, não 35°

Fonte: Dados da Pesquisa (2023).

#### 9.1.5 Desempenho dos estudantes na questão 5 no pré-teste e pós-teste

Já a quinta questão, de ambos os testes, foi direcionada para analisar o conhecimento dos estudantes sobre o conceito de variabilidade/dispersão das medidas apresentadas em gráficos que se aproximam de uma Curva Normal. No pré-teste, foi apresentado duas distribuições de temperaturas, respectivamente das cidades de João Pessoa e Recife durante um período. Sabendo que a média das temperaturas é aproximadamente a mesma para as duas distribuições, os estudantes deveriam indicar, através dos gráficos apresentados, qual cidade apresenta maior variabilidade das medidas das temperaturas. De modo semelhante, no pós-teste foi apresentado duas distribuições das medidas da palma da mão de uma amostra de pessoas e questionou-se a partir dos gráficos apresentados e de suas respectivas médias, qual a distribuição apresentava maior variabilidade dos comprimentos da palma da mão. A seguir, apresentamos (Quadro 12) o percentual por tipo de resposta para essa questão em ambos os testes.

Quadro 12 – Percentual por tipo de resposta da Questão 5

Categoria	Turma/Fase da Pesquisa			
	Lápis e papel		Geogebra	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Resposta Inadequada	44	29	60,5	31
Resposta Parcialmente Adequada	21	6	10,5	8
Resposta Adequada	35	65	29,0	61

Fonte: Dados da Pesquisa (2023).

A partir desses dados, conclui-se que no pré-teste a maioria dos estudantes, das duas turmas, respondeu de modo inadequado. Constatamos que a resposta mais frequente para essa categoria, em ambas das turmas, foi a de mencionar que a distribuição das temperaturas da cidade do Recife possuía maior variabilidade porque apresentava uma maior oscilação na frequência dos dados, observada no eixo vertical do gráfico. Destaca-se que esses estudantes demonstravam fragilidades na interpretação dos gráficos ao refletirem apenas sobre a variação da frequência dos dados e não se atentarem para as medidas das temperaturas registradas no eixo horizontal, o qual contempla a relação entre a média, centro da distribuição e a variabilidade/dispersão que, por sua vez, representa o grau do distanciamento das medidas para a média. Essa conclusão vai ao encontro dos resultados observados na quarta questão, na qual os estudantes também demonstraram lacunas no conhecimento matemático e estatístico relacionados com a análise do gráfico da Curva Normal. Nessa questão, a variabilidade é verificada em maior grau na distribuição da cidade de João Pessoa. As respostas inadequadas (Figuras 108 e 109) de um estudante da turma Lápis e Papel e a de um estudante da turma Geogebra, respectivamente, notabiliza essa conclusão.

Figura 108 – Resposta inadequada de um estudante para a 5º questão do pré-teste – turma Lápis e Papel

~~Recife, outras temperaturas apresentam maior frequência no gráfico de Recife do que no de João Pessoa.~~

Recife, outras temperaturas apresentam maior frequência no gráfico de Recife do que no de João Pessoa.

Fonte: Dados da Pesquisa (2023).

Figura 109 – Resposta inadequada de um estudante para a 5º questão do pré-teste – turma Geogebra

A de Recife. Porque a frequência das temperaturas individualmente é maior.

Fonte: Dados da Pesquisa (2023).

Observa-se que as respostas desses dois estudantes são semelhantes, ao pontuarem que a maior variabilidade está presente na distribuição das temperaturas da cidade do Recife, justificando com base na intensidade ou oscilação da frequência dos dados. Constatamos que essa concepção da variabilidade atrelada à frequência dos dados foi fortemente apresentada como resposta para a quinta questão, em ambas as turmas. Esse resultado, como dito anteriormente, evidenciou as crenças e as lacunas no conhecimento matemático e estatístico (Gal, 2002) dos estudantes sobre a variabilidade presente em distribuição de dados atrelada a média da mesma.

Já no pós-teste percebe-se um avanço na compreensão desse conceito nos estudantes das duas turmas, com índice de respostas superior a 60%. Assim, destaca-se que intervenção de ensino contribuiu para o aprimoramento do conhecimento dos estudantes sobre a relação entre as medidas de centralidade e variabilidade/dispersão presente no modelo da Curva Normal. Verificamos que no pós-teste a maioria dos estudantes, seja na turma do Lápis e Papel ou Geogebra, mencionou adequadamente que a distribuição da figura 1 possuía maior variabilidade com base no maior grau de variabilidade (distanciamento) das medidas em relação à média da distribuição, ou seja, ao valor central da Curva. Também verificamos respostas adequadas, nas duas turmas, a partir da indicação da amplitude maior verificada na figura 1. A resposta adequada (Figura 110) do mesmo estudante da turma de Lápis e Papel abordado anteriormente exemplifica essa conclusão.

Figura 110 – Resposta adequada de um estudante para a 5º questão do pós-teste – turma Lápis e Papel

*Figura 1, os dados são mais distantes do centro.*

$\begin{array}{r} 21,9 \\ 16,0 \\ \hline 5,9 \end{array}$	$\begin{array}{r} 27,0 \\ 21,9 \\ \hline 5,1 \end{array}$	$\begin{array}{r} 19,6 \\ 14,0 \\ \hline 5,6 \end{array}$	$\begin{array}{r} 25,0 \\ 19,4 \\ \hline 5,6 \end{array}$
---	---	---	---

Figura 1, os dados são mais distantes do centro

Fonte: Dados da Pesquisa (2023).

A partir dessa resposta, nota-se que o estudante passa a indicar a maior variabilidade por meio do grau do distanciamento dos dados para o centro da Curva Normal, ou seja, a média da distribuição. Destacamos que esse estudante procedeu com o cálculo do distanciamento entre os dois extremos de cada Curva Normal e chegou à conclusão que a maior variabilidade era na figura 1. Acreditamos que essa estratégia foi influenciada pela intervenção de ensino realizada nessa turma com o uso do lápis e papel, na qual discutimos as medidas de centralidade e variabilidade, através da relação entre o desvio-padrão e a média da Curva Normal e estudamos formas de se medir essa variabilidade, contemplando o cálculo do distanciamento entre os extremos e o centro da Curva através do referido recurso. No que diz respeito à outra turma, a resposta adequada (Figura 111) do mesmo estudante da turma Geogebra abordado anteriormente exemplifica essa mesma conclusão.

Figura 111 – Resposta adequada de um estudante para a 5º questão do pós-teste – turma Lápis e Papel

*A primeira distribuição. Porque os dados estão mais afastados da média do gráfico.*

A primeira distribuição. Porque os dados estão mais afastados da média do Gráfico

Fonte: Dados da Pesquisa (2023).

A resposta desse estudante também evidência o aprimoramento do conhecimento relativo a variabilidade presente no modelo da Curva Normal. Assim como o estudante da outra turma, há nessa resposta a menção de que a maior variabilidade se encontra na primeira distribuição (figura 1) com a justificativa adequada de que os dados estão mais afastados da média do gráfico, ou seja, o centro da distribuição. Assim, o Geogebra, enquanto recurso didático, também contribuiu para o aprimoramento desse conhecimento, a partir das manipulações e visualizações gráficas. Logo, destacamos que cada recurso utilizado na intervenção de ensino, proporcionou o desenvolvimento da aprendizagem e consequentemente o Letramento Estatístico (Gal, 2002) sobre relação entre as medidas de centralidade e variabilidade presente no modelo da Curva Normal.

#### 9.1.6 Desempenho dos estudantes na questão 6 no pré-teste e pós-teste

A sexta e última questão de ambos os testes teve o objetivo de analisar o conhecimento dos estudantes sobre o conceito de probabilidade associado a intervalos da Curva Normal e a relação entre a distribuição da média amostral e a média populacional, fundamentada no teorema do limite central. No pré-teste, essa questão apresentou um gráfico da distribuição das notas que os estudantes do Ensino Médio de uma escola pública de Pernambuco obtiveram na disciplina de Matemática. Nela, é informado que a distribuição é aproximadamente normal e a sua média é 6,7. No item A, perguntou-se qual a probabilidade aproximada de selecionarmos um estudante dessa escola, aleatoriamente, e ele possuir a nota de matemática maior ou igual a média. De modo semelhante, o pós-teste apresentou um gráfico com a distribuição aproximadamente normal das notas de estudantes em um teste de Matemática, cuja média é 6,4. No item A, questionamos qual a probabilidade aproximada de selecionarmos um estudante, aleatoriamente, e ele possuir a nota no teste menor ou igual que a média. Diante disso, no primeiro item dos dois testes os estudantes deveriam reconhecer que a probabilidade solicitada corresponde à metade da Curva, ou seja, tem probabilidade aproximada de 50%, pelo fato da distribuição está bem próxima da normalidade. A seguir, destacamos (Quadro 13) os resultados do primeiro item da 6º questão em ambos os testes.

Quadro 13 – Percentual por tipo de resposta da Questão 6º A

Categoria	Turma/Fase da Pesquisa			
	Lápis e papel		Geogebra	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Resposta Inadequada	91	9,0	90	16
Resposta Adequada	9,0	91	10	84

Fonte: Dados da pesquisa (2023).

Através desses dados podemos notar um avanço bastante significado, por parte dos estudantes, na compreensão do conceito de probabilidade abordado na questão. Isso indica que a intervenção de ensino nas duas turmas propiciou o desenvolvimento do Letramento Probabilístico, tanto em relação aos elementos cognitivos quanto ao de disposição (Gal, 2005) a partir da relação entre o conceito de probabilidade e a área de intervalos da Curva Normal. Na turma Lápis e Papel houve uma inversão nos índices das categorias de respostas do pré-teste para o pós-teste, e neste último observa-se que 91% dos estudantes responderam adequadamente. Na turma Geogebra o avanço também foi gritante, com 84% dos estudantes respondendo adequadamente esse item no pós-teste.

Ao analisarmos as respostas dos estudantes no pré-teste, constatamos que nas duas turmas, a maioria dos que responderam de modo inadequado indicou um valor acima de 50% para probabilidade solicitada. Essa conclusão, mais uma vez, evidencia o que apresentamos nas questões quatro e cinco, de que os estudantes participantes apresentam lacunas no que diz respeito aos elementos de conhecimento (Gal, 2002, 2005) sobre a Curva Normal e suas propriedades, incluindo a estimativa de probabilidades. As respostas (Figuras 112 e 113), respectivamente, de um estudante da turma Lápis e papel e de um estudante da turma Geogebra abordam essas conclusões.

Figura 112 – Resposta inadequada de um estudante para a 6º(A) questão do pré-teste – turma Lápis e Papel

Cerca de 67% .

Fonte: Dados da Pesquisa (2023).

Figura 113 – Resposta inadequada de um estudante para a 6º(A) questão do pré-teste – turma Geogebra

Mais de 50%. como é mostrado no gráfico a partir das quantidades.

Fonte: Dados da Pesquisa (2023).

Nota-se que as respostas indicam para uma probabilidade acima de 50%. Diante disso, constatamos que a maioria dos estudantes não se atentaram para a relação entre a metade da área da Curva Normal e a probabilidade solicitada, que corresponde a aproximadamente 50%, tendo em vista que a distribuição abordada é próxima do modelo de normalidade. Logo, demonstraram não dominar esse conhecimento, como pontuado anteriormente. Observamos que em algumas respostas, os estudantes tentaram indicar essa probabilidade pelo significado da razão clássica que envolve o cálculo da divisão entre os elementos do evento e os elementos do espaço amostral. No estudo 1 da presente pesquisa, verificamos que os livros didáticos do Ensino Médio também priorizam esse significado de probabilidade. Já no estudo 2, constatamos que a prova do ENEM vai na mesma direção. Logo, acreditamos que os estudantes do Ensino Médio são fortemente direcionados para o estudo da probabilidade envolvendo apenas esse significado, não explorando outros contextos, ideias, e formas para se estimar probabilidades (Gal,2005), a exemplo da Curva Normal.

Já no pós-teste, como já anunciado, a maioria dos estudantes, nas duas turmas, responderam adequadamente. Isso indica que eles analisaram e fizeram a leitura correta do gráfico da Curva Normal, indicando que a probabilidade solicitada era de 50% ou aproximadamente esse valor. As respostas adequadas (Figuras 114 e 115) do referido item no pós-teste dos mesmos estudantes abordados anteriormente no pré-teste evidenciam essa conclusão.

Figura 114 – Resposta inadequada de um estudante para a 6º(A) questão do pré-teste – turma Lápis e Papel

50%, já que é metade do valor.

50%, já que é metade do valor

Fonte: Dados da Pesquisa (2023).

Figura 115 – Resposta adequada de um estudante para a 6º(A) questão do pré-teste – turma Geogebra

<i>Cerca de 50%. Pois, o valor no gráfico fica na metade</i>
Cerca de 50%. Pois o valor no gráfico fica a metade

Fonte: Dados da Pesquisa (2023).

Podemos notar que ambos os estudantes indicam adequadamente o valor da probabilidade solicitada e justificam relacionando-o com a metade da área da Curva Normal, isto é, 50% dos dados. Diante desse resultado, ratificamos que a intervenção de ensino, independente do recurso utilizado, contribuiu para a construção do conhecimento relativo ao conceito de probabilidade no modelo da Curva Normal e, consequentemente, o desenvolvimento do Letramento Estatístico e Probabilístico dos estudantes (Gal, 2002,2005), tanto do que diz respeito aos elementos cognitivos que explicitam o conhecimento teórico, como também os elementos de disposição que evidenciam a postura crítica e as crenças dos estudantes mediante as justificativas apresentadas.

Já no item B da 6º questão dos dois testes, com base no gráfico da distribuição da nota de estudantes do Ensino Médio em Matemática e da média abordados na questão, apresentou-se um contexto em que o diretor da escola desejava selecionar uma amostra, ao acaso, de 30 desses estudantes. A partir disso, indagou-se, no pré-teste, se a probabilidade de que a média amostral fosse igual a 4,5 seria baixa ou alta. De modo semelhante, no pós-teste foi questionado se a probabilidade de que essa média amostral fosse 9,0 seria baixa ou alta. Em ambos os casos, os estudantes deveriam reconhecer que a média amostral tende a se aproximar da média populacional, e, consequentemente, compreender que a probabilidade solicitada para a nota é baixa por, justamente, está distante da média populacional abordada na questão. Os resultados (Quadro 14) por tipo de resposta para esse item nas duas turmas estão destacados a seguir.

Quadro 14 – Percentual por tipo de resposta da Questão 6º B

Categoria	Turma/Fase da Pesquisa			
	Lápis e papel		Geogebra	
	Pré-teste	Pós-teste	Pré-teste	Pós-teste
Resposta Inadequada	94%	9%	82%	3%

Resposta Parcialmente Adequada	3%	26%	5%	26%
Resposta Adequada	3%	65%	13%	71%

Fonte: Dados da pesquisa (2023).

Esses dados evidenciam que, no pré-teste, o índice que resposta inadequada foi bastante alto nas duas turmas. Constatamos que a maioria dos estudantes, antes da intervenção do ensino, possivelmente não tiveram um contato, mesmo que de modo informal, sobre a relação da média amostral com a média populacional estabelecida pelo teorema central do limite. Nas respostas, os estudantes indicaram que a probabilidade solicitada era baixa porque a frequência da nota 4,5 era baixa. Chamamos a atenção que esse tipo de resposta se classificou como inadequada porque os estudantes, nitidamente, confundiram a distribuição da variável nota apresentada no gráfico e que contempla diferentes valores e suas respectivas frequências, com a distribuição amostral das médias, a qual é formulada a partir dos dados populacionais e obedece ao Teorema Central do Limite. Esse tipo de afirmação evidencia as crenças dos estudantes relacionadas com o conhecimento estatístico e probabilístico (Gal, 2002,2005) que possuíam sobre o tema na primeira etapa do nosso estudo 4. As respostas inadequadas (Figuras 116 e 117), respectivamente, de um estudante da turma Lápis e Papel e de um estudante da turma Geogebra exemplificam essas conclusões.

Figura 116 – Resposta inadequada de um estudante para a 6º(B) questão do pré-teste – turma Lápis e Papel

É baixa. Porque poucos alunos tiraram a nota  
4,5 , de acordo com o gráfico.

É baixa. Porque poucos alunos tiraram a nota 4,5 , de acordo com o gráfico

Fonte: Dados da Pesquisa (2023).

Figura 117 – Resposta inadequada de um estudante para a 6º(B) questão do pré-teste – turma Geogebra

Baixa, pois a maior parte dos alunos tiraram nota maior que 4,5.

Baixa, pois a maior parte dos alunos tiraram nota maior que 4,5.

Fonte: Dados da Pesquisa (2023).

Já no pós-teste houve um aumento no índice de respostas adequadas nas duas turmas. Essa conclusão, notabiliza, mais uma vez, que a intervenção de ensino possibilitou a construção e aprimoramento, por parte dos estudantes, do conhecimento sobre a relação entre distribuição da média amostral e a média populacional, relacionando a Amostragem, eixo estruturador da Inferência Estatística com o modelo da Curva Normal, principal ferramenta de análise de dados no campo da Inferência Estatística (Batanero, Tauber e Sánchez, 2004).

Uma parcela dos estudantes das duas turmas respondeu parcialmente adequado por indicaram que a probabilidade solicitada era baixa, porém não apresentaram justificativa. No entanto, a maioria dos estudantes, também nas duas turmas, respondeu adequadamente, ao justificar que a média da amostral tende a ser próxima à da população e por isso a probabilidade solicitada é baixa porque a nota 9,0 está distante do centro da Curva, ou seja, afastava da média populacional que é de 6,4. Logo, evidencia-se que independente do recurso na intervenção de ensino, esses estudantes se apropriaram do conhecimento abordado e refletiram de modo adequado sobre a relação entre a média amostral e a populacional, uma habilidade importante para o desenvolvimento do Letramento Estatístico e probabilístico (Gal, 2002, 2005). As respostas adequadas (Figuras 118 e 119) dos mesmos estudantes apresentados anteriormente no pré-teste respectivamente, um estudante da turma Lápis e Papel e um estudante da turma Geogebra evidencia essa conclusão.

Figura 118 – Resposta adequada de um estudante para a 6º(B) questão do pré-teste – turma Lápis e Papel

*Baixa, pois a média amostral tende a ser próxima da média populacional que é 6,4.*

Baixa, pois a média amostral tende a ser próxima da média populacional que é 6,4.

Fonte: Dados da Pesquisa (2023).

Figura 119 – Resposta adequada de um estudante para a 6º(B) questão do pré-teste – turma Geogebra

*Baixa, porque a medida da média amostral tem que ser um valor próximo ou igual à média populacional.*

Baixa, porque a medida da média amostral tem que ser um valor próximo ou igual à média populacional

Fonte: Dados da Pesquisa (2023).

Apresentado o desempenho dos estudantes das turmas Lápis e Papel e Geogebra nos testes, reafirma-se que, em linhas gerais, os estudantes das duas turmas, independente do recurso empregado na intervenção, conseguiram construir e aprimorar seus conhecimentos sobre os conceitos relativos à Amostragem e à Curva Normal. Ou seja, passaram a compreender sobre os métodos de Amostragem, representatividade e variabilidade amostral, o modelo da Curva Normal, abarcando sua representação gráfica, medidas de centralidade e variabilidade e o cálculo de probabilidades. Além disso, aprenderam sobre a relação entre a distribuição da média amostral e a média populacional, relacionando a Amostragem com a Curva Normal, através do Teorema Central do Limite.

A partir dos resultados, nota-se que as lacunas observadas no pré-teste foram superadas pela maioria dos estudantes de ambas as turmas no pós-teste. Assim, evidencia-se que os momentos desenvolvidos na intervenção de ensino, seja com o Lápis e Papel ou o Geogebra contribuíram para o desenvolvimento do Letramento Estatístico e probabilístico (Gal, 2002,2005) dos estudantes participantes. Diante disso, na próxima seção, apresentamos os momentos desenvolvidos na intervenção de ensino com ênfase para a descrição e análise das atividades realizadas e das discussões mais pertinentes realizadas entre o pesquisador e os estudantes participantes em cada momento.

## 9.2 INTERVENÇÃO DE ENSINO

Como pontuamos na nossa metodologia, a intervenção de ensino constituiu a segunda etapa do nosso estudo 4 e teve o objetivo de analisar a contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada. Cabe destacar que as ações realizadas nessa etapa foram construídas mediante aos resultados observados nos estudos 1,2,3 e no pré-teste do estudo 4.

Relembramos ainda que a referida intervenção foi implementada em duas turmas do 3º ano do Ensino médio e contemplou quatro momentos. O primeiro momento, enquanto introdução à temática estudada, contemplou uma atividade comum para as duas turmas. A partir do segundo momento, as atividades foram adaptadas para o recurso a ser utilizado, sendo uma turma o lápis e papel e a outra o software Geogebra. Nas próximas seções, apresentamos a análise de cada momento realizado em cada uma das duas turmas, com foco para as atividades desenvolvidas

e as discussões emanadas dos momentos de socialização, buscando evidenciar as contribuições dos recursos empregados para a aprendizagem do tema supracitado.

### 9.2.1 1º Momento: Reflexões iniciais sobre a Amostragem e a Curva Normal

Na turma Lápis e Papel, o professor pesquisador deu início ao primeiro encontro da intervenção desenvolvendo ações que buscaram instigar os estudantes a pensarem e debaterem sobre os conceitos relativos à Amostragem e à Curva Normal. Para tanto, o professor pesquisador procedeu indagando aos estudantes a respeito das questões contidas no pré-teste e as temáticas que elas envolviam. A seguir, apresentamos (Quadro 15) um trecho desse diálogo.

Quadro 15 – Recorte do diálogo preliminar à Atividade 1 da turma Lápis e Papel

<b>Professor pesquisador:</b> <i>Bom, pessoal. Nessa aula iremos iniciar um estudo sobre alguns temas que foram abordados naquele teste que vocês resolveram na última aula. Quem aqui pode dizer um conceito da Estatística, da Matemática que estava presente naquele teste?</i>
<b>Estudante 18:</b> <i>Porcentagem</i>
<b>Estudante 2:</b> <i>Média, probabilidade</i>
<b>Estudante 27:</b> <i>Amostra</i>
<b>Estudante 9:</b> <i>Pesquisa estatística</i>
<b>Estudante 5:</b> <i>Pesquisa amostral</i>
<b>Professor Pesquisador:</b> <i>Ok. Inicialmente gostaria de debater com vocês sobre a pesquisa estatística. Qual o principal propósito de uma pesquisa estatística?</i>
<b>Estudante 14:</b> <i>Coletar dados para se descobrir algo</i>
<b>Estudante 3:</b> <i>Fazer entrevistas para saber a opinião das pessoas sobre alguma coisa.</i>
<b>Estudante 10:</b> <i>Acho que tem o objetivo de se concluir sobre algo, responder a algo através do que se coleta.</i>
<b>Professor Pesquisador:</b> <i>Exato. As pesquisas estatísticas são realizadas com o intuito de se compreender a realidade e a partir disso se tomar decisões. Como exemplo, o governo, empresas e instituições realizam pesquisas para compreender a nossa realidade e partir disso, tomar as decisões mais adequadas. A pesquisa estatística envolve algumas fases que iremos debater mais adiante. O estudante 5 falou de pesquisa amostral, então do que se trata uma pesquisa amostral?</i>
<b>Estudante 5:</b> <i>Pesquisa amostral é quando você pega pessoas como amostra e quer saber a opinião delas sobre alguma coisa.</i>
<b>Professor Pesquisador:</b> <i>Mas a pesquisa amostral só pode ser com pessoas?</i>

**Alguns estudantes:** Não

**Estudante 18:** Pode ser com animais também.

**Estudante 12:** Pode estabelecimentos ou lojas. Por exemplo, no comércio.

**Professor Pesquisador:** Isso. Os elementos amostrais podem ser pessoas, animais, objetos. Nessa aula de hoje iremos compreender mais a fundo sobre a pesquisa estatística, incluindo essas características da pesquisa amostral.

Fonte: Dados da pesquisa (2023).

A partir dos conceitos que foram explorados no pré-teste, o diálogo inicial com a turma Lápis e Papel foi direcionado para que os estudantes explicitassem suas concepções sobre o papel da pesquisa estatística. Em linhas gerais, eles apresentaram o conhecimento prévio de que uma pesquisa estatística envolve a coleta de dados com o intuito de se concluir ou responder alguma questão. Com a mediação do professor pesquisador refletiram formalmente que a pesquisa estatística tem o objetivo de diagnosticar a realidade para se tomar decisões (Guimarães e Gitirana, 2013). Nesse diálogo preliminar, destaca-se a resposta do estudante 5, o qual cita a pesquisa amostral envolvendo pessoas. Outros estudantes indicaram também que ela pode envolver outros elementos, como animais e objetos. No entanto, para alguns estudantes, esse entendimento de amostra ou população se remetendo unicamente a pessoas ficou evidente no pré-teste e repetiu-se nesse diálogo, explicitando suas crenças e experiência com relação a esse conceito (Gal, 2002, Reyes, 2019).

De modo semelhante, também fizemos esse diálogo introdutório (Quadro 16) na turma Geogebra com o intuito questionar os estudantes a respeito das questões contidas no pré-teste e instigá-los a pensar sobre as temáticas que elas envolviam, como foco inicial sobre o papel da pesquisa estatística. Abaixo, apresentamos um trecho desse diálogo.

Quadro 16 – Recorte do diálogo preliminar à Atividade 1 da turma Geogebra

**Professor Pesquisador:** Olá, pessoal. Na aula de hoje, iremos iniciar um estudo sobre alguns temas que foram abordados naquele teste que vocês resolveram na última aula. Inicialmente, quem aqui pode dizer um conceito da Estatística ou da Matemática que estava presente naquele teste?

**Alguns estudantes:** Média

**Estudante 26:** População e amostra

**Estudante 8:** Gráficos com medidas estatísticas

**Estudante 27:** pesquisa feita com estudantes

**Estudante 9:** Pesquisa estatística

**Estudante 14:** Pesquisa com amostras

**Professor Pesquisador:** Certo. Inicialmente é importante pensarmos um pouco sobre a pesquisa estatística. Qual o principal propósito de uma pesquisa estatística?

**Estudante 27:** Investiga sobre alguma questão importante para se concluir algo.

**Estudante 32:** Acho que é todo processo de se coletar informações para se chegar em uma conclusão.

**Professor Pesquisador:** Mais alguém pode falar sua opinião?

*Estudante 8: Eu acho que é o que os colegas falaram mesmo. Envolve o procedimento de coletar dados sobre uma determinada questão visando trazer conclusões sobre isso.*

**Professor Pesquisador:** Correto. As pesquisas estatísticas são realizadas com a finalidade de se compreender a realidade e a partir disso se tomar decisões. No nosso dia a dia, temos o exemplo do governo, empresas e instituições que realizam pesquisas para compreender a nossa realidade e partir disso, tomar as decisões. É importante destacar que a pesquisa estatística envolve algumas fases que iremos debater mais adiante. O estudante 14 mencionou que o teste havia pesquisas com amostras. Do que se trata, então, uma pesquisa amostral?

**Estudante 15:** Pesquisa feita com um grupo, uma parte.

**Estudante 13:** Pesquisa feita com amostras

**Professor Pesquisador:** E essas amostras podem ser constituídas de que tipo de elementos? Quem pode compor essas amostras?

**Alguns estudantes:** Pessoas

**Estudante 18:** Animais também.

**Estudante 27:** Pode ser também carros, por exemplo em um estudo em uma grande fábrica. Saber se algum carro tem algum tipo de falha.

**Professor Pesquisador:** Correto. Então os elementos amostrais podem ser pessoas, animais, objetos. Tudo depende do objetivo da pesquisa. Nessa aula de hoje iremos compreender mais sobre a pesquisa estatística, incluindo a pesquisa amostral.

Fonte: Dados da pesquisa (2023).

No diálogo realizado da turma Geogebra ficou evidente que os estudantes também detinham um conhecimento prévio sobre o papel da pesquisa estatística, incluindo a pesquisa amostral. Logo, refletiram na importância do processo de investigação estatística para compreender a realidade e se tomar decisões a partir do que os dados revelam. Nesse diálogo, destaca-se que de modo análogo a outra turma, quando o pesquisador questionou que tipo de elemento pode compor a amostra, um grupo de estudantes respondeu rapidamente “pessoas”, o que transparece ser a única ideia de elemento amostral que eles concebem. Outros estudantes indicaram que a

amostra pode envolver animais e objetos, a exemplo de carro. Assim, alguns estudantes, de ambas as turmas, precisavam avançar na compreensão desses conceitos. Para isso, demos início a primeira atividade (Figura 120).

Figura 120 – Primeira questão da atividade 1

1º) Para cada item abaixo, identifique a população e amostra relativa a cada pesquisa:

- a) O instituto Ibope Conecta fez uma pesquisa com 2.000 internautas das classes A, B, C e D, de todas as regiões do Brasil para saber o que mais incomoda os internautas brasileiros na internet.
- b) O Procon do Rio de Janeiro fez uma pesquisa em 344 postos da capital fluminense para verificar os preços dos combustíveis.

Fonte: Dados da pesquisa (2023).

Para a resolução da atividade 1 do primeiro dia da intervenção, os estudantes foram organizados em duplas. A medida que uma questão ia sendo resolvida, prosseguimos imediatamente com a socialização das respostas, tendo o envolvimento de toda a turma e o professor pesquisador, de modo que só prosseguimos com a resolução da segunda questão após a socialização da primeira e assim sucessivamente. Nesse sentido, a primeira questão da atividade 1 teve o objetivo de proporcionar aos estudantes a reflexão sobre o conceito de População e Amostra na Estatística. Nela, é apresentado duas situações de pesquisas estatísticas amostrais com contextos diferentes, uma apresenta elementos que são pessoas e a outra, postos de combustíveis. Apresentamos (Quadro 17), a seguir, o recorte do diálogo dessa questão na turma Lápis e Papel.

Quadro 17 – Recorte do diálogo da 1º questão da Atividade 1 - turma Lápis e Papel

**Professor Pesquisador:** Na primeira questão é abordado duas situações de pesquisas estatísticas amostrais e pergunta-se qual é a população e amostra de cada uma. A alternativa “a” informa sobre uma pesquisa que foi feita com internautas. Nessa pesquisa qual é a população?

**Alguns estudantes:** Internautas brasileiros

**Professor Pesquisador:** E qual é amostra?

**Estudante 9:** o número de internautas

**Estudantes:** 2000 internautas

**Professor Pesquisador:** Ok. Correto. A alternativa b apresenta uma pesquisa realizada com postos de combustíveis. Qual a população dessa pesquisa?

**Estudantes:** Os postos de combustíveis

**Professor Pesquisador:** E a amostra?

**Alguns estudantes:** Os 344 postos.

**Professor Pesquisador:** Exato. Então qual o conceito de população para vocês?

**Estudante 10:** População é o todo que queremos informação sobre

**Estudante 14:** População é o total

**Professor Pesquisador:** E qual o conceito de Amostra?

**Estudante 10:** Amostra é um grupo da população

**Estudante 9:** Amostra é uma parte da população

**Professor Pesquisador:** Ok. De modo mais formal, a população é conceituada como um conjunto de todos os elementos que possuem ao menos uma característica em comum, assim pode envolver pessoas, animais ou objetos. Por exemplo, a população de pernambucanos, a população de felinos de uma cidade e a população de livros de uma biblioteca. Já a amostra é conceituada como um subconjunto, ou seja, um grupo da população. Quem pode dizer um exemplo de população que não envolve pessoas?

**Estudante 3:** Se fizer uma pesquisa com animais de uma cidade, por exemplo a população dos gatos.

**Estudante 20:** Pode ser também com lojas, em fabricas, a população de certo produto.

**Professor Pesquisador:** Correto. Então é importante compreendermos que a partir desse conceito de população, na estatística, ela pode estar relacionada com um conjunto de animais, objetos, estabelecimentos, etc. que possuem pelo menos uma característica em comum. Diante disso, sabemos que a pesquisa realizada com amostra é chamada de pesquisa amostral. Alguém pode citar o nome da pesquisa que é feita com toda a população?

**Estudante 3:** Censo. Inclusive esse ano foi divulgado pelo IBGE o censo do Brasil.

**Professor Pesquisador:** Correto. A pesquisa censitária ou censo é o tipo de pesquisa que envolve toda a população. Mas qual o critério para definir se a pesquisa deve ser censitária ou amostral, ou seja, se deve ser investigado toda a população ou uma amostra?

-Silêncio

**Professor Pesquisador:** Quando o tamanho da população é grande, o custo e tempo para desenvolver a pesquisa são enormes, a exemplo da população brasileira. Ou ainda quando a pesquisa resulta na inutilização de um produto, como por exemplo a fabricação de fósforos. Assim, nem sempre é possível/recomendável realizar um censo. Nessas situações, se realiza a pesquisa amostral.

Fonte: Dados da pesquisa (2023).

Esse diálogo ilustra que os estudantes da turma Lápis e Papel souberam identificar adequadamente a população e amostras nas duas situações apresentadas da primeira questão. No entanto, foi possível perceber que eles não concebiam o conceito formal para esses dois termos, ao se remeterem a população apenas como “um todo” ou “o total” e amostra como “um grupo” ou “uma parte”. Porém, a intervenção do professor pesquisador propiciou a compreensão do conceito formal de população e amostra, com base em Triola (2008), o qual argumenta que a população remete ao conjunto de todos os elementos que possuem pelo menos uma característica em comum e a amostra se trata de um subconjunto da mesma.

Desta forma, ressaltamos a importância de, no campo estatístico, conceber que a população, além de pessoas, também pode ser formada por animais, objetos, estabelecimentos e etc. Destaca-se que alguns estudantes apresentaram situações que envolvem pesquisa amostral com elementos que não são pessoas, como a população de gatos de uma cidade e a população de produtos de um fábrica. Além disso, no final do diálogo, o estudante 3 respondeu adequadamente sobre a pesquisa censitária e deu o exemplo do IBGE (Instituto Brasileiro de Geografia e Estatísticas). Esses estudantes demonstraram ter conhecimento do contexto alinhado ao conhecimento estatístico sobre população e amostra (Gal,2002). Além disso, o professor pesquisador contribuiu para o entendimento sobre a diferença entre uma pesquisa censitária e amostral e os fatores que influenciam para determinar qual deve ser realizada, tendo em vista que nenhum estudante argumentou sobre esse tema.

Na turma do Geogebra, procedemos de forma análoga, debatendo com os estudantes sobre os conceitos de população e amostra explorados na primeira questão e as diferenças entre a pesquisa amostral e censitária. Apresentamos (Quadro 18), a seguir, um recorte do diálogo realizado nessa turma.

Quadro 18 – Recorte do diálogo da 1º questão da Atividade 1 – turma Geogebra

**Professor Pesquisador:** Na primeira questão é abordado duas situações de pesquisas estatísticas amostrais e pergunta-se qual é a população e amostra de cada uma. A primeira alternativa informa sobre uma pesquisa que foi feita com internautas. Nessa pesquisa qual é a população?

**Estudantes:** todos os internautas brasileiros

**Professor Pesquisador:** E qual é amostra?

**Estudantes:** 2000 internautas

**Professor Pesquisador:** Exato. A alternativa b apresenta uma pesquisa realizada com postos de combustíveis. Qual a população dessa pesquisa?

**Alguns estudantes:** todos postos de combustíveis da capital fluminense

**Professor Pesquisador:** E a amostra?

**Alguns estudantes:** Os 344 postos.

**Professor Pesquisador:** Exato. Então, diante dessas ideias, qual o conceito de população para vocês?

**Estudante 8:** É o geral

**Alguns estudantes:** É o todo

**Professor Pesquisador:** E qual o conceito de Amostra?

**Estudante 15:** Uma parte

**Estudante 19:** Uma parte desse todo

**Professor Pesquisador:** Ok. Conceituando de modo mais formal, a população é o conjunto de todos os elementos que possuem ao menos uma característica em comum, assim pode envolver pessoas, animais ou objetos. Por exemplo, a população de pernambucanos, a população de felinos de uma cidade e a população de livros de uma biblioteca. Já a amostra é conceituada como um subconjunto, ou seja, um grupo da população. Quem pode dizer um exemplo de população que não envolve pessoas?

**Estudante 3:** Todos os carros fabricados em uma determinada fábrica.

**Estudante 27:** Pode ser todas as máquinas de um laboratório

**Estudante 20:** A população de animais, como exemplo a de macacos que estava no teste

**Professor Pesquisador:** Correto. Então a partir desse conceito de população, na estatística, ela pode estar relacionada com um conjunto de animais, objetos, estabelecimentos, etc. que possuem pelo menos uma característica em comum. Diante disso, sabemos que a pesquisa realizada com amostra é chamada de pesquisa amostral. Alguém pode citar o nome da pesquisa que é feita com toda a população?

- Silêncio

**Professor Pesquisador:** A pesquisa censitária ou censo é o tipo de pesquisa que envolve toda a população. No brasil, temos a do IBGE realizada recentemente, na qual toda a população foi entrevistada. Mas qual o critério para definir se a pesquisa deve ser censitária ou amostral, ou seja, se deve ser investigado toda a população ou uma amostra?

-Silêncio

**Professor Pesquisador:** Bem, quando o tamanho da população é grande, o custo e tempo para desenvolver a pesquisa são enormes, a exemplo da população brasileira. Ou ainda quando a pesquisa resulta na inutilização de um produto, como por exemplo a fabricação de fósforos. Assim, nem sempre é possível/recomendável realizar um censo. Nessas situações, se realiza a pesquisa amostral.

Fonte: Dados da pesquisa (2023).

A turma do Geogebra se comportou de modo semelhante a outra turma. Nota-se que os estudantes souberem indicar adequadamente qual é a população e amostra nas pesquisas abordadas da primeira questão. Porém, também apresentaram

dificuldades para conceituar esses termos de modo mais formal, ao conceberem a população apenas como “o todo” e “o geral” e à amostra como “uma parte”. Assim como na outra turma, o professor pesquisador apresentou o conceito formal para esses termos, enfatizando que a população e consequentemente a amostra podem envolver elementos que não são pessoas. Além disso, alguns estudantes também demonstram conhecimento de contexto paralelamente ao conhecimento estatístico (Gal,2002) ao indicarem exemplos de população que não abarcam pessoas. Além disso, destaca-se que nenhum aluno da turma soube dizer o nome que é dado a pesquisa populacional, como também, não souberam apontar os fatores que influenciam para a determinação se uma pesquisa deve ser censitária ou amostral. Diante disso, o professor pesquisador entrou em ação apresentando essas definições e contribuindo para aprendizagem dos estudantes sobre esse tema.

Em continuidade, a segunda questão da primeira atividade (Figura 121) abordou uma situação envolvendo pesquisa estatística na área da saúde, mais precisamente, na fabricação de medicamentos.

Figura 121 – Segunda questão da atividade 1

2º O dipirona é um medicamento indicado para dor e febre. Um estudo clínico com 628 crianças verificou a eficácia desse medicamento. Os resultados indicam que, após o uso do dipirona, 82% das crianças tiveram melhora significativa (Wong *et al*, 2001).

- Qual a amostra utilizada nesse estudo?
- Qual a importância do processo de amostragem para a fabricação de medicamentos?
- A partir dos resultados do estudo relatado, você compraria esse medicamento para uma criança com quadro clínico de dor e febre? Por que?

Fonte: Dados da pesquisa (2023).

Após a resolução dessa questão, procedemos nas duas turmas com a reflexão sobre a amostra utilizada na pesquisa e a importância da estatística inferencial para as pesquisas na área de saúde, ao possibilitar a comprovação científica da eficácia de medicamentos. O quadro 19 apresenta um trecho do diálogo dessa questão na turma Lápis e Papel.

Quadro 19 – Recorte do diálogo da 2º questão da Atividade 1 - turma Lápis e Papel

**Professor Pesquisador:** A segunda questão aborda uma situação de pesquisa amostral envolvendo a eficácia de um medicamento. A alternativa a pergunta: qual a amostra utilizada nessa pesquisa?

**Estudantes:** 628 crianças

**Professor Pesquisador:** Exato. A segunda alternativa perguntou qual a importância do processo de Amostragem para a fabricação de medicamentos. O que vocês responderam?

**Estudante 9:** a partir de uma amostra é possível testar e comprovar a eficácia de medicamentos

**Estudante 3:** Determinar a eficácia de medicamentos e procurar possíveis contraindicações.

**Professor pesquisador:** Ok. Correto. Na última alternativa há uma pergunta se você compraria o medicamento abordado na questão a partir dos resultados do estudo.

**Estudantes:** Sim!

**Professor pesquisador:** Por que?

**Estudante 21:** Mais de 80% das crianças tiveram melhora

**Estudante 3:** Apresenta uma porcentagem alta de eficácia

**Estudante 13:** Apresenta uma taxa de eficácia relativamente alta

**Professor Pesquisador:** Certo. Então o que podemos concluir dessa questão?

**Estudante 9:** Que a pesquisa com amostras tem sua importância também na questão de fabricar medicamentos.

**Professor Pesquisador:** Exatamente. A pesquisa amostral, em nosso cotidiano, é realizada em várias áreas da ciência. Na área da saúde, incluindo a da farmácia que engloba fabricação de medicamentos, a pesquisa amostral possibilita testar um medicamento em uma amostra e se garantir a eficácia e a generalização, garantindo que todas as pessoas podem utilizar, respeitando também, como falou o estudante 3, as possíveis contraindicações que o estudo também revela. Tudo isso é possível por meio de uma pesquisa amostral, garantido de maneira mais rápida e prática resultados confiáveis, comparado a uma pesquisa com toda a população, que seria muito grande e demandaria muito tempo, custo e trabalho.

Fonte: Dados da pesquisa (2023).

O diálogo relativo à segunda questão na turma Lápis e papel revela a importância de se trabalhar a Educação Estatística com dados reais do nosso cotidiano que permitem os estudantes desenvolverem uma postura crítica paralelamente a mobilização dos elementos de conhecimento do Letramento Estatístico (Gal,2002). O tema abordado na questão derivou uma reflexão sobre a importância da pesquisa amostral para a o desenvolvimento da ciência, mais precisamente para a fabricação e eficácia de medicamentos. A discussão evidencia que os estudantes dessa turma responderam adequadamente sobre o tamanho da amostra utilizado no estudo, a importância da pesquisa amostral para a fabricação de medicamentos e por fim, apresentaram uma tomada de decisão sobre a compra do

medicamento a partir dos resultados do estudo amostral. Innabi (2007) e Reyes (2019) apontam que é importante no ensino da Amostragem abordar situações do cotidiano que exploram o processo de pesquisa amostral para que os estudantes reconheçam sua utilidade e aprofundem os conhecimentos. Essa prática, em sala de aula, propicia o desenvolvimento do Letramento Estatístico (Gal,2002) e, consequentemente, contribui para a formação de cidadãos mais cientes e emancipados quanto aos resultados de estudos científicos que são pautados por pesquisas amostrais. O quadro 20 apresenta o diálogo dessa questão na turma Geogebra.

Quadro 20 – Recorte do diálogo da 2º questão da Atividade 1 - turma Geogebra

<b>Professor Pesquisador:</b> Dando continuidade, a segunda questão aborda uma situação de pesquisa amostral sobre a eficácia de um medicamento. A alternativa a pergunta: qual a amostra utilizada nessa pesquisa?
<b>Estudantes:</b> 628 crianças
<b>Professor Pesquisador:</b> Certo. A segunda alternativa pergunta qual a importância do processo de Amostragem para a fabricação de medicamentos. O que vocês responderam?
<b>Estudante 15:</b> Para verificar a eficácia
<b>Estudante 27:</b> possibilita ver se o medicamento tem eficácia.
Professor pesquisador: Ok. Correto. Na última alternativa há uma pergunta se você compraria o medicamento abordado na questão a partir dos resultados do estudo. Qual a resposta de vocês?
<b>Estudantes:</b> Sim!
<b>Professor pesquisador:</b> Por que?
<b>Estudante 19:</b> Teve eficácia em uma grande parte das crianças
<b>Estudante 31:</b> Teve eficácia comprovada.
Professor Pesquisador: Certo. Então qual a conclusão que podemos ter dessa situação apresentada na questão?
<b>Estudante 8:</b> Que a Estatística também está presente nessa questão de medicamentos e então é possível com amostras tem uma conclusão sobre a eficácia dos medicamentos.
<b>Professor Pesquisador:</b> Correto. A pesquisa amostral, em nosso cotidiano, é realizada em várias áreas da ciência. Na área da saúde também, incluindo a área da farmácia que engloba fabricação de medicamentos. Então a pesquisa amostral possibilita testar um medicamento em uma amostra e se garantir a eficácia e a generalização, garantindo que todas as pessoas podem utilizar, respeitando as possíveis contraindicações que o estudo vai revelar também. Tudo isso é possível por meio de uma pesquisa amostral, garantido de maneira mais rápida e prática resultados confiáveis, o que não seria o mesmo com uma pesquisa com toda a população, que seria muito grande e demandaria muito tempo, custo e trabalho para chegar as conclusões.

Fonte: Dados da pesquisa (2023).

A discussão realizada na turma Geogebra revela conclusões semelhantes as observadas na outra turma. Os estudantes responderam adequadamente ao tamanho da amostra, como também, sobre a importância da pesquisa amostral para a fabricação de medicamentos e apresentaram a tomada adequada de decisão sobre a compra do medicamento abordado da questão, mediante aos resultados do estudo. Essa prática dos estudantes ilustra a atitude e postura crítica paralela ao conhecimento estatístico mobilizados, importantes fatores para o desenvolvimento do Letramento Estatístico (Gal, 2002). Assim como na outra turma, o professor pesquisador ainda contribuiu sobre a importância da pesquisa amostral para a ciência comparada a pesquisa populacional, tendo em vista nem sempre é viável se realizar uma pesquisa censitária por conta do tamanho da população e/ou a inacessibilidade aos elementos populacionais.

Na sequência, demos continuidade à atividade 1 com a resolução da terceira questão, a qual teve o objetivo de aprofundar o conhecimento relativo à Amostragem. Essa questão abordou os resultados de uma pesquisa amostral relativa à opinião de jovens brasileiros sobre a ciência e a tecnologia. Além de apresentar os resultados da pesquisa, também informou o tamanho da amostra, o perfil das pessoas que foram entrevistadas, o nível de confiança e a margem de erro da pesquisa.

Figura 122 – Terceira questão da atividade 1

2º) O instituto Nacional de Ciência e tecnologia em Comunicação (INCT-CPCT) realizou uma pesquisa com 2206 jovens brasileiros para saber o que eles pensam sobre ciência e tecnologia. A pesquisa foi realizada com jovens do sexo masculino e feminino, de 15 a 24 anos e de todas as regiões do Brasil. O nível de confiança é de 95% e a margem de erro de 2%. A Figura, a seguir, apresenta os principais resultados da pesquisa.



- Qual o tamanho da amostra?
- Na sua opinião, por que o pesquisador entrevistou uma amostra e não a população toda?
- Na sua opinião, por que as características como o sexo e a região geográfica foram importantes para a seleção da amostra?
- O que significa a margem de erro de 2% nessa pesquisa?

Fonte: Dados da pesquisa (2023).

Antes da resolução dessa questão, o professor pesquisador com o auxílio de uma apresentação em slide, abordou e debateu com os estudantes, nas duas turmas, uma sistematização teórica sobre o conceito de representatividade amostral, de variabilidade nos tipos, natural, amostral e de medida (Reyes, 2019), e os métodos probabilísticos mais comuns em pesquisas (Triola, 2008): Amostragem Aleatória Simples, Amostragem Sistemática, Amostragem Estratificada e Amostragem por Conglomerados. Tomamos essa iniciativa por verificar que no pré-teste a maioria dos estudantes apresentou profundas lacunas na aprendizagem desses conceitos. Os mesmos foram abordados de modo semelhante a como estão apresentados e discutidos no capítulo 1 do presente estudo.

Nesse momento de discussão teórica foi possível perceber que a maioria dos estudantes não conheciam sobre a representação amostral e os fatores que estão inerentes a ela, isto é, a variabilidade dos dados, o tamanho da amostra e o método de seleção amostral. A partir disso, iniciamos a apresentação do conceito de variabilidade nos seus três tipos, conforme Reyes (2019): natural, amostral e de medida. Na abordagem sobre a variabilidade de medida, ou seja, a margem de erro, também estudamos sobre o intervalo de confiança, e o significado do nível de confiança de 95% abordado na questão.

Posteriormente, debatemos sobre a relação entre a variabilidade natural e o tamanho da amostra para se realizar a pesquisa. Nessa direção, refletimos que quando a população é mais homogênea, ou seja, que tenha uma menor variabilidade, por exemplo um lote de peças, menor o tamanho necessário da amostra para representá-la. Do contrário, uma população com maior variabilidade, a exemplo da população brasileira, maior é esse tamanho, pois será necessário conter todas as características da população para que a amostra seja, de fato, representativa. No entanto, chamamos a atenção que essa relação entre a variabilidade e o tamanho da amostra é resultado de oscilações aleatórias e não é estabelecida de forma proporcional.

Em seguida, estudamos e debatemos (Quadro 21) sobre os métodos probabilísticos, destacando os seus tipos e o caráter aleatório que os compõem. Logo, refletimos que os métodos probabilísticos são aqueles em que a probabilidade de cada elemento da população pertencer à amostra é conhecida e diferente de zero (Moore, 1995; Triola, 2008). Nesse momento, foi perceptível que os estudantes não conheciam, nem haviam estudado sobre totalidade desses métodos.

Quadro 21 – Recorte do diálogo da 3º questão da Atividade 1 - turma Lápis e Papel

**Professor Pesquisador:** Qual a opinião de vocês sobre esses métodos probabilísticos e as respectivas formas de seleção?

**Alguns estudantes:** Não sabia sobre isso

**Estudante 5:** Achei bem interessante

**Estudante 21:** Achei interessante. A aleatória simples, que é basicamente um sorteio, eu não achava que poderia garantir a representação. Mas agora entendo.

**Professor Pesquisador:** Certo. Vamos agora debater sobre a 3 questão, seguindo o que é questionado em cada alternativa. Qual foi o tamanho da amostra?

**Estudantes:** 2206

**Professor Pesquisador:** Ok. Por que o pesquisador entrevistou uma amostra e não a população toda?

**Estudante 21:** Seria muito difícil, caro, entrevistar toda a população.

**Estudante 13:** Porque é inviável.

**Estudante 30:** Seria muito trabalhoso

**Professor pesquisador:** Ok. E por que as características como o sexo e a região geográfica foram importantes para a seleção da amostra?

**Estudante 21:** Para informar que a pesquisa não foi feita com um grupo fechado de um só lugar. Foram pessoas dos dois sexos e de várias regiões do país

**Estudante 15:** Uma forma de variar bem para garantir a representação

**Estudante 12:** Se fosse em uma região só não estaria correto

**Professor pesquisador:** Isso que vocês afirmam está relacionado a qual conceito que discutimos há pouco?

**Estudantes:** Variabilidade!

**Professor Pesquisador:** Exato. Vimos que a variabilidade em uma pesquisa amostral é um fator importante a ser considerado e está associada ao grau dispersão dos dados e a quantidade de fatores para caracterizar uma população. Nessa pesquisa, as regiões geográficas e o sexo foram características importantes para garantir a variabilidade e, consequentemente, está relacionada ao método de Amostragem e ao tamanho da amostra, para garantir a representação.

**Professor Pesquisador:** O que significa a margem de erro de 2% nessa pesquisa?

**Estudante 15:** É a variação

**Estudante 21:** Que pode ser um valor a mais ou a menos. 2% a mais ou 2% a menos do que resultado que der.

**Professor pesquisador:** Certo. Como vimos, a margem de erro é a variabilidade de medição, ou seja, representa a diferença entre o resultado amostral e o verdadeiro resultado da população, porque dificilmente uma amostra vai apresentar resultados idênticos a sua população. Esse conceito está relacionado com o nível de confiança e o intervalo de confiança. Na atividade, é indicado que o nível de confiança é de 95%, o que significa que

*há uma probabilidade de 95% de que o intervalo estabelecido contenha o parâmetro populacional. Em outras palavras, se repetíssemos a seleção de amostras muitas vezes, em 95% dessas repetições, o intervalo resultante incluiria o verdadeiro valor da população. O intervalo de confiança representa uma faixa de valores na qual esperamos que esse valor populacional esteja contido. O centro do intervalo é a estimativa pontual, ou seja, a estatística obtida na amostra, e a ele é adicionada a margem de erro, que, neste caso, é de 2%, indicando a variação permitida para mais ou para menos em relação a essa estimativa pontual.*

Fonte: Dados da pesquisa (2023).

A partir de discussões, destaca-se que os estudantes, de modo geral, afirmaram que não conheciam sobre os métodos probabilísticos ou nunca tinha estudado sobre eles. No entanto, a abordagem realizada possibilitou a aprendizagem sobre os mesmos. O diálogo também evidenciou a construção do conhecimento dos estudantes sobre a importância e a necessidade da realização de uma pesquisa amostral, em comparação com a pesquisa censitária quando se tem uma população muito grande. Respostas apresentadas pelos estudantes, tais como “seria inviável” ou “seria muito difícil, caro, entrevistar toda a população” elucidam essa conclusão. Como também, a compreensão sobre a variabilidade natural como um fator importante para se garantir a representação amostral e o significado do nível de confiança, intervalo de confiança e da margem de erro, enquanto variabilidade de medida. Destaca-se que o estudante 21 relata, ao seu modo, o significado da margem de 2% na pesquisa abordada. Assim, a atividade realizada possibilitou aprimoramento do conhecimento estatístico e matemático (Gal,2002) sobre os conceitos abordados. De modo semelhante, foram discutidos (Quadro 22) os conceitos abordados nessa questão na turma Geogebra.

Quadro 22 – Recorte do diálogo da 3º questão da Atividade 1 - Geogebra

**Professor Pesquisador:** Qual a opinião de vocês sobre esses métodos probabilísticos e as respectivas formas de seleção?

**Estudante 12:** Eu achei interessante. Eu não conhecia todos eles, basicamente só o aleatório simples, que é tipo um sorteio.

**Estudante 10:** Eu nunca tinha estudado sobre isso. Achei interessante, principalmente a estratificada.

**Estudante 15:** Eu também nunca tinha visto isso.

**Professor Pesquisador:** Ok. Vamos agora debater sobre a 3 questão, seguindo o que é questionado em cada alternativa. Qual foi o tamanho da amostra?

**Estudantes:** 2206

**Professor Pesquisador:** Ok. Por que o pesquisador entrevistou uma amostra e não a população toda?

**Estudante 8:** Porque seria inviável entrevistar a população todinha

**Estudante 13:** Como a população é grande, é mais fácil pesquisar através de uma amostra.

**Estudante 25:** Mais prático e rápido fazer com uma amostra.

**Professor pesquisador:** Ok. E por que as características como o sexo e a região geográfica foram importantes para a seleção da amostra?

**Estudante 3:** Para garantir a representatividade

**Estudante 13:** Para garantir a participação de todos os seguintes sociais possível. Fazendo com que tenha a representação

**Professor pesquisador:** Isso está relacionado a qual conceito que discutimos há pouco?

**Estudantes:** Variabilidade!

**Professor Pesquisador:** Exato. Como vimos, a variabilidade em uma pesquisa amostral é um fator importante a ser considerado e está associada ao grau dispersão dos dados e a quantidade de fatores para caracterizar uma população. Nessa pesquisa, as regiões geográficas e o sexo foram características importantes para garantir a variabilidade e, consequente, aliada ao método de Amostragem e ao tamanho da amostra, a representação.

**Professor Pesquisador:** O que significa a margem de erro de 2% nessa pesquisa?

**Estudante 25:** Variação dos dados

**Estudante 3:** Os resultados podem variar 2% acima ou 2% abaixo. Nesse resultado de 60% podia ser 58% ou 62%.

**Professor pesquisador:** Perfeito. Como vimos há pouco, a margem de erro é a variabilidade de medição, ou seja, representa a diferença entre o resultado amostral e o verdadeiro resultado da população, porque dificilmente uma amostra vai apresentar resultados idênticos a sua população. Além disso, vocês devem ter notado que a indicação do nível de confiança de 95%. Alguém poderia dizer o que significa essa informação?

-Silêncio

**Professor Pesquisador:** Nesse exemplo é indicado que o nível de confiança é de 95%, ou seja, há uma probabilidade de 95% de que o intervalo estabelecido contenha o parâmetro populacional. Assim, se repetíssemos a seleção de amostras muitas vezes, em 95% dessas repetições, o intervalo resultante incluiria o verdadeiro valor da população. O intervalo de confiança representa um intervalo de valores na qual esperamos que esse valor populacional esteja contido. O centro do intervalo é justamente a estimativa pontual, ou seja, a estatística obtida na amostra, a exemplo do 60% que o estudante 3 mencionou. A ele é adicionada justamente margem de erro, que, neste caso, é de 2%, indicando a variação permitida para mais ou para menos em relação a essa estimativa pontual.

Fonte: Dados da pesquisa (2023).

O diálogo realizado na turma Geogebra evidencia conclusões semelhantes às observadas na outra turma. Inicialmente, os estudantes, afirmaram que passaram a conhecer a totalidade dos métodos probabilísticos e suas características a partir daquele momento. Além disso, a intervenção de ensino também possibilitou a aprendizagem sobre a importância da pesquisa amostral, compara a pesquisa censitária em contextos de populações grandes, como também, o conceito de variabilidade, nos tipos, natural, amostral e de medida (Reyes, 2019) e o quanto ela está presente em processos de pesquisas amostrais para se garantir a representatividade. A fala do estudante 3 evidencia a compreensão da margem de erro da questão ao apontar o intervalo determinado por ela em uns dos resultados da pesquisa mencionada na questão. Além disso, os estudantes dessa turma também avançaram na compreensão dos conceitos de nível de confiança e intervalo de confiança. Ao final da terceira questão, destacamos que a intervenção se ensino até aqui, possibilitou a construção e o aprimoramento tanto dos elementos de conhecimento como de disposição (Gal,2002) dos estudantes de ambas as turmas sobre a Amostragem e os conceitos que ela abarca. Nesse viés, Batanero, Begué e Gea (2018) argumentam que a Amostragem é um conceito complexo porque reúne vários outros conceitos subjacentes, como os explorados nessa questão. Logo, se faz necessário abordá-los para que os estudantes possam avançar na aprendizagem sobre a Amostragem.

Dando prosseguimento, a quarta questão da primeira atividade teve o objetivo de propiciar discussões e reflexões sobre os métodos de Amostragem não probabilísticos e o caráter de não representação amostral dos mesmos. Essa questão abordou uma enquete realizada por um site de notícia sobre o perfil do consumidor brasileiro e questionou se os resultados dessa pesquisa poderiam ser generalizados para toda a população.

Figura 123 – quarta questão da atividade 1

4º) Um site de notícia brasileiro lançou uma enquete, apresentada na imagem abaixo, para saber o perfil do consumidor brasileiro. Você considera que esse tipo de pesquisa pode ser representativo (generalizado) para toda população do Brasil? Justifique sua resposta.

<b>Você já comprou alguma coisa por impulso?</b>	
<input type="radio"/>	<input type="radio"/>
Sim	Não

Fonte: Dados da pesquisa (2023).

Nessa direção, antes da resolução da questão o professor pesquisador, com o auxílio de uma apresentação em slide, assim como na abordagem dos métodos probabilísticos, abordou nas duas turmas uma sistematização teórica sobre as características dos métodos por conveniência, resposta voluntária, julgamento e por quotas (Triola, 2008), destacando a importância deles para determinados tipos de pesquisa, como a de exploração e o caráter não representativo. Nessa direção, apresentamos (Quadro 23) o recorte do diálogo realizado na turma Lápis e Papel.

Quadro 23 – Recorte do diálogo da 4º questão da Atividade 1 - turma Lápis e Papel

<b>Professor Pesquisador:</b> <i>Sobre esses métodos não probabilísticos e suas formas de seleção, qual a opinião de vocês?</i>
<b>Estudante 10:</b> <i>Também não conhecia</i>
<b>Estudante 5:</b> <i>Achei também interessante. Apesar de não gerar a representação.</i>
<b>Estudante 15:</b> <i>Essa forma por conveniência eu achei a mais interessante. A pessoa acaba escolhendo os que estão mais perto, mais acessíveis, nesse caso foi o exemplo do rapaz que entrevistou no refeitório naquele teste</i>
<b>Professor Pesquisador:</b> <i>Exato. Vamos seguir respondendo a quarta questão. O que vocês responderam?</i>
<b>Alguns estudantes:</b> <i>Não pode ser representativo.</i>
<b>Professor pesquisador:</b> <i>Por que?</i>
<b>Estudante 15:</b> <i>Porque não abrange todo mundo</i>
<b>Estudante 13:</b> <i>Porque nem todo mundo acessa esse site para responder. Só quem responde é quem acompanha o site. Não tem variabilidade</i>
<b>Professor pesquisador:</b> <i>Isso mesmo. E qual o tipo de método não-probabilístico utilizado nessa pesquisa?</i>
<b>Alguns estudantes:</b> <i>Resposta voluntária</i>
<b>Professor pesquisador:</b> <i>Perfeito. Mas por que?</i>
<b>Estudante 3:</b> <i>Porque a resposta depende da decisão das pessoas se vão participar ou não.</i>
<b>Estudante 15:</b> <i>Depende do desejo das pessoas.</i>
<b>Professor pesquisador:</b> <i>Ok. Correto</i>

Fonte: Dados da pesquisa (2023).

A sistematização teórica sobre os métodos não probabilísticos realizados nas duas turmas possibilitou a reflexão sobre o caráter não aleatório para seleção dos elementos amostrais e dos vieses de não representação amostral presentes nesses métodos. O diálogo realizado na turma Lápis e Papel indica, inicialmente, que os estudantes não conheciam os métodos não probabilísticos, assim como os probabilísticos. Além disso, notabiliza o entendimento dos estudantes sobre o caráter

de não representação amostral e a compreensão sobre as características deles. O estudante 15, ao longo do diálogo, conseguiu lembrar de uma pesquisa abordada no pré-teste que se tratava do método de conveniência. Outros responderam adequadamente que o método abordado na questão era o de resposta voluntária. Isso evidencia a apropriação, por parte dos estudantes, do conhecimento estatístico (Gal, 2002) sobre as características desses métodos. De modo análogo, destacamos (Quadro 24) o diálogo realizado na turma Geogebra.

Quadro 24 – Recorte do diálogo da 4º questão da Atividade 1 - Geogebra

<b>Professor Pesquisador:</b> Sobre esses métodos não probabilísticos gostaria de ouvir a opinião de vocês sobre essas formas de seleção
<b>Estudante 21:</b> Como o senhor falou, eles não garantem a representação, mas também são utilizadas. Esse método aí de resposta voluntária a gente ver muito nas redes sociais.
<b>Professor Pesquisador:</b> Por que uma enquete em rede social não pode ser representativa estatisticamente?
<b>Estudante 21:</b> Por que só quem pode responder é quem é seguidor da página
<b>Professor Pesquisador:</b> Mais alguém pode dar uma opinião?
<b>Estudante 12:</b> Esses eu também nunca tinha estudado
<b>Estudante 13:</b> Também não conhecia
<b>Professor pesquisador:</b> Ok. Vamos agora para a resposta da quarta questão. O que vocês responderam?
<b>Estudantes:</b> Não pode ser representativa
<b>Professor pesquisador:</b> Por que?
<b>Alguns estudantes:</b> Só quem responde é quem acompanha o site.
<b>Estudante 3:</b> Nem todo mundo vai responder
<b>Professor Pesquisador:</b> Exato. Então qual esse tipo de método dessa pesquisa?
<b>Estudantes:</b> Resposta voluntária
<b>Professor Pesquisador:</b> Por que é por resposta voluntária?
<b>Estudante 21:</b> Só quem pode responder é quem acompanha o site e tem a vontade de participar
<b>Professor Pesquisador:</b> Exato

Fonte: Dados da pesquisa (2023).

O debate realizado entre o pesquisador e os estudantes da turma Geogebra evidencia que eles também não conheciam os métodos não probabilísticos. Destacamos que o estudante 21, após a sistematização teórica, mencionou as enquetes realizadas por redes sociais como um exemplo do método resposta voluntária, demonstrando conhecimento de contexto (Gal,2002) sobre o tema aliado ao

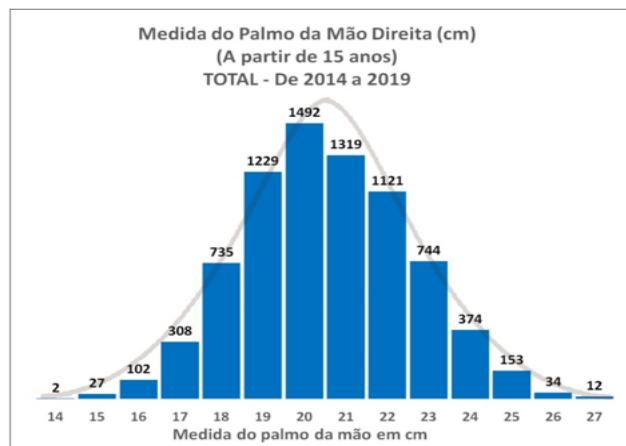
conhecimento estatístico. Além disso, podemos notar que os estudantes responderam adequadamente a quarta questão indicando o método utilizado e seu caráter não representativo. A partir disso, destacamos que a intervenção de ensino proporcionou aos estudantes de ambas as turmas a construção do conhecimento estatístico (Gal,2002) sobre os métodos de Amostragem.

Na sequência, os estudantes prosseguiram para a quinta questão (Figura 124) da atividade 1, a qual teve o objetivo de introduzir o modelo da Curva Normal com foco para a sua representação gráfica e as medidas de centralidade e dispersão que estão presentes nesse modelo. Tendo em vista que a maioria dos estudantes demonstrou não conhecer esse modelo, recorremos a mesma estratégia na abordagem dos métodos de Amostragem, ou seja, antes da resolução da questão procedemos com uma sistematização teórica da Curva Normal, a partir do gráfico da quarta questão do pré-teste, o qual abordou a distribuição da pressão arterial de um conjunto de pessoas.

Nessa direção, discutimos a ideia de que determinados fenômenos na natureza seguem um determinado padrão, em que a disposição gráfica dos dados obedece a uma Curva Normal, em que os valores extremos têm menor frequência e os valores mais próximos da média, localizada no centro, ocorrem em maior número de vezes. Dessa forma, buscamos ressaltar a importância desse modelo para análise na Inferência Estatística pelo fato do mesmo modelar de uma variedade de fenômenos do nosso cotidiano.

Figura 124 – Quinta questão da atividade 1

5º) A imagem, a seguir, apresenta a distribuição das medidas do palmo da mão direita (cm) de 7652 pessoas, entre homens e mulheres, a partir de 15 anos, que participaram de um estudo. Com base nesses dados responda aos itens dessa questão.



- a) Na sua opinião, essa distribuição de dados, estabelece, aproximadamente, algum tipo de simetria?
- b) Uma pessoa que tem a medida da palma da mão igual a 22 cm corresponde ao valor da média da palma da mão de todos os participantes dessa pesquisa? Justifique sua resposta.
- c) Uma fábrica de luvas de borracha analisou esses dados e decidiu fabricar a mesma quantidade de luvas para cada tamanho. A escolha foi correta? Justifique sua resposta.
- d) Duas pessoas desse grupo foram escolhidas ao acaso, a primeira possui a medida do palmo da mão igual a 17 cm. Já a segunda tem essa medida igual a 22 cm. Qual das duas possui a medida com maior desvio-padrão em relação à média de todo o grupo?

Fonte: Dados da pesquisa (2023).

Além disso, destacamos que formato da Curva Normal depende dos parâmetros média ( $\mu$ ) e desvio padrão ( $\sigma$ ) estabelecidos em cada situação. Assim, foi debatido que média corresponde ao valor central, e que o mesmo coincide com o valor da mediana e moda do conjunto. Já o desvio-padrão indica a dispersão do conjunto. Diante disso, refletimos quanto maior o Desvio Padrão, maior será a dispersão e a amplitude dos dados, e, consequentemente, a Curva Normal será mais comprimida, com menos dados concentrados em torno da média. Do contrário, quanto menor o Desvio Padrão, menor será a dispersão do conjunto e a amplitude dos dados e, consequente, a Curva Normal terá mais dados concentrados próximos a média.

Nessa etapa, constatamos que a maioria dos estudantes, de ambas as turmas, não tinham estudado ou conheciam o modelo da Curva Normal. Um fato curioso observado se deu na abordagem sobre o “outlier”, ou seja, um ponto fora da curva que tem comportamento discrepante em relação à média e que pode surgir em uma Curva Normal. Os estudantes de ambas as turmas alegaram que conheciam a expressão “ponto fora da curva” para tratar de algo que é incomum ou raro de acontecer, no cotidiano, mas que não imaginavam que essa expressão tinha relação com o modelo da Curva Normal. Após a abordagem da referida sistematização teórica, procedemos com a resolução da quinta questão. Dessa forma, destacamos (Quadro 25) um trecho do diálogo realizado na turma Lápis e Papel.

Quadro 25 – Recorte do diálogo da 5º questão da Atividade 1 – turma Lápis e Papel

**Professor Pesquisador:** Sobre essa quinta questão, qual a resposta da alternativa a?

<b>Estudantes:</b> Tem simetria
<b>Professor Pesquisador:</b> Em relação a qual valor?
<b>Estudante 3:</b> A média
<b>Estudante 15:</b> O valor do centro aí que é a média
<b>Professor Pesquisador:</b> Na alternativa b: A pessoa que tem a medida da palma da mão igual a 22 cm corresponde ao valor da média dessa distribuição?
<b>Estudantes:</b> Não!
<b>Professor Pesquisador:</b> Por que?
<b>Estudante 3:</b> O valor da média está entre 20 e 21 que é a média e também a moda e mediana
<b>Estudante 15:</b> 22 não é o valor central
<b>Professor Pesquisador:</b> Correto. Seguindo com a alternativa C, uma fábrica de luvas de borracha analisou esses dados e decidiu fabricar a mesma quantidade de luvas para cada tamanho. A escolha foi correta?
<b>Estudantes:</b> Não!
<b>Professor Pesquisador:</b> Por que?
<b>Estudante 12:</b> Porque tem diferentes valores, maior quantidade próximo da média e vai diminuindo ao se afastar dela.
<b>Professor Pesquisador:</b> Qual medida indica a intensidade desse afastamento em relação à média?
<b>Alguns estudantes:</b> Desvio-padrão
<b>Professor pesquisador:</b> Exato. Então quando se aumenta o desvio-padrão a frequência dos dados vai diminuindo de forma simétrica, acima e abaixo da média, caracterizando o gráfico em curva
<b>Professor pesquisador:</b> Com base nisso, qual a resposta da letra d?
<b>Alguns estudantes:</b> A de 17 cm
<b>Professor pesquisador:</b> Por que?
<b>Estudante 8:</b> A pessoa que tem 17 cm está mais diante da média que é o centro.
<b>Estudante 15:</b> Se ela está mais distante tem maior desvio-padrão
<b>Professor pesquisador:</b> Exato. Correto

Fonte: Dados da pesquisa (2023).

Esse diálogo elucida a consolidação do conhecimento construído por parte dos estudantes da turma de Lápis e Papel sobre os fundamentos teóricos do modelo da Curva Normal referentes às medidas de centralidade e variabilidade/dispersão, como também, a representação gráfica. A partir da sistematização teórica e da resolução da atividade, os estudantes passaram a compreender a característica do gráfico, ou seja, seu formato em curva, a média como valor central da distribuição, a qual coincide com a mediana e a moda, e a relação entre a média e o desvio-padrão, no qual este indica o grau de dispersão, ou distanciamento dos dados da distribuição em torno da média.

Destacamos que além de mobilizarem os elementos de conhecimento (Gal, 2002), os estudantes explicitaram uma postura crítica ao responderem a alternativa c, emitindo o julgamento quanto à tomada de decisão da fábrica. Logo, atividades semelhantes a essa, que explorem dados reais e estimulem os estudantes a pensar estatisticamente e também exercerem a criticidade, podem contribuir para o desenvolvimento do Letramento Estatístico (Gal, 2002). Na turma Geogebra, procedemos de forma semelhante, discutindo (Quadro 26) os fundamentos teóricos da Curva Normal a partir da atividade e da sistematização teórica.

Quadro 26 – Recorte do diálogo da 5º questão da Atividade 1 - turma Geogebra

**Professor Pesquisador:** Vamos agora para a quinta questão, qual a resposta da alternativa a?

**Alguns Estudantes:** Tem simetria

**Professor Pesquisador:** Como é essa simetria? Com base em que valor?

**Estudante 12:** tem simetria porque os dados têm menos quantidade nos extremos e aumenta a quantidade próximo à média. Então é em relação a média a simetria.

**Professor Pesquisador:** Perfeito. Em relação ao ponto central, que é a média. Na alternativa b: A pessoa que tem a medida da palma da mão igual a 22 cm corresponde ao valor da média dessa distribuição?

**Estudantes:** Não!

**Professor Pesquisador:** Por que?

**Estudante 3:** o ponto central está entre 20 e 21

**Estudante 15:** 22 não é a média. Está uma pouco mais à direita

**Professor Pesquisador:** Correto. A alternativa C diz que uma fábrica de luvas de borracha analisou esses dados e decidiu fabricar a mesma quantidade de luvas para cada tamanho. A escolha foi correta?

**Estudantes:** Não!

**Professor Pesquisador:** Por que?

**Estudante 11:** Não faz sentido porque tem 14 medidas e quantidades diferentes.

**Professor Pesquisador:** Qual seria atitude correta então?

**Estudante 21:** Fabricar mais a quantidade que é a média que também é a moda, que tem maior quantidade.

**Estudante 27:** De forma proporcional. Fabricar mais a que tem mais quantidade

**Professor Pesquisador:** Correto. Em relação a diminuição da frequência a medida que se afasta da média, qual medida que indica a intensidade desse afastamento em relação da média?

**Alguns estudantes:** Desvio-padrão

**Professor pesquisador:** Exato. Então, revisando o que estudamos, quando se aumenta o desvio-padrão a frequência dos dados vai diminuindo de forma simétrica, acima e abaixo da média, caracterizando o gráfico da Curva Normal

**Professor pesquisador:** A partir disso, qual a resposta da letra d? Qual das duas possui a medida com maior desvio-padrão em relação à média?

**Alguns estudantes:** A de 17 cm

**Professor pesquisador:** Por que?

**Estudante 8:** Está mais distante

**Estudante 15:** A que tem 17 cm está mais longe do centro

**Professor pesquisador:** Exato. Correto

Fonte: Dados da pesquisa (2023).

Esse diálogo evidencia que os estudantes da turma Geogebra também consolidaram o conhecimento sobre a representação gráfica da Curva Normal e as medidas de centralidade e dispersão presentes nesse modelo. Essa conclusão pode ser observada a partir das respostas adequadas dos estudantes responderam adequadamente quando perguntados sobre a simetria do gráfico, média da distribuição e a relação entre esta e o desvio-padrão, enquanto medida de variabilidade. Assim, como na outra turma, também mobilizaram uma postura crítica para julgar a tomada de decisão da fábrica, na alternativa c da quinta questão. Diante disso, essa questão possibilitou a introdução ao modelo da Curva Normal, nas duas turmas, a aprendizagem dos conceitos abordados.

Ao final da análise da atividade 1, ratificamos que a intervenção de ensino no primeiro encontro, possibilitaram aos estudantes das duas turmas, a construção e aprimoramento do conhecimento sobre os conceitos de população, amostra, representatividade e tamanho amostral, variabilidade, métodos de Amostragem e o modelo da Curva Normal, com foco na representação gráfica e nas medidas de centralidade e dispersão abarcadas por esse modelo. Logo, essa primeira etapa propiciou, aos estudantes, o desenvolvimento tanto dos elementos de conhecimento quanto de disposição do Letramento Estatístico (Gal,2002).

#### 9.2.2 2º Momento: Etapas do Ciclo Investigativo e Variabilidade Amostral

A atividade do segundo momento da nossa intervenção foi direcionada para a vivencia de algumas etapas do ciclo investigativo (amostra, coleta de dados, registro e representação dos dados e a análise e interpretação) e a reflexão, com uma simulação prática, sobre a variabilidade na Amostragem, nos tipos natural, amostral e

de medida (Reyes, 2019). Os estudantes seguiram organizados em duplas, e a partir desse momento, as atividades foram conduzidas e adaptadas ao recurso utilizado por cada turma, sendo uma o Lápis e Papel e a outra, o Geogebra.

### **9.2.2.1 Turma Lápis e Papel**

Na turma Lápis e papel, a primeira questão solicitou que os estudantes procedessem com a coleta das alturas de todos os estudantes da turma, e em seguida, a elaboração de um banco de dados, em uma folha de ofício, com todas as 34 medidas das alturas. Recapitulamos que nessa etapa, ao coletar esses dados, consideramos cada turma como uma população de interesse e a escolha dessa variável se deu pelo fato de ser contínua e a Curva Normal descrever comportamento de variáveis com essa característica. A segunda questão, por sua vez, solicitou que os estudantes a partir da média e o desvio-padrão das alturas fornecidos, respondessem qual o significado dessas medidas. Relembramos também que nessa questão o foco não era o cálculo em si dessas duas medidas, mas sim, o significado delas, algo que é mais primordial para a aprendizagem dos estudantes. Logo, considerando que o banco de dados continha 34 valores para a altura, o professor pesquisador apresentou aos estudantes, com o auxílio de um projetor de Datashow, o valor da média e do desvio-padrão dos 34 dados, utilizando o software Excel como ferramenta para o cálculo. Tal procedimento foi adotado para dar mais viabilidade a intervenção de ensino e possibilitar que os estudantes conhecessem um recurso que possibilita o cálculo dessas medidas, principalmente quando se tem uma grande quantidade de dados.

**Figura 125 – Primeira e Segunda questão da atividade 2**

- 1º) Caros estudantes, para darmos início a nossa atividade, realize o seguinte procedimento:
- Faça uma pesquisa com seus colegas da classe e colete a altura (em cm) de cada um deles
  - Organize os dados em uma lista com o número de identificação dos alunos e suas respectivas alturas
- 2º) A média da altura dos estudantes de sua classe é igual a \_\_\_\_\_. O desvio-padrão é igual a \_\_\_\_\_. O que essas medidas significam?

Fonte: Dados da pesquisa (2023).

Dando continuidade, a terceira questão teve o objetivo de a partir de uma simulação prática, compreender o conceito de variabilidade e refletir sobre a relação entre o tamanho da amostra e a variabilidade para a representatividade amostral. Para tanto, os estudantes, com o banco de dados em mãos, procederam com a seleção aleatória de amostras com diferentes tamanhos para em seguida, calcular a média de cada amostra. Na sequência responderam as indagações da questão (Figura 126)

Figura 126 – Terceira questão da atividade 2

3º) A partir da lista com os dados das alturas de todos os estudantes, selecione, aleatoriamente, amostras de estudantes de acordo com cada tamanho destacados a seguir, e em seguida, calcule a média das alturas de cada amostra.

Tamanho da Amostra	Média da altura (cm)
Amostra 1: 4 estudantes	
Amostra 2: 8 estudantes	
Amostra 3: 12 estudantes	
Amostra 4: 16 estudantes	

- a) Compare as medidas das suas médias das amostras com as de outra dupla de colegas. Os resultados são iguais ou diferentes?
- b) De acordo com o tamanho da amostra, qual delas tem a média mais próxima da média de populacional (todos os alunos da turma)?
- c) Na sua opinião, para uma amostra representativa há alguma relação entre o tamanho da amostra e a margem de erro?
- d) Se essa pesquisa fosse realizada com toda a escola, quais características da população deveriam ser consideradas para a seleção de uma amostra representativa?

Fonte: Dados da pesquisa (2023).

Nessa questão, os diferentes tamanhos da amostra e suas respectivas médias, possibilitou o entendimento e a observação da variabilidade amostral e como a variabilidade de medida nos resultados amostrais diminui ao passo que o tamanho da amostra aumenta. Consequentemente, observamos que a precisão das estimativas estatísticas também aumenta, ou seja, foi possível refletir como o tamanho da amostra afeta a representatividade das amostras e a variabilidade de medida (margem de erro) das estimativas estatísticas. Ben-Zvi, Bakker e Makar (2015) defendem que a prática de simulações é pertinente no processo de ensino e aprendizagem de Amostragem

porque propicia, de maneira prática, o entendimento de conceitos que estão subjacentes a essa área, a exemplo da variabilidade. Além disso, a alternativa “d” possibilitou a reflexão sobre a variabilidade natural (diferentes características dos elementos amostrais) e sua importância para a representação amostral, tópico de conhecimento que se mostrou defasado nas concepções apresentadas pelos estudantes no pré-teste. Apresentamos na sequência (Quadro 27) um trecho do diálogo com os estudantes da turma Lápis e Papel na socialização da segunda e terceira questão.

Quadro 27 – Recorte do diálogo da 2º e 3º questão da Atividade 2 - turma Lápis e Papel

**Professor Pesquisador:** Vamos comentar sobre a 2º questão. A média das alturas, como vimos, é aproximadamente 168 cm. O desvio-padrão é igual, aproximadamente a 8 cm. Começando pela média, o que ela significa nessa situação?

**(Silêncio)**

**O professor Pesquisador:** O que essa medida, ou seja, a média que é 168 cm significa?

**Estudante 23:** A média da altura turma

**Professor pesquisador:** Certo. Mas o que podemos concluir a partir dela? O que ela significa?

**Estudante 3:** É uma medida de centralidade. Que nesse caso representa essa turma. A concentração das alturas está para esse valor.

**Professor Pesquisador:** Isso. Ela é uma medida representativa para esse conjunto de dados. Então, a média representa um valor central que equilibra esse conjunto. Ela fornece uma indicação geral de onde os valores estão concentrados, sendo bastante indicada quando temos distribuição de dados simétricas, ou seja, quando os valores estão distribuídos de forma homogênea em torno do centro, como vimos na Curva Normal. Porém ela é sensível, é afetada quando temos dados extremos, chamados também de (outliers) que estudamos na última aula. Quando isso acontece, outras medidas são mais indicadas, como a mediana. Temos também a moda, como medida de centralidade que é mais indicada para representar um conjunto quando se deseja identificar valores mais frequentes, geralmente em dados discretos que são contáveis e assumem valores inteiros.

**Professor Pesquisador:** E o valor do desvio-padrão? O que ele significa nessa situação?

**Estudante 15:** A dispersão

**Estudante 3:** O grau de distanciamento para a média dos valores. Quanto maior mais distante está e quanto menor o valor do desvio-padrão significa que está variando menos em relação à média

**Professor Pesquisador:** Exato. É uma medida de dispersão que justamente fornece uma ideia de quão longe os valores individuais estão da média do

conjunto de dados. Um desvio-padrão maior indica uma maior dispersão dos valores em relação à média, enquanto um desvio-padrão menor indica uma menor dispersão. Então é importante tanto entender a centralidade de um conjunto quanto a dispersão para compreender o seu comportamento.

**Professor pesquisador:** Na questão 3, vocês compararam os resultados das médias com as outras duplas, foram iguais?

**Estudantes:** Não! Diferentes!

**Professor pesquisador:** Por que?

**Estudante 21:** Porque as amostras se diferem

**Professor pesquisador:** E essa diferença se dá por conta de que fator na amostra?

**Estudante 15:** A variabilidade

**Professor pesquisador:** Isso. Estudamos três tipos de variabilidade, duas delas foi a natural e a amostral, que respondem essa questão. Naturalmente os valores das alturas que compõem a amostra que vocês escolheram são diferentes e isso implica em médias diferentes também. Então é a variabilidade.

**Professor pesquisador:** Prosseguindo para a próxima alternativa, qual das amostras tem a média mais próxima da média populacional?

**A maioria dos estudantes:** A amostra 4

**Professor pesquisador:** Por que vocês acham que foi amostra 4 e o que podemos concluir disso?

**Estudante 12:** Porque é amostra maior. Tem mais gente

**Estudante 8:** Então quando a amostra é maior a variabilidade diminui. Se aproxima dos resultados da população

**Professor Pesquisador:** Exatamente. Quando aumentamos o tamanho da amostra, a variabilidade nesse caso aí é a de medida, vai diminuir. Essa variabilidade também é chamada como?

**Alguns estudantes:** Margem de erro.

**Professor pesquisador:** Isso. Maior tamanho da amostra significa que as estatísticas da amostra estarão mais próximas da população, diminuindo a margem de erro. Então com base nisso, qual a resposta da alternativa c?

**Alguns estudantes:** Tem relação!

**Estudante 21:** Quando o tamanho da amostra é maior a variabilidade diminui

**Professor Pesquisador:** Ok! Exato. O tamanho da amostra sendo maior, menor será a margem de erro. Consequentemente, se obtém uma estimativa mais precisa e aumentando a confiança nos resultados. Então, com base nisso, na última alternativa, quais características da população deveriam ser consideradas para a seleção de uma amostra representativa?

**Alguns estudantes:** Sexo, turma, idade...

**Professor Pesquisador:** Correto. Então a amostra tem que garantir a variabilidade natural, que são justamente as características da população que precisam estar na amostra para ela ser representativa.

Fonte: Dados da pesquisa (2023).

O diálogo realizado no segundo momento na turma Lápis e papel revela o processo de aprendizagem dos temas abordados nas questões. De início, destaca-se que os estudantes demonstraram dificuldades em refletir e comunicar os significados da média e o desvio-padrão abordados na questão. Esse fato confirma que muitas vezes, os estudantes, da Educação Básica se atentam e dominam apenas o processo do cálculo dessas medidas, mas não refletem para além disso (Gonçalves, 2014; Oliveira e Rosa, 2020). No entanto, a partir da intervenção do professor pesquisador eles avançaram na compreensão desses significados e ainda tiveram a oportunidade de refletir os contextos em que cada medida de centralidade (média, mediana e moda) é mais indicada para representar um conjunto.

Na terceira questão, a qual abordou o conceito de variabilidade, foi possível aprofundar, de forma prática, as influências desse conceito na representação amostral e consolidar o que já havíamos estudado no momento anterior. De início, destaca-se que os estudantes acharam interessante e instigador o fato das médias amostrais darem valores diferentes quando comparadas entre as duplas. A partir disso, começaram a refletir que tal fato tinha relação com a escolha aleatória dos elementos amostrais. Na discussão, ficou evidente o entendimento sobre a influência da variabilidade natural e amostral para a diferentes médias e a relação entre o tamanho da amostra e a variabilidade de medida (margem de erro). A resposta do estudante 12 de que amostra 4 tem média mais próxima da média populacional porque é “amostra maior e tem mais gente”, como também, a resposta do estudante 8 de que “quando a amostra é maior a variabilidade diminui, se aproxima dos resultados da população”, evidencia a aprendizagem sobre essa relação entre o tamanho da amostra e variabilidade e quanto isso implica na representação amostral. Alguns estudantes afirmaram que já tinham se deparado com situações que abordam a margem de erro em pesquisas amostrais na mídia, a exemplo de pesquisas eleitorais, mas que até antes desse estudo não comprehendiam o seu significado.

Cabe frisar que a discussão entre o professor pesquisador e os estudantes, a partir da simulação realizada, solidificou o entendimento de que a variabilidade de medida (margem de erro) diminui quando o tamanho da amostra é maior. No entanto, a discussão também reforçou que essa relação não é tida de forma diretamente proporcional, ou seja, não são lineares, culminando em um estágio que se aumentar ou não o tamanho da amostra, será insignificante para alterar a margem de erro. Em acréscimo, algumas duplas constataram que a média amostral que mais se aproximou da média populacional não teria sido a quarta, ou seja, a maior. Sobre

esse fato, debatemos que como a escolha dos elementos foi aleatória, algumas amostras podem ter tido pouca variabilidade natural, sendo composta, por exemplo, só por alturas de pessoas do sexo masculino ou feminino, ou com poucos de cada um deles o que, naturalmente acabou influenciando na média amostral. Em acréscimo, ainda refletimos sobre a proporção de uma característica da população e a proporção correspondente nas amostras retiradas, bem como o efeito do tamanho da amostra em tal variabilidade. Para tanto, verificamos que na turma há 34 estudantes, sendo 19 meninas. Logo, a proporção de meninas em relação ao total de alunos da turma é igual a aproximadamente 0,55. A partir disso, o professor pesquisador solicitou que os estudantes calculassem qual a amostra, entre as quatro construídas na questão, continha esse tipo de proporção mais próxima da populacional. A maioria dos estudantes responderam que foi a amostra 4, a de maior tamanho. Mais uma vez, refletimos sobre o efeito do tamanho da amostra para a variabilidade.

Encerrando a atividade, o debate relativo a última alternativa revela a apropriação do conhecimento sobre a variabilidade natural que precisa constar na amostra para que ela se torne um reflexo da população. As repostas da maioria dos estudantes que indica o sexo, a idade e a turma como características importantes a serem consideradas para a seleção amostral, sinaliza essa aprendizagem, tendo em vista que a maioria deles no pré-teste, em uma questão semelhante apontou apenas uma característica, o que não garante a representatividade.

Diante dessas conclusões, destacamos que o recurso de Lápis e Papel permitiu o cálculo das médias amostrais e a reflexão de seus significados. Embora esse processo tenha demandado mais tempo, quando comparado com o software Geogebra, tal atividade com o Lápis e Papel, contribui para o desenvolvimento do letramento Estatístico (Gal,2002). Os estudantes estavam motivados a irem além da aplicação do algoritmo para obter o valor da média, ou seja, estavam instigados a pensar estatisticamente, refletindo no porquê e para que o cálculo foi realizado. Ao longo da atividade ficou explícito que os estudantes mobilizaram diferentes elementos do Letramento Estatístico (Gal, 2002), desde o conhecimento de contexto, a indicarem exemplos do cotidiano, como as pesquisas eleitorais que apresentam margem de erro, ao conhecimento matemático e estatístico relacionados com as medidas de centralidade e dispersão, bem como, o tamanho da amostra e a variabilidade amostral. Paralelamente a isso, mobilizaram habilidades de letramento para se comunicar sobre a estatística e se portaram de maneira crítica para refletir

sobre a análise dos dados e aos questionamentos da atividade. A resposta (Figura 127) exemplifica a respostas de uma dupla para a 3<sup>a</sup> questão da segunda atividade exemplifica essa conclusão.

Figura 127 – Reposta de uma dupla para a 3<sup>º</sup> questão da atividade 2

Tamanho da Amostra	Média da altura (cm)
Amostra 1: 4 estudantes	$163,25 \rightarrow \frac{653}{4}$
Amostra 2: 8 estudantes	$172,5 \rightarrow \frac{1380}{8}$
Amostra 3: 12 estudantes	$169,25 \rightarrow \frac{2031}{12}$
Amostra 4: 16 estudantes	$168,25 \rightarrow \frac{2652}{16}$

- a) Compare as medidas das suas médias das amostras com as de outra dupla de colegas. Os resultados são iguais ou diferentes?

*Diferentes, pois as amostras foram escolhidas de forma aleatória.*

- b) De acordo com o tamanho da amostra, qual delas tem a média mais próxima da média de populacional (todos os alunos da turma)?

*Amostra 4*

- c) Na sua opinião, para uma amostra ser representativa há alguma relação entre o tamanho da amostra e a margem de erro?

*Sim. Quanto maior o tamanho da amostra, menor a margem de erro.*

- d) Se essa pesquisa fosse realizada com toda a escola, quais características da população deveriam ser consideradas para a seleção de uma amostra representativa?

*Idade, turma, sexo*

- a) Diferentes, pois as amostras foram escolhidas de forma aleatória
- b) Amostra 4
- c) Sim. Quanto maior o tamanho da amostra, menor margem de erro
- d) Idade, turma, sexo

Fonte: Dados da pesquisa (2023).

### 9.2.2.2 Turma Geogebra

Na turma Geogebra, procedemos com a mesma atividade, contudo, os comandos das questões foram adaptados para o recurso empregado. Destacamos que os estudantes também foram organizados em duplas e cada uma dispunha de um notebook/tablet para o trabalho com o software Geogebra. Além disso, o professor pesquisador orientou e mediou as ações a serem realizadas no software para a resolução das questões (Figuras 128 e 129), as quais tiveram os mesmos objetivos de aprendizagem da outra turma.

Figura 128 – Primeira e Segunda questão da atividade 2 – Turma Geogebra

- 1º) Caros estudantes, para darmos início a nossa atividade, realize o seguinte procedimento:
- Faça uma pesquisa com seus colegas da classe e colete a altura (em cm) de cada um deles.
  - Insira os dados na planilha do Geogebra com o número de identificação dos alunos e suas respectivas alturas.
- 2º) Utilizando a ferramenta análise univariada, determine a média, e o desvio-padrão das medidas das alturas dos estudantes da sua turma. O que cada uma dessas medidas significa?

Fonte: Dados da pesquisa (2023).

Figura 129 – Terceira questão da atividade 2 – Turma Geogebra

- 3º) A partir da lista com os dados das alturas de todos os estudantes, selecione, aleatoriamente, no Geogebra, amostras de estudantes de acordo com cada tamanho destacados a seguir. Em seguida, utilizando a ferramenta análise Multivariada no Geogebra, calcule a média das alturas de cada amostra.

Tamanho da Amostra	Média da altura (cm)
Amostra 1: 4 estudantes	
Amostra 2: 8 estudantes	
Amostra 3: 12 estudantes	
Amostra 4: 16 estudantes	

- Compare as medidas das suas médias das amostras com as de outra dupla de colegas. Os resultados são iguais ou diferentes?
- De acordo com o tamanho da amostra, qual delas tem a média mais próxima da média de populacional (todos os alunos da turma)?
- Na sua opinião, para uma amostra representativa há alguma relação entre o tamanho da amostra a margem de erro?
- Se essa pesquisa fosse realizada com toda a escola, quais características da população deveriam ser consideradas para a seleção de uma amostra representativa?

Fonte: Dados da pesquisa (2023).

Nessa direção, para a resolução da primeira e segunda questão, os estudantes coletaram os dados da altura de seus pares e inseriram na planilha de dados do Geogebra. Em seguida, procederam com cálculo da média e do desvio-padrão utilizando a ferramenta “análise univariada” do software, como também, responderam ao questionamento sobre o significado dessas medidas. Dando continuidade, na

terceira questão os estudantes selecionaram aleatoriamente quatro amostras de diferentes tamanhos a partir do banco de dados e realizaram os cálculos de suas respectivas médias utilizando a ferramenta de análise multivariada, a qual possibilitou a visualização simultânea das estatísticas das quatro amostras e da população (todos os estudantes da turma). Destacamos (Quadro 28) a seguir um recorte do diálogo com essa turma relacionado à segunda atividade.

Quadro 28 - Recorte do diálogo da 2º e 3º questão da Atividade 2 - turma Geogebra

<b>Professor Pesquisador:</b> Vamos comentar sobre a 2º questão. Através da análise univariada que vocês fizeram aí no Geogebra, qual o valor da média das alturas aproximadamente?
<b>Estudantes:</b> 169 cm
<b>Professor Pesquisador:</b> Ok. O que a média significa nessa situação?
<b>Estudante 23:</b> O valor da soma dividido pela quantidade
<b>Estudante 23:</b> O valor que resume a turma
<b>Estudante 12:</b> Esse o valor, no caso, representa essa turma
<b>Professor Pesquisador:</b> Devemos compreender que ela é uma medida representativa para esse conjunto de dados. Isso quer dizer que a média fornece uma indicação geral de onde os valores estão concentrados, sendo bastante indicada quando temos distribuição de dados simétricas, ou seja, quando os valores estão distribuídos de forma homogênea em torno do centro, como acontece na Curva Normal. Porém, nem sempre a média é a mais indicada para representar um determinado conjunto de dados. Tudo depende do contexto de onde os dados foram coletados. A média é sensível, a dados extremos, chamados também de (outliers) que estudamos anteriormente. Quando isso acontece, outras medidas são mais indicadas, como a mediana. Já a moda, como medida de centralidade que é mais indicada para representar um conjunto quando se deseja identificar valores mais frequentes, geralmente em dados discretos que são contáveis e assumem valores inteiros.
<b>Professor Pesquisador:</b> E qual foi o valor do desvio-padrão aproximadamente?
<b>Estudantes:</b> 8,3 cm
<b>Professor pesquisador:</b> E o que ele significa nessa situação?
<b>Estudante 21:</b> O quanto é disperso os valores em relação à média
<b>Estudante 25:</b> Indica o grau de afastamento para a média desse conjunto
<b>Professor Pesquisador:</b> Exato. O desvio-padrão indica o quanto distante os valores individuais estão da média do conjunto de dados. Então, um desvio-padrão maior indica uma maior dispersão dos valores em relação à média, enquanto um desvio-padrão menor, mais próximo de zero, indica uma menor dispersão.
<b>Professor pesquisador:</b> Na terceira questão, vocês fizeram a análise multivariada no Geogebra para analisar quatro amostras de tamanho

diferentes. Ao comparar os resultados das médias com as outras duplas, os resultados foram iguais ou diferentes?

**Estudantes:** Diferentes!

**Estudante 8:** Bem diferentes

**Professor Pesquisador:** Por que deram diferentes?

**Estudante 21:** As amostras não são iguais porque os dados mudam

**Estudante 30:** As amostras acabam sendo diferentes

**Professor pesquisador:** Ok. Então as amostras são diferentes por conta das escolhas aleatórias dos elementos. E essa diferença se dá por conta de que fator na amostra?

**Estudante 10:** A variabilidade

**Estudante 16:** A variabilidade presente nas amostras

**Professor pesquisador:** Perfeito. Na última aula estudamos três tipos de variabilidade: a natural, a amostral e a de medida. A natural nos diz que naturalmente os dados são diferentes em cada amostra e, consequente, geram estatísticas amostrais diferentes que é a variabilidade amostral

**Professor pesquisador:** Prosseguindo para a próxima alternativa, ao analisar a ferramenta multivariada qual das amostras vocês observaram que tem a média mais próxima da média populacional?

**A maioria dos estudantes:** A amostra 4

**Professor pesquisador:** Ok. Por que vocês acham que foi amostra 4 e o que podemos concluir disso?

**Estudante 18:** Acho que é porque é a maior

**Estudante 29:** A amostra 4 é a que tem mais pessoas

**Professor Pesquisador:** Então o que pode concluir sobre o fato de a média da maior amostra ser a mais próxima da média populacional

**Estudante 21:** Nesse caso, quando o tamanho é maior está mais próximo, a variabilidade diminui

**Alguns estudantes:** Maior tamanho gera menor variabilidade

**Professor Pesquisador:** Correto. Quando aumentamos o tamanho da amostra, a variabilidade diminui. Essa variabilidade que relaciona o distanciamento do resultado amostral para o populacional é chamada como?

**Alguns estudantes:** Margem de erro.

**Estudante 13:** A de medida. Também chamada de margem de erro

**Professor pesquisador:** Então prosseguindo, para uma amostra ser representativa, há alguma relação entre o tamanho da amostra e a variabilidade? Qual a resposta da alternativa c?

**Alguns estudantes:** Tem relação!

**Estudante 21:** Quando o tamanho aumenta, diminui a variabilidade e vice-versa.

**Professor Pesquisador:** Ok! A última alternativa perguntou quais características da população deveriam ser consideradas para a seleção de uma amostra representativa?

**Alguns estudantes:** A turma, a idade, o sexo...

**Estudante 16:** A turma, a idade, o sexo, a cidade onde mora, porque aqui na escola tem alunos de várias cidades.

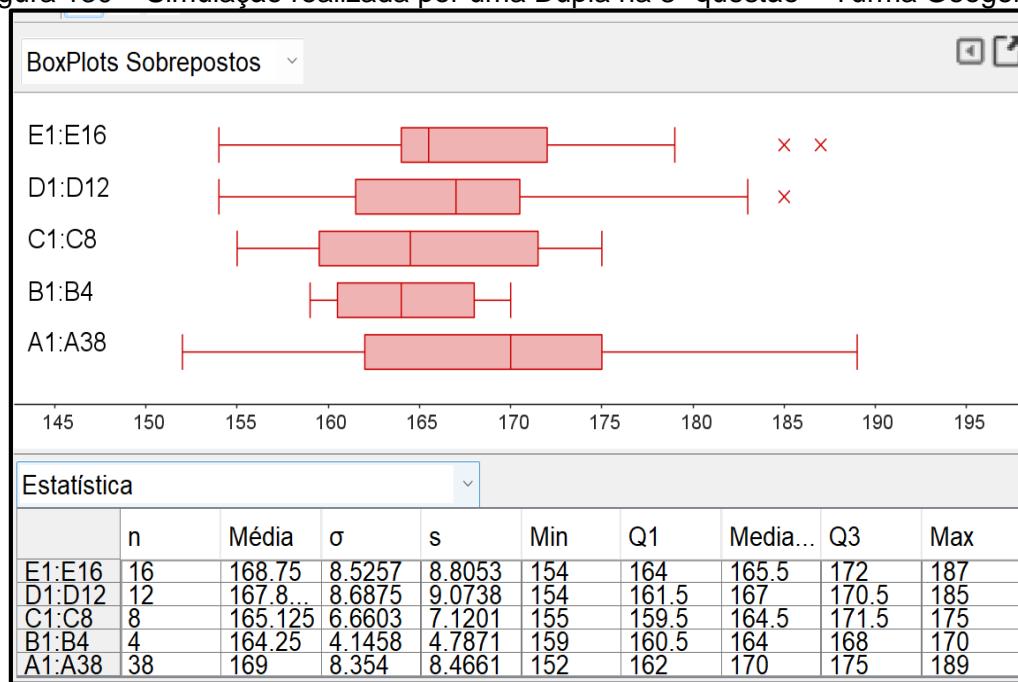
**Professor Pesquisador:** Correto. Então a amostra tem que garantir a variabilidade natural, que tem relação com as características da população e precisam ser consideradas na amostra para ela ser representativa.

Fonte: Dados da pesquisa (2023).

O diálogo desenvolvido na turma Geogebra também ilustra a consolidação, por parte dos estudantes, sobre os conceitos estudados. Foi perceptível que, de modo semelhante a outra turma, os estudantes inicialmente também apresentaram dificuldades para refletir e emitir opiniões sobre o significado da média na segunda questão. A resposta do estudante 23 que o significado da média é “o valor da soma dividido pela quantidade” evidencia que na Educação Básica esse conceito é fortemente trabalhado com ênfase na aplicação do seu algoritmo.

No que diz respeito à terceira questão, a discussão notabiliza a construção do conhecimento estatístico sobre a influência da variabilidade natural para a determinação de amostras diferentes e consequentemente, a variabilidade amostral para a ocorrência de médias distintas em cada amostra. As respostas dos estudantes aos questionamentos do professor pesquisador demonstram o entendimento sobre a relação entre o tamanho da amostra e a variabilidade de medida (margem de erro) e a influência dessa relação para a representatividade amostral. Nessa turma, os estudantes, em linhas gerais, afirmaram que já tinham se vivenciado situações que abordaram a margem de erro em pesquisas amostrais, mas que até então não compreendiam o conhecimento por trás desse conceito. No entanto, através da simulação prática realizada com o software, conseguiram avançar na aprendizagem desse tema. Nesse contexto, Begué et al (2017) defendem que simulações práticas envolvendo a variabilidade são pertinentes para o seu entendimento, por parte dos estudantes. A figura 130 apresenta a simulação realizada por uma dupla com a ferramenta análise multivariada do Geogebra que possibilita a análise simultânea dos dados das 4 amostras e da população. É possível notar que a maior amostra é a que mais se aproxima da média populacional, explicitando a referida relação.

Figura 130 – Simulação realizada por uma Dupla na 3º questão – Turma Geogebra



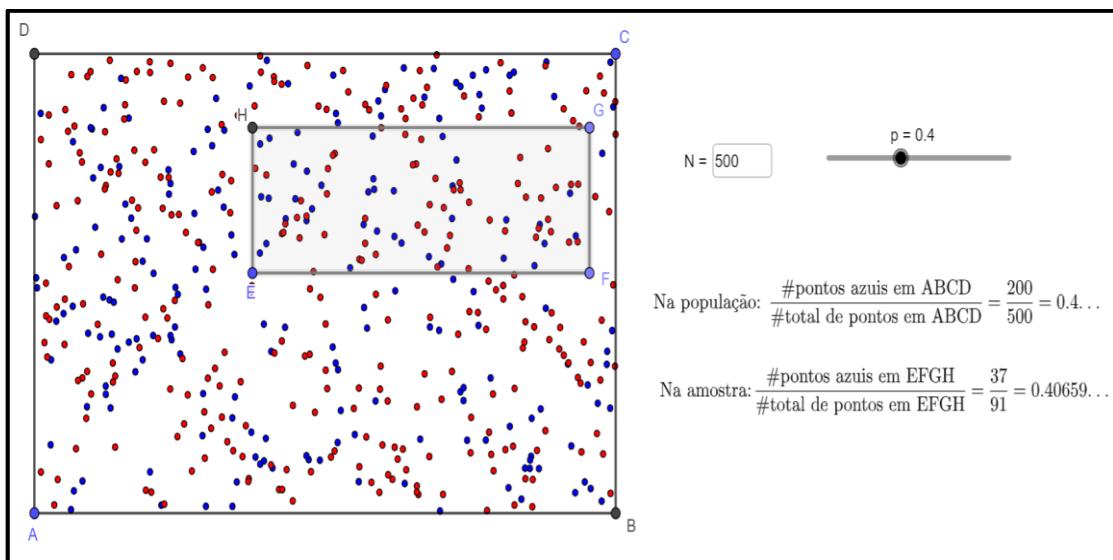
Fonte: Dados da pesquisa (2023).

No decorrer da discussão, constatamos que a maioria das duplas verificou em suas simulações que a amostra cuja a média mais se aproximou da média populacional foi a amostra 4, a de maior tamanho. Logo, reforçamos a compreensão de que quanto maior o tamanho da amostra menor a variabilidade e vice-versa. No entanto, de modo análogo ao observado na outra turma, algumas duplas afirmaram que a média mais próxima à populacional não teria sido a da maior amostra. Assim, também refletimos que a escolha aleatória dos elementos amostrais pode em alguns casos gerar uma baixa variabilidade natural, e, por conseguinte, afetar as estatísticas. Além disso, de igual modo a explanação na turma Lápis e papel, chamamos a atenção de que a relação entre o tamanho da amostra não é dada de forma linear.

Em continuidade, a discussão relativa da última alternativa revela o aprendizado sobre a variabilidade natural da população que precisa ser refletida na amostra. Os estudantes mencionaram algumas características que notabilizam o entendimento sobre a quantidade de fatores que necessitavam serem abarcados na amostra para verdadeiramente caracterizar a população abordada na atividade. Isso demonstra um avanço em relação as concepções apresentadas no pré-teste, tendo em vista que, a maioria dos estudantes de ambas as turmas, responderam apenas uma característica populacional que deveria estar na amostra em uma questão semelhante, o que não garantiria a representatividade.

Após a socialização da terceira questão, ainda refletimos sobre a proporção de uma característica da população e a proporção correspondente nas amostras retiradas, bem como o efeito do tamanho da amostra em tal variabilidade. Para tanto, utilizamos uma ferramenta disponível no Geogebra, desenvolvida por Bortolossi (2020). Nela, é possível observar que a proporção de bolas azuis vista na amostra (retângulo menor) se aproxima da proporção desses elementos na população (retângulo maior) à medida que o tamanho da amostra, equivalente a área do retângulo aumenta. A partir disso, reforçamos o entendimento sobre a relação entre o tamanho da amostra e a variabilidade. Exemplificamos a seguir (Figura 131) a simulação realizada por uma dupla.

Figura 131 – Simulação realizada por uma Dupla – Turma Geogebra



Fonte: Dados da pesquisa (2023).

A partir dessas conclusões nessa etapa da intervenção de ensino, destacamos que o software Geogebra se revelou como um recurso significativo para o trabalho com a Estatística em sala de aula. Seus recursos possibilitaram o tratamento de dados de forma robusta e dinâmica, a exemplo da análise univariada que permitiu o cálculo das medidas estatísticas de forma ágil e a análise multivariada que propiciou a análise dos dados de forma simultânea das amostras e da população, incluindo a visualização dinâmica do comportamento dos dados com a exibição de diferentes medidas estatísticas e representações gráficas, como o Boxplot. Deste modo, o Geogebra contribuiu para um olhar analítico com riqueza de detalhes sobre a relação entre o tamanho da amostra e a variabilidade, nos tipos natural, amostral e de medida. Além do mais, os estudantes se mostraram motivados ao trabalhar com o software e

mencionaram que nunca tiveram o contato com o mesmo para estudos sobre a Estatística.

Por fim, os resultados observados indicam que os estudantes da turma Geogebra explicitaram diferentes elementos do Letramento Estatístico (2002), tanto os de conhecimento como os de disposição que os permitiram pensar estatisticamente para resolverem as questões e consolidar a aprendizagem. Diante disso, destacamos que a realização de atividades que vão além do domínio de fórmulas e números, contemplando a compreensão do contexto, a criticidade e a aplicação prática dos conceitos estatísticos em situações do mundo real, contribuem para o desenvolvimento do Letramento Estatístico (Gal, 2002) dos estudantes.

Em paralelo, enfatizamos que a vivência de etapas do ciclo investigativo, em ambas as turmas, possibilitou que os estudantes se portassem como produtores de dados e aprendizagem de conceitos que permearam as etapas vivenciadas. Logo, a realização de ciclos de pesquisas estatísticas na Educação Básica é um importante método para possibilitar a formação de sujeitos críticos frente a questionamentos relacionados a seus contextos, notabilizando a Estatística como uma ferramenta para compreensão e transformação da realidade (Guimarães e Gitirana (2013); Gonçalves e Lima, 2017).

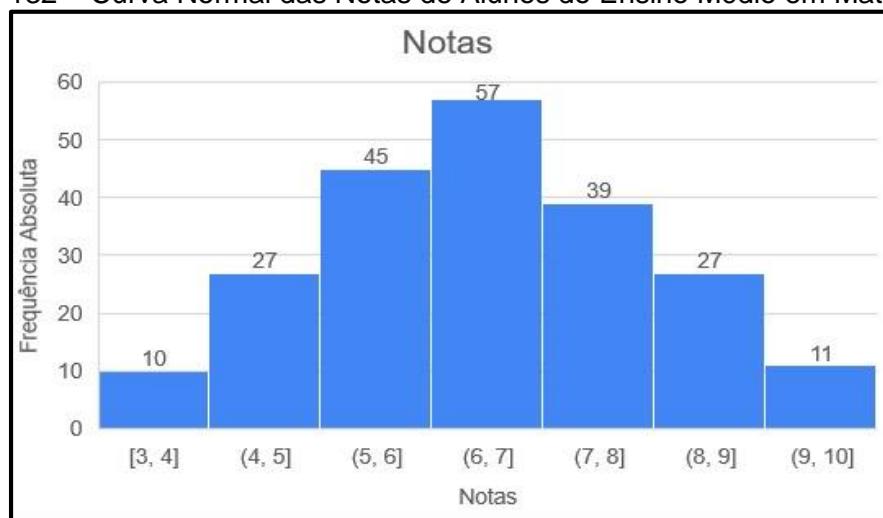
#### 9.2.3 3º Momento: Caracterização da Curva Normal e o cálculo de probabilidades

O terceiro momento da intervenção de ensino foi direcionado para aprofundamento do conhecimento sobre a representação gráfica do modelo da Curva Normal já vista no primeiro momento, como também, para o estudo das medidas estatísticas e o cálculo de probabilidades no referido modelo, através dos dados das alturas coletados no momento anterior.

Antes dos estudantes procederem com a resolução da terceira atividade, o professor pesquisador apresentou, nas duas turmas, uma sistematização teórica sobre a probabilidade relacionada a área sob a Curva Normal. Mais especificamente, essa sistematização buscou elucidar para os estudantes uma propriedade fundamental para o cálculo de probabilidades na Curva Normal, isto é, que a Curva Normal tem área igual 1. Paralelamente, nesse momento também buscamos promover o conhecimento sobre o gráfico de histogramas, tendo em vista que Curva Normal modela distribuições contínuas e sua representação gráfica fica mais caracterizada nesse tipo de gráfico. Esse procedimento foi adotado com base nos

resultados do pré-teste, o qual mostrou, em ambas as turmas, uma lacuna no que diz respeito ao conceito de probabilidade presente no modelo da Curva Normal. Para o desenvolvimento dessa sistematização utilizamos um exemplo real de uma Curva Normal em histograma que apresenta (Figura 132) a distribuição das notas em Matemática de 216 alunos do Ensino Médio durante uma unidade letiva.

Figura 132 – Curva Normal das Notas de Alunos do Ensino Médio em Matemática



Fonte: Dados da pesquisa (2023).

A partir desse exemplo, refletimos com os estudantes que esse tipo de gráfico, já observado no primeiro momento da intervenção, é um histograma, sendo classificado como uma representação gráfica de distribuição de frequências em barras, construída com intervalos de classes em um conjunto de dados contínuos, no qual não há espaços entre as barras. O eixo horizontal do histograma representa os intervalos (classes) dos dados, enquanto o eixo vertical representa a frequência com que os dados se enquadram em cada classe, e essas podem ser absolutas ou relativas, (Ferreira, 2015).

Na sequência, para compreender que a área da Curva Normal é igual a 1, os estudantes foram convidados a observar a área de cada retângulo, nos quais a base representa a amplitude do intervalo de classe. No nosso exemplo, a base de todos os retângulos tem medida igual 1. Já a altura dos retângulos corresponde as respectivas frequências relativas de cada intervalo. Por exemplo, o intervalo [3,4] possui a frequência relativa igual a  $\frac{10}{216} = 0,0462$ , aproximadamente. A seguir são apresentadas (Quadro 29) a frequência de cada intervalo da Curva Normal abordada no exemplo.

Quadro 29 – Frequências Relativas dos intervalos da Curva Normal

Intervalo de Classe	Frequência Relativa
[3,4]	0,0462
(4,5]	0,1250
(5,6]	0,2083
(6,7]	0,2638
(7,8]	0,1805
(8,9]	0,1250
(9,10]	0,0509
	<b>Soma = 0,9997</b>

Fonte: Dados da Pesquisa (2023).

Tendo em vista que área de um retângulo é dada pelo produto de sua base pela sua altura, no nosso exemplo, como a base de cada retângulo é igual a 1, a área de cada um deles correspondeu ao valor da respectiva frequência relativa. Dando prosseguimento, ao somarmos as áreas dos 7 retângulos que compõem a Curva Normal, verificamos que esse valor é igual a, 0,9997 ou seja, muito próximo do valor teórico para a área da curva que é igual a 1, haja vista que a Curva Normal exemplificada é uma proximidade do modelo normal. Deste modo, constatamos que de fato, a área sob a Curva Normal é igual a 1. Após essa abordagem teórica, os estudantes, mais uma vez em duplas, foram convidados a responder a terceira atividade.

#### 9.2.3.1 Turma Lápis e Papel

A primeira questão da terceira atividade (Figura 133) teve o objetivo de refletir sobre a relação entre a média, mediana e moda, as quais, no modelo normal, estão no centro da distribuição. Inicialmente, cada dupla procedeu com o cálculo da moda e mediana e junto com a média já fornecida na atividade anterior, elaboraram o esboço do histograma das alturas da população (turma). Na sequência, juntamente como o professor pesquisador, refletiram sobre a localização dessas medidas no gráfico.

Figura 133 – Primeira questão da terceira atividade – Turma Lápis e Papel

- 1º) A partir dos dados coletados na atividade anterior, realize os seguintes procedimentos.
- Determine a mediana e a moda das medidas das alturas coletadas.
  - Esboce um gráfico de histogramas para as medidas das alturas (em cm) coletadas. Localize, no gráfico, a média, moda e mediana das medidas das alturas. Você identifica alguma relação entre elas?

Fonte: Dados da pesquisa (2023).

Já por meio da segunda questão (Figura 134), demos início ao cálculo de probabilidades. Para tanto, considerando que o recurso utilizado pela turma foi o Lápis e Papel, o professor pesquisador procedeu com uma abordagem sobre o método da distribuição padronizada, abordada no capítulo 1 deste estudo. Esse método, possibilita, com a utilização do Lápis e Papel, o cálculo da área de intervalos específicos da Curva Normal, consequente, esse resultado equivale a probabilidade de a variável em questão assumir um valor no respectivo intervalo. Ao se apropriarem dos procedimentos do referido método, os estudantes resolveram a segunda questão.

Figura 134 – Segunda questão da terceira atividade – Turma Lápis e Papel

- 2º) Com base no esboço do gráfico realizado na atividade anterior, responda:
- Se escolhermos uma pessoa dentre todos os estudantes dessa turma, ao acaso, qual a probabilidade aproximada de que ela possua a altura maior que a média do grupo?
  - Se o professor escolher um aluno dessa turma, ao acaso, qual a probabilidade aproximada de que o aluno escolhido tenha a altura entre 168 cm e 178 cm?
  - Qual a probabilidade aproximada de um aluno dessa turma ter a altura entre 158 cm e 188 cm?
  - Com base nas propriedades teóricas do Modelo Normal, há alguma relação entre as respostas dos itens “b” e “c”?

Fonte: Dados da pesquisa (2023).

Após a resolução das questões, o professor pesquisador mediou o momento de socialização (Quadro 30) com o intuito de emergir reflexões sobre o tema juntamente com os estudantes.

Quadro 30 – Recorte do diálogo da 2º e 3º questão da Atividade 3 - turma Lápis e Papel

**Professor Pesquisador:** *Iniciando a primeira questão, qual o valor da moda e da mediana?*

**Estudante 9:** *A mediana é 168 cm. A moda é 159 cm*

**Professor Pesquisador:** *Todos encontraram a mesma resposta?*

**Estudantes:** Sim!

**Professor Pesquisador:** Ok. Na alternativa seguinte, vocês esboçaram o gráfico de Histograma e localizaram essas três medidas nele. Primeiramente, o gráfico tem alguma próxima com o modelo da Curva Normal

**Alguns estudantes:** Sim!

**Estudante 3:** O meu ficou próximo sim. Parecido com os exemplos que o senhor mostrou.

**Professor Pesquisador:** Certo. Mas há alguma relação entre essas três medidas?

**Estudante 21:** A média e a mediana são iguais, mas a moda deu diferente.

**Estudante 12:** Isso. A moda foi única que não deu o mesmo valor.

**Professor Pesquisador:** Ok. Porque isso aconteceu?

-Silêncio

**Professor Pesquisador:** Nem todas distribuições de variáveis contínuas são totalmente simétricas, obedecendo ao modelo normal. Isso é comum de acontecer. Por algum motivo, como um outlier, ou até mesmo o tamanho da população que não é suficiente grande para caracterizar uma curva, a distribuição pode ser assimétrica como a nossa aqui. No entanto, a partir desses gráficos que vocês construíram com os intervalos de classe no histograma, podemos observar que há uma proximidade do modelo da Curva Normal.

**Estudante 13:** Nesse caso se tivéssemos uma população bem maior poderia estar mais no formato da curva, né?

**Professor Pesquisador:** Exato. À medida que o tamanho da população aumenta, a representação gráfica da distribuição para variáveis contínuas tenderá a se aproximar de uma Curva Normal mais "perfeita" ou seja, simétrica.

**Professor Pesquisador:** Na primeira alternativa da segunda questão, qual a probabilidade aproximada de selecionar uma pessoa do grupo e ela ter a altura maior ou igual a média?

**Estudantes:** 50%

**Professor pesquisador:** Por que?

**Estudante 15:** Porque equivale à metade da área da Curva, 50%.

**Professor pesquisador:** Correto. Como estudamos, metade da Curva equivale a 50%. Como nosso gráfico é uma aproximação ao modelo normal, podemos dizer que é esse valor aproximadamente.

**Professor pesquisador:** Na alternativa b, vocês utilizaram o método da distribuição padronizada. Qual o valor da probabilidade?

**Estudantes:** 39,44%

**Professor pesquisador:** Por que? Como vocês realizaram o cálculo?

**Estudante 25:** Esse valor é igual a 178 menos 168 dividido por 8. Isso dá 1,25, que na tabela é igual a 39,44%

**Professor Pesquisador:** Perfeito. Os demais também fizeram assim?

**Estudantes:** Sim!

**Professor pesquisador:** E na letra c, qual a probabilidade?

**Estudantes:** 39,44% também

**Professor Pesquisador:** Por que deram iguais?

**Alguns estudantes:** O intervalo é o mesmo

**Estudante 21:** O intervalo entre 168 e 178 e entre 168 e 158 são idênticos

**Professor pesquisador:** Exatamente. Isso já responde a alternativa d. Considerando nosso gráfico próximo à normalidade, estudamos que os intervalos com mesma amplitude acima e abaixo da média tem a mesma área, e são simétricos em relação à média. Então tem a mesma probabilidade.

Fonte: Dados da pesquisa (2023).

Esse debate evidencia a consolidação da aprendizagem, por parte dos estudantes, sobre os fundamentos teóricos relacionados ao gráfico da Curva Normal e ao cálculo de probabilidades. Inicialmente os estudantes constataram que ao média e da mediana coincidiram, porém, o valor da moda diferiu desses duas. Além disso, o histograma construído pela maioria das duplas constatou que a representação gráfica apresentava uma proximidade com o modelo da Curva Normal. Nesse momento, o professor pesquisador entrou em cena para discutir com estudantes que a distribuição das alturas daquela turma era próxima de uma Curva Normal, porém possuía uma assimetria.

Logo, foi ressaltado que como o tamanho da população é relativamente pequeno, ou seja, 34 dados, a Curva Normal não ficou simétrica, algo comum para essa quantidade de dados. Assim, reforçamos as características desse modelo e refletimos que em uma distribuição normal, à medida que o tamanho da população aumenta, a representação gráfica da distribuição tenderá a se aproximar de uma Curva Normal mais "perfeita" ou simétrica. A resposta (Figura 135) de uma dupla para a primeira questão da terceira atividade notabiliza essa conclusão.

Figura 135 – Reposta de uma dupla para a 1º questão da atividade 3

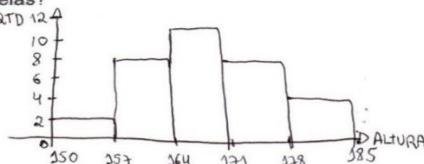
1º) A partir dos dados coletados na atividade anterior, realize os seguintes procedimentos.

a) Determine a mediana e a moda das medidas das alturas coletadas.

$$\text{Mediana} \rightarrow \text{posição: } \frac{35}{2} = 17,5^{\circ} \rightarrow M_e = 168 \text{ cm}$$

$$\text{Moda} \rightarrow 159 \text{ cm}$$

b) Esboce em um gráfico de barras para as medidas das alturas (em cm) coletadas. Localize, no gráfico, a média, moda e mediana das medidas das alturas. Você identifica alguma relação entre elas?



$$\text{Mediana} \rightarrow \text{posição } 35/2 = 17,5^{\circ} \rightarrow M_e = 168 \text{ cm}$$

$$\text{Moda} \rightarrow 159 \text{ cm}$$

Fonte: Dados da pesquisa (2023).

Em continuidade, o diálogo realizado também ilustra a construção do conhecimento sobre o cálculo de probabilidades associado a intervalos da Curva Normal. A resposta dos estudantes para a alternativa “a” da segunda questão, indicando que a probabilidade solicitada era de 50% porque equivalia a metade da Curva torna notável essa construção, tendo em vista que a maioria respondeu de modo inadequado a uma questão semelhante no pré-teste. Além disso, nas demais alternativas dessa questão, mostraram compreender propriedades mais específicas respondendo e justificando adequadamente o valor da probabilidade solicitada. A resposta (Figura 136) de uma dupla para a segunda questão da terceira atividade evidencia essa conclusão.

Figura 136 – Reposta de uma dupla para a 2º questão da atividade 3

- a) Se escolhermos uma pessoa dentre todos os estudantes dessa turma, ao acaso, qual a probabilidade de que ela possua a altura maior ou igual que a média do grupo?  
50%, metade da Curva
- b) Se o professor escolher um aluno dessa turma, ao acaso, qual a probabilidade que o aluno escolhido tenha a altura entre 168 cm e 178 cm?  
 $\frac{178-168}{8} = 1,25 \rightarrow 39,94\%$
- c) Qual a probabilidade de um aluno dessa turma ter a altura entre 158 cm e 168 cm?  
 $\frac{168-158}{8} = 1,25 \rightarrow 39,94\%$
- d) Com base nas propriedades teóricas do Modelo Normal, há alguma relação entre as respostas dos itens “b” e “c”?  
mesma probabilidade, pois os intervalos são iguais
- a) 50% metade da Curva  
b)  $178-168/2 = 1,25 \rightarrow 39,44\%$   
 $168-158/2 = 1,25 \rightarrow 39,44\%$   
d) mesma probabilidade, pois os intervalos são iguais

Fonte: Dados da pesquisa (2023).

Diante disso, destacamos que o recurso Lápis e papel possibilitou alcançarmos o objetivo de aprendizagem da atividade. No esboço do histograma das alturas, esse recurso contribuiu para o desenvolvimento das habilidades de representação gráfica dos estudantes, contemplando as características dos eixos, barras e outros elementos gráficos. Na aplicação do cálculo de probabilidades através da distribuição padronizada, o Lápis e Papel possibilitou desenvolver um entendimento sobre o conceito de probabilidade associado à média, desvio padrão e intervalos da Curva Normal, relacionando os cálculos com a representação gráfica. Nesse viés, Herrera, Monroy e Hernandez (2018) apontam que esse tipo de recurso pode ser proveitoso

para o processo de ensino de aprendizagem da Curva Normal, desde que envolvam situações práticas para o entendimento da representação gráfica e propriedades desse modelo.

Os resultados dessa atividade indicam que os estudantes, ao resolverem tarefas relacionadas ao conceito de probabilidade com Lápis e Papel, as quais vão além da simples aplicação de cálculos, como frequentemente ocorre nas abordagens centradas na razão clássica (Lima e Borba, 2019), podem aprofundar reflexões sobre esse conceito, especialmente quando ancorada em ações práticas com dados reais e contextos próximos. Esse tipo de abordagem pode contribuir positivamente para o desenvolvimento do Letramento Estatístico e Probabilístico (Gal, 2002,2005). Observamos que o estudo sobre modelo da Curva Normal desempenhou um papel significativo na ampliação das concepções dos estudantes em relação à inferência estatística, proporcionando uma compreensão mais aprofundada das características desse modelo, evidente em diversos fenômenos do nosso cotidiano.

#### **9.2.3.1 Turma Geogebra**

Na turma Geogebra, a terceira atividade (Figura 137) teve o mesmo objetivo de aprendizagem da outra turma, porém as questões foram adaptadas para a utilização do software Geogebra. Na primeira questão foi solicitado que os estudantes determinassem a medida da mediana e da moda das alturas (cm) dos estudantes dessa turma utilizando ferramentas do referido software. Posteriormente, os estudantes procederam com a construção do histograma das alturas no software, e a partir desta, realizaram a análise da representação gráfica e da localização das medidas de centralidade anteriormente calculadas.

**Figura 137 – Primeira questão da terceira atividade – Turma Geogebra**

- 1º) A partir dos dados coletados na atividade anterior, realize os seguintes procedimentos:
- Com a ferramenta análise univariada e o comando “moda”, determine, respectivamente, a mediana e a moda das medidas das alturas coletadas.
  - Utilizando a ferramenta de análise univariada no Geogebra, faça a construção de um histograma das medidas das alturas (em cm) coletadas. Localize, no gráfico, a média, moda e mediana das medidas das alturas. Você identifica alguma relação entre elas?

Fonte: Dados da pesquisa (2023).

Na sequência, a segunda questão (Figura 138) foi direcionada para o cálculo de probabilidades na Curva Normal. Para tanto, os estudantes recorreram a ferramenta “Calculadora de Probabilidades” do Geogebra, a qual possibilita, através dos parâmetros média e desvio-padrão dos dados, anteriormente já calculados pelos estudantes, determinar a área de intervalos da Curva Normal, e, consequentemente, a probabilidade de a variável em questão assumir valor em um determinado intervalo de interesse.

Figura 138 – Segunda questão da terceira atividade – Turma Geogebra

- 2º Com base na construção do gráfico realizado na atividade anterior, responda:
- Se escolhermos uma pessoa dentre todos os estudantes dessa turma, ao acaso, qual a probabilidade de que ela possua a altura maior que a média do grupo?
  - Se o professor escolher um aluno dessa turma, ao acaso, qual a probabilidade que o aluno escolhido tenha a altura entre 169 cm e 180 cm?
  - Qual a probabilidade de um aluno dessa turma ter a altura entre 158 cm e 169 cm?
  - Com base nas propriedades teóricas do Modelo Normal, há alguma relação entre as respostas dos itens “b” e “c”?

Fonte: Dados da pesquisa (2023).

De modo análogo a outra turma, após a resolução dessa atividade, o professor pesquisador mediou com os estudantes um momento de socialização (Quadro 31) para a reflexão dos temas abordados.

Quadro 31 – Recorte do diálogo da 2º e 3º questão da Atividade 3 - turma Geogebra

- Professor Pesquisador:** Qual a valor para a mediana e para moda vocês encontraram?
- Estudantes:** Mediana e a moda são 170 cm
- Professor Pesquisador:** Todas as duplas encontraram a mesma resposta?
- Estudantes:** Sim!
- Professor Pesquisador:** Ok. Vocês esboçaram o gráfico de Histograma com a ferramenta análise univariada no Geogebra. O gráfico é próximo ao modelo da Curva Normal?
- Alguns estudantes:** Sim!
- Estudante 21:** Fica uma Curva Normal mais ou menos
- Estudante 28:** Tem uma aproximação sim

**Professor Pesquisador:** Certo. Em relação as três medidas, média, moda e mediana, há alguma relação entre essas três medidas?

**Estudante 18:** A mediana e a moda são iguais a média deu diferente.

**Estudante 25:** A mediana e a moda são 170 cm, são iguais. A média deu 169 cm, ficou próximo, mas foi diferente.

**Professor Pesquisador:** Certo. Duas medidas iguais e uma próxima, mas diferente. Porque então isso aconteceu na distribuição?

**Estudante 21:** Por conta dos dados. Interferiram no valor

**Professor Pesquisador:** Nessa ideia. Devemos compreender que distribuições de variáveis contínuas nem sempre são totalmente simétricas, obedecendo ao modelo normal. Isso é comum de acontecer em vários fenômenos. Podem existir alguns motivos para isso, como um outlier, ou até mesmo o tamanho da população que não é suficiente grande para determinar uma curva. A distribuição pode ser assimétrica como essa aqui. No entanto, podemos notar que o histograma que foi construído por vocês no Geogebra tem uma aproximação com a Curva Normal. Então, com base nisso um fator importante para compreendermos é que à medida que o tamanho da população aumenta, a representação gráfica da distribuição para variáveis contínuas tenderá, ou seja, tem uma tendência a se aproximar de uma Curva Normal mais caracterizada, ou seja, simétrica.

**Professor Pesquisador:** Na segunda questão, utilizando a calculadora de probabilidades, qual a resposta para a primeira alternativa?

**Estudantes:** 50%

**Professor pesquisador:** Por que?

**Estudante 15:** Porque equivale à metade da Curva Normal

**Professor pesquisador:** Correto. Metade da área da Curva equivale a 50% de probabilidade, aproximadamente, tendo em vista que o nosso gráfico é uma aproximação ao modelo normal,

**Professor pesquisador:** Na alternativa b, qual a probabilidade encontrada?

**Estudantes:** 40,75%

**Professor pesquisador:** Certo. E qual a probabilidade solicitada na letra C?

**Estudantes:** Também 40,75%

**Professor Pesquisador:** Por que deram iguais?

**Alguns estudantes:** O intervalo é idêntico

**Professor Pesquisador:** Como podemos concluir isso? Já respondendo a alternativa d.

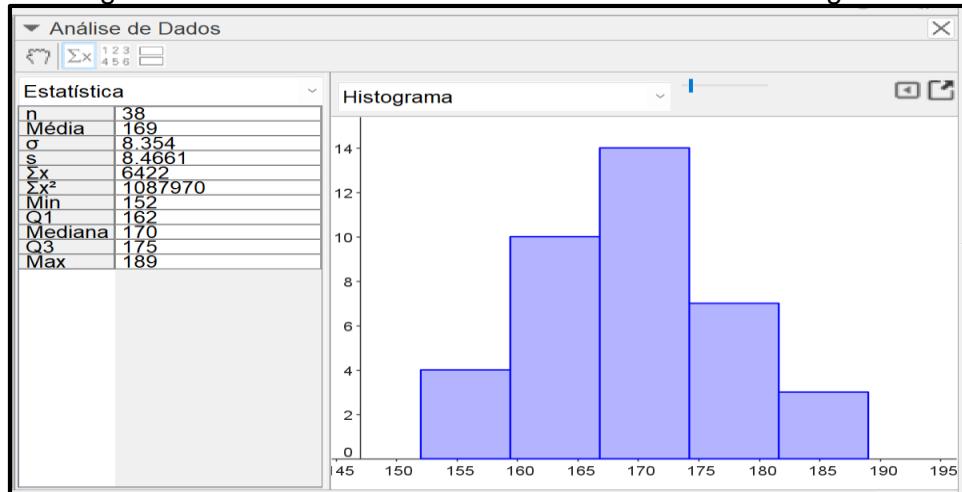
**Estudante 25:** O intervalo de 169 a 180 tem a mesma diferença que é de 11 para o intervalo 169 para 158. Então são iguais. E a região azul no Geogebra fica igual

**Professor pesquisador:** Correto. Os intervalos têm a mesma amplitude, que é de 11 cm e são simétricos em relação à média. Logo, a área considerada é a mesma, então tem a mesma probabilidade.

Fonte: Dados da pesquisa (2023).

O diálogo realizado na turma Geogebra ilustra um avanço na aprendizagem dos estudantes referente ao modelo da Curva Normal, contemplando as medidas de centralidade e dispersão e o cálculo de probabilidades. Na atividade, os estudantes utilizaram diferentes ferramentas no software Geogebra que envolveram o esboço do histograma das alturas, a determinação do valor da mediana e da moda e o cálculo de probabilidades para intervalos específicos da Curva Normal. Nessa etapa emergiram conclusões estatísticas semelhantes as observadas na outra turma. O Histograma se aproximou de uma distribuição normal, com a moda e a mediana assumindo o mesmo valor, porém a média não. A partir disso, foi debatido com os estudantes que a distribuição dos dados da turma não era totalmente simétrica, mas próxima do modelo de normalidade. Também foi refletivo que o gráfico da Curva Normal tende a ser simétrico e apresentar uma curva mais caracterizada a medida que a população aumenta. A utilização da ferramenta univariada (Figura 139) para o esboço do histograma das alturas é exemplificada a seguir.

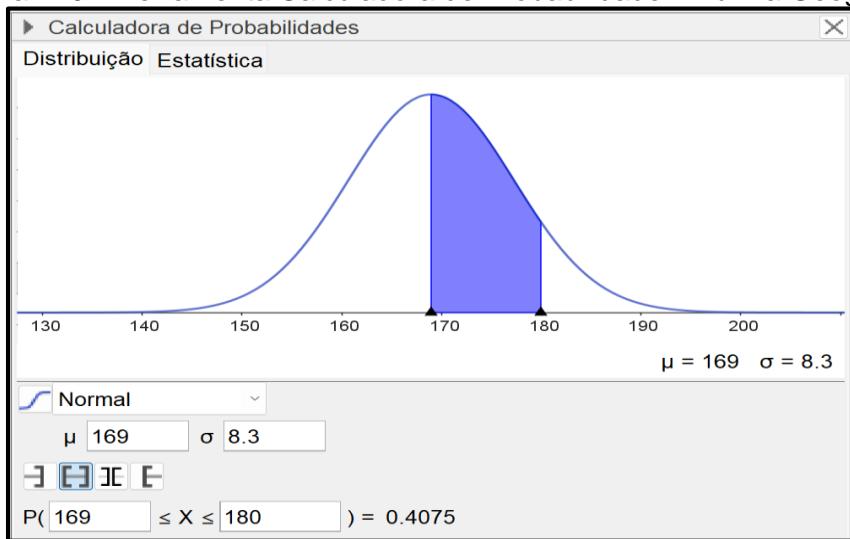
Figura 139 – Ferramenta análise univariada – Turma Geogebra



Fonte: Dados da pesquisa (2023).

Além disso, o debate também evidencia a aprendizagem dos estudantes sobre o conceito de probabilidade na Curva Normal. Utilizando a ferramenta “calculadora de probabilidades” no Geogebra, os estudantes procederam com respostas adequadas, associando a área do intervalo sob a Curva Normal ao valor da probabilidade solicitada nas questões. Deste modo, destacamos que a atividade vivenciada por meio da utilização do software Geogebra contribuiu para a construção do conhecimento sobre o tema e consequentemente, o desenvolvimento do Letramento Estatístico e Probabilístico dos estudantes. A seguir exemplificamos (Figura 140) a utilização da calculadora de probabilidades para a resposta da alternativa b da segunda questão.

Figura 140 – Ferramenta Calculadora de Probabilidade – Turma Geogebra



Fonte: Dados da pesquisa (2023).

Diante dessas conclusões, é pertinente ressaltar que o software Geogebra se revelou como um recurso significativo para a temática explorada nessa atividade, assim como constatamos na atividade anterior. Para o trabalho com a Curva Normal, suas ferramentas possibilitaram a construção rápida e dinâmica do histograma das alturas simultaneamente às análises estatísticas, o que torna a análise de dados enriquecedora. A calculadora de probabilidades possibilitou de maneira interativa e ágil a determinação das probabilidades por meio da manipulação de intervalos na representação gráfica da Curva Normal. Corroborando com essas conclusões, Tauber (2001) e Monroy e Herrera (2019) afirmam que a utilização de softwares para o ensino e aprendizagem da Curva Normal pode servir de elemento motivador por permitir um trabalho de forma mais dinâmica e exploratória, desde que tenha a adequada mediação do professor. Logo, a utilização de softwares como o Geogebra pode potencializar as ações docentes e discentes em favor do desenvolvimento do Letramento Estatístico e Probabilístico (Gal, 2002,2005).

#### 9.2.4 4º Momento Teorema Central do Limite: Relacionando a Amostragem com a Curva Normal

O quarto e último momento de nossa intervenção de ensino teve como objetivo promover a aprendizagem sobre o Teorema Central do Limite, com foco para a relação abarcada por ele entre a Amostragem e o modelo da Curva Normal. Para tanto, a quarta atividade foi desenvolvida em 4 etapas e os estudantes foram organizados em trios, nas duas turmas.

### 9.2.4.1 Turma Lápis e Papel

Na turma "Lápis e Papel", os estudantes inicialmente foram divididos em trios para realizar da quarta atividade. Cada trio representou uma população de interesse, e cada integrante forneceu a sua altura em centímetros. A primeira etapa consistiu no cálculo da média populacional, na qual cada trio determinou a média das alturas de seus membros. Na segunda etapa, os estudantes formaram todas as amostras possíveis de tamanho 2, retiradas com reposição, considerando que a população era composta por 3 integrantes. Assim, foram geradas 9 amostras distintas.

Na terceira etapa, cada trio calculou a média das 9 amostras formadas, e em seguida, determinaram a média das 9 médias obtidas. A partir disso, compararam o resultado obtido nessa etapa com o valor da média populacional determinada inicialmente. Na última etapa, os estudantes fizeram um esboço do histograma para as 9 médias, e partir disso, refletiram sobre a representação gráfica formada. Após a atividade o professor pesquisador conduziu um debate (Quadro 32) com a turma para a discussão dos conceitos abordados.

Quadro 32 – Recorte do diálogo da atividade 4 - turma Lápis e Papel

<p><b>Professor Pesquisador:</b> Vocês calcularam a média das três alturas do trio que vocês pertencem. Depois formaram as 9 amostras possíveis de tamanho 2 retiradas da população, ou seja, do trio. Em seguida, vocês calcularam a média de cada amostra e, depois, a média das 9 médias. Ao comparar o resultado da média das 9 médias com o resultado da média populacional a que conclusão vocês chegaram?</p> <p><b>Estudantes:</b> Tem o mesmo valor.</p> <p><b>Estudante 3:</b> Que a média das médias das amostras é igual à média da população.</p> <p><b>Professor Pesquisador:</b> Certo. Então a média da distribuição das médias amostrais, ou seja, a média das 9 médias amostrais tem resultado igual à média da população. Além disso, há alguma relação entre o gráfico construído e o modelo da Curva Normal?</p> <p><b>Estudantes:</b> Sim!</p> <p><b>Estudante 15:</b> O meu se aproxima da Curva Normal</p> <p><b>Estudante 31:</b> O meu também.</p> <p><b>Professor Pesquisador:</b> Certo. Então concluímos que a média distribuição das médias amostrais tem resultado igual à média da população e que o gráfico da distribuição dessas médias é uma Curva Normal. Vocês sabem que teorema na Estatística traz essas conclusões?</p> <p><b>Estudantes:</b> Não!</p>
---

**Professor Pesquisador:** É o teorema Central o Limite. Ele nos diz que a média da distribuição das médias amostrais será igual à média da população original. Além disso, se a representação gráfica dos dados de uma população é uma Curva Normal, a distribuição amostral das médias também será. Já se o comportamento da população for desconhecido ou aproximadamente normal, a distribuição das médias amostrais, geralmente para amostras com tamanho maior ou igual a 30, poderá ser uma Curva Normal ou se aproximar da normalidade à proporção que o tamanho da amostra aumenta.

**Professor Pesquisador:** Esse teorema é uma ferramenta muito importante para a Estatística Inferencial porque permite os estatísticos usem o modelo da Curva Normal e suas propriedades que estudamos ao longo dessas aulas para fazer inferências, tirar conclusões sobre a população, independentemente da forma da distribuição populacional dos dados. Esse teorema é base para outros conceitos como a margem de erro, intervalo de confiança que também discutimos, e testes de hipóteses. Qual a opinião de vocês sobre isso?

**Estudante 15:** Interessante. Não conhecia.

**Estudante 23:** Também achei interessante

Fonte: Dados da pesquisa (2023).

A discussão realizada na turma Lápis e Papel durante a quarta atividade revela que o resultado da média da distribuição das médias amostrais foi igual à média populacional calculada pelos estudantes. Além disso, constataram que a representação gráfica dessa distribuição se aproximou a uma Curva Normal. Essas conclusões obedecem aos fundamentos teóricos do Teorema Central do Limite. A partir disso, o professor pesquisador e os estudantes refletiram mais a fundo sobre esse teorema, debatendo que se a representação gráfica dos dados de uma população é uma distribuição normal, ou seja, uma Curva Normal, a distribuição amostral das médias igualmente será. Como também, se o comportamento da população for desconhecido ou aproximadamente normal, a distribuição das médias amostrais poderá ser uma Curva Normal ou se aproximar da normalidade à proporção que o tamanho da amostra aumenta.

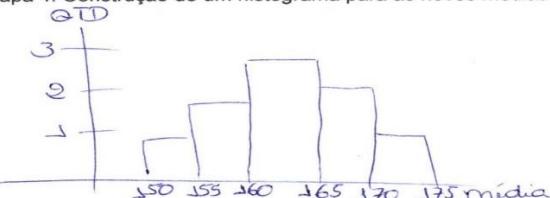
Consequentemente, foi discutido sobre a importância desse teorema para a Estatística Inferencial e o quanto o modelo da Curva Normal é útil para permitir os estatísticos realizarem inferências, o que notabiliza esse modelo como a principal ferramenta de análise de dados nos processos de inferências estatísticas (Batanero, Tauber e Sanchez, 2004). A atividade realizada proporcionou a aprendizagem de mais um importante tema, relacionando a Amostragem e a Curva Normal. Como exemplo, apresentamos a resposta (Figura 141) de um trio de estudantes para essa atividade.

Figura 141 – Resposta de um trio para a quarta atividade

c) Pergunta: Há alguma relação entre a média encontrada nessa etapa 3 com a média populacional obtida na etapa 1?

Sim. Mesmo valor. A média de (170, 162, 154) é 162 e a média das nove médias amostrais também.

d) Etapa 4: Construção de um histograma para as nove médias encontradas na Etapa 3



Pergunta: Há alguma relação entre o gráfico construído e o modelo da Curva Normal?

O gráfico tem a característica de uma Curva normal.

c) Sim. Mesmo valor. A média de (170, 162, 154) é 162 e a média das 9 médias amostrais também.

d) O gráfico tem a característica de uma Curva Normal

Fonte: Dados da pesquisa (2023).

Diante desses resultados, enfatizamos que o recurso Lápis e Papel permitiu alcançarmos o objetivo de aprendizagem da atividade e consequentemente contribuiu para o aprimoramento do Letramento Estatístico e Probabilístico (Gal, 2002, 2005) dos estudantes. Esse recurso possibilitou o entendimento de como as amostras se comportam em relação à média populacional e como a distribuição das médias amostrais obedece ao modelo normal, graficamente. No estudo 1 dessa pesquisa, observamos que os Livros Didáticos priorizam a abordagem da Estatística descritiva e pouco exploram a Estatística Inferencial. No entanto, essa atividade, com o uso de Lápis e Papel, representa um exemplo de como o professor, em sala de aula, pode trabalhar a Estatística para além da mera aplicação de cálculos e análises descritivas, ou seja, oportunizando a reflexão de conceitos e representações gráficas da inferência estatística que estão muito presentes em nosso cotidiano.

#### 9.2.4.2 Turma Geogebra

De forma semelhante, na turma Geogebra os estudantes também foram divididos em trios e cada um deles constituiu de uma população de interesse. A partir disso, as demais etapas foram adaptadas para o uso do Software Geogebra. Na primeira, os componentes de cada trio forneceram suas alturas e inseriram na tabela do Geogebra. A partir disso, calcularam a média populacional com o uso da ferramenta de análise univariada. Na etapa seguinte extraíram da população (trio)

todas as amostras possíveis de tamanho ( $n=2$ ) com reposição de elementos, totalizando 9 amostras e também as inseriam na tabela.

Em continuidade, utilizando a ferramenta análise univariada mais uma vez, os estudantes determinaram a média das 9 amostras e em seguida a média dessas 9 médias amostrais. A partir desse resultado esboçaram no software o gráfico de histograma para as 9 médias obtidas e refletiram sobre a sua representação gráfica em comparação como o modelo da Curva Normal. Após a resolução da atividade, o professor pesquisador procedeu com um debate (Quadro 33) junto aos estudantes para a discussão dos conceitos abordados.

Quadro 33 – Recorte do diálogo da atividade 4 - turma Geogebra

**Professor Pesquisador:** Na tabela do Geogebra, vocês inseriam os dados das alturas do trio, depois calcularam a média desses três valores com a ferramenta análise univariada. Depois formaram as 9 amostras possíveis de tamanho 2 retiradas da população e calcularam a média de cada amostra. Na sequência calcularam a média das 9 médias. Ao comparar o resultado da média dessas 9 médias com o resultado da média populacional a que conclusão vocês chegaram?

**Estudantes:** São iguais

**Professor Pesquisador:** Certo. Então podemos concluir que a média da distribuição das médias amostrais, ou seja, a média das 9 médias amostrais tem resultado igual à média da população. E sobre o histograma construído através da análise multivariada, há alguma relação entre o gráfico construído e o modelo da Curva Normal?

**Estudantes:** Sim!

**Estudante 21:** Está com a ideia da Curva Normal

**Professor Pesquisador:** Certo. Além de concluir que a média distribuição das médias amostrais tem resultado igual à média da população, o gráfico da distribuição dessas médias é uma Curva Normal. Alguém sabe o nome do teorema na Estatística que traz essas conclusões?

**Estudantes:** Não!

**Professor Pesquisador:** É o teorema Central do Limite. Esse teorema afirma que a média da distribuição das médias amostrais será igual à média da população. E se a representação gráfica dos dados de uma população é uma curva normal, a distribuição amostral das médias também será. Por outro lado, se o comportamento da população for desconhecido ou aproximadamente normal, a distribuição das médias amostrais, geralmente para amostras com tamanho maior ou igual a 30, poderá ser uma Curva Normal ou se aproximar da normalidade à proporção que o tamanho da amostra aumenta.

**Professor Pesquisador:** E Por que esse teorema é muito importante? Porque fundamenta a Estatística Inferencial, permitindo os estatísticos usem o modelo da Curva Normal e suas propriedades que estudamos para realizar

*inferências, tirar conclusões sobre a população, independentemente da forma da distribuição populacional dos dados. É a partir desse teorema que são aplicados outros conceitos, como a margem de erro, intervalo de confiança que também estudamos, e testes de hipóteses. O que vocês acharam dessa atividade?*

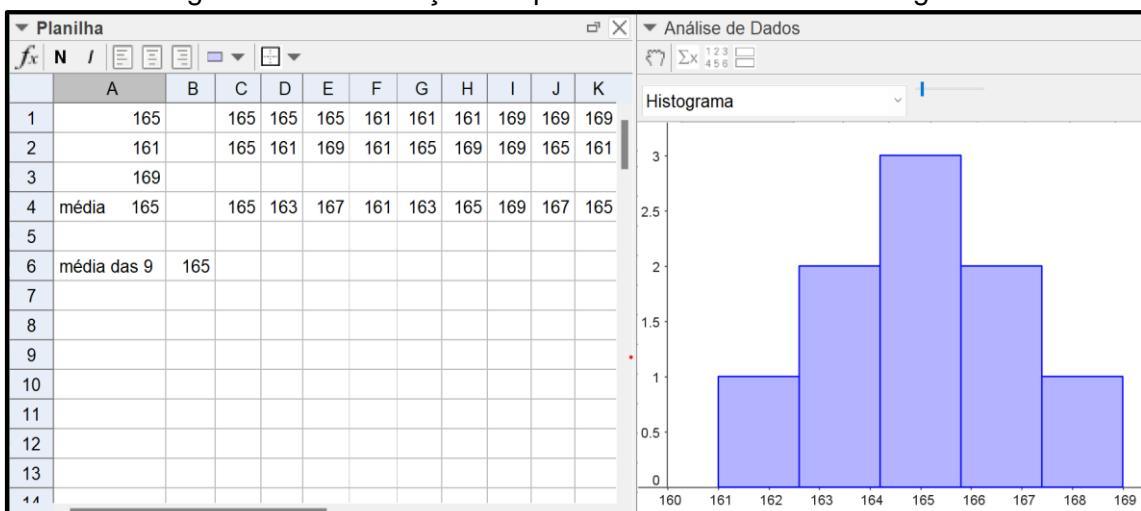
**Estudante 15:** *Eu achei interessante*

**Estudante 23:** *Achei legal o fato da média das médias da 9 amostra ser igual a primeira média, a da população.*

Fonte: Dados da pesquisa (2023).

A socialização realizada na turma Geogebra demonstra que os resultados práticos encontrados pelos estudantes utilizando software Geogebra vão ao encontro dos preceitos do Teorema Central do Limite. A partir disso, assim como realizado na outra turma, o professor pesquisador abordou as principais conclusões desse teorema, como foco para o entendimento de que a média da distribuição das médias amostrais é igual à média populacional, como também, que se a representação gráfica dos dados de uma população é uma distribuição normal, a distribuição amostral das médias também será. Já se o comportamento da população for desconhecido ou aproximadamente normal, a distribuição das médias amostrais poderá ser uma Curva Normal ou se aproximar da normalidade à proporção que o tamanho da amostra aumenta. Ao realizarmos essa atividade, ficou evidente que os estudantes não conheciam esse teorema e suas conclusões, porém, mediante a discussão coletiva, conseguiram avançar na aprendizagem dessa temática. A seguir, destacamos a resposta (Figura 142) de um trio de estudantes para essa questão utilizando o Geogebra.

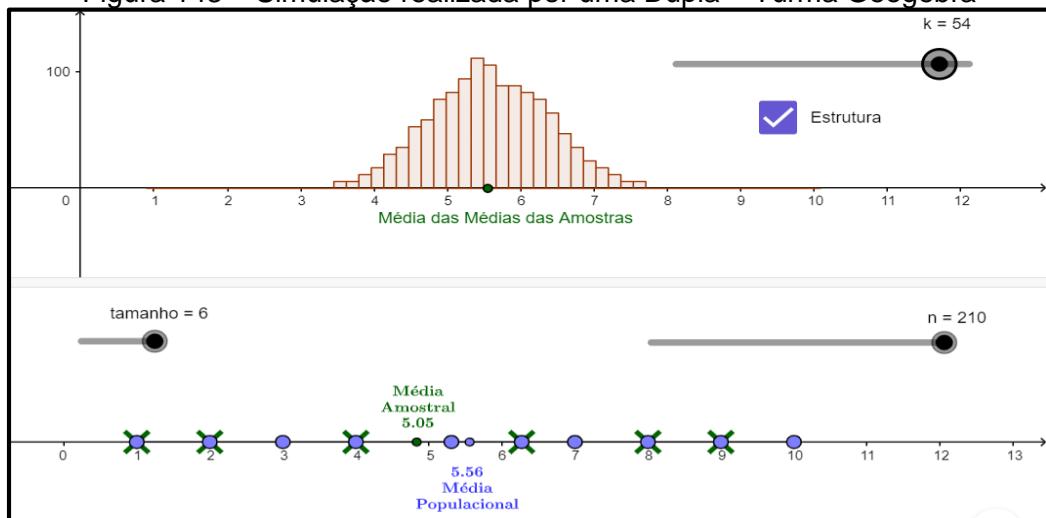
Figura 142 – Resolução da quarta atividade – Turma Geogebra



Fonte: Dados da pesquisa (2023).

Após a resolução da quarta atividade, os estudantes da turma Geogebra ainda manipularam, no software, outra ferramenta didática, desenvolvida por Bortolossi (2015) que ilustra as consequências do teorema central do limite. Através das manipulações dos controles deslizantes, os estudantes reforçam de a conclusão que quanto maior o tamanho da amostra, menor é a variabilidade de medida (Reyes, 2019) e, consequentemente, mais concentrada a média amostral está em torno da média de todas as médias amostrais que, por sua vez, coincide com a média populacional. Logo, a amostra selecionada terá uma probabilidade maior de que sua média esteja mais próxima da média populacional. Em acréscimo, refletimos que o histograma construído, obedece, graficamente, ao modelo da Curva Normal. Abaixo, destacamos (Figura 143) a manipulação da ferramenta realizada por uma dupla estudantes.

Figura 143 – Simulação realizada por uma Dupla – Turma Geogebra



Fonte: Dados da pesquisa (2023).

Essas conclusões reforçam que a Inferência Estatística pode ser abordada na Educação Básica, tanto com o lápis e papel, como também com o Geogebra, que se revelou nessa turma como um recurso adequado para a abordagem da Amostragem articulada com a Curva Normal, permitindo a automatização do tratamento dos dados e a visualização dinâmica e interativa do comportamento das amostras e da representação gráfica dos dados (Batanero, Begué e Gea, 2018). Essa prática contribuiu para o aprimoramento do Letramento Estatístico e Probabilístico (Gal, 2002, 2005) dos estudantes.

Diante das conclusões expostas para os quatros momentos com suas respectivas atividades, ratificamos que a intervenção de ensino realizada, nas duas turmas, desempenhou um papel significativo no aprendizado de conceitos

relacionados à Amostragem e à Curva Normal, bem como na articulação entre esses conceitos. Os recursos, tanto o lápis e papel, quanto o software Geogebra contribuíram, cada um em sua perspectiva, para o desenvolvimento das atividades e a construção do conhecimento. Desta forma, destacamos que na Educação Básica é plenamente possível a abordagem de temas relacionados a Inferência Estatística, seja utilizando métodos tradicionais como o lápis e papel, ou explorando as potencialidades de softwares como o Geogebra. A realização do processo de ensino e aprendizagem dessa temática de maneira significativa possibilitará o aprimoramento e ampliação do Letramento Estatístico e Probabilístico dos estudantes (Gal, 2002,2005).

## CONSIDERAÇÕES FINAIS

A Amostragem e a Curva Normal desempenham papéis cruciais na Estatística Inferencial, proporcionando bases para generalizar resultados que permitem os estatísticos extraírem conclusões significativas e confiáveis sobre populações a partir de amostras representativas. A Base Nacional Comum Curricular (BNCC, 2018) orienta que no Ensino Médio o estudante realize e analise criticamente pesquisas amostrais, bem como, compreenda as medidas de centralidade, de dispersão e o cálculo de probabilidade em diferentes contextos, o que faz recomendar de modo explícito e implícito a abordagem da Amostragem e da Curva Normal nessa etapa de escolarização.

Reconhecendo a importância do ensino dessa temática e com base nas concepções do Letramento Estatístico e Probabilístico (Gal, 2002, 2005) e do Ciclo Investigativo de uma pesquisa estatística (Guimarães e Gitirana, 2013), a realização dessa pesquisa teve o objetivo de investigar o que estudantes do 3º ano do Ensino Médio sabem e podem aprender sobre Amostragem, Curva Normal e suas relações.

Para alcançar tal objetivo, realizamos, 4 estudos, dos quais os três primeiros deram base para a realização do último. No Estudo 1, diante da importância dos livros didáticos para o processo de ensino e aprendizagem, analisamos o que tem sido proposto sobre Amostragem, Curva Normal e suas relações nos materiais para o ensino de Matemática (Livro de Projetos e Livros de Conhecimento) aprovados pelo PNLD 2021. Os resultados apontam que, tanto os livros de projetos quanto aos de conhecimento, apresentam interessantes propostas para o trabalho de tópicos da Estatística e da Probabilidade, principalmente na abordagem dos resultados de pesquisas estatísticas e realização das mesmas.

No que diz respeito, especificamente, aos livros de projetos, constatamos que a maioria das situações exploram dados reais e, assim, possibilitam uma maior reflexão dos estudantes sobre o contexto social ao qual estão inseridos e, consequentemente, uma maior compreensão de mundo. Dessa forma, essa perspectiva contribui para o aprimoramento do Letramento Estatístico e Probabilístico (GAL, 2002, 2005) dos estudantes. No entanto, também observamos uma predominância de situações que envolvem o conceito de população/amostra se referindo apenas a pessoas, pouca discussão sobre a seleção e representatividade da amostra, uma maior ênfase para as análises descritivas dos dados amostrais e pouco abordagem para o modelo da Curva Normal.

Já os livros de conhecimento, as obras de Matemática e suas Tecnologias exploram mais dados fictícios com ênfase para aplicação do cálculo de medidas estatísticas. Além disso, também há uma predominância de situações que envolvem o conceito de população/amostra se referindo apenas a pessoas. Verificamos ainda que nos tópicos sobre Amostragem não são abordados todos os tipos de métodos para a seleção de amostra e pouco se discute os aspectos para a representatividade amostral. Com relação a Curva Normal, a partir de dados fictícios, há uma ênfase para a aplicação de técnicas operatórias para o cálculo das medidas de centralidade, dispersão e probabilidade. Assim, pouco se explora os significados dos conceitos estatísticos e probabilísticos presentes nesse modelo, bem como, como ele é aplicado e modelado em fenômenos do nosso cotidiano.

Os livros de Ciências Humanas e Matemática, por sua vez, abordam apenas o conceito de Amostragem, se assemelhando aos livros de projetos, com uma predominância para a análise e realização de pesquisas estatísticas, o termo população/amostra, em sua maioria, se referindo apenas a pessoas e pouca discussão sobre a seleção e representatividade da amostra. Logo, isso pode implicar em uma concepção, por parte dos estudantes, que o termo população é relativo apenas a pessoas na Estatística, bem como, não reconhecer os critérios para a representatividade amostral.

No Estudo 2, investigamos o que foi proposto nas questões da prova de Matemática e suas tecnologias do Exame Nacional do Ensino Médio, em sua edição de 2022, sobre a Amostragem, Curva Normal e suas relações, uma vez que esse exame orienta as propostas de ensino nas escolas. Para tanto, inicialmente fizemos uma análise na Matriz de Referência desse exame a qual reúne os eixos cognitivos, competências e habilidades exigidas aos candidatos participantes. Assim, direcionamos nosso olhar para a área de Matemática e suas tecnologias, e mais especificamente, elencamos quais dessas normativas tem relação com estatística e probabilidade. A partir desse entendimento, analisamos as 45 questões de Matemática do caderno de prova amarelo da edição do ENEM em 2022.

Os resultados apontam que, de modo explícito, há apenas uma competência que cita o conceito de Amostragem relacionada à compreensão do caráter aleatório e não-determinístico dos fenômenos naturais e sociais e da utilização instrumentos adequados para medidas, determinação de amostras e cálculos de probabilidade para interpretar informações de variáveis apresentadas em uma distribuição estatística. Já a Curva Normal é contemplada implicitamente nessa competência ao citar

“distribuição estatística” e em habilidades voltadas para a interpretação de gráfico ou tabelas para fazer inferências, a resolução de problemas com dados apresentados em representação gráficos ou tabelas, a análise das informações expressas nessas representações para a construção de argumentos, o cálculo de medidas de tendência central ou de dispersão, na resolução de situações-problema que envolva conhecimentos de estatística e probabilidade e na utilização de desses conhecimentos como recurso para a construção de argumentação.

Com relação as questões, verificamos que nenhuma delas aborda o conceito de Amostragem e nem da Curva Normal. As questões que exploram a Estatística têm uma abordagem voltada apenas para o cálculo de medidas de centralidade e as que exploram a probabilidade abordagem unicamente o significado de probabilidade enquanto razão clássica. Assim, conclui-se que a ênfase dessas questões está relacionada com o tratamento e aplicação de técnicas operatórias e nenhuma delas relaciona a Estatística e a Probabilidade, bem como a Inferência Estatística. Dessa forma, a proposta dessas questões limita a Estatística e Probabilidade a exploração descritiva de alguns conceitos e seus significados, o que não atende à demanda descrita nos documentos oficiais para o ensino e aprendizagem dessas áreas na Educação Básica

No Estudo 3, identificamos os conhecimentos iniciais de estudantes do 3º ano do Ensino Médio sobre Amostragem, Curva Normal e suas relações. Esse estudo teve por objetivo realizar uma diagnose com vistas a revelar o nível de aprendizagem de turmas desse nível de escolaridade Médio sobre a referida temática, como também validar nosso instrumento de análise. Para tanto, os estudantes responderam a um questionário diagnóstico, pelo qual ficou evidenciado que a maioria dos estudantes demonstrou domínio sobre o significado amostra e sua relação com a população, porém apresentaram lacunas sobre a compreensão da representatividade da amostra, incluindo a variabilidade amostral e os métodos de Amostragem, como também, no modelo da Curva Normal, contemplando sua representação gráfica e propriedades. Em acréscimo, evidenciou-se fragilidades na compreensão da relação entre a distribuição das médias amostrais com os parâmetros populacionais, pautada no Teorema Central do Limite que relaciona a Amostragem com a Curva Normal.

Com base nesses resultados, realizamos o quarto estudo que teve o objetivo de analisar a contribuição dos diferentes recursos (Lápis e papel e o software Geogebra) para a aprendizagem de Amostragem e Curva Normal de forma articulada a partir de um processo interventivo. Acreditamos que a utilização de diferentes

recursos amplia o leque de possibilidades didáticas para o ensino, bem como, pode enriquecer o processo de aprendizagem dos estudantes, atendendo às suas necessidades individuais, estimulando a criatividade e o pensamento crítico. Esse estudo foi realizado em duas turmas de estudantes do 3º ano do Ensino Médio da rede pública de ensino da cidade de Nazaré da Mata. Para alcançar esse objetivo, pautamos esse estudo em três etapas: pré-teste, a intervenção de ensino e o pós-teste. O pré-teste teve como objetivo investigar os conhecimentos prévios dos estudantes sobre o tema do nosso estudo e pós-teste visou analisar as contribuições da proposta de intervenção de ensino. A intervenção de ensino foi desenvolvida em quatro momentos distintos, na qual uma turma teve como recurso o lápis e papel e a outra o software Geogebra.

Na etapa do pré-teste, verificamos que as turmas apresentaram desempenhos equivalentes e lacunas no conhecimento sobre a Amostragem e a Curva Normal, principalmente sobre a variabilidade e representatividade amostral, métodos de Amostragem, o gráfico da Curva Normal e suas propriedades. No pós-teste, ambas as turmas apresentaram um desempenho significativamente superior, evidenciando que a intervenção de ensino contribuiu para a aprendizagem dos estudantes sobre a Amostragem e a Curva Normal, seja com o apoio do lápis e papel ou do software Geogebra.

Durante a intervenção foram realizadas várias atividades práticas envolvendo contextos significativos para os estudantes: pesquisas amostrais com dados reais, refletindo sobre o tamanho da amostra e a variabilidade dos dados; o modelo da Curva Normal com ênfase para a sua representação gráfica e as medidas de centralidade e dispersão; a variabilidade de medida (margem de erro) diferentes estatísticas obtidas por diferentes amostras retiradas de uma mesma população e uma atividade prática sobre o teorema central do limite. Essas atividades realizadas em pequenos grupos e refletidas na turma com a mediação do professor pesquisador proporcionaram a compreensão da relação entre a média populacional e sua relação com a média da distribuição das médias amostrais e seu gráfico que obedece ao gráfico da Curva Normal. Assim, os estudantes se apropriaram das conclusões desse teorema que fundamenta a inferência estatística.

A partir do pós-teste, pudemos constatar que os estudantes apresentaram melhoras significativas em relação ao desempenho nas questões tanto de Amostragem como de Curva Normal e a relação entre ambos.

Diante dos resultados da pesquisa, destacamos que tanto o Lápis e Papel quanto o Geogebra contribuíram para a aprendizagem dos estudantes sobre a Amostragem e a Curva Normal. O software Geogebra possibilitou o tratamento dos dados de forma robusta por disponibilizar ferramentas que permitiram os estudantes uma análise multivariada, como também, a visualização dinâmica do comportamento dos dados, a medida em que eles eram tratados, contemplando diferentes medidas estatísticas e representações gráficas. Além disso, a análise dos dados foi realizada de forma mais ágil, o que proporcionou um ambiente interativo e instigador. Já o lápis e papel também se revelou como um recurso significativo que de igual modo, promoveu a construção de conhecimentos sobre esses conceitos. Os estudantes ao fazerem os cálculos manualmente, por exemplo, podem desenvolver um entendimento profundo sobre os significados dos conceitos estatísticos e probabilísticos, pois estão envolvidos diretamente nos cálculos e na manipulação dos dados. Assim, o recurso do Lápis e Papel possibilitou o desenvolvimento de habilidades analíticas que proporcionaram a compreensão do significado dos conceitos abordados.

Logo, enfatizamos que esses dois recursos podem contribuir para a aprendizagem e, consequentemente, para o desenvolvimento do Letramento Estatístico e Probabilístico (Gal, 2002 e 2005) desde que, como nesse estudo, sejam propostas atividades desafiadoras. Essas não podem se resumir a mera aplicação de cálculos e procedimentos práticos, mas devem provocar a mobilização de elementos de conhecimento e de disposição (Gal, 2002, 2005) por parte dos estudantes. Outro fator crucial é que essa prática, independente do recurso, deve ser mediada de modo adequado pelo professor, o qual deve considerar os objetivos a serem alcançados e as diferentes aprendizagens dos estudantes a partir das potencialidades e desafios de cada recurso didático utilizado.

Diante do exposto, defendemos que o ensino sobre Amostragem e Curva normal de forma articulada, por meio do ciclo investigativo e da utilização de diferentes recursos (Lápis e papel e o Geogebra), propicia a aprendizagem e o desenvolvimento do Letramento Estatístico e Probabilístico de estudantes do 3º ano do Ensino Médio. Consequentemente, acreditamos os resultados dos estudos que compõem essa pesquisa podem promover novas discussões e reflexões sobre o ensino dessa temática, como também favorecer a qualidade das ações docente e discente, em sala de aula, ampliando a potencialidade das tarefas didático-pedagógica e de aprendizagem.

Considerando esta Tese, concluímos apontando como perspectiva de investigações futuras, a realização de estudos que explorem formações continuadas de professores da Educação Básica, com vistas à ampliação de seus conhecimentos e, por conseguinte, a potencialização do processo de ensino e aprendizagem da Amostragem e Curva Normal de forma articulada em sala de aula. Tal investimento pode ampliar o leque de possibilidades didáticas para esse processo contemplando diferentes recursos como softwares, lápis e papel e materiais manipuláveis. Destacamos que o ensino da Inferência Estatística contemplando esses dois conceitos favorece aos estudantes a compreensão de diversos fenômenos do nosso cotidiano e a tomadas de decisões, bem como, o desenvolvimento do Letramento Estatístico e Probabilístico.

## REFERÊNCIAS

- ANDERSON, D. R.; SWEENEY, D. J.; WILLIAMS, T. A. **Estatística aplicada à administração e economia.** 2º ed. São Paulo: Pioneira. 2002
- ALCHORNE, I.; CARVALHO, S. **Vivências: Projeto de Vida.** 1º Ed. Scipione. São Paulo, 2021.
- ANDRADE, D. M.; BATISTA, J. O.; SANTOS, D. S. S.; MOCROSKY, L. F. A influência do Novo ENEM na relação de ensino e aprendizagem matemática: o caso da Escola Jorceli da Silva Sestari. **Ensino e Multidisciplinaridade**, v. 6, n. 1, p. 30-47, 2020.
- ARAÚJO, A. F. Q. **A inter-relação entre a Estatística e a Probabilidade: um estudo com professores de Matemática do Ensino Médio sobre a curva normal.** 189f. Dissertação (Mestrado em Educação Matemática e Tecnológica) – Universidade Federal de Pernambuco, 2020.
- BATANERO, C. **Significados de la probabilidad en la educación secundaria. Revista Latinoamericana de Investigacion en Matemática Educativa.** Vol. 8. Nº3. México. pp.247-263. 2005.
- BARBERINO, M. R. B.; MAGALHÃES M. N. uma proposta para desenvolver o letramento estatístico através de projetos. **VYDIA**, v. 36, n. 2, p. 363-375, jul./dez. Santa Maria, 2016.
- BATANERO, C.; TAUBER, L.; SÁNCHEZ, V. Students' reasoning about the normal distribution. In D. Ben-Zvi and J. Garfield (Eds.), **The challenge of developing statistical literacy, reasoning and thinking**, 2004. p.257–276.
- BATANERO, C.; BEGUE, N.; GEA, M. M. Como desenvolver o senso de amostragem nos alunos? Ingrith (Ed.), **Memórias do III Encontro Colombiano de Educação Estocástica** Bogotá, Colômbia, 11-22, 2018.
- BAYER, A.; ECHEVESTE, S.; BITTENCOURT, H. R.; ROCHA, J. Preparação do formando em Matemática-licenciatura plena para lecionar estatística no Ensino Fundamental e Médio. In: **Encontro Nacional de Pesquisa em Educação em Ciências**. Bauru, Brasil, 2005.
- BEGUÉ, N.; GEA, M.; BATANERO, C.; BELTRÁN P. Meaning of sampling for secondary school students. **Anais do ICERI.** 1430-1435, 2017.
- BEGUÉ, N., BATANERO, C., GEA, M.M. y BELTRÁN, P. **Comprensión de la representatividad y variabilidad muestral por estudiantes de educación secundaria.** Yupana, 12, 8-22, 2020.
- BELTRAMI, E. **O que é aleatório? Chance e ordem na matemática e na vida.** Nova York: Copernicus/Springer-Verlag, 1999.
- BENNETT, D. J. Aleatoriedade. Cambridge, MA: Harvard Gal, I. Ligações entre alfabetização e numeramento. **Jornal universitário**, 1998.

BEN-ZVI, D.; BAKKER, A.; MAKKAR, K. Learning to reason from samples. **Educational Studies in Mathematics**, v. 88, n. 3. 291-303, 2015.

BITTENCOURT, H. R., VIALI, L., **Contribuições para o ensino da distribuição normal ou curva de Gauss em cursos de Graduação**. In: SIPEM (Seminário Internacional de Pesquisa em Educação Matemática), 2006. Águas de Lindóia. Anais do III SIPEM, 2006.

BITTAR, M. A Teoria Antropológica do Didático como ferramenta metodológica para análise de livros didáticos. **Zetetiké**, Campinas, SP, v.25, n. 3, set./dez, 2017. 364-387.

BUENO, C. O. C. **Ação na escola e na comunidade-Projetos integradores**. 1. ed. FTD. São Paulo, 2021.

BRASIL. **Portaria n.438**, de 28 de maio de 1998. Institui o Exame Nacional do Ensino Médio. Disponível em [https://www.camara.leg.br/proposicoesWeb/prop\\_mostrarIntegra?codteor=1121401](https://www.camara.leg.br/proposicoesWeb/prop_mostrarIntegra?codteor=1121401). Acesso em 10 de novembro de 2022.

BRASIL. Ministério de Educação e Cultura. **LDB - Lei nº 9394/96**, Brasília, 1996.

BRASIL. **Matriz de Referência do ENEM**. Ministério da Educação. Brasília, 2009.

BRASIL, Ministério da Educação e da Secretaria do Ensino Médio. **Parâmetros Curriculares Nacionais**. Brasília, 1998.

BRASIL, Ministério da Educação e da Secretaria do Ensino Médio. **Parâmetros Curriculares Nacionais do Ensino Médio**. Brasília, 2000.

BRASIL, Ministério da Educação e da Secretaria do Ensino Médio. **PCN+ Ensino Médio: Orientações Educacionais complementares aos Parâmetros Curriculares Nacionais - Ciências da Natureza, Matemática e suas Tecnologias**. Brasília, 2002.

BRASIL, Ministério da Educação e da Secretaria de Educação. **Base Nacional Comum Curricular**. Brasília, 2018.

BRASIL, Ministério da Educação e da Secretaria de Educação. **Guia do Programa Nacional do Livro e do Material Didático**. Brasília, 2021.

BONJORNO, J. R.; GIOVANNI JÚNIOR, J. R.; CÂMARA, P. R. S. **Prisma Matemática**. 1º Ed. FTD, São Paulo, 2021.

BORTOLOSSI, H. J. **Efeito do Tamanho da Amostra na Variabilidade da Estatística Amostral**, 2015. Disponível em: [Efeito do Tamanho da Amostra na Variabilidade da Estatística Amostral – GeoGebra](https://www.geogebra.org/m/343433). Acesso em 10 de Janeiro de 2023.

BORTOLOSSI, H. J. **Amostragem: Tamanho da amostra**, 2020. Disponível em: [Amostragem: Tamanho da Amostra – GeoGebra](https://www.geogebra.org/m/343433). Acesso em 15 de Janeiro de 2023.

CAZARES, S. I. Comprensión que estudiantes universitarios desarrollan sobre el muestreo y variabilidad muestral desde una perspectiva informal. In J. M. Contreras; M. M. Gea; M. M. López- Martín; E. Molina-Portillo (Eds.), **Actas del Tercer Congreso International Virtual de Educación Estadística**, 2019.

CAZORLA, I.M; SANTANA, E. R. S. **Tratamento da informação para o ensino fundamental e médio**. Ed. Via Litterarum. Bahia, 2006.

CAZORLA, I., MAGINA, S. GITIRANA, V. GUIMARÃES, G. **Estatística para os anos iniciais do ensino fundamental** [livro eletrônico] . -- 1. ed. - Brasília : Sociedade Brasileira de Educação Matemática - SBEM, 2017.

COPE, B., KALANTZIS, M. & New London Group. **Multiliteracies: Literacy learning and the design of social futures**. New York: Routledge, 2000.

CONRE-3. **Amostras-média-dominó**. 2017. Disponível em: [EXEMPLO - amostras-média-dominó - FRENTE E VERSO.pdf \(dropbox.com\)](#) Acesso em 30 de Janeiro de 2023.

CONRE-3. **Estação 1a - medida do palmo da mão direita** 2020. Disponível em: [material para a tenda estatística - dropbox](#) Acesso em 21 de março de 2023.

DANCEY, C. P. & REIDY, J. **Estatística sem matemática para psicologia**. Tradução Lori Viali. Porto Alegre: Artmed, 2006.

DANTE, L. R.; VIANA. F. **Matemática em Contextos**. 1º Ed. Ática. São Paulo, 2021.

DANZA, H. C.; SILVA, M. A. M. **Projeto de vida: Construindo o futuro**. 1º Ed. Ática. São Paulo, 2021.

DE LEONARDO, F. M. **Moderna em Projetos: Matemática e suas tecnologias**. 1º Ed. Moderna. São Paulo, 2021a.

DE LEONARDO, F. M. **Conexões: Matemática e suas tecnologias**. 1º Ed. Moderna. São Paulo, 2021b.

DUARTE, L. R. **A utilização do software Geogebra no ensino da distribuição normal de probabilidade: uma aproximação entre a geometria dinâmica e a educação estatística**. 130f. (Dissertação de Mestrado em ensino de ciências e Matemática) – Pontifícia Universidade Católica de Minas Gerais, 2010.

FERREIRA, V. **Estatística básica**. Vol 1, Ed.. Seses, Rio de Janeiro, 2015

FILHO A. D. P.; COSTA, N. M. L. **Introdução ao ensino da curva normal: uma experiência por meio de jogos com estudantes do ensino médio**. Editora Atena, Paraná, 2021.

FREITAS, L. M. T. F. **Interação Matemática**. 1º Ed. Editora do Brasil, São Paulo, 2021

FURTADO, P.; FACALANI, J. GONÇALVES E. P. **Vamos Juntos, Profe! Projetos Integradores.** Ed. Saraiva, São Paulo, 2021.

GAL, I. Adult's Statistical Literacy: Meanings, Components, Responsibilities – **Internacional Statistical Review**, Australia, v. 70, 2002. p. 1-33

GAL, I. Towards 'probability literacy' for all citizens. In: Jones, G.A (ed.), **Exploring probability in school: Challenges for teaching and learning.** USA: Springer, 2005. p. 39-63.

GAL, I. Understanding statistical literacy: About knowledge of contexts and models. In: Tercer Congreso Internacional Virtual de Educación Estadística. **Actas.** Granada: Universidad de Granada, p. 1-15, 2019.

GAY, M. R. G. **Práticas na escola: matemática e suas tecnologias.** 1º Ed. Moderna. São Paulo, 2021.

GIL, A. C. **Como elaborar projetos de pesquisa.** 4. ed. São Paulo: Atlas, 2008..

GOMES, T. M. **O todo é a soma das partes, mas uma parte representa o todo? Compreensão de Estudantes do 5º e 9º ano sobre Amostragem.** [Dissertação de Mestrado em. Centro de Educação - Universidade Federal de Pernambuco], 2013.

GOMES, T. M.; GUIMARÃES, L. G. Compreensão dos estudantes do ensino fundamental sobre seleção de uma amostra representativa. **Com a palavra, o professor**, v. 3, n. 6, p.132-149, 2018.

GONÇALVES, P. **Uma abordagem da distribuição normal através da resolução de uma situação problema com a utilização do software geogebra.** 102f. [Dissertação de Mestrado profissional em Matemática], Universidade Federal de Goiás, 2014.

GONÇALVES, P. G. F., & LIMA, R. A. S. V. o ensino de estatística por meio da pesquisa: uma experiência a luz da modelagem matemática. **Holos**, 2, 190–198, 2017.

GONZÁLEZ, Y. K. OJEDA, A. M. PALACIOS, J. L. **Comprensión de Profesores de la Distribución Normal.** Acta Latinonamerica de Matemática Educativa. VOL 31, número 2, 2018.

GUIMARÃES, G. L.; GITIRANA, V. Estatística no Ensino Fundamental: a pesquisa como eixo estruturador. In: BORBA, R. E.; MONTEIRO, C. E. (Org.). **Processos de ensino e aprendizagem em Educação Matemática.** Recife: UFPE, 2013, p. 93-132.

GUIMARÃES, G.; CAVALCANTI, M.; EVANGELISTA, B. Ensino e aprendizagem de escalas representadas em gráficos: alunos do ensino regular e EJA dos anos iniciais. **Revista de Matemática, Ensino e Cultura.** v.15. 43 – 59, 2020.

HERRERA, J. S., MONROY, J. C. V.; HERNÁNDEZ, U. S. **Un acercamiento a la metodología lesson study para la enseñanza de la distribución normal.** En L. J. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. AguilarGonzález, P. Alonso, F. J. García

García y A. Bruno (Eds.), Investigación en Educación Matemática XXII (pp. 525-534). Gijón: SEIEM, 2018.

INNABI, H. A. Factors Considered by Secondary Students When Judging the Validity of a Given Statistical Generalization. **International Electronic Journal of Mathematics Education**, 2(3), 168-186, 2007.

INEP. **Caderno de Provas do ENEM do ano de 2022**. Disponível em: [2022 — Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira | Inep \(www.gov.br\)](http://www.gov.br/instituto-nacional-de-estudos-e-pesquisas-educacionais-anisio-teixeira/inep/caderno-de-provas-do-enem-do-ano-de-2022) Acesso em 15 de julho de 2021.

JACOB, B.; DOERR, H. M. **Students' informal inferential reasoning when working with the sampling distribution**. Anais do 8º Cerme, Turquia, 2013.

KAHNEMAN, D.; e TVERSKY, A. Subjective probability: A judgment of representativeness. **Cognitive Psychology**, 3(3), 430-454, 1972.

LARSON, Ron; FARBER, Betsy. **Estatística aplicada**. São Paulo: Pearson Education do Brasil, 6º ed. 2015.

LIMA, E. T.; BORBA, E. S. R. B. Articulando os raciocínios combinatório e probabilístico a partir da resolução de problemas na EJA. **Educação Matemática. Matem Pesquisa**. São Paulo, v.21, n.1, pp.136-159, 2019.

LIMA, O. A. **Distribuição Normal: Uma introdução voltada ao Ensino Médio por simulações via planilha eletrônica e exercícios interativos**. Dissertação de Mestrado. PUC, São Paulo, 2009.

LUNA, L. C.; GUIMARÃES, G. L. Aprendizagem de amostragem nos PCN e na BNCR e a influência nos livros didáticos. **Revista de Educação, Ciências e Matemática**, v. 11, n. 1, 2021. 1-20.

LUNA, L. C. **Ensino e aprendizagem de amostragem nos anos finais do Ensino Fundamental**. 261f. Tese (Doutorado em Educação Matemática e Tecnológica) – Universidade Federal de Pernambuco, Recife, 2023.

MACIEL, D. B. M. **Tipologia das condições de letramento estatístico dos estudantes brasileiros de economia: o que revela o ENADE?** 216f Tese (Doutorado em Educação Matemática e Tecnológica) – Universidade Federal de Pernambuco, Recife, 2023.

MALHOTRA, Naresh K. **Introdução à pesquisa de marketing**. Editora Prentice Hall, São Paulo, 2005.

MARTINS, G. E. **Desvio padrão amostral**. Revista de Ciência Elementar, Lisboa, v.1, n.1, p. 1, out./dez. 2013.

MATTAR, F. **Pesquisa de marketing**. Ed. Atlas, São Paulo, 1996.

MELÃO, W. S.; SOARES, M. T. C. Implicações do novo ENEM na perspectiva dos Professores de Matemática do ensino médio. **Série-Estudos**. n.33, 2012.

MELO, R. A. F. P. **Uma Análise Sobre o Estudo de Amostragem no Ensino Médio**, 80f. Dissertação (Mestrado) – Universidade Federal de São Paulo. Instituto de Ciência e Tecnologia. Programa de Pós-Graduação em Matemática em Rede Nacional, 2017.

MOORE, D. S. **A Estatística básica e sua prática**. 3<sup>a</sup> ed. New York: W. H. Freeman and Company, 1995

MORRONE, M. L. O Exame Nacional do Ensino Médio: democratização do ensino superior e regulação do sistema escolar. In: Simpósio Brasileiro de Política e Administração da Educação. **Anais** São Paulo, 2011.

MONROY, J. C. V.; HERRERA, J. S. **Análisis de las respuestas de estudiantes debachillerato a problemas sobre la distribución normal**. En J. M. Contreras, M. M. Gea, M. M. López- Martín y E. Molina-Portillo (Eds.), **Actas del Tercer CongresoInternacional Virtual de Educación Estadística**, 2019.

MOORE, D. S. **A Estatística básica e sua prática**. 3<sup>a</sup> ed. New York: W. H. Freemanand Company, 1995.

MULLER, D. A.; NUNES, L. N. Ensino de estatística no ensino médio noturno pela prática de uma pesquisa de campo. **Educação Matemática. Matem Pesquisa**. São Paulo, v.18, n.3, pp. 1245-1263, 2016.

OLIVEIRA, A. F. de ; ROSA, D. E. G. . A estatística no ensino médio: em busca da contextualização. **Zetetike**, Campinas, SP, v. 28, p. e020006, 2020. DOI: 10.20396/zet.v28i0.8657024.

OLIVEIRA, T. S. O ENEM: breves considerações sobre importância avaliativa e reforma educacional. **Educação Por Escrito**, Porto Alegre, v. 7, n. 2, p. 278-288, jul.-dez. 2016.

OLIVEIRA, P. F. C. **Identidade em Ação - Ciências Humanas e Sociais Aplicadas e Matemática**. 1<sup>º</sup> Ed. Moderna, São Paulo, Brasil.

OTOBELLI, E. S. **Questões do Enem: uma possibilidade de (re)construção de conhecimentos**. 214f. [Dissertaçâode Mestrado profissional em Ensino DE Ciências Matemática], Universidade de Caxias do Sul, 2018.

PEREIRA, F. A.; SOUZA, F. S. O Exame Nacional do Ensino Médio e a construção do letramento e pensamento estatístico. **Educação Matemática Pesquisa**, São Paulo, v. 18, n. 3, p. 1319-43, 2016

PERNAMBUCO. **Parâmetros para a Educação Básica do Estado de Pernambuco**, Parâmetros na sala de aula Matemática para o Ensino Fundamental e Médio, Secretaria de Educação, Pernambuco, 2013.

PERNAMBUCO. **Curriculum do Ensino Médio**, Secretaria de Educação, Pernambuco, 2021.

PINTO, M. G. M. o Enem e a Legislação Educacional Brasileira: avanços e recuos do exame na efetivação das orientações da ldb e dos pcns **Anais do VIII Colóquio Internacional Educação e Contemporaneidade**, Sergipe, 2014.

REYES, K. A. R. **Comprendión del muestreo por alumnos chilenos de educación secundaria**. 2019. 276f. TESE (Doutorado em Ciências da Educação) Universidade de Granada, Espanha, 2019.

REYES, K. A. R Ruz, F., Molina-Portillo, E. y Contreras, J. M. Comprendión del concepto de muestra por estudiantes chilenos de cuarto año medio de educación secundaria. En J. M. Contreras, M. M. Gea, M. M. López-Martín y E. Molina-Portillo (Eds.), **Actas del Tercer Congreso Internacional Virtual de Educación Estadística**, 2019.

SALOMÃO. J. F. **Conexões - Ciências Humanas e Sociais Aplicadas e Matemática**. 1º Ed. Moderna, São Paulo, 2021.

SÃO PAULO. **Caderno do Professor**. Secretaria de Educação do estado de São Paulo, vol 2. 2014

SANTOS, J. C. P. A. **Da Escola para o Mundo: Projetos Integradores**. 1º Ed. Ática. São Paulo, 2021.

SELKE, R. C.; HORONATO, A.; FUGITA, F. TONETT, M. A. GARCIA, C. M. **Dimensões**. 1º Ed. FTD, São Paulo, 2021.

SHAUGHNESSY, J.; ZECHMEISTER, E.; ZECHMEISTER, J. **Metodologia de pesquisa em psicologia**. 9. Ed. Porto Alegre: AMGH, 2012

SILVA, S. C.; PINTO, S. S. **Estatística**, volume II, Rio Grande: Ed. da FURG, 2021.

SILVA, A. R. O.; CARVALHO, J. I. F. **Probabilidade Subjetiva no Ensino Médio: uma proposta à luz do letramento probabilístico** IX CIBEM (Congresso Ibero Americano de Educação Matemática). São Paulo, 2022.

SILVEIRA, F. L. ; BARBOSA, M. C. B.; SILVA, R. Exame Nacional do Ensino Médio (ENEM): Uma análise crítica. **Revista Brasileira de Ensino de Física**. São Paulo, v. 37,n. 1, 1101,2015.

SMOLE, K. C. S.; DINIZ, M. I. S. V. **Ser Protagonista: Matemática e suas Tecnologias**. 1º Ed. Edições SM. São Paulo, 2021.

SOUZA, J. R. **Ver o mundo: projetos integradores** 1. ed. FTD , São Paulo, 2021.

SOARES, D. J. M. SOARES, T. E. A. e SANTOS, W. Análise da qualidade psicométrica da prova de matemática do Exame Nacional do Ensino Médio brasileiro de 2018. **Revista Actualidades Investigativas en Educación**, 21(1), 1-29, 2021.

SOUZA, L., SILVA, L. e PIRES, J. **O raciocínio de professores de matemática sobre amostragem e argumentos persuasivos nas redes sociais**. En A. Salcedo

y D. Díaz-Levicoy (Ed.), *Formación del Profesorado para Enseñar Estadística: Retos y Oportunidades* (pp. 349-366). Centro de Investigación en Educación Matemática y Estadística. Universidad Católica del Maule, 2022.

SOUZA, F. M.; VAZ, V. A. **Ser Protagonista Ciências Humanas e Sociais Aplicadas em Diálogo Com a Matemática.** 1º Ed. Edições SM. São Paulo, 2021

TAUBER, L. **La Construcción del Significado de La Distribución Normal a partir de Actividades de Análisis de Datos.** Tese de Doutorado. 276f Universidad de Servilla, Servilla, 2001.

TEIXEIRA. L. A. **Diálogo: Matemática e suas tecnologias.** 1º Ed. Moderna, São Paulo, 2021.

THOMAZ, D. **Do livro didático ao aluno: transposição didática na aula de matemática do ensino médio diurno e noturno.** 213f. [Dissertação de Mestrado em Educação Matemática], Universidade do Mato Grosso, 2013

TRIOLA, M. F. **Introdução à estatística:** atualização da tecnologia. (Tradução e revisão técnica: Ana Farias e Vera Flores). 10 ed. Rio de Janeiro: LTC. 2008.

WALPOLE, R. E.; MYERS R. H.; MYERS, S.L.; YE, K. **Probabilidade e Estatística para Engenharia e Ciências,** Pearson Universidades; 8ª edição, São Paulo, 2008

WATSON, J.; KELLY, B. A. Cognition and instruction: Reasoning about bias in sampling. **Mathematics Education Research Journal**, v. 17, n. 1, p. 25–57, 2005.

WONG A.; SIBBALD A.; FERRERO F.; PLAGER M.; SANTOLAYA M.E.; ESCOBAR A.M.; CAMPOS, S.; BARRAGÁN, S.; DE LEÓN GONZÁLEZ, M.; KESSELRING, G. L. **Antipyretic effects of dipyrone versus ibuprofen versus acetaminophen in children: results of a multinational, randomized, modified doubleblind study..** 40(6):313-24, 2001.

ZAMMATARO, A. F. D.; NETO, E. A. R. **Diálogo - Ciências Humanas e Sociais Aplicadas e Matemática.** 1º Ed. Moderna, São Paulo, 2021.