

---

**Universidade Federal de São Paulo**

Instituto de Ciência e Tecnologia

---



**Mestrado Profissional em Matemática  
em Rede Nacional - PROFMAT**

**Probabilidade e Estatística: uma proposta  
de abordagem gráfica utilizando o GeoGebra**

**Jayrton Silva Carvalho**

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Juliana Garcia Cespedes

São José dos Campos

Maio, 2021



**PROFMAT**

Título: *Probabilidade e Estatística: uma proposta de abordagem gráfica utilizando o GeoGebra*

Dissertação apresentada ao Instituto de Ciência e Tecnologia da UNIFESP, campus São José dos Campos/SP, como parte dos requisitos exigidos para a obtenção do título de Mestre pelo Programa de Mestrado Profissional em Matemática em Rede Nacional – PROFMAT.

**São José dos Campos**  
**Maio, 2021**

Carvalho, Jayrton

**Probabilidade e Estatística: uma proposta de abordagem gráfica utilizando o GeoGebra**, Jayrton Silva Carvalho – São José dos Campos, 2021.

XCII, 92f.

Dissertação (Mestrado) – Universidade Federal de São Paulo. Instituto de Ciência e Tecnologia. Programa de Pós-Graduação em Matemática em Rede Nacional (PROFMAT).

Probability and Statistics: a proposal for a graphical approach using GeoGebra

1. Probabilidade. 2. Estatística Descritiva. 3. Tecnologia da Informação e Comunicação (TIC). 4. Abordagem Gráfica. 5. GeoGebra.

UNIVERSIDADE FEDERAL DE SÃO PAULO  
INSTITUTO DE CIÊNCIA E TECNOLOGIA

Mestrado Profissional em Matemática em Rede Nacional  
PROFMAT

**Chefe de departamento:**

Prof. Dr. Marcelo Cristino Gama

**Coordenador do Programa de Pós-Graduação:**

Prof. Dr. Angelo Calil Bianchi

JAYRTON SILVA CARVALHO

PROBABILIDADE E ESTATÍSTICA: UMA PROPOSTA DE  
ABORDAGEM GRÁFICA UTILIZANDO O GEOGEBRA

A handwritten signature in black ink, appearing to read 'Cespedes', with a large, stylized initial 'C' on the left.

**Presidente da banca:** Prof<sup>ª</sup>. Dr<sup>ª</sup>. Juliana Garcia Cespedes

**Banca examinadora:**

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Roseli Aparecida Leandro

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Cláudia Aline Azevedo dos Santos Mesquita

Prof<sup>ª</sup>. Dr<sup>ª</sup>. Luzia Pedroso de Oliveira

**Data da Defesa:** 7 de maio de 2021

*"A Matemática é um conjunto de teorias gerais e abstratas que se aplicam em situações concretas."*

***Elon Lages Lima.***

## AGRADECIMENTOS

---

Quero eternizar neste documento a gratidão por minha mãe. Uma mulher batalhadora, destemida e exemplo de retidão. Professora respeitada e enfermeira dedicada. Recordo-me que não descansou para que seus filhos não sofressem dificuldades semelhantes às vividas por ela. Por esses e muitos outros motivos, eu escrevo em destaque que, **a senhora MARIA SOUSA SILVA é a base da minha vida e a quem eu agradeço pelo bom filho, esposo e pai que sou e pelo profissional que me tornei.**

Existem outras mulheres que sempre estiveram comigo neste caminho e às quais também registro o meu reconhecimento: minha esposa, WALDÊNIA SUZZY OLIVEIRA SILVA CARVALHO, e a nossa filha, ANNA LETÍCIA SILVA CARVALHO. Ressalto com orgulho que esta carregará para sempre um pedaço daquela, um gesto de dedicação que revelo apenas para mostrar a imensidão do amor materno e despertar entre nós a importância da doação de órgãos. Ademais, o incentivo, o apoio e a força que emanam dessas mulheres foram o que me fez prosseguir neste curso. Por tudo isso, não é exagero dizer que, mesmo que seus nomes não estejam impressos no diploma, elas estão mais presentes nesse documento do que eu.

Em especial à Prof<sup>a</sup> Dr<sup>a</sup> Juliana Garcia Cespedes, agradeço por me orientar sempre com paciência e dedicação, pelos conselhos e contribuições para a realização deste trabalho, pelo aprendizado que me proporcionou em Estatística. Externo e eternizo a minha gratidão ao mesmo tempo que reconheço a importância da sua participação nesta obra.

Sou grato às Prof<sup>a</sup> Dr<sup>a</sup> Roseli Aparecida Leandro, Cláudia Aline Azevedo dos Santos Mesquita, Luzia Pedroso de Oliveira e aos Prof Dr Gilmar Lopes Dias e Luís Felipe Bueno por terem aceitado o convite para compor a banca examinadora desta dissertação. As suas presenças muito me envaidece e enaltece a importância deste trabalho.

Ainda, às Prof<sup>a</sup> Dr<sup>a</sup> Daiana Oliveira dos Santos, Grasielle Cristiane Jorge e Kelly Cristina Jorge Sakamoto e aos Prof Dr Marcelo Cristino Gama, Pedro Levit Kaufmann e Luís Felipe Bueno. A capacidade, a dedicação, o respeito e a cordialidade com o estudante praticados por cada uma das senhoras e dos senhores os fazem, para mim, fontes de inspiração para que eu permaneça no caminho acadêmico.

Aos colegas da turma PROFMAT 2019: Geisa Feltrin Santana de Alvarenga, Jéssica Mara Campos Cunha Ferreira, Letícia Carolina Barbosa, Gabriel Silva Delgado, José Mar-

ques da Fonseca Júnior, Leonardo Romancini Leite, Marcio Oliveira de Moraes Júnior, Robert Henrique Gonçalves Silva, Vagner Silva Lima, é com muita satisfação que torno pública a minha gratidão a todos vocês, reconhecendo o apoio, o incentivo e o exemplo de bons profissionais e verdadeiros amigos que são. Que a amizade, o respeito e a admiração iniciadas no curso permaneçam por muitos anos. Não sumam!

Ao criador do céu e da terra, reverencio-o escrevendo aqui suas palavras: "deem graças em todas as circunstâncias, pois esta é a vontade de Deus". Portanto, dou graças a ti por esta conquista e peço que não permita que eu me afaste da tua luz.



## RESUMO

---

O discreto uso pedagógico do GeoGebra relacionado ao ensino de probabilidade e estatística descritiva, a escassez de material gratuito que oriente usuários acerca dos recursos do aplicativo e a necessidade docente de se apropriar de conhecimento sobre o programa motivaram a realização de uma pesquisa exploratória na bibliografia disponível, sobre como construir ferramentas pedagógicas digitais utilizando o GeoGebra, que resultou nesta dissertação, cujo objetivo é apresentar um material de apoio contendo instruções sobre o programa, com o propósito de contribuir para a construção de aplicativos que abordem graficamente assuntos de probabilidade e estatística descritiva. Apresenta-se uma Proposta Didática destinada a estudantes do 3º ano do Ensino Médio, contendo uma abordagem significativa, que possibilita a compreensão de conceitos estatísticos, o desenvolvimento do senso crítico, o aprendizado ativo e o conhecimento prático, a partir de uma atividade cooperativa com utilização de tecnologia. Espera-se que este trabalho ofereça formação complementar a estudantes, professoras e professores, com reflexos no desenvolvimento do aprendizado e no trabalho docente realizado em sala de aula.

**Palavras-chave:** 1. Probabilidade. 2. Estatística Descritiva. 3. Tecnologia da Informação e Comunicação (TIC). 4. Abordagem Gráfica. 5. GeoGebra.

## ABSTRACT

---

The little pedagogical use of GeoGebra related to the teaching of probability and descriptive statistics, the low supply of free material that guides users of the software about the application's features and the need for teachers to learn about the program motivated the realization of exploratory research in existing bibliography, indirectly, about how to build digital pedagogical tools using GeoGebra, which resulted in this masters dissertation, whose objective is to present a support material containing instructions about the program, with the purpose of contributing to the construction of pedagogical support applications, with a graphical approach to probability and descriptive statistics for high school. A Didactic Proposal for students of the 3rd year of High School is presented, containing a significant approach, which enables the understanding of statistical concepts, the development of critical sense, active learning and practical knowledge, based on a cooperative activity with the use of technology. It is expected that this paper will offer complementary training to students and teachers, with repercussions on the development of learning and the teaching work carried out in the classroom.

**Keywords:** 1. Probability. 2. Descriptive Statistic. 3. Information and Communication Technology (ICT). 4. Graphical Approach. 5. GeoGebra.

## LISTA DE FIGURAS

---

Figura 1	Tipos de variáveis.	16
Figura 2	Representação gráfica dos conceitos de população e amostra.	16
Figura 3	Gráfico de dispersão unidimensional da variável moradores, da amostra da Escola SP.	26
Figura 4	Gráfico de barras verticais da variável moradores, da amostra da Escola SP.	27
Figura 5	Histograma dos dados da variável altura dos elementos da amostra da Escola SP.	28
Figura 6	<i>Boxplot</i> dos dados da variável altura somente dos alunos da amostra da Escola SP.	28
Figura 7	Identificação de associação entre as variáveis altura e sexo da amostra de alunos da Escola SP, utilizando <i>boxplot</i>	31
Figura 8	<i>Link</i> para <i>download</i> do GeoGebra Clássico 5.	32
Figura 9	Janelas de diálogo de instalação.	33
Figura 10	Interface padrão do programa GeoGebra.	34
Figura 11	Ícones da barra de <i>menus</i> .	35
Figura 12	Ferramentas agrupadas em uma caixa de ferramentas.	35
Figura 13	Janela de álgebra.	36
Figura 14	Janela de visualização.	37
Figura 15	Janela planilha.	38
Figura 16	Menu ferramentas.	38
Figura 17	Janela de configuração da ferramenta controle deslizante.	39
Figura 18	Janela de configuração da ferramenta texto.	40
Figura 19	Janela de configuração da ferramenta caixa para exibir/esconder objetos.	40
Figura 20	Janela de configuração da ferramenta campo de entrada.	41
Figura 21	Gráfico de dispersão unidimensional da variável moradores, da amostra da Escola SP.	55
Figura 22	Histograma dos dados da variável altura dos alunos e alunas da amostra da Escola SP.	57
Figura 23	<i>Boxplot</i> dos dados da variável altura apenas dos alunos da amostra da Escola SP.	58
Figura 24	Apresentação da ferramenta de apoio pedagógico para análise de conjunto de dados.	62
Figura 25	Representações gráficas exibidas pela ferramenta de apoio pedagógico para análise de conjunto de dados.	63

- Figura 26 Ilustração dos intervalos de classificação da massa do indivíduo, segundo o índice de massa corporal. 71
- Figura 27 *Boxplot* da variável IMC, segundo a variável conceito. 72

## LISTA DE TABELAS

---

Tabela 1	Trabalhos que exemplificam o uso do GeoGebra, publicados entre 2013 e 2018, disponíveis na Plataforma Sucupira.	12
Tabela 2	Artigo disponível no repositório de dissertações do PROFMAT, publicado em 2020, que exemplica o uso de comandos do GeoGebra.	13
Tabela 3	Artigo com aplicações do GeoGebra a estatística descritiva, publicado em 2012 e disponível no site do Instituto Politécnico de Leiria.	13
Tabela 4	Informações coletadas de uma amostra de 30 alunos e alunas sobre sexo, altura, massa, se residem no mesmo bairro onde se localiza a escola, a quantidade de moradores na mesma residência e opinião sobre o lanche servido na cantina da Escola SP.	14
Tabela 5	Tabela primitiva dos dados da variável massa dos alunos e alunas da amostra da Escola SP.	17
Tabela 6	Rol dos dados da variável massa, da amostra de estudantes da Escola SP.	17
Tabela 7	Distribuição de frequência dos dados da variável moradores, da amostra de estudantes da Escola SP.	18
Tabela 8	Distribuição de frequência em intervalos de classe dos dados da variável altura dos estudantes da amostra da Escola SP.	18
Tabela 9	Distribuição dos dados da variável altura dos estudantes da amostra da Escola SP utilizando as frequências absoluta, relativa, acumulada e acumulada relativa.	22
Tabela 10	Distribuição de frequência dos dados da variável altura dos estudantes da amostra da Escola SP, segundo os atributos da variável sexo.	30
Tabela 11	Distribuição das proporções dos dados da variável altura dos estudantes da amostra da Escola SP, segundo os atributos da variável sexo.	30
Tabela 12	Medidas resumo da variável altura dos alunos e alunas da amostra da Escola SP, segundo a variável sexo.	31
Tabela 13	Distribuição de frequência acumulada dos dados da variável moradores, da amostra da Escola SP.	53
Tabela 14	Distribuição de frequência acumulada em intervalos de classe dos dados da variável altura, da amostra de alunos e alunas da Escola SP.	54

- Tabela 15      Dados coletados de uma amostra de 30 alunos e alunas da escola pública de tempo integral pesquisada sobre sexo, altura, massa e conceito dado ao almoço servido na cantina.      70
- Tabela 16      Medidas resumo para a variável IMC dos alunos e alunas da amostra, segundo a variável conceito, na escola pesquisada.      71
- Tabela 17      Medidas resumo para a variável IMC dos alunos e alunas da amostra, segundo a variável sexo, na escola pesquisada.      73

## SUMÁRIO

---

1	INTRODUÇÃO	3
2	REVISÃO DE LITERATURA	7
3	PROBABILIDADE E ESTATÍSTICA DESCRITIVA	14
3.1	Variável	15
3.2	População e Amostra	16
3.3	Tabela Primitiva e Rol	17
3.4	Distribuição de Frequência	17
3.5	Classe	18
3.6	Amplitudes	19
3.7	Tipos de Frequência	21
3.8	Medidas de Posição	22
3.9	Medidas de Dispersão	24
3.10	Representação Gráfica de uma Distribuição	26
3.11	Simetria	28
3.12	Associação entre Variáveis	29
4	O GEOGEBRA	32
4.1	Apresentação	32
4.2	Download e Instalação	32
4.3	Interface	33
4.4	Ferramentas	38
4.5	Comandos Internos	41
4.6	Representação Tabular de uma Distribuição de Dados	52
4.7	Representação Gráfica de uma Distribuição	54
5	PROPOSTA DIDÁTICA	59
5.1	Justificativa	59
5.2	Público	59
5.3	Problema	59
5.4	Objetivo Geral	60
5.5	Objetivos Específicos	60
5.6	Metodologia	61
5.7	Recursos Didáticos	62
5.8	Ferramenta de Apoio	62
5.9	Protocolo de Construção	63

5.10 Atividade	70
5.11 Avaliação	73
6 CONSIDERAÇÕES FINAIS	75
REFERÊNCIAS BIBLIOGRÁFICAS	77



## INTRODUÇÃO

---

Há relatos do uso da estatística pelos hebreus para recenseamento demográfico: "Deus disse a Davi: vá e faça a contagem do povo de Israel e de Judá." [3]. Contudo, foi apenas no século XVI que um italiano da província de Pavia, médico, astrólogo, professor e jogador compulsivo, escreveu um tratado com 32 capítulos curtos, intitulado *Liber de Ludo Aleae* (Livro de Jogos de Azar), o primeiro na história a tratar sobre a aleatoriedade e considerado o alicerce para a compreensão da natureza da incerteza [19]. Nele, Gerolamo Cardano introduziu uma regra geral, que mais tarde seria aperfeiçoada e se tornaria a Lei do Espaço Amostral. O livro de jogos de azar foi publicado após a morte de Cardano, em 1663.

A estatística faz parte do cotidiano, ela está nos jornais, nas revistas, nas tabelas de esportes, nos jogos da loteria, no cálculo do valor de seguros, nos exames de laboratório, nas pesquisas e nos diversos programas de televisão. Sites de notícias e publicações nas redes sociais abusam do seu uso, às vezes a utilizam de maneira totalmente equivocada. Se uma informação não for corroborada por números, parece não ter veracidade. Por isso, algumas vezes, vê-se estatísticas enviesadas, que conduzem a conclusões equivocadas ou falsas. A estatística possibilita um entendimento sobre o mundo que não se poderia tê-lo utilizando outro método, ela dá sentido a dados desordenados e auxilia na tomada de decisão com base na análise de fatos.

Além disso, a estatística é definida como uma ciência que se atém ao planejamento e execução da coleta de dados, sua organização, análise e interpretação com o objetivo de inferir sobre a população da qual os dados foram extraídos [20]. Ou ainda, é função-chave da estatística usar os dados que se tem para fazer conjecturas sobre perguntas para as quais não se tem informação completa; em resumo, é usar dados conhecidos para fazer inferência sobre algo desconhecido [23].

A estatística possui dois ramos: o inferencial e o descritivo. A estatística inferencial utiliza-se da probabilidade como ferramenta básica para, partindo de uma amostra, se chegar a conclusões sobre uma população, da qual retirou-se a amostra analisada. Enquanto que a estatística descritiva vale-se de técnicas, tabelas e gráficos para organizar, selecionar, resumir e representar dados coletados de uma população [16].

Nesta pesquisa, fez-se uma leitura de documentos normativos da educação brasileira e verificou-se existir previsão de se utilizar programas de matemática dinâmica para a conquista de objetivos e competências descritas nessas normas, exigidas para estudantes do Ensino Médio.

Com isso, buscou-se conhecer sobre o uso de tecnologias no ensino de matemática e encontrou-se publicações de bons trabalhos e orientações para a sua utilização em sala de aula.

Investigou-se a literatura sobre o uso de aplicativos no ensino de probabilidade e estatística e encontrou-se diretrizes internacionais e trabalhos de eminentes pesquisadores brasileiros sobre o Ensino de Estatística (EE) que sugerem a adoção de aplicativos didáticos e recomendam utilizar o GeoGebra em detrimento de programas mais técnicos de análise estatística.

Apesar da importância, da boa avaliação e da viabilidade do uso do GeoGebra para o ensino de probabilidade e estatística, existem poucas instruções sobre como construir objetos pedagógicos utilizando o aplicativo.

Todavia, professores necessitam buscar formação, conhecimento e treinamento, sobretudo na rede mundial de computadores, com o objetivo de compreender os instrumentos pedagógicos e se adaptar ao seu manuseio à medida que forem explorados [21].

O modesto uso pedagógico do GeoGebra vinculado ao estudo de probabilidade e estatística descritiva, a escassez de material detalhado e gratuito que oriente usuários do programa para a utilização do aplicativo relacionado-o a esses assuntos, e a necessidade docente de se apropriar de conhecimento sobre essa ferramenta motivaram a pesquisa e a redação desta dissertação.

O trabalho está organizado da seguinte forma: no capítulo 2 encontra-se um panorama sobre o assunto abordado no corpo desta dissertação, desde documentos que regulamentam a educação no Brasil às pesquisas e publicações consultadas acerca do ensino de estatística. A seguir, no capítulo 3, apresenta-se um resumo da teoria estatística, iniciando pelo conceito de variável, passando pelas medidas resumo, simetria, análise entre duas variáveis, representação tabular e representação gráfica do conjunto de dados. Busca-se detalhar o uso do programa GeoGebra, na versão Clássico 5, com aplicação no ensino de probabilidade e estatística descritiva, no capítulo 4. A Proposta Didática, elaborada para alunos do 3º ano do Ensino Médio, é apresentada no capítulo 5, onde expõe-se uma ferramenta digital de apoio pedagógico para análise didática de conjuntos de dados e ensino de probabilidade e estatística descritiva, disponível gratuitamente na *internet*, construída utilizando instruções contidas no corpo desta dissertação.

## DEFINIÇÃO DO PROBLEMA

Observando a previsão do uso de programas que inter-relacionem estatística, geometria e álgebra para a consecução de objetivos e habilidades contidos em documentos normativos do Ensino Médio da educação brasileira, realizou-se uma pesquisa exploratória em bibliografias existentes na rede mundial de computadores, sobre como construir objetos pedagógicos no aplicativo GeoGebra, utilizando o manual oficial *on-line* do pro-

grama [14], na versão GeoGebra Clássico 5 para Windows, em busca de respostas ao questionamento:

Como utilizar as funcionalidades do GeoGebra para a criação de ferramentas digitais de apoio ao estudo, a análise, ao ensino e aprendizagem de probabilidade e estatística descritiva?

A escolha do GeoGebra para a realização deste trabalho deve-se a motivos que serão descritos no capítulo 2; no entanto, ratifica-se que:

Independentemente das ferramentas usadas, é essencial ver a utilização da tecnologia não unicamente como uma maneira de produzir uma solução estatística, mas como uma forma de permitir que o aluno possa investigar ideias conceituais, focar a aprendizagem da estatística sobre diferentes áreas de aplicação e, assim, perceber sua utilidade. ([17], p. 6)

Além disso, os conhecimentos do docente e dos estudantes devem ser considerados no momento da escolha do programa para evitar que o seu uso torne-se fator de desmotivação e impeça de se atingir metas e objetivos [2].

#### OBJETIVO GERAL

O propósito deste trabalho é apresentar um material que contenha instruções sobre o GeoGebra, com a finalidade de contribuir para a construção de ferramentas digitais de apoio pedagógico, que abordem graficamente assuntos de probabilidade e estatística descritiva, a fim de se alcançar as habilidades descritas nas normas da educação brasileira.

#### OBJETIVOS ESPECÍFICOS

A fim de contribuir com professores, inclusive usuários incipientes, para o uso ativo e o desenvolvimento de produtos educacionais utilizando o GeoGebra, com foco no ensino e aprendizagem de probabilidade e estatística descritiva, têm-se os seguintes objetivos específicos:

1. Descrever como realizar o *download* e a instalação do aplicativo no sistema operacional Windows;
2. Explicar o *layout* do programa, a sua configuração e funcionamento em geral;
3. Apresentar as funções e a utilização das Barra de Menus, Barra de Ferramentas e Barra de Entrada, das Janela de Visualização, Janela de Álgebra e Janela Planilha;
4. Apresentar, detalhar e exemplificar as configurações dos diversos comandos, limitando-se aos relacionados à probabilidade e estatística descritiva, suas respectivas sintaxes, além das instruções que resultam nas várias formas de representação de um conjunto de dados por intermédio de tabelas e gráficos, no GeoGebra;

5. Descrever, demonstrar com ilustrações e disponibilizar ferramenta desenvolvida no programa, que possa ser aplicada em propostas didáticas utilizando uma abordagem gráfica de assuntos de probabilidade e estatística descritiva, com o GeoGebra.

Com a finalidade de facilitar a comunicação e o entendimento da redação desta dissertação, as palavras programa, ferramenta e aplicativo referem-se ao GeoGebra. O professor também será citado como analista, pesquisador e usuário do aplicativo. Enquanto que, as formas de se utilizar cada comando do GeoGebra, descritas no capítulo 4, serão tratadas por sintaxe, instrução, ordem ou estrutura.

□

## REVISÃO DE LITERATURA

---

### PANORAMA

Com a intenção de realizar uma pesquisa e produzir um trabalho relacionado ao ensino de probabilidade e estatística que pudesse ter impacto na prática pedagógica e reflexos positivos em sala de aula, trazendo complementação e aprofundamento do ensino e aprendizado desses assuntos, iniciou-se a busca com uma leitura dos Parâmetros Curriculares Nacionais (PCN) [8] e verificou-se existir previsão de se utilizar programas de matemática dinâmica para a conquista de objetivos e o desenvolvimento de competências descritas nessa norma. Observou-se nos PCN, na parte dedicada às Ciências da Natureza, Matemática e suas Tecnologias, proposta para o Ensino Médio, a indicação de objetivos formativos que compreendem o desenvolvimento do conhecimento prático, alinhados às necessidades da vida contemporânea com o propósito de dar significado aos conteúdos abordados.

Alguns dos objetivos formativos, listados no corpo dos PCN, excitaram o interesse do autor e indicaram um rumo a seguir nesta pesquisa, são eles:

1. Interpretar e utilizar diferentes formas de representação (tabelas, gráficos, expressões, ícones...);
2. Analisar qualitativamente dados quantitativos representados gráfica ou algebricamente relacionados a contextos sócio-econômicos, científicos ou cotidianos.
3. Compreender o caráter aleatório e não determinístico dos fenômenos naturais e sociais e utilizar instrumentos adequados para medidas, determinação de amostras e cálculo de probabilidades. ([8], p. 12)

A seguir, passou-se a examinar a Base Nacional Comum Curricular (BNCC) [7] em busca de orientações que complementassem os objetivos relacionados, que se tornaram as diretrizes iniciais para a realização deste trabalho. A BNCC especifica competências gerais e por áreas, que são definidas como uma mobilização de conceitos, procedimentos, práticas, atitudes e valores para a solução de problemas do cotidiano, do exercício da cidadania e do mundo do trabalho, que devem ser desenvolvidas pelos estudantes nas diversas fases da educação [7]. Essa norma declara que estudar estatística no Ensino Médio contribui para uma formação científica sólida dos estudantes, que se tornarão cidadãos críticos e reflexivos capazes de identificar informações equivocadas, obtidas pelo uso inadequado do processo de amostragem, divulgadas nos diversos meios de comunicação.

Verificou-se as Competências Gerais para a Educação Básica e encontrou-se alinhamento entre o procedimento transcrito a seguir e as diretrizes iniciais, ao se estabelecer que a representação de tabelas e gráficos, a determinação de amostras e o cálculo de probabilidades poderiam ocorrer por meio de tecnologia digital, a fim de produzir conhecimento e resolver problemas do cotidiano do estudante:

Compreender, utilizar e **criar tecnologias digitais** de informação e comunicação de forma crítica, significativa, reflexiva e ética nas diversas práticas sociais (incluindo as escolares) **para** se comunicar, acessar e disseminar informações, **produzir conhecimento, resolver problemas** e exercer protagonismo e autoria na vida pessoal e coletiva. ([7], p. 9, grifo do autor)

Assim, às diretrizes iniciais para a realização deste trabalho juntou-se a ideia de criar tecnologias digitais que, utilizadas em sala de aula, possibilitem a solução de problemas, a produção de conhecimento e a sua comunicação. Porém, a conexão com o estudo de probabilidade e estatística ainda se mostra tênue.

A área de Matemática e suas Tecnologias deve garantir que estudantes do Ensino Médio desenvolvam competências específicas definidas e, para isso, indica habilidades que deverão ser alcançadas pelos alunos nessa etapa educacional [7]. Passou-se à análise das Competências Específicas de Matemática e suas Tecnologias para o Ensino Médio e das suas respectivas Habilidades.

Os conceitos e aplicações da probabilidade e as ferramentas da estatística são instrumentos utilizados em pesquisas e estudos não apenas na área das Ciências Naturais e Matemática como também o são na área das Ciências Humanas, e isso está descrito com clareza na Competência Específica 2, que diz:

Articular conhecimentos matemáticos ao propor e/ou participar de ações para investigar desafios do mundo contemporâneo e tomar decisões éticas e socialmente responsáveis, com base na análise de problemas de urgência social, como os voltados a situações de saúde, sustentabilidade, das implicações da tecnologia no mundo do trabalho, entre outros, recorrendo a conceitos, procedimentos e linguagens próprios da Matemática. ([7], p. 526)

Para o desenvolvimento dessa competência, presupõe-se que a seguinte habilidade deve ser alcançada por estudantes do Ensino Médio [7]:

(EM13MAT202) Planejar e executar pesquisa amostral usando dados coletados ou de diferentes fontes sobre questões relevantes atuais, incluindo ou não, **apoio de recursos tecnológicos**, e comunicar os resultados por meio de relatório contendo **gráficos** e interpretação das **medidas de tendência central e das de dispersão**. ([7], p. 526, grifo do autor)

A Competência Específica 3 prevê que o estudante desenvolva argumentação consistente a partir da construção de modelos para resolução de problemas em diversos contextos, a interpretação dos resultados obtidos a partir de estratégias, conceitos e procedimentos matemáticos. E transcreve algumas das habilidades a serviço dessa competência, correlatas às intenções deste trabalho, que se juntam às suas diretrizes iniciais:

(EM13MAT310) Resolver e elaborar problemas de contagem envolvendo diferentes tipos de agrupamento de elementos, por meio dos princípios multiplicativo e aditivo, recorrendo a estratégias diversas como o diagrama de árvore.

(EM13MAT311) Resolver e elaborar problemas que envolvem o cálculo da probabilidade de eventos aleatórios, identificando e descrevendo o espaço amostral e realizando contagem das possibilidades.

(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). ([7], p. 529)

A Competência Específica 4 propõe as seguintes habilidades, entre outras, para que estudantes construam e desenvolvam o raciocínio matemático, a partir da compreensão e utilização de diferentes tipos de representação matemática:

(EM13MAT408) 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.

(EM13MAT409) Interpretar e comparar conjuntos de dados estatísticos por meio de diferentes **diagramas e gráficos, como o histograma, o de caixa (box-plot)**, o de ramos e folhas, reconhecendo os mais eficientes para sua análise. ([7], p. 531, grifo do autor)

Destaca-se que essas habilidades citam o estudo de estatística com apoio de Tecnologias da Informação e Comunicação (TIC) e sugerem o uso de programas que inter-relacionem estatística, geometria e álgebra.

O uso de programas, recursos tecnológicos ou tecnologias digitais são termos afetos às TIC e, portanto, buscou-se conhecer sobre a utilização das TIC no ensino de matemática e encontrou-se descrições de bons resultados e orientações para a sua utilização em sala de aula.

O uso das TIC no ensino é uma forma de inclusão digital e não significa que o docente deva abandonar as antigas metodologias, mas sim incorporar novas ferramentas para facilitar e enriquecer o aprendizado, visto que, o perfil dos alunos e as demandas do mercado de trabalho estão em constante mudança [21].

As tecnologias invadiram as escolas e salas de aula há muito tempo, desde o uso pedagógico do projetor multimídia às atuais lousas digitais, passando pelo antigo videocassete, calculadoras, retroprojetor com tubo de luz, televisão e programas de TV (Telecurso 2000), computador de mesa e *notebook*, mesa digitalizadora, projetor, internet, vídeos (canais YouTube), blogs, laboratório de informática, *tablets*, *smartphones*, óculos de realidade virtual, entre muitos outros; tecnologias que estão (ou estiveram) à disposição dos professores como ferramentas do processo de ensino e que são formas de atrair os alunos

ao aprendizado, mas que é necessário utilizá-las de maneira adequada, de forma que façam diferença no tratamento da informação, sem banalização [21].

A utilização das tecnologias nas aulas de Matemática pode promover mudanças na dinâmica da sala de aula e também nas formas de ensinar e de aprender os conteúdos. Para tanto, os professores precisam compreender e ter clareza das possibilidades e também dos limites das tecnologias. ([9], p. 101)

Os educadores devem adotar o uso de computadores e programas, além de outros recursos tecnológicos, que devem ser acompanhados de reformulação e modernização dos conteúdos, pois de outra maneira não se aproximarão da realidade dos estudantes [11].

Somente a introdução de computadores nas escolas não provoca mudanças nas práticas docentes e no processo de ensino, o professor precisa participar de forma ativa do processo de conhecimento do aluno, como mediador, orientador da aprendizagem [9].

O aprendizado de alunos e professores se completam, ou se complementam, pois "quem ensina aprende ao ensinar e quem aprende ensina ao aprender", esse processo deve sofrer contínuo aperfeiçoamento por meio de pesquisas [13]. Portanto, é necessário que professores busquem formação contínua, treinamentos para conhecer e utilizar novas TIC [21].

Dessa forma, fortaleceu-se a ideia de desenvolver uma pesquisa que, além de oferecer suporte à formação do professor sobre o uso de tecnologias digitais, possa influenciar no trabalho docente em sala de aula, abordar o ensino de probabilidade e estatística no Ensino Médio com a utilização ou criação de programas, tendo como propósito produzir conhecimento, solucionar problemas da realidade do estudante, desenvolver argumentação consistente e formação cidadã crítica.

A *American Statistical Association* (ASA), entidade internacional que reúne profissionais, professores e pesquisadores em estatística, publicou em 2005 um relatório intitulado *Guidelines for Assessment and Instruction in Statistics Education* (GAISE) [1] citando recomendações para aumentar a eficácia no ensino de estatística e contendo orientações sobre ensino e avaliação na educação estatística, revisadas e ratificadas no relatório expedido em 2016 [2], enumeradas a seguir:

1. Ensinar o pensamento estatístico.
2. Concentrar-se na compreensão conceitual.
3. Integrar dados reais com um contexto e um propósito.
4. Promover o aprendizado ativo.
5. Usar a tecnologia para explorar conceitos e analisar dados.
6. Usar as avaliações para melhorar o ensino e avaliar o aprendizado do aluno.



Com o objetivo de desenvolver o letramento, o raciocínio e o pensamento estatísticos, o ensino de estatística mudou o foco de o que, por que, a quem e quando ensinar e passou a desenvolver excelentes atividades dedicadas a como ensinar, desenvolvidas a partir das sugestões citadas [1];

Destaca-se a orientação para o uso de tecnologia, referida no relatório como uma diversidade de *hardwares* e *softwares*, a fim de explorar conceitos e analisar dados, permitindo que os alunos façam análises com mais facilidade, se concentrem na interpretação dos resultados, dos conceitos e desenvolvam o pensamento estatístico.

Recomenda-se que professores utilizem tecnologia para realizar cálculos que permitam que os alunos deem ênfase nas interpretações dos números, analisem conjuntos de dados reais e estimulem o aprendizado do estudante com perguntas do tipo "o que acontece se"[2].

Além disso, ao se escolher ferramentas tecnológicas para utilização no ensino de estatística, deve-se priorizar o uso dos dispositivos interativos, que apresentem ligação dinâmica entre análise gráfica e numérica do conjunto de dados, seja funcional em diversas plataformas e haja suporte para uso em *tablets* e *smartphones* [2].

Ao se fazer uma escolha por uma determinada tecnologia de desenvolvimento, questões como o tempo, a facilidade de atualização, a portabilidade, a curva de aprendizagem e o poder de expressão das ferramentas de programação devem ser consideradas. ([5], p. 1)

Contudo, é importante salientar que:

Essas formas de utilização do computador podem ser focadas pelo professor, sendo que cada uma delas tem seu limite e sua potencialidade, sendo essencial que o docente tenha claro seu objetivo, tenha conhecimentos técnicos profundos do *software* utilizado, conheça seus limites e potencialidades, planeje com muito cuidado as atividades a serem desenvolvidas, tente prever algumas dificuldades dos alunos e tenha compreensão das possibilidades de abordar aquele conteúdo matemático. ([9], p. 106)

Destaca-se a seguinte ponderação sobre a escolha da ferramenta tecnológica para uso didático: "Na adoção de *softwares* para atividades acadêmicas, é relevante considerar que a maioria deles foi desenvolvida com o propósito de 'fazer' estatística e não de 'aprender' estatística"([17], p. 4).

O aplicativo GeoGebra foi desenvolvido na dissertação de mestrado do austríaco Markus Honhenwarter para uso acadêmico por estudantes e professores; é um programa de matemática de código aberto; está à disposição de usuários não comerciais; é gratuito, interativo e dinâmico, permite a manipulação direta dos dados com resposta instantânea das medidas do conjunto; integra álgebra, cálculo e geometria em diferentes janelas; está disponível para os sistemas iOS, Android, Chromebook, Windows, Mac OS e Linux, além de possuir uma versão de realidade aumentada para *tablets* e *smartphones*, um fórum para discussões e uma rede social onde milhares de usuários espalhados pelo mundo compartilham seus trabalhos desenvolvidos neste aplicativo [14].

Solicita-se considerar o uso do GeoGebra nas disciplinas de estatística preferencialmente a outros *softwares* mais técnicos de estatística [6].

Buscou-se publicações acadêmicas sob a ótica instrucional do uso do aplicativo GeoGebra para o desenvolvimento de objetos pedagógicos, para uso em sala de aula, direcionados ao ensino de probabilidade e estatística descritiva.

É comum o uso do GeoGebra para o estudo de funções e geometria no ambiente escolar, de acordo com levantamento realizado no banco de teses e dissertações da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES); consultados os documentos disponíveis na Plataforma Sucupira de 2013 a 2018, encontrou-se 1030 trabalhos que relacionavam o programa ao estudo de matemática; desses, apenas 16 associavam o aplicativo ao ensino e aprendizagem de probabilidade e estatística [17].

Passou-se a realizar uma leitura de cada um desses 16 artigos com o escopo de verificar quantos deles exemplificavam formas de uso dos comandos do GeoGebra para a construção de objetos pedagógicos destinados ao estudo dos temas desta dissertação. Verificou-se que somente três deles apresentavam como utilizar o GeoGebra e, de maneira superficial, tais artigos continham exemplos relacionados com os próprios objetos de estudo. Esses trabalhos encontram-se descritos na Tabela 1.

Tabela 1: Trabalhos que exemplificam o uso do GeoGebra, publicados entre 2013 e 2018, disponíveis na Plataforma Sucupira.

<b>Autor</b>	<b>Instituição</b>	<b>Título</b>	<b>Nível</b>	<b>Ano</b>
Medeiros, Sérgio da Silva	Universidade Federal do ABC - PROFMAT	Cadeias de Markov ocultas.	Mestrado Profissional	2017
Gonçalves, Paulo Henrique Rodrigues	Universidade Federal de Goiás - PROFMAT	Uma abordagem da Distribuição Normal através da resolução de uma situação problema com a utilização do software GeoGebra	Mestrado Profissional	2014
Duarte, Lucas Rodrigues	Pontifícia Universidade Católica de Minas Gerais	A utilização do software GeoGebra no ensino da distribuição normal de probabilidade: uma aproximação entre a geometria dinâmica e a educação estatística	Mestrado Acadêmico	2013

Fonte: Próprio autor

Decidiu-se verificar a existência de possíveis trabalhos, que relacionem o aplicativo GeoGebra ao estudo de probabilidade e estatística descritiva, publicados após 2018. Primeiramente, foi realizada uma pesquisa no repositório de dissertações do Programa de Mestrado Profissional em Matemática em Rede Nacional (PROFMAT); nessa base de dados encontrou-se 105 registros de dissertações que continham a palavra "Estatística" no

seu título, outras 139 dissertações incluíam no título a palavra "Probabilidade"; refinou-se a busca sobre esse resultado para atingir somente as que continham, além das palavras pesquisadas anteriormente, a palavra "GeoGebra", restando apenas seis publicações. Dessas, cinco haviam sido analisadas e descritas no levantamento realizado no banco de teses da CAPES e somente uma fora publicada após 2018, trata-se do trabalho indicado na Tabela 2 que, apesar de no seu título citar probabilidade e GeoGebra, o autor apresenta no seu conteúdo comandos e procedimentos no GeoGebra para a construção geométrica de um cubo, não se referindo, portanto, a probabilidade e estatística descritiva:

Tabela 2: Artigo disponível no repositório de dissertações do PROFMAT, publicado em 2020, que exemplica o uso de comandos do GeoGebra.

<b>Autor</b>	<b>Instituição</b>	<b>Título</b>	<b>Nível</b>	<b>Ano</b>
Moura, Tarcísio Junqueira	Universidade Federal de Goiás - PROFMAT	Probabilidade e Jogos Digitais: uma experiência com o <i>software</i> GeoGebra no ensino médio	Mestrado Profissional	2020

Fonte: Próprio autor

Em seguida, buscou-se no *site* Google Acadêmico (ferramenta de busca do Google, que permite pesquisar trabalhos acadêmicos, literatura escolar, jornais de universidades e artigos) outros trabalhos com finalidades correlatas à proposta de pesquisa desta dissertação, utilizando para isso as palavras "Estatística", "Probabilidade" e "GeoGebra". Encontrou-se o seguinte artigo, cujo propósito é apresentar o Geogebra para estudo de estatística descritiva, sem qualquer menção a comandos relacionados a probabilidade.

Tabela 3: Artigo com aplicações do GeoGebra a estatística descritiva, publicado em 2012 e disponível no site do Instituto Politécnico de Leiria.

<b>Autor</b>	<b>Instituição</b>	<b>Título</b>	<b>Nível</b>	<b>Ano</b>
Martins, Maria Alice da Silva	Instituto Politécnico de Leiria - Portugal	Estatística no Ensino Básico e Secundário	Mestrado Acadêmico	2012

Fonte: Próprio autor

□

PROBABILIDADE E ESTATÍSTICA DESCRITIVA

---

Este capítulo contém revisão de literatura dos assuntos de probabilidade e estatística descritiva que serão abordados na seção sobre o aplicativo GeoGebra e utilizados na Proposta Didática.

Tabela 4: Informações coletadas de uma amostra de 30 alunos e alunas sobre sexo, altura, massa, se residem no mesmo bairro onde se localiza a escola, a quantidade de moradores na mesma residência e opinião sobre o lanche servido na cantina da Escola SP.

Nº	Sexo	Altura	Massa	Bairro	Moradores	Opinião
1	F	1,59	56,6	Sim	3	B
2	F	1,61	63,8	Sim	5	B
3	F	1,61	70,2	Sim	4	O
4	F	1,62	60,5	Não	3	B
5	M	1,66	72,0	Sim	5	R
6	F	1,67	64,8	Não	2	R
7	F	1,67	68,1	Não	4	O
8	F	1,67	72,1	Sim	6	B
9	F	1,67	74,5	Sim	2	O
10	M	1,68	66,6	Não	2	R
11	F	1,69	70,0	Sim	5	B
12	F	1,70	80,7	Não	3	R
13	M	1,72	80,3	Sim	3	B
14	F	1,73	70,3	Não	3	O
15	M	1,73	88,8	Não	4	B
16	M	1,73	92,3	Sim	2	B
17	M	1,74	76,5	Não	4	B
18	M	1,74	90,0	Não	3	O
19	M	1,74	100,6	Sim	3	B
20	M	1,75	63,0	Sim	3	B
21	M	1,75	70,3	Não	4	O
22	M	1,75	80,7	Sim	3	R
23	M	1,75	98,9	Não	4	O
24	F	1,76	84,0	Sim	3	O
25	M	1,76	92,6	Sim	4	O
26	M	1,78	88,6	Não	4	R
27	M	1,80	84,1	Sim	5	O
28	M	1,81	88,2	Não	2	O
29	M	1,82	100,2	Não	6	B
30	M	1,88	106,3	Não	3	B

Fonte: Próprio autor

Para ilustrar os conceitos apresentados, optou-se por definir a seguinte situação problema, cujos dados encontram-se descritos na Tabela 4. Selecionou-se uma amostra de 30 estudantes de uma escola pública fictícia, denominada Escola SP, e perguntou-se o sexo (M - masculino ou F - feminino), a altura (em metro), a massa (em quilograma), se reside no mesmo bairro onde se localiza a escola (Sim ou Não), a quantidade de pessoas que moram na mesma residência que o aluno e uma opinião sobre o lanche servido na cantina da escola (O - ótimo, B - bom, R - ruim).

### 3.1 VARIÁVEL

Define-se variável como a coleção dos resultados possíveis de um fenômeno [10]. Ou, trata-se de uma ou mais características de interesse à pesquisa que estão presentes nos elementos da amostra [4].

Observa-se na amostra dos estudantes da Escola SP que existem as variáveis sexo, altura, massa, bairro, moradores e opinião. O conjunto de respostas referente a uma variável é chamado de dados [22].

#### *Tipos de variáveis*

O conjunto de dados das variáveis sexo, bairro e opinião é formado, neste caso, por palavras que indicam uma qualidade dos indivíduos consultados; são dados literais, categóricos, que expressam um atributo; esse tipo de variável é chamado de variável qualitativa [4].

A variável qualitativa divide-se em dois grupos: nominal e ordinal.

Variável qualitativa nominal é uma variável literal que não apresenta ordenação entre as possíveis realizações [20]. É o caso das variáveis sexo e bairro, em que não há uma ordem específica entre as respostas apresentadas pelos alunos e alunas. A variável bairro, além de qualitativa nominal, é binária, por assumir apenas duas categorias: Sim ou Não.

Variável qualitativa ordinal é a variável cujas possíveis respostas podem ser arranjadas em alguma ordem [22]. É o exemplo da variável opinião, cujos conceitos atribuídos ao lanche servido pela cantina da escola podem ser ordenados em R (ruim), B (bom) ou O (ótimo).

Além da variável qualitativa, tem-se a variável quantitativa, cujos dados são representados por algarismos. É o caso das variáveis altura, massa e moradores.

A variável quantitativa é classificada em uma de duas categorias: discreta ou contínua.

Variável quantitativa discreta é caracterizada quando o número de valores possíveis é ou um número finito ou uma quantidade enumerável [22]. Tem-se como exemplo a variável moradores.

Variável quantitativa contínua é identificada por valores que estão em um intervalo de números reais que resultam de uma mensuração [20], pertencem a este grupo as variáveis altura e massa.

A classificação das variáveis está resumida na Figura 1, a seguir:

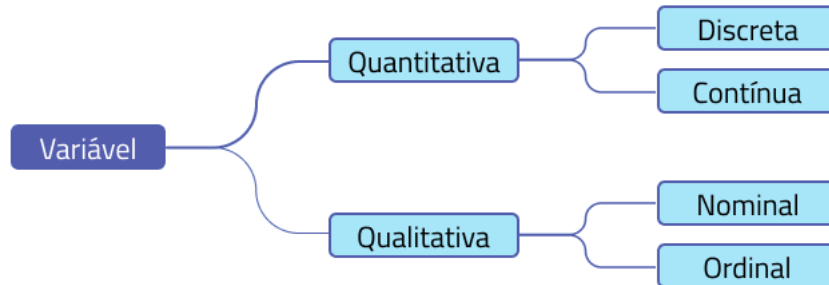


Figura 1: Tipos de variáveis.

Fonte: Próprio autor

### 3.2 POPULAÇÃO E AMOSTRA

População é a coleção completa de todos os elementos a serem estudados e que possuem, pelo menos, uma característica em comum [10] [22].

Uma população pode dar origem a vários subconjuntos denominados amostras, cada um desses subconjuntos é formado por elementos que apresentam um mesmo atributo em estudo, a variável.

Resume-se graficamente os conceitos de população e amostra na Figura 2. O conjunto formado por todos os alunos e alunas da Escola SP está representado pelo círculo maior, a população. A amostra (círculo menor) retirada da população é o subconjunto formado pelos 30 alunos e alunas, cujos dados coletados encontram-se organizados na Tabela 4.

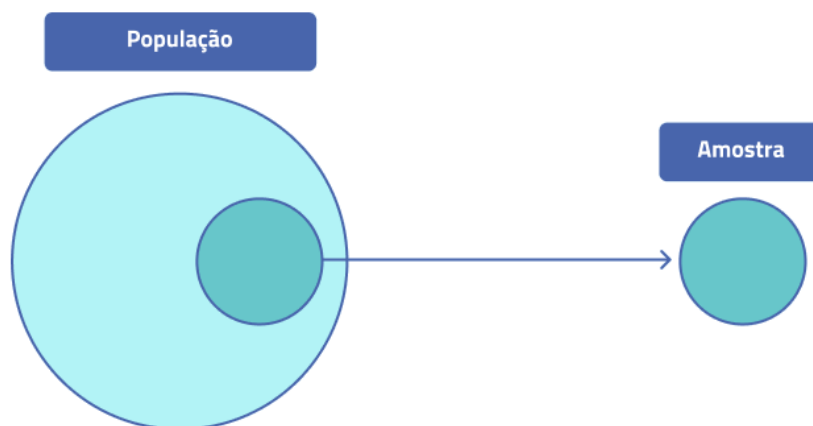


Figura 2: Representação gráfica dos conceitos de população e amostra.

Fonte: Próprio autor

## 3.3 TABELA PRIMITIVA E ROL

Um objetivo da estatística é, com base nos dados de uma amostra, fazer generalizações acerca da população da qual extraiu-se a amostra analisada [22].

Para que seja cumprido esse objetivo, após a seleção da amostra, os dados devem ser organizados cuidadosamente, para em seguida serem analisados. Nesse ponto, chega-se às definições de tabela primitiva e rol.

Os dados dispostos numa tabela, assim como são coletados, é o que se denomina tabela primitiva [10].

Representa-se uma tabela primitiva formada pelas massas dos 30 alunos e alunas que compõem a amostra da Escola SP na Tabela 5, da forma como se apresenta no GeoGebra, o aplicativo usa ponto para separar a parte inteira da parte fracionária.

Tabela 5: Tabela primitiva dos dados da variável massa dos alunos e alunas da amostra da Escola SP.

56.6	63.8	70.2	60.5	72	74.5	64.8	72.1	68.1	70
80.7	66.6	80.3	88.8	92.3	70.3	90	76.5	100.6	70.3
80.7	98.9	84	63	92.6	88.6	84.1	88.2	100.2	106.3

Fonte: Tabela 4

Rol é a tabela obtida após a ordenação (crescente ou decrescente) dos dados contidos na tabela primitiva [10]. A Tabela 6 é o rol dos dados da variável massa, ordenados de forma crescente. Com os dados organizados é mais fácil observar, por exemplo, os valores que mais se repetem, o maior ou o menor deles, ou ainda, a quantidade de alunos e alunas que se encontram acima ou abaixo de determinado valor.

Tabela 6: Rol dos dados da variável massa, da amostra de estudantes da Escola SP.

56.6	60.5	63	63.8	64.8	66.6	68.1	70	70.2	70.3
70.3	72	72.1	74.5	76.5	80.3	80.7	80.7	84	84.1
88.2	88.6	88.8	90	92.3	92.6	98.9	100.2	100.6	106.3

Fonte: Tabela 4

## 3.4 DISTRIBUIÇÃO DE FREQUÊNCIA

Para analisar uma variável, é importante que o pesquisador organize os dados de modo que lhe proporcione entendimento sobre o comportamento da variável na amostra. Apresentar dados em tabelas é um dos modos que se pode fazer essa organização.

Distribuição de frequência é uma lista com os valores dos dados organizados individualmente ou em grupo, com suas frequências correspondentes na amostra [22].

Verifica-se na Tabela 7 a distribuição de frequência da variável moradores, que estar associada à quantidade de pessoas que moram na mesma residência do aluno que compõe

a amostra da Escola SP. As quantidades estão descritas em ordem crescente na coluna Valor e associadas ao número de ocorrência de cada uma delas na coluna Frequência.

Tabela 7: Distribuição de frequência dos dados da variável moradores, da amostra de estudantes da Escola SP.

Valor	Frequência
2	5
3	11
4	8
5	4
6	2

Fonte: Tabela 4

### 3.5 CLASSE

Os dados ainda podem ser organizados em grupos de intervalos denominados classes de frequência ou, de forma resumida, classe [22]. As classes são indicadas simbolicamente por um índice  $i$ , sendo  $i = 1, 2, 3, \dots, k$ ; com  $k$  igual ao número total de classes [10]. Essa organização apresenta a vantagem de simplificar a exibição dos dados, porém há a desvantagem de se perder detalhes da distribuição.

Tabela 8: Distribuição de frequência em intervalos de classe dos dados da variável altura dos estudantes da amostra da Escola SP.

Intervalo	Contagem
1.55 – 1.6	1
1.6 – 1.65	3
1.65 – 1.7	7
1.7 – 1.75	8
1.75 – 1.8	7
1.8 – 1.85	3
1.85 – 1.9	1

Fonte: Tabela 4

A distribuição de frequência das alturas dos alunos e alunas da amostra da Escola SP está agrupada em intervalos de classe na Tabela 8, conforme é exibido no GeoGebra. Nota-se que foram perdidas algumas informações com a distribuição de frequência em intervalos de classe. Não se distingue, por exemplo, a altura que mais se repete entre os elementos da amostra ou quantos deles têm estatura de 1,75m.

#### *Limites de classe*

Os limites das classes são o menor e o maior valor que podem pertencer ao intervalo [16]; portanto, o menor valor é o limite inferior da classe, representado por  $(l_i)$ , enquanto



que o maior valor é o limite superior da classe, indicado por  $(L_i)$ , sendo  $i = 1, 2, 3, \dots, k$ ; e  $k$  igual ao número total de classes [10].

As distribuições de frequência em intervalos de classe devem ser apresentadas com intervalos semiabertos, no extremo inferior ou no superior [15]. As classes podem ser organizadas conforme a necessidade, o objetivo do estudo ou a análise dos dados, não devendo apresentar limites sobrepostos. A representação tabular (Tabela 8) gerada pelo GeoGebra apresenta intervalo fechado no limite inferior e aberto no limite superior. Assim, um aluno ou aluna com estatura de 1,65m está incluído ou incluída na 3ª classe.

Na tabela, as estaturas dos elementos da amostra da situação problema foram agrupadas em setes intervalos de classe, os números extremos indicados nos intervalos são os limites da classe:

A primeira classe ( $i = 1$ ) tem 1,55m como limite inferior ( $l_1 = 1,55$ ) e 1,60m como limite superior ( $L_1 = 1,60$ ).

Observa-se que os limites de classes não necessariamente são valores que pertencem à distribuição dos dados. Nota-se que o limite superior da sétima classe ( $L_7 = 1,90$ ) é um valor que não existe entre os dados.

#### *Ponto médio da classe*

É o ponto que divide o intervalo de classe em duas partes iguais [10]; o ponto médio da classe ( $m_i$ ) é determinado pela média aritmética entre o limite inferior e o limite superior da classe. Este ponto também é chamado de marca da classe, por ser o valor que representa a classe nos cálculos estatísticos [16].

$$m_i = \frac{l_i + L_i}{2} \quad (1)$$

com  $i = 1, 2, 3, \dots, k$ ; e  $k$  o número total de classes.

Assim, o ponto médio da primeira classe ( $i = 1$ ) é dado por:

$$m_1 = \frac{1,55 + 1,60}{2} = 1,58$$

### 3.6 AMPLITUDES

#### *Amplitude de classe*

É o tamanho da classe, a medida do comprimento do intervalo que a define [10]. A amplitude da classe ( $h_i$ ) é dada pela diferença entre limite superior e o limite inferior da classe.

$$h_i = L_i - l_i \quad (2)$$

com  $i = 1, 2, 3, \dots, k$ ; e  $k$  igual ao número total de classes.

Observa-se na Tabela 8 que os intervalos das classes possuem a mesma amplitude. Calculando a amplitude da primeira classe ( $h_1$ ) da distribuição das alturas dos elementos da amostra em intervalos de classe, tem-se:

$$h_1 = L_1 - l_1 = 1,60 - 1,55 = 0,05$$

As amplitudes das classes da distribuição das alturas dos elementos da amostra da Escola SP têm 0,05m ou 5cm.

#### *Amplitude de distribuição*

Amplitude de distribuição (AD) é a diferença entre o limite superior da última classe e o limite inferior da primeira classe.

Dado que as classes são representadas pelo índice  $i$ , com  $i = 1, 2, 3, \dots, k$ ; com  $k$  igual ao número total de classes [10]. E sendo o limite inferior da classe representado por ( $l_i$ ) e o limite superior indicado por ( $L_i$ ), tem-se que a amplitude de distribuição é dada por:

$$AD = L_k - l_1 \quad (3)$$

em que  $L_k$  é o limite superior da última classe e  $l_1$  é o limite inferior da primeira classe.

Calculando a amplitude da distribuição das estaturas da amostra dos elementos da Escola SP, organizadas em intervalos de classe na Tabela 8, tem-se:

$$AD = L_7 - l_1 = 1,90 - 1,55 = 0,35$$

#### *Amplitude amostral*

Amplitude amostral (AA) é a diferença entre os valores máximo e mínimo pertencentes à amostra de dados [16].

Para o conjunto  $\{x_1, x_2, \dots, x_n\}$ , a amplitude amostral é dada por:

$$AA = x_{max} - x_{min} \quad (4)$$

em que  $x_{max}$  é o maior valor pertencente ao conjunto e  $x_{min}$  é o menor valor.

Agora, obtendo a amplitude amostral das massas dos alunos e alunas da Escola SP, cujos dados encontram-se organizados na Tabela 6, tem-se:

$$AA = x_{max} - x_{min}$$

com  $x_{max} = 106,3$  e  $x_{min} = 56,6$ . Logo:

$$AA = 106,3 - 56,6 = 49,7$$

### 3.7 TIPOS DE FREQUÊNCIA

#### *Frequência simples ou absoluta*

Frequência simples ou absoluta ( $f_i$ ) é o número de ocorrências de dados numa determinada classe [16].

A soma das frequências simples é igual à quantidade total de dados da amostra ( $n$ ) [10]:

$$\sum_{i=1}^k f_i = n \quad (5)$$

com  $i = 1, 2, 3, \dots, k$ ; e  $k$  igual ao número total de classes.

#### *Frequência relativa*

Frequência relativa ( $fr_i$ ) é a fração, ou proporção, de dados contidos numa classe [16]. A frequência relativa de uma determinada classe é dada pela divisão entre frequência absoluta ( $f_i$ ) da classe considerada e o tamanho da amostra ( $n$ ):

$$fr_i = \frac{f_i}{n} \quad (6)$$

com  $i = 1, 2, 3, \dots, k$ .  $k$  igual ao número total de classes.

#### *Frequência acumulada*

A frequência acumulada ( $F_i$ ) para uma classe é o somatório da frequência simples daquela classe com as frequências simples de todas as classes anteriores a ela [22], é dada por:

$$F_i = \sum_{i=1}^w f_i \quad (7)$$

com  $i$  variando de 1 até  $w$ , sendo  $w$  igual ao índice da classe considerada.

A frequência acumulada indica a quantidade de ocorrências até a classe observada. Logo, a frequência acumulada da última classe ( $F_k$ ) é igual ao tamanho total da amostra ( $n$ ), sendo  $k$  o número total de classes.

*Frequência acumulada relativa*

Frequência acumulada relativa ( $Fr_i$ ) de uma classe é a frequência acumulada da classe ( $F_i$ ) dividida pelo tamanho total da amostra ( $n$ ) [10].

$$Fr_i = \frac{\sum_{i=1}^w f_i}{n} \quad (8)$$

com  $i$  variando de 1 até  $w$ ; sendo  $w$  igual ao índice da classe considerada.

A Tabela 9 apresenta uma organização das alturas dos alunos e alunas, que compõem a amostra da Escola SP, distribuídas utilizando a frequência absoluta, frequência relativa, frequência acumulada e frequência acumulada relativa.

Tabela 9: Distribuição dos dados da variável altura dos estudantes da amostra da Escola SP utilizando as frequências absoluta, relativa, acumulada e acumulada relativa.

$i$	Classe	$f_i$	$fr_i$	$F_i$	$Fr_i$
1	1,55 † 1,60	1	0,03	1	0,03
2	1,60 † 1,65	3	0,10	4	0,13
3	1,65 † 1,70	7	0,23	11	0,36
4	1,70 † 1,75	8	0,28	19	0,64
5	1,75 † 1,80	7	0,23	26	0,87
6	1,80 † 1,85	3	0,10	29	0,97
7	1,85 † 1,90	1	0,03	30	1
Total		n = 30	1,00		

Fonte: Tabela 4

## 3.8 MEDIDAS DE POSIÇÃO

Medida de posição ou medida de centro é um valor central do conjunto de dados [22]. As medidas de posição mais comuns são a moda, a média, a mediana e os quantis.

*Moda*

Moda ( $M_o$ ) é o valor que apresenta o maior número de ocorrência numa série de dados quantitativos discretos. Em um conjunto de dados qualitativos, a moda é a observação que possuir a maior frequência simples.

Pode haver mais de uma observação que apresenta o maior número de frequência; quando isso ocorre, classifica-se o conjunto de dados da seguinte forma:

- quando não houver valor que se repita com maior frequência no conjunto, diz-se que não há moda ou que o conjunto de dados é amodal.

- quando dois valores apresentam a mesma maior frequência, cada valor é uma moda e o conjunto de dados é bimodal;

- quando mais de dois valores ocorrem com a mesma maior frequência no conjunto de dados, cada um deles é uma moda e diz-se que o conjunto de dados é multimodal. [22]

Observa-se na Tabela 6 que não há valores que se repetem entre as massas dos elementos da amostra da Escola SP; portanto, o conjunto de dados Massa é amodal. Da mesma forma, verificando o conjunto de dados Opinião, Tabela 4, nota-se que a menção mais citada pelos alunos e alunas foi B, com 13 observações; logo, a moda desse conjunto é B. O conjunto de dados Altura (Tabela 4) trata-se de um conjunto bimodal, cujas modas são 1,67 e 1,75.

### *Média*

A média aritmética é a soma dos valores dos dados observados dividida pelo número total de observações [16].

Para a amostra  $\{x_1, x_2, \dots, x_n\}$  a média ( $\bar{x}$ ) é dada por:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} = \frac{1}{n} \sum_{i=1}^n x_i \quad (9)$$

### *Mediana*

A mediana ( $Md$ ) é o valor que está no centro dos dados de um conjunto ordenado, dividindo-o em duas partes com quantidades iguais de valores [16]. Quando o conjunto de dados tem um número ímpar de observações, a mediana é o elemento central. Se a quantidade de observações do conjunto é um número par, a mediana é obtida fazendo a média dos dois elementos das posições centrais [20].

Potanto, considerando uma amostra cujos dados dispostos em ordem crescente sejam indicados por:

$$x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq \dots \leq x_{(n-1)} \leq x_{(n)}$$

A mediana é dada por:

Para  $n$  ímpar:

$$Md = x_{(\frac{n+1}{2})} \quad (10)$$

Para  $n$  par:

$$Md = \frac{x_{(\frac{n}{2})} + x_{(\frac{n}{2}+1)}}{2} \quad (11)$$

Considerando a amostra obtida dos alunos e alunas da Escola SP e observando os dados da variável altura, organizados em ordem crescente na Tabela 4, obtém-se a mediana das alturas fazendo a média aritmética entre a 15<sup>a</sup> e a 16<sup>a</sup> observações, dado que  $n$  é par e igual a 30:

$$Md = \frac{x_{(15)} + x_{(16)}}{2} = \frac{1,73 + 1,73}{2} = 1,73$$

Indicando que, entre os alunos e alunas que compõem a amostra, 50% deles têm altura menor que ou igual a 1,73m; enquanto que a outra metade tem estatura maior que ou igual a 1,73m.

### *Quartil*

Os quartis são valores que dividem um conjunto de dados ordenados em quatro partes iguais [16]; tem-se, portanto, três quartis.

O primeiro quartil ( $q_1$ ) é o valor situado no conjunto de dados de forma que 25% dos valores do conjunto são menores que ele.

O valor de  $q_1$  é obtido por [14]:

Para  $n$  ímpar

$$q_1 = \frac{n + 1}{4} \quad (12)$$

Para  $n$  par

$$q_1 = \frac{n + 2}{4} \quad (13)$$

O segundo quartil ( $q_2$ ) é o valor que divide o conjunto de dados em dois com iguais quantidades de elementos, o seu valor coincide com o da mediana ( $q_2 = M_d$ ) [10].

O terceiro quartil ( $q_3$ ) é o valor situado no conjunto de dados, de modo que 75% dos valores menores que ele, enquanto os outros 25% dos dados são maiores que  $q_3$  [10].

O valor de  $q_3$  é obtido por [14]:

Para  $n$  ímpar

$$q_3 = \frac{3n + 3}{4} \quad (14)$$

Para  $n$  par

$$q_3 = \frac{3n + 2}{4} \quad (15)$$

### 3.9 MEDIDAS DE DISPERSÃO

As medidas de posição, apesar de serem valores representativos de um conjunto de dados, por si só não possuem a capacidade de exprimir o quanto um conjunto de dados se dispersa em torno de um valor central. A medida da dispersão dos dados é importante para indicar o padrão de homogeneidade do conjunto, para isso utiliza-se, entre outras ferramentas, a distância interquartil, a variância e o desvio padrão.

*Distância interquartil*

Distância interquartil ( $D_q$ ) é a diferença entre o terceiro e o primeiro quartil [20], indica a amplitude da porção central da distribuição dos dados [16]. A distância interquartil é dada por:

$$D_q = q_3 - q_1 \quad (16)$$

O conceito de distância interquartil é útil para identificar os valores extremos de uma distribuição de dados. Os valores extremos, atípicos, discrepantes ou *outliers*, são observações que destoam no conjunto de dados [20].

O limite inferior (LI) da distribuição é obtido a partir do primeiro quartil subtraindo uma vez e meia a distância interquartil ( $D_q$ ). Observações abaixo de LI são consideradas discrepantes.

$$LI = q_1 - (1,5)D_q \quad (17)$$

Da mesma forma, acha-se o limite superior (LS) somando uma vez e meia a distância interquartil ( $D_q$ ) ao terceiro quartil da distribuição. Valores acima do LS são considerados discrepantes.

$$LS = q_3 + (1,5)D_q \quad (18)$$

Considerando os dados da variável altura dos indivíduos da amostra da Escola SP (Tabela 4), tem-se que:

$$D_q = 1,75 - 1,67 = 0,08$$

$$LI = 1,67 - (1,5)0,08 = 1,55$$

$$LS = 1,75 + (1,5)0,08 = 1,87$$

Portanto, valores abaixo de 1,55 e acima de 1,87 são observações atípicas desse conjunto de dados.

*Variância*

A variância é uma medida de dispersão dos dados em torno da média [20]. Considerando o conjunto de dados da variável quantitativa X,  $\{x_1, x_2, \dots, x_n\}$ , a variância amostral ( $S^2$ ) é dada por:

$$S^2(X) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (19)$$

com  $i = 1, 2, \dots, n$

*Desvio padrão*

O desvio padrão mede a variação dos dados em relação à média do conjunto de valores [16]. O desvio padrão amostral ( $S$ ) do conjunto de dados da variável quantitativa  $X$ ,  $\{x_1, x_2, \dots, x_n\}$ , é calculado fazendo a raiz quadrada da variância.

$$S(X) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (20)$$

com  $i = 1, 2, \dots, n$

## 3.10 REPRESENTAÇÃO GRÁFICA DE UMA DISTRIBUIÇÃO

Os dados devem ser apresentados de modo que sejam facilmente compreendidos pelos leitores e pesquisadores; a representação gráfica de um conjunto de dados possui a vantagem de informar, de forma breve e sucinta, sobre a distribuição e variabilidade dos dados representados [20].

*Gráfico de dispersão unidimensional*

Também referido como diagrama ou gráfico de pontos, é um gráfico no qual cada valor do conjunto de dados é indicado por um ponto no eixo horizontal [22]. Valores com mais de uma observação são indicados pelo empilhamento dos respectivos pontos.

Esta representação permite observar como os dados estão distribuídos, identificar valores específicos, além de observar valores incomuns ou extremos no conjunto de dados [16].

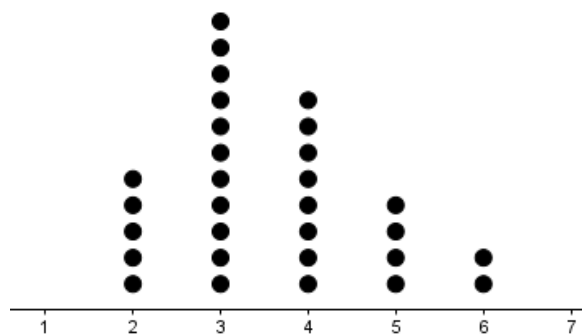


Figura 3: Gráfico de dispersão unidimensional da variável moradores, da amostra da Escola SP.

Fonte: Tabela 4

A Figura 3, construída no GeoGebra, é a representação gráfica unidimensional das quantidades de moradores na mesma residência do aluno ou aluna da amostra da Escola SP (Tabela 4).



### Gráfico de barras

É a representação de um conjunto de dados por intermédio de retângulos (colunas ou barras), em que uma das dimensões é proporcional à magnitude a ser representada e a outra dimensão arbitrária, porém igual para todas as barras [20].

Na Figura 4 estão representadas em um gráfico de barras verticais (colunas), construído no GeoGebra, as quantidades de moradores na mesma residência do aluno ou aluna da amostra da Escola SP.

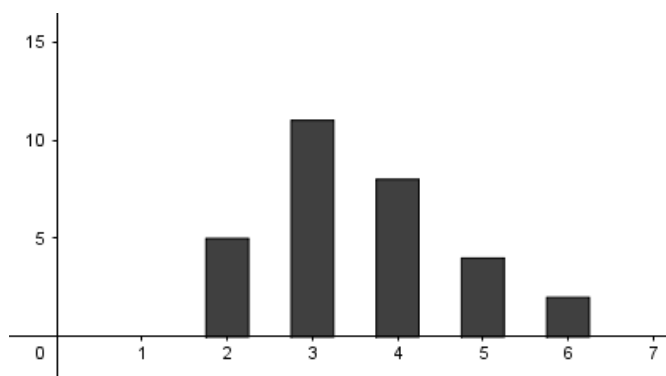


Figura 4: Gráfico de barras verticais da variável moradores, da amostra da Escola SP.

Fonte: Tabela 4

### Histograma

É um gráfico de barras verticais contíguas, cujas bases se localizam sobre o eixo horizontal, de modo que seus pontos médios coincidam com os pontos médios dos intervalos de classe; as barras possuem largura igual a amplitude dos intervalos de classe e altura proporcional à frequência da classe que representa [10].

A seguir, na Figura 5, construída no GeoGebra, tem-se o histograma dos dados da variável altura da amostra de alunos e alunas da situação problema que ilustra esta dissertação; observa-se que os dados foram distribuídos em sete intervalos de classes, todos com a mesma amplitude (0,05m), variando entre 1,55 e 1,90m.

### Boxplot

É uma representação que utiliza a mediana, o primeiro quartil, o terceiro quartil e os valores extremos da distribuição para expressar graficamente uma ideia de posição, dispersão, assimetria, caudas e observações discrepantes de um conjunto de dados [20].

O *boxplot* é composto por um retângulo limitado à direita pelo primeiro quartil ( $q_1$ ) e na porção esquerda pelo terceiro quartil ( $q_3$ ). Na parte interna do retângulo encontra-se um traço na vertical indicando a mediana ( $M_d$ ) do conjunto. As caudas se prolongam

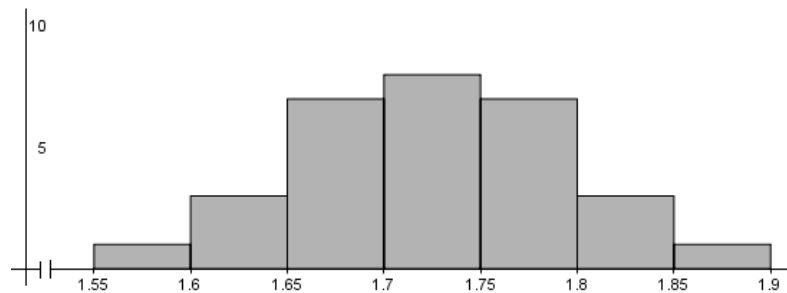


Figura 5: Histograma dos dados da variável altura dos elementos da amostra da Escola SP.

Fonte: Tabela 4

para ambos os lados sem que ultrapassem o limite inferior (LI) e limite superior (LS), enquanto que os valores atípicos (*outliers*) são representadas por sinais característicos.

Para ilustrar um *boxplot* e os elementos que o constituem, a Figura 6, construída no GeoGebra, representa a distribuição dos dados da variável altura apenas dos estudantes do sexo masculino, da amostra da Escola SP. Observa-se o retângulo central delimitado à esquerda pelo primeiro quartil ( $q_1 = 1,73m$ ) e à direita pelo terceiro quartil ( $q_3 = 1,78m$ ); logo, tem-se distância interquartil ( $D_q$ ) igual a  $0,05m$ . Ao centro do retângulo encontra-se a representação gráfica da mediana ( $M_d = 1,75m$ ). As caudas se alongam para a esquerda partindo do primeiro quartil até o limite inferior ( $LI = 1,66m$ ) e para a direita, a partir do terceiro quartil até o limite superior ( $LS = 1,82m$ ). Valores fora desse intervalo são atípicos, como a estatura  $1,88m$  representada na ilustração por um sinal que lembra a letra xis.

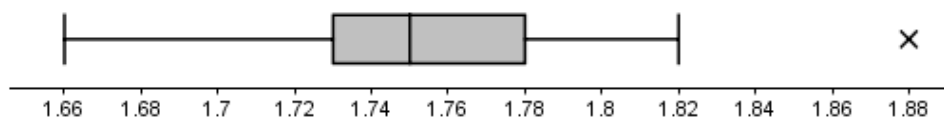


Figura 6: *Boxplot* dos dados da variável altura somente dos alunos da amostra da Escola SP.

Fonte: Tabela 4

### 3.11 SIMETRIA

Uma distribuição é simétrica quando os elementos que a compõem estão dispostos uniformemente em torno da média do conjunto. Graficamente, se uma linha vertical puder ser traçada sobre a média da distribuição e as metades resultantes são imagens espelhadas (ou aproximadamente espelhadas), tem-se uma distribuição simétrica [16].

Considerando-se os pontos  $x_{(1)}, q_1, q_2, q_3$  e  $x_{(n)}$ , uma distribuição simétrica deve apresentar [20]:

$$q_2 - x_{(1)} \approx x_{(n)} - q_2; \quad (21)$$

$$q_2 - q_1 \approx q_3 - q_2; \quad (22)$$

$$q_1 - x_{(1)} \approx x_{(n)} - q_3 \quad (23)$$

Além dessas características, uma distribuição simétrica apresenta valores próximos para as medidas resumo da média, mediana e moda [10]:

$$\bar{x} \approx M_d \approx M_o \quad (24)$$

### 3.12 ASSOCIAÇÃO ENTRE VARIÁVEIS

Na fase de análise dos dados, o pesquisador pode avaliar o grau de relacionamento entre duas variáveis em busca de padrões que autorizem fazer inferências, emitir opiniões. Nessa análise verifica-se se os resultados de uma variável são influenciados pela presença (ou ausência) de outra variável. Dependendo dos tipos de variáveis envolvidas, existem diferentes critérios para se observar essa relação, que podem ser realizadas por intermédio de métodos gráficos: histogramas, *boxplot*, diagrama de ramo e folhas; ou de medidas numéricas: tabela de dupla entrada e medidas resumo [20]. Descreve-se, a seguir, uma análise envolvendo variáveis qualitativa e quantitativa, cuja utilização está apresentada na Proposta Didática.

#### *Qualitativa e Quantitativa*

Associação entre duas variáveis, uma qualitativa e a outra quantitativa, é percebida analisando o que ocorre com a variável quantitativa quando os dados são organizados de acordo com as categorias da variável qualitativa [20].

Organizando os dados da variável quantitativa altura dos estudantes da amostra da Escola SP (Tabela 4) de acordo com os atributos da variável qualitativa sexo, numa tabela de dupla entrada, tem-se a distribuição conjunta da Tabela 10, cujo objetivo é conhecer o grau de dependência entre as variáveis. Tabelas de dupla entrada permitem duas ordens de classificação: uma na horizontal e outra na vertical [10].

Percebe-se que a distribuição dos dados da variável altura, segundo o atributo feminino da variável sexo, está deslocada para a esquerda; enquanto que para o atributo masculino, os dados estão deslocados para a direita. Isso pode indicar a existência de associação entre as variáveis altura e sexo dos estudantes da amostra da Escola SP.

Tabela 10: Distribuição de frequência dos dados da variável altura dos estudantes da amostra da Escola SP, segundo os atributos da variável sexo.

	Feminino	Masculino	Total
1,55 † 1,60	1	0	1
1,60 † 1,65	3	0	3
1,65 † 1,70	5	2	7
1,70 † 1,75	2	6	8
1,75 † 1,80	1	6	7
1,80 † 1,85	0	3	3
1,85 † 190	0	1	1
Total	12	18	30

Fonte: Tabela 4

A associação entre as variáveis também pode ser observada fazendo a distribuição das proporções dos dados da variável altura segundo os atributos da variável sexo, a Tabela 11 apresenta essa distribuição.

Tabela 11: Distribuição das proporções dos dados da variável altura dos estudantes da amostra da Escola SP, segundo os atributos da variável sexo.

	Feminino	Masculino	Total
1,55 † 1,60	8%	0%	3%
1,60 † 1,65	25%	0%	10%
1,65 † 1,70	42%	11%	23%
1,70 † 1,75	17%	33%	27%
1,75 † 1,80	8%	33%	23%
1,80 † 1,85	0%	17%	10%
1,85 † 190	0%	6%	3%
Total	100%	100%	100%

Fonte: Tabela 4

Caso não houvesse associação entre as variáveis em análise, observando a coluna dos totais, esperava-se que 23% dos alunos tivessem altura entre 1,65 e 1,70m, independente do sexo. Porém, observando-se as colunas Feminino e Masculino, vê-se que as proporções divergem do valor esperado para essa faixa de estatura: 42% dos estudantes da terceira classe são do sexo feminino e 11% deles do sexo masculino, o que também indica que pode haver associação entre as variáveis.

Medidas resumo podem indicar a existência de associação entre as variáveis. Na tabela a seguir estão as medidas resumo da variável quantitativa altura, de acordo com as qualidades da variável sexo.

A variância ( $S$ ) da Amostra indica a dispersão dos dados globalmente. Se a variância da variável quantitativa, calculada separadamente segundo os atributos da variável qualitativa, for menor do que a variância global, significa que pode existir uma associação entre as variáveis em análise [20]. Isso ocorre ao se comparar a variância da variável quantitativa segundo os atributos Feminino e Masculino da variável qualitativa, na coluna em destaque, na Tabela 12.

Tabela 12: Medidas resumo da variável altura dos alunos e alunas da amostra da Escola SP, segundo a variável sexo.

Sexo	$n$	$\bar{x}$	$S^2$	$S$	$x_1$	$q_1$	$q_2$	$q_3$	$x_n$
Feminino	12	1,67	0,05	<b>0,002</b>	1,59	1,61	1,67	1,69	1,76
Masculino	18	1,76	0,05	<b>0,002</b>	1,66	1,73	1,75	1,78	1,88
Amostra	30	1,72	0,07	<b>0,004</b>	1,59	1,67	1,73	1,75	1,88

Fonte: Tabela 4

A identificação da associação entre duas variáveis pode ser realizada graficamente. A Figura 7 apresenta um diagrama *boxplot* para a distribuição da variável quantitativa segundo os atributos da variável qualitativa; nele percebe-se a distribuição dos dados da variável altura, segundo a qualidade Feminino, deslocado para a esquerda, enquanto que a distribuição dos dados da variável quantitativa, de acordo com o atributo Masculino, estão deslocados para a direita; isso indica que a variável qualitativa pode exercer influência sobre os resultados da variável quantitativa.

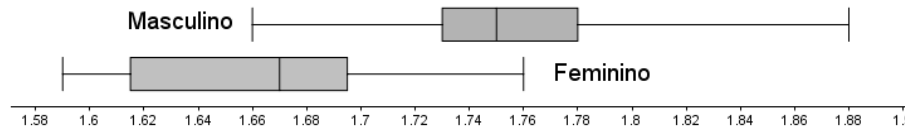


Figura 7: Identificação de associação entre as variáveis altura e sexo da amostra de alunos da Escola SP, utilizando *boxplot*

Fonte: Tabela 4

□

## O GEOGEBRA

---

### 4.1 APRESENTAÇÃO

O GeoGebra é um programa de código aberto, de fácil utilização, destinado a apoiar o ensino e aprendizagem de Ciências, Tecnologia, Matemática e Engenharia, em todos os níveis; nele as representações geométrica e algébrica de um mesmo objeto estão conectadas de forma dinâmica; possibilita o desenvolvimento de páginas *web* interativas, materiais didáticos gráficos; possui uma rede social com milhares de trabalhos compartilhados por usuários espalhados pelo mundo, além de um fórum onde essas pessoas dirimem dúvidas e discutem o uso do aplicativo; encontra-se disponível para o sistema operacional Windows, iOS, Android, Mac, Chromebook e Linux [14]. Nesta dissertação, trabalhou-se com a versão GeoGebra Clássico 5 para Windows, o aplicativo possui ainda as opções GeoGebra Clássico 6, Calculadora 3D e Realidade Aumentada para *smartphone*, todos com *download* gratuito.

### 4.2 DOWNLOAD E INSTALAÇÃO

Para instalar o GeoGebra Clássico 5 no computador, o usuário deve acessar o endereço <https://www.geogebra.org/download> e clicar no *link* da Figura 8 para que o arquivo instalador seja baixado para a sua máquina. Para este trabalho, baixou-se o arquivo GeoGebra-Windows-Installer-5-5-574-0.

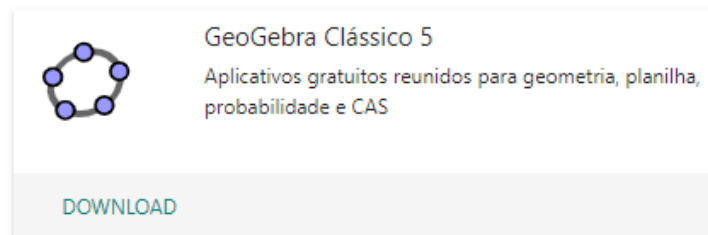


Figura 8: *Link* para *download* do GeoGebra Clássico 5.

Fonte: Próprio autor

Durante a instalação do aplicativo, algumas janelas de diálogo se sucedem (Figura 9) para que o usuário selecione as configurações desejadas. Na primeira janela de diálogo o usuário deve selecionar o idioma que pretende que seja exibido na tela do GeoGebra. Neste trabalho, escolheu-se o idioma Português Brasileiro. A tela seguinte trata dos termos da Licença de Uso. Para seguir com a instalação do programa, o usuário deve ler o termo

e estar de acordo com o uso não comercial do aplicativo. Na janela seguinte o usuário pode indicar o tipo de configuração de acordo com as suas necessidades. Para o objetivo desta dissertação, selecionou-se a opção *Standard*, o desenvolvedor recomenda essa opção para a maioria dos usuários do programa, selecionando-a todos os principais recursos do GeoGebra são instalados. Após finalizada a instalação, um atalho para o programa é criado na área de trabalho do computador.

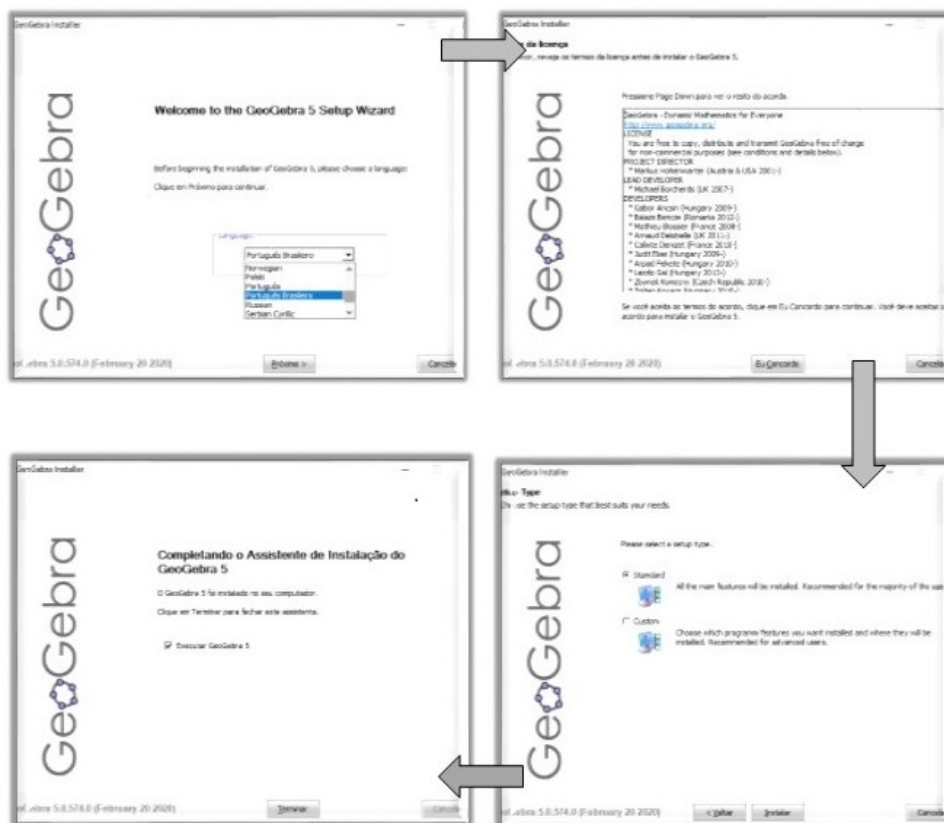


Figura 9: Janelas de diálogo de instalação.  
Fonte: Próprio autor

### 4.3 INTERFACE

O GeoGebra possui uma área de interação padrão (Figura 10), que pode ser alterada de acordo com as necessidades e objetivos do usuário. A nomenclatura de alguns dos segmentos que constituem a interface do aplicativo não difere das designações dadas a partes homólogas de outros programas.

#### *Cabeçalho*

O cabeçalho está assinalado na Figura 10 pela seta acompanhada pelo número 1. No cabeçalho encontram-se o logotipo do GeoGebra e o nome dado ao arquivo, na extremidade

oposta encontram-se os botões padrões para minimizar, maximizar e fechar a janela do aplicativo.

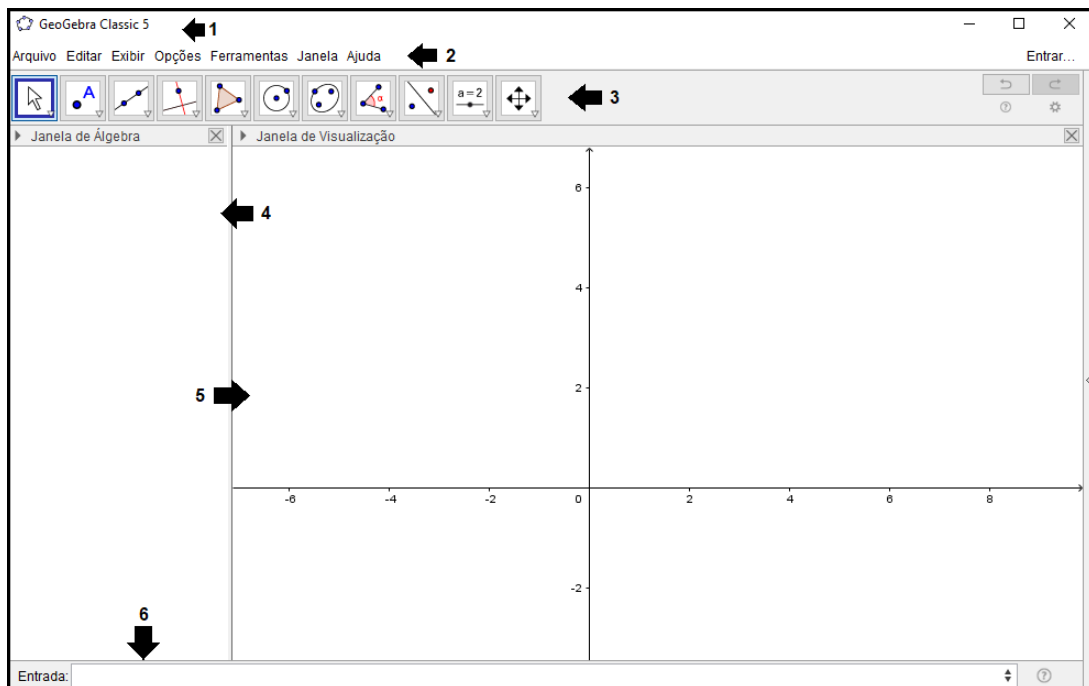


Figura 10: Interface padrão do programa GeoGebra.  
Fonte: Próprio autor

### Barra de menus

Seguindo a numeração destacada na imagem da interface (Figura 10), a barra de *menus* encontra-se indicada pelo número 2. Os *menus* Arquivo e Editar possuem funcionalidades idênticas às encontradas em outros aplicativos. Os *menus* Exibir, Opções e Ferramentas possuem itens específicos do GeoGebra e merecem destaque, Figura 11. O *menu* Exibir apresenta opções para alterar a aparência da interface do GeoGebra, mostrar outras janelas de trabalho disponíveis e alterar as preferências do *layout* do aplicativo. No *menu* Opções encontram-se ícones para alterar configurações internas do programa, como: idioma, tamanho da fonte, quantidade de casas decimais, unidades de medida e estilos de apresentação de ângulos e coordenadas, exibição de dicas, entre outras. No *menu* Ferramentas pode-se configurar a barra de ferramentas ou ainda, de acordo com as necessidades do usuário, criar uma nova ferramenta para o aplicativo.

### Barra de ferramentas

A barra de ferramentas, indicada pelo número 3 na Figura 10, possui ícones próprios do aplicativo; para cada ícone representado nessa barra existe uma caixa de ferramentas com atalhos agrupados de acordo com a natureza dos objetos gráficos resultantes (Figura



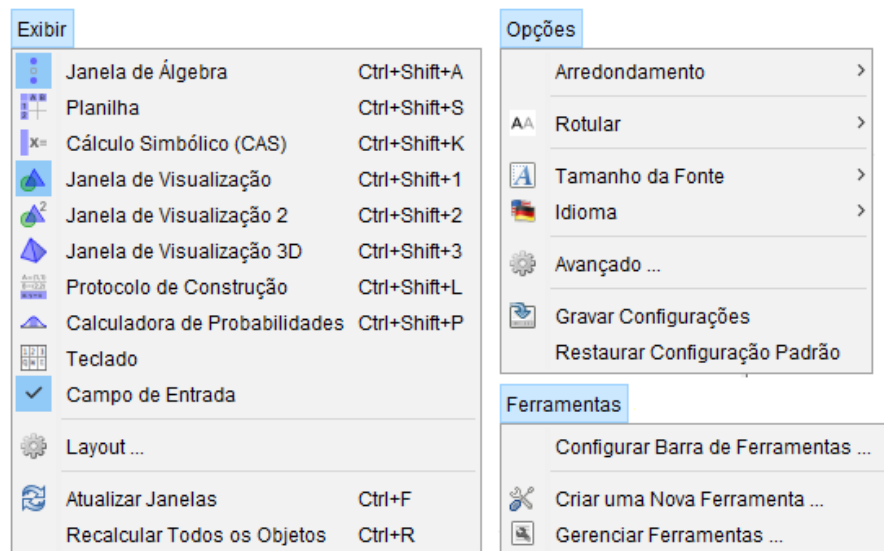


Figura 11: Ícones da barra de *menus*.  
Fonte: Próprio autor

12). Pode-se configurar a barra de ferramentas no *menu* Ferramentas da barra de *menus*, organizando ou alterando a exibição das ferramentas conforme a intenção do usuário. Para cada uma das ferramentas do GeoGebra, o aplicativo tem um comando interno predefinido que possibilita a criação do objeto sem a necessidade do uso da ferramenta.



Figura 12: Ferramentas agrupadas em uma caixa de ferramentas.  
Fonte: Próprio autor

### *Janela de álgebra*

A janela de álgebra está sinalizada pela seta e o algarismo 4 na Figura 10. O GeoGebra exibe na janela de álgebra os códigos, fórmulas, números, coordenadas ou textos relativos aos objetos criados pelo usuário no aplicativo, Figura 13. Pode-se usar esses

códigos relacionando-os a outros objetos, em um mesmo arquivo, criando conectividade entre eles.

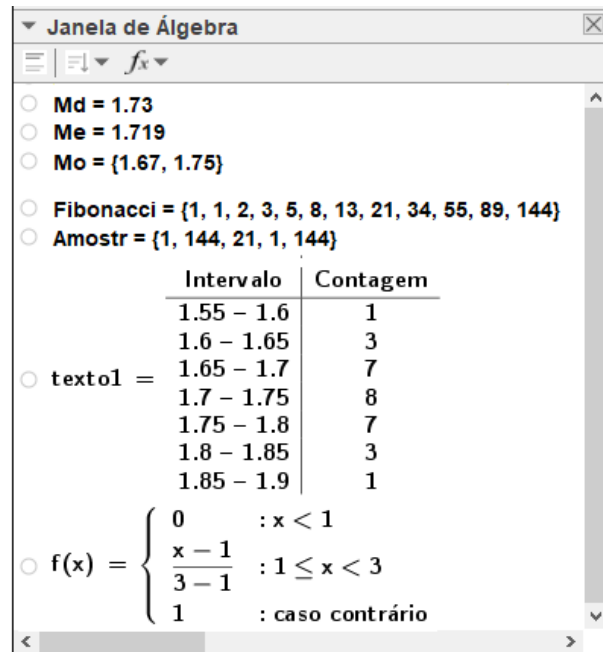


Figura 13: Janela de álgebra.

Fonte: Próprio autor

Na Figura 13, os seguintes objetos estão representados algebricamente:

- A mediana, a média e as modas (bimodal) do conjunto de dados altura, da amostra de alunos e alunas da Escola SP (Tabela 4).

$$Md = 1,73$$

$$Me = 1,71$$

$$Mo = \{1,67, 1,75\}$$

- Uma lista composta pelos doze primeiros números da sequência de Fibonacci.

$$Fibonacci = \{1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144\}$$

- Uma amostra contendo cinco elementos selecionados aleatoriamente, com reposição, da lista nomeada Fibonacci.

$$Amostr = \{1, 144, 21, 1, 144\}$$

- Uma tabela de distribuição de frequência absoluta em intervalos de classe da variável altura da amostra de alunos e alunas da Escola SP, nomeada automaticamente pelo sistema como texto1.

- Uma função de distribuição empírica de uma distribuição uniforme,  $f(x)$ .

### *Janela de visualização*

A janela de visualização, indicada na Figura 10 pelo número 5, é o local onde são exibidos os objetos gráficos construídos no GeoGebra. Na Figura 14, sobre o eixo das abscissas, tem-se a representação do gráfico de barras verticais do conjunto de dados da variável moradores; sob o eixo das abscissas, observa-se um *boxplot* referente ao mesmo conjunto de valores.

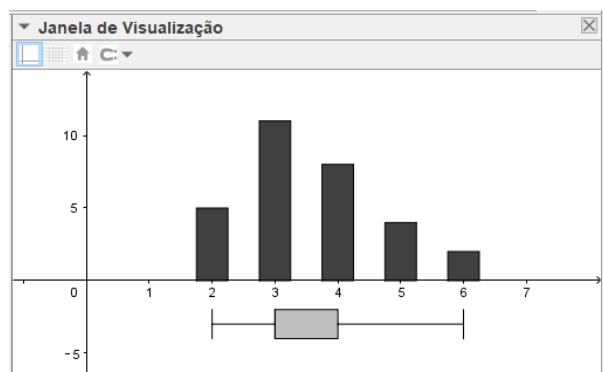


Figura 14: Janela de visualização.

Fonte: Próprio autor

O GeoGebra também possui uma janela de visualização 3D, que possibilita a exibição de sólidos geométricos construídos no aplicativo.

Os objetos construídos no GeoGebra podem ser modificados a partir da representação algébrica exibida na janela de álgebra ou pela representação gráfica na janela de visualização. Por serem representações de um mesmo objeto, as alterações realizadas em uma das janelas resultam em mudanças na representação exibida na outra janela.

### *Barra de entrada*

A barra de entrada, caixa de entrada ou apenas entrada, é o local onde digitam-se os comandos para a construção de objetos no GeoGebra. Ela está indicada na Figura 10 pelo algarismo 6.

### *Janela planilha*

A janela planilha (Figura 15) não é exibida no modo padrão do GeoGebra; para exibi-la, deve-se clicar no *menu* Exibir → Planilha.

A aparência e nomenclatura das células são idênticas às do Microsoft Excel ou do Libre Office Calc.

Observa-se que as informações digitadas nessas células associam-se a objetos algébricos que são indicados na janela de álgebra e a objetos gráficos traçados na janela de visualização, que podem receber tratamento estatístico posterior.

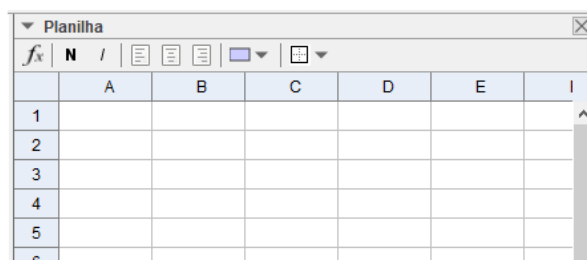


Figura 15: Janela planilha.

Fonte: Próprio autor

## 4.4 FERRAMENTAS

Nesta seção apresenta-se algumas das ferramentas do GeoGebra, descreve-se suas funcionalidades, características e formas de utilização.

*Controle deslizante*

O controle deslizante, ou *slider*, é um seletor gráfico que pode ser associado a um número; com esta ferramenta é possível dar dinamicidade a um objeto criado no GeoGebra. O ícone e atalho para a construção de um controle deslizante estão indicados na Figura 16.

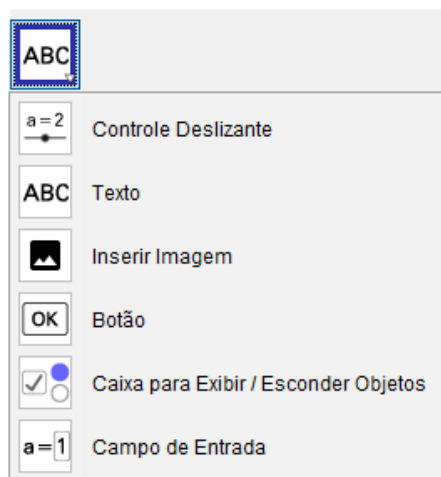


Figura 16: Menu ferramentas.

Fonte: Próprio autor

Um modo para se construir um controle deslizante é clicar no atalho para a ferramenta e em seguida no local da janela de visualização em que se deseja posicioná-lo. Em seguida é aberta a caixa de diálogo da Figura 17, onde o usuário poderá nomear o controle deslizante criado, definir um intervalo de variação e um incremento; na aba Controle Deslizante, o usuário poderá fixar o *slider* na posição horizontal ou vertical; na aba Animação terá a oportunidade de determinar a velocidade e o modo de animação do controle. Caso o

usuário selecione a caixa Aleatório (F9), o controle deslizante assumirá valores ao acaso, dentro do intervalo definido, quando for animado.

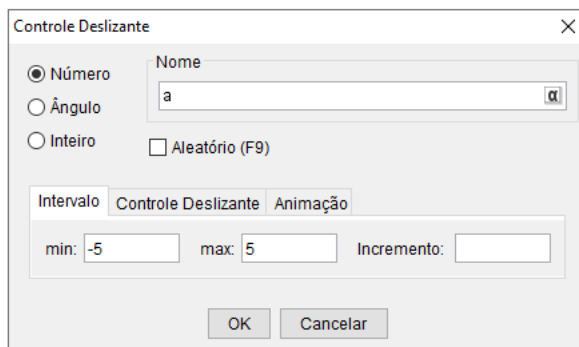


Figura 17: Janela de configuração da ferramenta controle deslizante.  
Fonte: Próprio autor

### *Texto*

Observa-se que o aplicativo permite a inclusão de textos estáticos, dinâmicos ou mistos na janela de visualização.

Um texto pode ser criado pressionando o ícone de mesmo nome indicado na Figura 16 e, em seguida, no local onde se deseja posicioná-lo na janela de visualização. O texto criado pode ser um objeto fixo, como um título ou um enunciado; pode mover-se conforme o deslocamento de um objeto gráfico relacionado, ou ainda, pode ter seus caracteres exibidos ou não de acordo com o resultado de operação algébrica envolvendo o objeto gráfico que ele representa.

O usuário pode escrever o texto utilizando linguagem LaTeX, símbolos ou referindo-se a objetos já existentes na construção mencionando seus rótulos. Um texto pode ser editado quanto à cor, tipo, tamanho e estilo da fonte. A janela de diálogo da Figura 18 auxilia o usuário na criação de um texto.

### *Botão*

O aplicativo permite a criação de botões de comando. Ressalta-se que um botão numa construção do GeoGebra pode ter a função de iniciar uma animação, a execução de ações ou de comandos definidos durante a criação do botão.

Para adicionar um botão a uma construção no GeoGebra, basta clicar no ícone botão da Figura 16 e no local onde deseja tê-lo na janela de visualização.

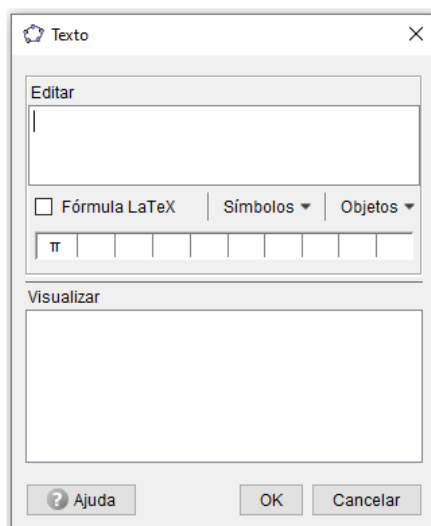


Figura 18: Janela de configuração da ferramenta texto.  
Fonte: Próprio autor

#### *Caixa para exibir/esconder objetos*

A ferramenta caixa para exibir/esconder objetos cria um campo de valor booleano (caixa de opção) que exibe objetos ao assumir o valor verdadeiro (*true*) e os oculta quando o valor assumido for falso (*false*). O ícone para esta ferramenta está indicado na Figura 16.

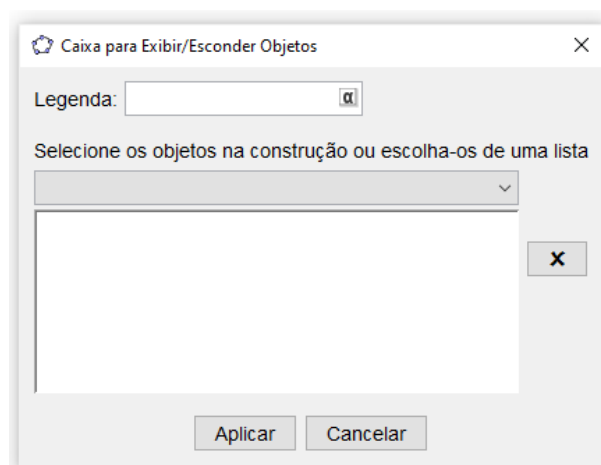


Figura 19: Janela de configuração da ferramenta caixa para exibir/esconder objetos.  
Fonte: Próprio autor

Ao clicar no ícone da ferramenta e no local da janela de visualização onde se deseja posicionar uma caixa para exibir/esconder objetos, abre-se uma janela de diálogo (Figura 19) onde insere-se uma legenda e seleciona-se um ou mais objetos que serão exibidos/escondidos ao se clicar na caixa de opção. Quando a caixa estiver selecionada, o seu valor booleano é verdadeiro e os objetos serão exibidos; quando a caixa não estiver selecionada, o valor booleano será falso e os objetos ocultados.

### *Campo de entrada*

A ferramenta inserir campo de entrada pode ser acessada pelo ícone Campo de Entrada indicado na Figura 16. Trata-se de um campo onde o usuário poderá incluir conteúdo alfanumérico, que estará vinculado a outro objeto da construção. Ao clicar no ícone e no local da janela de visualização onde se deseja construir um campo de entrada, abre-se a janela de diálogo da Figura 20, onde o usuário indica um nome para o objeto que estar sendo criado (Legenda) e seleciona a que outro objeto os dados digitados no campo de entrada serão vinculados.

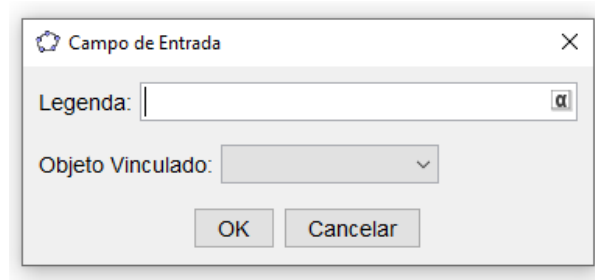


Figura 20: Janela de configuração da ferramenta campo de entrada.  
Fonte: Próprio autor

### 4.5 COMANDOS INTERNOS

O aplicativo não possui ferramenta para cada uma das suas inúmeras possibilidades, algumas funções são acessadas por meio de comandos que são digitados na barra de entrada. A seguir, relaciona-se alguns comandos que podem ser utilizados em construções para estudo de probabilidade e estatística descritiva, descreve-se sua estrutura e exemplifica uma forma de utilização.

#### *Ponto*

Nas construções realizadas no GeoGebra, em que se abordam probabilidade e estatística descritiva, os valores de um conjunto de dados podem ser indicando por pontos no eixo Ox.

Marca-se um ponto na janela de visualização digitando suas coordenadas entre parênteses, separadas por vírgula, na barra de entrada. O ponto de abscissa 1 e ordenada 3 indica que há três ocorrências do valor um no conjunto de dados e deve ser digitado da seguinte forma na barra de entrada:

$$(1, 3)$$

O GeoGebra nomeia o ponto com uma letra maiúscula do alfabeto latino, exibe uma representação algébrica na janela de álgebra e uma representação gráfica na janela de visualização.

Para se criar um ponto com um nome específico, deve-se digitar na barra de entrada o nome do ponto seguido por um sinal de igualdade e as coordenadas entre parênteses:

$$M = (1, 3)$$

### *Lista*

Um conjunto de dados é representado no GeoGebra por uma lista, que é criada digitando seu nome, seguido por um sinal de igualdade e a descrição dos dados, separados por vírgula, entre chaves.

Nome da Lista = { <valores do conjunto de dados> }

**Exemplo 4.1.** - *A lista das alturas dos elementos que compõem a amostra da Escola SP está descrita a seguir:*

$ListaA = \{1.61, 1.75, 1.67, 1.81, 1.66, 1.75, 1.67, 1.80, 1.69, 1.70, 1.75, 1.68, 1.76, 1.72, 1.75, 1.59, 1.76, 1.82, 1.73, 1.73, 1.78, 1.74, 1.62, 1.61, 1.74, 1.67, 1.73, 1.67, 1.74, 1.88\}$

- *A lista das quantidades de pessoas que residem na mesma residência que os alunos que compõem a amostra da Escola SP está descrita a seguir:*

$Moradores = \{3, 5, 4, 3, 5, 2, 4, 6, 2, 2, 5, 3, 3, 3, 4, 2, 4, 3, 3, 3, 4, 3, 4, 3, 4, 4, 5, 2, 6, 3\}$

### *Ordenar*

As observações do conjunto de dados devem ser organizadas, em ordem crescente ou decrescente, antes de serem analisadas. Para dispor os dados de uma lista em ordem crescente deve-se usar o comando *Ordenar*, com a seguinte sintaxe:

Ordenar(<Nome da Lista>)

**Exemplo 4.2.** - *Os dados da ListaA, do exemplo anterior, são ordenados utilizando a sintaxe Ordenar(ListaA), que resulta em uma nova lista com os dados dispostos em ordem crescente, renomeada para Altura:*



$Altura = \{1.59, 1.61, 1.61, 1.62, 1.66, 1.67, 1.67, 1.67, 1.67, 1.68, 1.69, 1.70, 1.72, 1.73, 1.73, 1.73, 1.74, 1.74, 1.74, 1.75, 1.75, 1.75, 1.75, 1.76, 1.76, 1.78, 1.80, 1.81, 1.82, 1.88\}$

### *Reverter*

Para organizar os dados em ordem decrescente, utiliza-se o comando *Reverter*, com a seguinte estrutura:

```
Reverter(<Nome da Lista>)
```

Observa-se que este comando apenas inverte a posição dos dados na lista. Para organizar os dados em ordem decrescente, o comando *Reverter* deve ser utilizado fazendo referência a uma lista com os elementos dispostos em ordem crescente.

**Exemplo 4.3.** - *Para arranjar os dados da lista *Altura* em ordem decrescente, utiliza-se a sintaxe *Reverter(Altura)*.*

### *Primeiros*

O comando *Primeiros*, cuja estrutura está descrita a seguir, retorna os  $n$  primeiros elementos da lista indicada na sintaxe.

```
Primeiros(<Lista>, <Número de Elementos>)
```

**Exemplo 4.4.** - *Para se obter os valores do conjunto de dados *Altura* que encontram-se abaixo da mediana, considerando que a lista *Altura* possui 30 elementos, utiliza-se a sintaxe *Primeiros(Altura, 15)*, cujo Resultado está descrito a seguir:*

$Resultado = \{1.59, 1.61, 1.61, 1.62, 1.66, 1.67, 1.67, 1.67, 1.67, 1.68, 1.69, 1.70, 1.72, 1.73, 1.73\}$

### *Últimos*

O comando *Últimos* funciona da mesma forma que o comando *Primeiros*, com a diferença que este seleciona as últimas  $n$  observações da coleção de dados, conforme descrito na estrutura do comando.

```
Últimos(<Lista>, <Número de Elementos>)
```

**Exemplo 4.5.** - *Agora, para se obter os valores do conjunto de dados *Altura* que encontram-se acima da mediana, considerando que a lista *Altura* possui 30 elementos, utiliza-se a*

*sintaxe*  $\acute{U}ltimos(Altura, 15)$ , cujo *Resultado* esta descrito a seguir:

*Resultado* = {1.73, 1.74, 1.74, 1.74, 1.75, 1.75, 1.75, 1.75, 1.76, 1.76, 1.78, 1.80, 1.81, 1.82, 1.88}

#### *Parte da lista*

Observa-se que o comando *Parte da Lista* retorna parte de uma colecao de valores; podendo ser, inclusive, uma porcao central do conjunto de dados, conforme as estruturas a seguir:

```
ParteDaLista(<Lista>, <Posicao Inicial>)
```

```
ParteDaLista(<Lista>, <Posicao Inicial>, <Posicao Final>)
```

A sintaxe  $ParteDaLista(\langle Lista \rangle, \langle Posicao Inicial \rangle)$  devolve uma colecao de valores a partir do elemento da posicao da lista indicada pelo usuario na sintaxe, inclusive, ate o seu ultimo elemento.

**Exemplo 4.6.** - *Considerando que a lista *Altura* contem 30 observacoes e desejando conhecer as ultimas cinco, pode-se obte-las utilizando a sintaxe  $ParteDaLista(Altura, 26)$ , cujo *Resultado* e o seguinte:*

*Resultado* = {1.78, 1.80, 1.81, 1.82, 1.88}

A instrucao  $ParteDaLista(\langle Lista \rangle, \langle Posicao Inicial \rangle, \langle Posicao Final \rangle)$  permite que se obtenha uma parcela de uma lista de valores, partindo do elemento indicado no campo  $\langle Posicao Inicial \rangle$  ate o valor que ocupa a  $\langle Posicao Final \rangle$ , inclusive, indicados na sintaxe do comando.

**Exemplo 4.7.** - *Desejando obter os dez elementos centrais do conjunto de dados *Altura*, utiliza-se a sintaxe  $ParteDaLista(Altura, 11, 20)$ , que relaciona os dados do 11o ao 20o valor da lista *Altura*, cujo *Resultado* e o seguinte:*

*Resultado* = {1.69, 1.70, 1.72, 1.73, 1.73, 1.73, 1.74, 1.74, 1.74, 1.75}

#### *Comprimento*

O comando *Comprimento* devolve o comprimento do objeto indicado na sua sintaxe:

```
Comprimento( <Objeto> )
```

Para o propósito desta dissertação, este comando devolve a quantidade de elementos existentes na Lista indicada no campo <Objeto>.

### *Elemento*

O comando *Elemento* devolve o valor que se encontra em determinada posição no conjunto de dados, utilizando a seguinte instrução:

```
Elemento(<Lista>, <Posição do Elemento>)
```

**Exemplo 4.8.** - *Obtém-se, por exemplo, o 15º elemento do conjunto de dados Altura utilizando a sintaxe Elemento(Altura, 15), cujo resultado é 1,73.*

### *Inserir*

Este comando insere valores no conjunto de dados utilizando a seguinte instrução:

```
Inserir(<Lista>, <Lista>, <Posição>)
```

A primeira lista citada na sintaxe do comando será incluída na segunda lista, a partir da posição indicada.

### *Anexar*

Este comando anexa um elemento ao conjunto de dados, utilizando a seguinte instrução:

```
Anexar(<Lista>, <Objeto>)
```

O comando *Anexar* acrescenta o segundo elemento indicado na sintaxe, ao primeiro. Na sintaxe Anexar(<Lista>, <Objeto>), o valor inserido no campo <Objeto> é anexado ao conjunto indicado no campo <Lista>.

### *Remover*

Ao contrário do comando *Inserir*, descrito anteriormente, o comando *Remover* retira valores de uma lista, conforme a seguinte instrução:

```
Remover(<Lista>, <Lista>)
```

Da primeira lista citada na sintaxe do comando são excluídos os elementos descritos na segunda lista.

**Exemplo 4.9.** - Observando a amostra dos alunos e alunas da Escola SP (Tabela 4), obtém-se a seguinte lista formada apenas pelas alturas das alunas:

$$Altura_F = \{1.59, 1.61, 1.61, 1.62, 1.67, 1.67, 1.67, 1.67, 1.69, 1.70, 1.73, 1.76\}$$

Utilizando a sintaxe `Remover(Altura, Altura_F)`, retira-se do conjunto de dados `Altura` os valores referentes apenas às alturas das alunas, restando o conjunto dos valores relativos às alturas dos alunos, que foi renomeado para `Altura_M`, descrito a seguir:

$$Altura_M = \{1.66, 1.68, 1.72, 1.73, 1.73, 1.74, 1.74, 1.74, 1.75, 1.75, 1.75, 1.75, 1.76, \\ 1.78, 1.80, 1.81, 1.82, 1.88\}$$

### *Elementos únicos*

Este comando retorna uma lista com os dados de um conjunto de valores, sem repetição, ordenados de modo crescente, mesmo que na lista original não o estejam, utilizando a seguinte sintaxe:

```
ElementosÚnicos(<Lista>)
```

### *Embaralhar*

Este comando embaralha os elementos que compõem uma lista, com a seguinte instrução:

```
Embaralhar(<Lista>)
```

### *Escolher elemento aleatoriamente*

Utiliza-se este comando para selecionar aleatoriamente um elemento de uma amostra, com a seguinte sintaxe:

```
EscolherElementoAleatoriamente(<Lista>)
```

A escolha é realizada considerando que cada elemento possui a mesma probabilidade de ser selecionado, mesmo que haja repetição de elementos na amostra.

### *Sequência*

O comando *Sequência* permite construir uma lista, com os elementos ordenados, utilizando uma das seguintes sintaxes:

Sequência(<Valor Final>)

Sequência(<Valor Inicial>, <Valor Final>)

Sequência(<Valor Inicial>, <Valor Final>, <Incremento>)

A instrução Sequência (<Valor Final>) retorna uma série de números inteiros, em ordem crescente, a partir da unidade até o valor indicado no campo <Valor Final>, inclusive.

A estrutura Sequência (<Valor Inicial>, <Valor Final>) devolve uma lista de números inteiros, em ordem crescente ou decrescente, a partir do <Valor Inicial> até o <Valor Final>, inclusive, indicados na sintaxe.

A ordem Sequência(<Valor Inicial>, <Valor Final>, <Incremento>) exhibe uma lista de números reais, em ordem crescente ou decrescente, partindo do <Valor Inicial> até o <Valor Final>, inclusive, variando de acordo com o <Incremento> indicado na sintaxe.

**Exemplo 4.10.** - A lista  $K = \{1.55, 1.60, 1.65, 1.70, 1.75, 1.80, 1.85, 1.90\}$  pode ser obtida com a seguinte sintaxe do comando Sequência:

$$K = \text{Sequência}(1.55, 1.90, 0.05)$$

#### Número aleatório

Este comando, diferente do descrito anteriormente, retorna um número escolhido aleatoriamente entre os números inteiros de um intervalo fechado, de acordo com as seguintes sintaxes:

NúmeroAleatório(<Mínimo(Inteiro)>, <Máximo(Inteiro)>)

NúmeroAleatório(<Mínimo(Inteiro)>, <Máximo(Inteiro)>, <Fixo(*true|false*)>)

A instrução NúmeroAleatório(<Mínimo(Inteiro)>, <Máximo(Inteiro)>) fornece um número inteiro, escolhido aleatoriamente no intervalo entre os valores mínimo e máximo indicados na instrução, sempre que for solicitado ao GeoGebra, o que pode ocorrer por intermédio de um botão de comando, ou quando a construção for atualizada.

NúmeroAleatório(<Mínimo(Inteiro)>, <Máximo(Inteiro)>, <Fixo(*true|false*)>), essa instrução permite que a escolha seja realizada uma única vez, mesmo que a construção seja atualizada não haverá uma nova escolha; para isso deve-se indicar *true* no campo <Fixo(*true|false*)> da sintaxe.

*Amostra*

O comando *Amostra* devolve um conjunto de dados (amostra), cujos valores são escolhidos aleatoriamente, com ou sem reposição, de uma lista (população) indicada pelo usuário, de acordo com as seguintes sintaxes:

```
Amostra(<Lista>, <Tamanho>)
```

```
Amostra(<Lista>, <Tamanho>, <Com Substituição>)
```

A instrução *Amostra*(*<Lista>*, *<Tamanho>*) informa ao GeoGebra que se deseja uma amostra, extraída sem reposição da população indicada no campo *<Lista>*, e dimensão descrita no espaço *<Tamanho>*.

A sintaxe *Amostra*(*<Lista>*, *<Tamanho>*, *<Com Substituição>*) permite que se obtenha uma amostra com o *<Tamanho>* indicado, extraída aleatoriamente da população citada no campo *<Lista>*, com ou sem reposição. Caso seja digitado a palavra *false* no campo *<Com Substituição>*, a seleção será sem reposição; se utilizar *true*, a escolha será com reposição.

*Classes*

O comando *Classes* possui as seguintes estruturas que fornecem classes de frequência de uma distribuição de dados, conforme parâmetros indicados nas sintaxes:

```
Classes(<Lista de Dados>, <Quantidade de Classes>)
```

```
Classes(<Lista de Dados>, <Início>, <Largura das Classes>)
```

A instrução *Classes*(*<Lista de Dados>*, *<Quantidade de Classes>*) fornece os limites de classe para uma amostra indicada no campo *<Lista de Dados>*, considerando a quantidade de classes desejada pelo usuário inserida no campo *<Quantidade de Classes>*. Neste caso, o limite inferior da primeira classe de frequência é o menor valor presente na amostra, enquanto que o maior valor do conjunto de dados é o limite superior da última classe.

**Exemplo 4.11.** - *Utilizando os dados da amostra dos alunos e alunas da Escola SP, pode-se obter os limites de sete intervalos de classe para o conjunto de dados Altura, utilizando a seguinte sintaxe *Classes*(*Altura*, 7), cujo Resultado é:*

$$\text{Resultado} = \{1.59, 1.63, 1.67, 1.71, 1.76, 1.80, 1.84, 1.88\}$$

*Significa que a primeira classe ( $i = 1$ ) varia entre 1,59 a 1,63, a segunda classe ( $i = 2$ ) de 1,63 a 1,67, sucessivamente até a sétima classe, que vai de 1,84 a 1,88.*

A estrutura `Classes(<Lista de Dados>, <Início>, <Largura das Classes>)` devolve uma lista contendo os limites de intervalos de classe conforme os parâmetros indicados pelo usuário na sintaxe do comando. O limite inferior da primeira classe é apontado no campo <Início>, a amplitude das classes deve ser inserida no espaço <Largura das Classes>, a amostra considerada deve ser indicada no campo <Lista de Dados>.

**Exemplo 4.12.** - *Agora, pretende-se obter os limites dos intervalos de classe para o conjunto de dados *Altura*, devendo o limite inferior da primeira classe ser igual a 1,55 e as classes apresentarem amplitude de 0,05. Para isso, utiliza-se a sintaxe `Classes(Altura, 1.55, 0.05)`, cujo resultado é a lista descrita a seguir, renomeada para *Classe*:*

$$\text{Classe} = \{1.55, 1.60, 1.65, 1.70, 1.75, 1.80, 1.85, 1.90\}$$

*Conforme os limites das classes da distribuição de frequência apresentada na Tabela 8.*

### *Moda*

O valor que ocorre com maior frequência no conjunto de dados é identificado utilizando o comando *Moda*, com a seguinte sintaxe:

```
Moda(<Lista de Números>)
```

O resultado é apresentado sob a forma de uma lista, visto que a amostra pode apresentar mais de uma moda.

**Exemplo 4.13.** - *Para o conjunto de dados *Altura*, a moda é obtida utilizando a sintaxe `Moda(Altura)`, cujo resultado é o conjunto  $M_o = \{1.67, 1.75\}$ .*

### *Média*

O cálculo da média aritmética dos valores que compõem um conjunto de dados é realizado pelo comando *Média*, com a seguinte instrução:

```
média(<Lista de Dados Brutos>)
```

**Exemplo 4.14.** - *A média aritmética dos valores do conjunto de dados da variável *Altura*, da amostra coletada entre os alunos e alunas da Escola SP, é obtida no GeoGebra com a sintaxe `média(Altura)`, cujo resultado é 1,72.*

### *Mediana*

A mediana de um conjunto de valores é obtida com a seguinte sintaxe do comando *Mediana*:

```
Mediana(<Lista de Dados Brutos>)
```

Utiliza-se a instrução *Mediana*(<Lista de Dados Brutos>) quando se pretende determinar a mediana de um conjunto de dados indicado no campo <Lista de Dados Brutos>.

**Exemplo 4.15.** - *A mediana do conjunto de dados Altura é obtida utilizando a sintaxe *Mediana*(*Altura*), cujo resultado é  $Md = 1,73$ .*

#### *Desvio padrão*

O desvio padrão de uma distribuição é obtido por intermédio do comando *dp*, segundo a seguinte sintaxe:

```
dp(<Lista dos Dados Brutos>)
```

Calcula-se o desvio padrão de um conjunto de dados utilizando o comando *dp*(<Lista dos Dados Brutos>), indicando a coleção de valores no campo <Lista dos Dados Brutos>.

**Exemplo 4.16.** - *O desvio padrão do conjunto de valores da variável *Altura* é obtido fazendo *dp*(*Altura*), que retorna o valor  $S = 0.07$ .*

#### *Percentil*

Para se obter os percentis (ou o quantil de ordem  $p$ ) de uma amostra, utiliza-se a seguinte sintaxe do comando *Percentil*:

```
Percentil(<Lista de Números>, <Porcentagem>)
```

No campo <Lista de Números> deve conter o nome do conjunto de dados para o qual se deseja calcular o percentil; no espaço <Porcentagem> é necessário haver a indicação do percentual relacionado à posição que se pretende calcular, devendo ser um valor entre 0,01 (1% =  $P_1$ ) e 1 (100% =  $P_{100}$ ).

**Exemplo 4.17.** - *Utilizando a lista de dados da variável *Altura* obtida da amostra de alunos e alunas da Escola SP para ilustrar este comando, tem-se os seguintes percentis:*

$$\text{Percentil}(\textit{Altura}, 0.1) = P_{10} = 1,62$$

$$\text{Percentil}(\textit{Altura}, 0.25) = P_{25} = q_1 = 1,67$$

$$\text{Percentil}(\textit{Altura}, 0.5) = P_{50} = q_2 = 1,73$$

$$\text{Percentil}(\textit{Altura}, 0.75) = P_{75} = q_3 = 1,75$$

$$\text{Percentil}(\textit{Altura}, 0.9) = P_{90} = 1,81$$



*Quartil*

Observa-se que o 1º quartil ( $q_1$ ) e o 3º quartil ( $q_3$ ) de uma amostra de dados são obtidos pelas seguintes sintaxes do comando *Quartil*:

```
Quartil1(<Lista dos Dados Brutos>)
```

```
Quartil3(<Lista dos Dados Brutos>)
```

O primeiro quartil do conjunto de dados é obtido pela sintaxe `Quartil1(<Lista de Dados Brutos>)`. Para a obtenção do terceiro quartil da coleção de valores, utiliza-se a sintaxe `Quartil3(<Lista de Dados Brutos>)`.

**Exemplo 4.18.** - *Utilizando o conjunto de dados da variável Altura para ilustrar o comando *Quartil*, tem-se o seguinte:*

*Primeiro Quartil:  $Quartil1(Altura) = 1,67$*

*Terceiro Quartil:  $Quartil3(Altura) = 1,75$*

Mesmos valores obtidos com o comando *Percentil* para  $P_{25}$  e  $P_{75}$ , ilustrado anteriormente.

*Frequência*

O comando *Frequência* retorna na janela de álgebra as quantidades de observações dos valores presentes no conjunto de dados, de acordo com o interesse do analista e a sintaxe utilizada:

```
Frequência(<Lista dos Dados Brutos>)
```

```
Frequência(<Acumulada(true|false)>, <Lista dos Dados Brutos>)
```

```
Frequência(<Lista dos Limites das Classes>, <Lista dos Dados Brutos>)
```

```
Frequência(<Acumulada(true|false)>, <Lista dos Limites das Classes>, <Lista dos  
Dados Brutos>)
```

A ordem `Frequência(<Lista de Dados Brutos>)` retorna uma lista contendo as frequências absolutas dos valores constantes no conjunto indicado no campo `<Lista de Dados Brutos>`.

**Exemplo 4.19.** - *O comando *Frequência(Moradores)* devolve uma lista composta pelas frequências absolutas dos valores do conjunto de dados moradores, conforme observa-se na Tabela 7. A lista das frequências está descrita a seguir e renomeada para *Absoluta*:*

$$Absoluta = \{5, 11, 8, 4, 2\}$$

A sintaxe ( $\langle \text{Acumulada}(true|false) \rangle$ ,  $\langle \text{Lista dos Dados Brutos} \rangle$ ) devolve uma lista com as frequências acumuladas dos valores constantes na amostra indicada no campo  $\langle \text{Lista dos Dados Brutos} \rangle$ , caso o usuário insira a palavra *true* no espaço  $\langle \text{Acumulada}(true|false) \rangle$ .

**Exemplo 4.20.** - A instrução *Frequência(true, Altura)* devolve a seguinte lista com as frequências acumuladas do conjunto de dados altura, renomeada para *Acumulada*:

$$Acumulada = \{1, 3, 4, 5, 9, 10, 11, 12, 13, 16, 19, 23, 25, 26, 27, 28, 29, 30\}$$

A sintaxe *Frequência*( $\langle \text{Lista dos Limites das Classes} \rangle$ ,  $\langle \text{Lista dos Dados Brutos} \rangle$ ) exhibe a frequência absoluta dos dados do conjunto indicado no espaço  $\langle \text{Lista dos Dados Brutos} \rangle$  organizados segundo o intervalo de classe descrito no campo  $\langle \text{Lista dos Limites das Classes} \rangle$ .

**Exemplo 4.21.** - Utilizando-se os intervalos de classe definidos na lista *Classe* (Exemplo 4.12) e a lista *Altura*, obtém-se uma lista com a distribuição apresentada na Tabela 8, utilizando a sintaxe:

$$Frequência(Classe, Altura)$$

Para se obter uma lista das frequências acumuladas de uma distribuição de dados em intervalos de classe, utiliza-se a sintaxe *Frequência*( $\langle \text{Acumulada}(true|false) \rangle$ ,  $\langle \text{Lista dos Limites das Classes} \rangle$ ,  $\langle \text{Lista dos Dados Brutos} \rangle$ ).

**Exemplo 4.22.** - A instrução *Frequência(true, Classe, Altura)* exhibe na janela de álgebra uma lista com as frequências acumuladas dos valores do conjunto de dados *Altura*, conforme Tabela 14.

#### 4.6 REPRESENTAÇÃO TABULAR DE UMA DISTRIBUIÇÃO DE DADOS

##### *Tabela de frequência*

Os dados de uma amostra podem ser resumidos em uma tabela de frequência, exibida na janela de visualização do GeoGebra, utilizando uma das seguintes sintaxes do comando *Tabela De Frequências*:

```
TabelaDeFrequências(⟨Lista de Dados Brutos⟩, ⟨Fator de Escala (opcional)⟩)
```

```
TabelaDeFrequências(⟨Acumulada(true|false)⟩, ⟨Lista de Dados Brutos⟩)
```

```
TabelaDeFrequências(<Lista dos Limites das Classes>, <Lista dos Dados Brutos>)
```

```
TabelaDeFrequências(<Acumulada(true|false)>, <Lista dos Limites das Classes>, <Lista dos Dados Brutos>)
```

Observa-se que a instrução `TabelaDeFrequências(<Lista de Dados Brutos>, <Fator de Escala (opcional)>)` devolve na janela de visualização uma tabela contendo duas colunas, que apresentam os dados dispostos em ordem crescente na coluna da esquerda, com o título "Valor", e à direita uma coluna intitulada "Frequência" contendo as quantidades de cada valor na lista de dados indicada na sintaxe. O campo `<Fator de Escala>` é opcional; caso ele seja utilizado, deve ser preenchido com um número real que multiplica a quantidade de observação de cada valor do conjunto de dados.

**Exemplo 4.23.** - *A distribuição de frequência dos valores da variável moradores, da amostra de alunos e alunas da Escola SP (Tabela 7), página 18, é obtida com a utilização de uma das seguintes sintaxes:*

```
TabelaDeFrequências(Moradores,1)
```

```
TabelaDeFrequências(Moradores)
```

`TabelaDeFrequências(<Acumulada(true|false)>, <Lista de Dados Brutos>)`. Essa sintaxe devolve uma tabela contendo uma coluna com os valores constantes na amostra e outra coluna com a frequência dos valores, podendo essa frequência ser absoluta ou acumulada.

**Exemplo 4.24.** - *Utilizando o conjunto de dados moradores para ilustração, se o usuário optar pela sintaxe `TabelaDeFrequências(false, Moradores)`, o resultado será o da Tabela 7.*

*Porém, se o usuário escolher a instrução `TabelaDeFrequências(true, Moradores)`, o GeoGebra retorna a Tabela 13, com uma distribuição de frequência acumulada.*

Tabela 13: Distribuição de frequência acumulada dos dados da variável moradores, da amostra da Escola SP.

Valor	Frequência
2	5
3	16
4	24
5	28
6	30

Fonte: Tabela 4

Na Análise Exploratória dos Dados (AED), o analista pode ter que distribuir os valores em intervalos de classes. Isso é possível no GeoGebra utilizando a instrução `TabelaDeFrequências(<Lista dos Limites das Classes>, <Lista dos Dados Brutos>)`.

**Exemplo 4.25.** - *Utilizando os intervalos de classe obtidos no Exemplo 4.12, que resultou na lista nomeada Classe, os valores da lista Altura e a seguinte sintaxe, obtém-se a distribuição das frequências absolutas em intervalos de classe do conjunto de dados altura, da amostra da Escola SP, constante na Tabela 8, página 18:*

*TabelaDeFrequências(Classe, Altura).*

A sintaxe `TabelaDeFrequências(<Acumulada(true|false)>, <Lista dos Limites das Classes>, <Lista dos Dados Brutos>)` permite ao usuário construir uma tabela com distribuição de frequência acumulada em intervalos de classe. Para isso, o usuário deve, além de indicar a lista dos limites das classes e a lista dos dados, inserir `true` no campo `<Acumulada(true|false)>` da estrutura do comando.

Tabela 14: Distribuição de frequência acumulada em intervalos de classe dos dados da variável altura, da amostra de alunos e alunas da Escola SP.

Intervalo	Contagem
1.55 – 1.6	1
1.6 – 1.65	4
1.65 – 1.7	11
1.7 – 1.75	19
1.75 – 1.8	26
1.8 – 1.85	29
1.85 – 1.9	30

Fonte: Tabela 4

**Exemplo 4.26.** - *A distribuição de frequência acumulada em intervalos de classe (Tabela 14) dos dados da variável altura, observados na amostra de alunos e alunas da Escola SP, é obtida utilizando a sintaxe:*

*TabelaDeFrequências(true, Classe, Altura).*

## 4.7 REPRESENTAÇÃO GRÁFICA DE UMA DISTRIBUIÇÃO

Nesta subseção estão descritos alguns comandos do GeoGebra para construção de representações gráficas de uma distribuição de dados.

*Gráfico de dispersão unidimensional*

O GeoGebra constrói o gráfico de dispersão unidimensional com a seguinte sintaxe do comando *Diagrama De Pontos*:

```
DiagramaDePontos(<Lista de Dados Brutos>, <Empilhar Pontos Adjascentes (opcional)>, <Fator de Escala (opcional)>)
```

Essa sintaxe permite construir um gráfico de dispersão unidimensional para os dados da amostra indicada no campo <Lista de Dados Brutos>; o usuário não é obrigado a preencher o campo <Empilhar Pontos Adjascentes (opcional)>, que permite que os pontos correspondentes a uma mesma observação sejam exibidos empilhados verticalmente, sem espaçamento entre eles, caso insira a palavra *true*; o campo <Fator de Escala (opcional)> permite que o espaçamento vertical entre os pontos de uma mesma observação seja aumentado ou reduzido, caso seja preenchido com um número real.

**Exemplo 4.27.** - *O gráfico unidimensional da distribuição da variável moradores, da amostra da Escola SP, Figura 3, página 26, que apresenta espaçamento vertical entre os pontos representativos de uma mesma quantidade, foi obtido utilizando a lista Moradores, com a sintaxe:*

```
DiagramaDePontos(Moradores)
```

*Agora, a representação gráfica da distribuição da variável moradores, indicada na Figura 21 por um gráfico de dispersão unidimensional com os pontos referentes a uma mesma estatura empilhados, foi construída utilizando a sintaxe:*

```
DiagramaDePontos(Moradores, true)
```

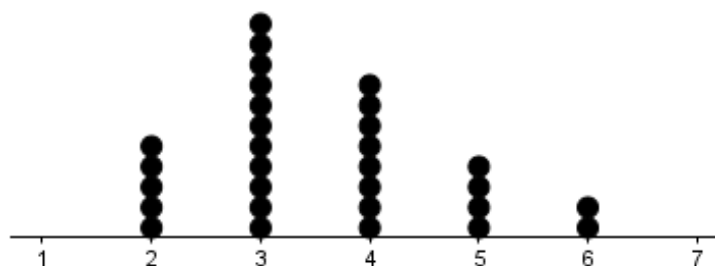


Figura 21: Gráfico de dispersão unidimensional da variável moradores, da amostra da Escola SP.

Fonte: Tabela 4

### Gráfico de barras

O comando *Diagrama de Barras* possui as seguintes sintaxes para a construção de gráficos de barras verticais, no GeoGebra:

```
DiagramaDeBarras(<Lista de Dados Brutos>, <Largura das Barras>, <Fator de Escala Vertical (opcional)>)
```

A sintaxe `DiagramaDeBarras(<Lista de Dados Brutos>, <Largura das Barras>, <Fator de Escala Vertical (opcional)>)` permite ao usuário construir um gráfico de barras verticais representando os dados da amostra indicada no campo `<Lista de Dados Brutos>`. A largura das barras pode ser configurada inserindo um valor no intervalo  $]0,1]$  no campo `<Largura das Barras>`. As alturas das barras serão proporcionais ao valor indicado no campo `<Fator de Escala Vertical (opcional)>`.

**Exemplo 4.28.** - *O gráfico de barras verticais da variável moradores, da amostra da Escola SP, ilustrado na Figura 4, página 27, foi obtido com a seguinte instrução:*

```
DiagramaDeBarras(Moradores,0.5,1)
```

### *Histograma*

O GeoGebra possui o comando *Histograma* para construir uma representação gráfica de barras contíguas, cujas bases são proporcionais aos intervalos das classes e as alturas proporcionais à frequência absoluta ou à densidade das classes, cuja sintaxe é:

```
Histograma(<Lista dos Limites das Classes>, <Lista dos Dados Brutos>)
```

```
Histograma(<Acumulada (true|false)>, <Lista dos Limites das Classes>, <Lista dos Dados Brutos>, <Usar Densidade (true|false)>, <Fator de Escala da Densidade (opcional)>)
```

A sintaxe `Histograma(<Lista dos Limites das Classes>, <Lista dos Dados Brutos>)` constrói um histograma de uma coleção de valores, especificando os limites das classes na sintaxe. Observa-se que, por padrão, utilizando esta sintaxe, as alturas das barras são dadas pelas densidades das classes.

A instrução `Histograma(<Acumulada (true | false)>, <Lista dos Limites das Classes>, <Lista dos Dados Brutos>, <Usar Densidade (true | false)>, <Fator de Escala da Densidade (opcional)>)` permite construir um histograma no qual as alturas das barras são proporcionais à frequência absoluta das classes.

**Exemplo 4.29.** - *O histograma dos dados da variável altura dos alunos e alunas da amostra da Escola SP, Figura 22, foi obtido utilizando a lista Classe (Exemplo 4.12) e a seguinte sintaxe:*

```
Histograma(false, Classe, Altura, false, 1)
```

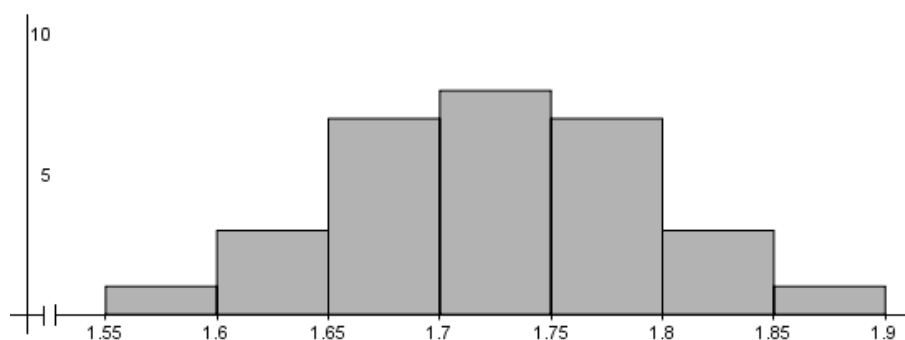


Figura 22: Histograma dos dados da variável altura dos alunos e alunas da amostra da Escola SP.

Fonte: Tabela 4

### *Boxplot*

O GeoGebra constrói um *boxplot* representativo de um conjunto de dados, expondo suas medidas de posição, dispersão e valores discrepantes, utilizando uma das seguintes sintaxes:

```
BoxPlot(<Posição Vertical>, <Escala Vertical>, <Lista de Dados Brutos>)
```

```
BoxPlot(<Posição Vertical>, <Escala Vertical>, <Lista de Dados Brutos>, <Outliers (true | false)>)
```

A sintaxe `BoxPlot(<Posição Vertical>, <Escala Vertical>, <Lista de Dados Brutos>)` produz um diagrama de caixa, disposto na horizontal, na posição da ordenada indicada no campo `<Posição Vertical>`, com largura da caixa o dobro do valor indicado no espaço `<Escala Vertical>`, correspondente aos valores da `<Lista de Dados Brutos>`, sem assinalar possíveis valores atípicos presentes na distribuição.

Para exibir os valores atípicos do conjunto de dados utiliza-se a instrução `BoxPlot(<Posição Vertical>, <Escala Vertical>, <Lista de Dados Brutos>, <Outliers (true | false)>)`. Neste caso, a indicação da palavra *true* no campo `<Outliers (true | false)>` faz com que os valores discrepantes da distribuição sejam indicados por um "x" na representação gráfica, caso existam.

**Exemplo 4.30.** - O *boxplot* dos dados da variável altura considerando somente os alunos da amostra da Escola SP (Figura 23), foi obtido utilizando a lista  $Altura_M$  (Exemplo 4.9) e a sintaxe:

```
BoxPlot(1, 0.25, Altura_M, true)
```

Observa-se que o *boxplot* está construído na horizontal, segundo os dados da lista  $Altura_M$ , na posição vertical da ordenada 1, com a caixa (retângulo) apresentando largura igual a 0,5 (o dobro de 0,25 incluído na sintaxe) e indicando o valor discrepante do conjunto de dados.

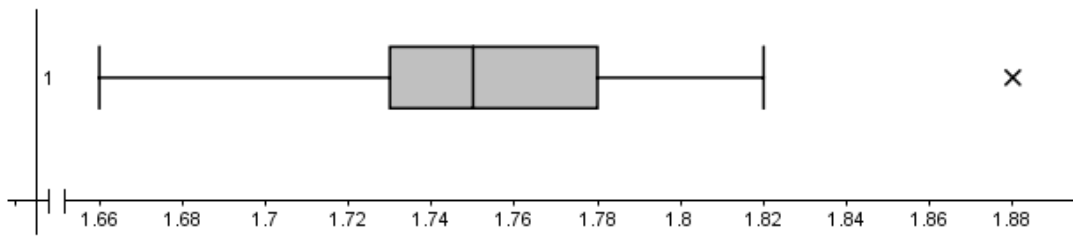


Figura 23: *Boxplot* dos dados da variável altura apenas dos alunos da amostra da Escola SP.

Fonte: Tabela 4

□



## PROPOSTA DIDÁTICA

---

Nesta seção apresenta-se uma proposta didática que trata do ensino de probabilidade e estatística descritiva utilizando ferramenta de apoio pedagógico desenvolvida no aplicativo GeoGebra com o auxílio de instruções contidas nesta dissertação. Reforça-se que a ferramenta apresentada não é o modo único de se trabalhar esse assunto utilizando o GeoGebra. As atividades em sala de aula devem depender dos alunos e do professor, dos seus conhecimentos matemáticos prévios e do interesse do grupo [11].

### 5.1 JUSTIFICATIVA

A proposta didática se justifica pela necessidade de desenvolver criticidade, permitir a compreensão dos conceitos estatísticos e suas possibilidades de aplicação; a identificação de informações estatísticas equivocadas por meio de reflexões, partindo de atividades que possibilitem o conhecimento prático, com significado de conteúdo, alinhado às necessidades do cotidiano e a utilização de tecnologia.

### 5.2 PÚBLICO

A atividade que se apresenta nesta proposta didática destina-se a estudantes do 3º ano do Ensino Médio.

### 5.3 PROBLEMA

A problematização sugerida para investigação, formulação de hipóteses e busca de solução é a seguinte:

Supondo que uma escola pública de tempo integral possua turmas de 1º, 2º e 3º ano do Ensino Médio, totalizando 200 estudantes, que sirva almoço a esses alunos e alunas e, que ao deixarem o refeitório após o almoço, os estudantes opinam sobre a qualidade da refeição servida indicando conceito ótimo, bom ou ruim. O professor de matemática propõe aos estudantes de uma das turmas do 3º ano, composta por 20 discentes, que realizem uma pesquisa entre os alunos e alunas do Ensino Médio, com o interesse de identificar soluções para os problemas relacionados a seguir; o tempo e recursos escassos não permitem que sejam consultados todos os estudantes, por isso deve-se selecionar uma amostra dos discentes:

1. O conceito dado pelos alunos e alunas do Ensino Médio ao almoço servido na cantina da escola dependente do índice de massa corporal (IMC) do estudante?
2. Que tipo de alimentação (mais calórica ou menos calórica) a nutricionista da escola deve servir no almoço, considerando a classificação do peso dos alunos e alunas pesquisados?
3. Que tipo de atividade o professor de Educação Física da escola deve adotar nas suas aulas para as turmas do Ensino Médio, considerando a classificação do peso dos alunos e alunas pesquisados?
4. Considerando o IMC dos estudantes do Ensino Médio da escola pesquisada, que grupo apresenta maior regularidade, quanto ao valor da massa corporal, os alunos ou as alunas?

Os resultados devem ser apresentados em relatório contendo tabelas e gráficos, para justificar a análise estatística, as conjecturas realizadas e as hipóteses formuladas para a obtenção da solução aos problemas sugeridos.

#### 5.4 OBJETIVO GERAL

O objetivo da proposta didática é propiciar condições para que o estudante compreenda conceitos estatísticos e desenvolva capacidade de raciocínio, de resolução de problemas e de comunicação, bem como o espírito crítico, criativo e cooperativo, que possam resultar em aprendizagem real e significativa.

#### 5.5 OBJETIVOS ESPECÍFICOS

Alinhada às habilidades descritas na BNCC [7] para a área de Matemática e suas Tecnologias no Ensino Médio e às orientações contidas nos relatórios da ASA [1] [2], a proposta didática apresentada tem as seguintes finalidades específicas:

1. Planejar e executar pesquisa amostral utilizando dados reais com um contexto e propósito;
2. Resolver problemas que envolvam cálculo e interpretação das medidas de tendência central e de dispersão;
3. Interpretar tabelas de frequências com base nos dados obtidos na pesquisa amostral;
4. Interpretar e comparar conjuntos de dados estatísticos por meio de diferentes diagramas e gráficos; e
5. Usar tecnologia para explorar conceitos e analisar dados.

## 5.6 METODOLOGIA

Planejou-se a execução da atividade em cinco aulas, conforme descrição a seguir:

Tempo: 2 aulas. Realiza-se revisão teórica dos seguintes assuntos de estatística: variáveis, população e amostra, amostragem, frequências e distribuições, média, mediana e desvio padrão, gráfico de dispersão unidimensional, histograma e *boxplot*.

Tempo: 1 aula. Apresenta-se a ferramenta de apoio que irá auxiliar os estudantes na análise dos dados. Realiza-se uma leitura e explica-se o problema e o que se deseja como solução. Apresenta-se sucintamente as fases da pesquisa (planejamento e coleta, organização, análise e interpretação dos dados e apresentação dos resultados). Divide-se igualmente os 20 estudantes da turma em quatro grupos, numerados de 1 a 4 de acordo com o número da questão que couber ao grupo buscar uma solução. Escolhe-se dois alunos de cada grupo para compor uma equipe de coleta de dados. Orienta-se a equipe de coleta de dados a separar todos os estudantes do Ensino Médio em estratos, conforme sexo e turma (Feminino - 1º ano, Masculino - 1º ano, Feminino - 2º ano, ..., Masculino - 3º ano) e, após isso, selecionar de forma aleatória, proporcionalmente ao tamanho da amostra (30 alunos), os integrantes desses estratos que serão consultados sobre altura (em metro), massa (em quilograma), sexo (M - masculino ou F - feminino) e conceito sobre o almoço servido na escola (O - ótimo, B - bom, R - ruim).

Tempo: 1 aula. Conduz-se a coleta dos dados dos 30 alunos que compõem a amostra, realizada pela equipe de coleta de dados, com o auxílio de uma balança e uma fita métrica, sob observação e orientação do professor.

Após a realização da coleta, os grupos separadamente se reúnem em horários alternativos, durante uma semana, e os integrantes discutem e analisam os dados obtidos, fazem conjecturas, formulam hipóteses e buscam soluções para as questões, com orientação do professor e uso da ferramenta de apoio desenvolvida no GeoGebra (os estudantes recebem a ferramenta pronta, não a constroem).

Tempo: 1 aula. Apresentação dos resultados. Cada grupo faz a apresentação do seu relatório para os demais grupos na sala de aula, destacando as hipóteses sugeridas, as soluções encontradas e os fundamentos estatísticos que as baseiam, com o tempo máximo de 10 minutos para cada grupo.

## 5.7 RECURSOS DIDÁTICOS

Para esta proposta, pretende-se utilizar os seguintes recursos didáticos: quadro, giz ou marcador, apagador, livro didático, computador, projetor, impressora, ferramenta de apoio didático desenvolvida no GeoGebra, fita métrica, balança, caderno e material para anotações.

## 5.8 FERRAMENTA DE APOIO

Apresenta-se a ferramenta de apoio desenvolvida no GeoGebra para a realização da atividade sugerida nesta proposta didática.

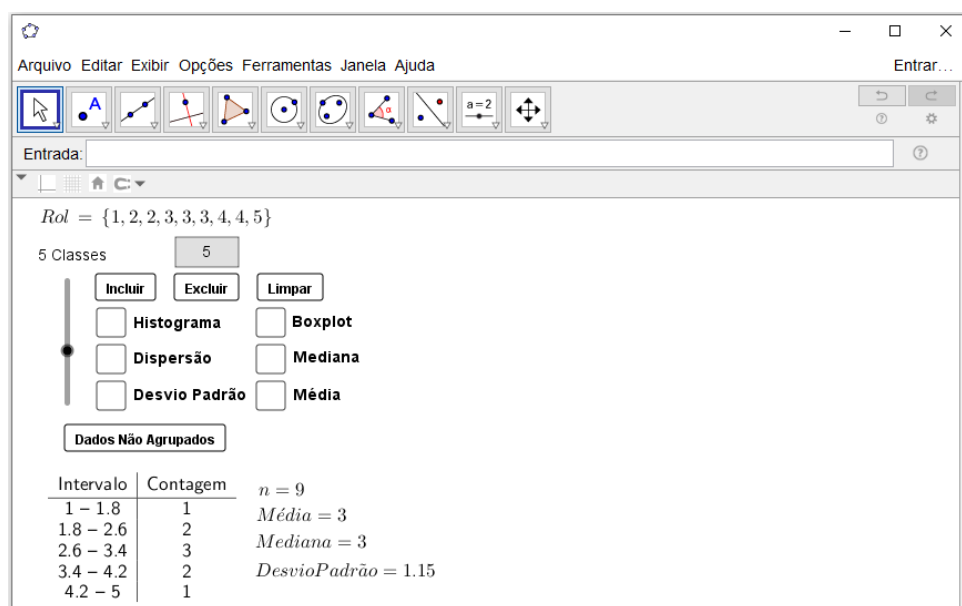


Figura 24: Apresentação da ferramenta de apoio pedagógico para análise de conjunto de dados.

Fonte: Próprio autor

A ferramenta foi desenvolvida com propósito didático, a fim de permitir que estudantes do Ensino Médio compreendam conceitos estatísticos a partir da representação gráfica de medidas de resumo e de dispersão, obtidas pela análise de conjuntos de dados formados por até 30 elementos, cujos valores podem variar entre 0 e 50, inclusive.

Observa-se a existência de um painel de controle à esquerda na janela de visualização, enquanto que à direita existe uma área para exibição dos objetos gráficos.

Na parte superior do painel de controle há um campo de entrada, onde o usuário poderá digitar o valor que deseja incluir no conjunto de dados  $Rol$ . A inclusão do valor digitado no campo de entrada no conjunto ocorrerá ao se clicar no botão <Incluir>; clicando no botão <Excluir>, exclui-se do conjunto de dados o valor digitado no campo de entrada; clicando no botão <Limpar>, exclui-se todos os valores do conjunto de dados.

A ferramenta exibe representações gráficas da distribuição dos valores do conjunto na área à direita na janela de visualização. Para isso, o usuário deve clicar na caixa de seleção referente ao gráfico desejado: Histograma, Gráfico de Dispersão Unidimensional, *Boxplot*, Desvio Padrão, Média e Mediana.

Verifica-se que é possível alternar a apresentação dos dados do conjunto entre uma tabela de distribuição de frequência absoluta e uma tabela de distribuição de frequência em intervalos de classe, ao se clicar no botão <Dados Agrupados>. O número de classes em que os dados serão exibidos na distribuição de frequência em intervalos de classe, ou no histograma, pode ser selecionado no controle deslizante vertical.

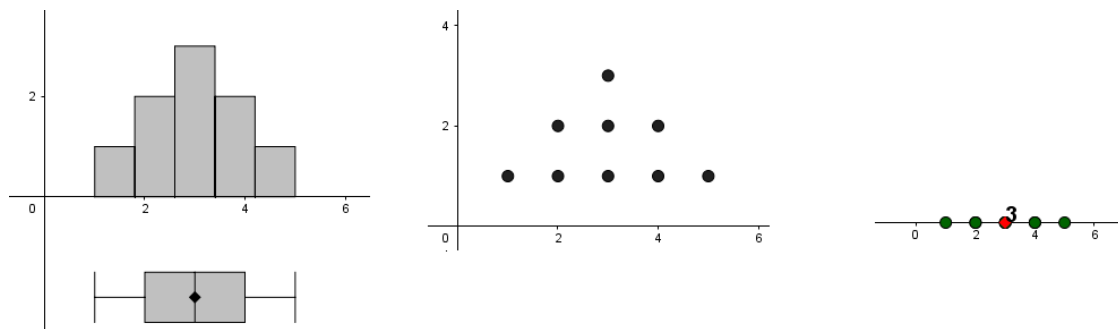


Figura 25: Representações gráficas exibidas pela ferramenta de apoio pedagógico para análise de conjunto de dados.

Fonte: Próprio autor

Ilustra-se, na Figura 25, representações gráficas do conjunto de dados  $Rol = \{1, 2, 2, 3, 3, 3, 4, 4, 5\}$  exibidas pela ferramenta de apoio.

## 5.9 PROTOCOLO DE CONSTRUÇÃO

Registra-se que os passos seguidos durante a construção de um objeto no GeoGebra pode ser acessado no *menu* Exibir  $\rightarrow$  Protocolo de Construção, ou por intermédio do atalho Ctrl + Shift + L. As representações gráficas utilizadas para análise dos dados desta proposta podem ser construídas com as informações constantes no corpo da dissertação. No entanto, detalha-se os procedimentos realizados durante a construção da ferramenta de apoio pedagógico apresentada:

1. Criar um controle deslizante denominado  $a$ , variando entre 15 e 50, com incremento 0,1. Configurar para que este objeto permaneça oculto;
2. Construir um campo de entrada, denominado  $ct1$ , vinculado ao controle deslizante  $a$ ;
3. Criar uma lista com o nome  $Dados$  digitando a seguinte sintaxe na barra de entrada e clicando <Enter>:

$$Dados = \{ \}$$

4. Construir três botões com os nomes *btincluir*, *btexcluir*, *btlimpar*;
5. Digitar a seguinte sintaxe na barra de entrada e clicar <Enter>. Esta ação cria uma variável de nome *b*, que retorna a quantidade de elementos existentes no conjunto *Dados*.  

$$b = \text{Comprimento}(\text{Dados})$$
6. Configurar o botão *btincluir* para adicionar o valor digitado no campo de entrada (*ct1*) ao conjunto *Dados*, caso esse conjunto tenha até 30 elementos.  
 Para isso, clicar com o botão direito do *mouse* sobre o botão *btincluir*, depois clicar na aba <Propriedades>, em seguida na aba <Programação> e digitar o seguinte comando na aba <Ao Clicar>:  

$$\text{Se}[b \leq 30, \text{DefinirValor}[\text{Dados}, \text{Anexar}[\text{Dados}, a]]]$$
7. Configurar o botão *btexcluir* para excluir do conjunto *Dados* o valor digitado no campo de entrada *ct1*, caso o valor exista entre os elementos do conjunto.  
 Para isso, clicar com o botão direito do *mouse* sobre o botão *btexcluir*, clicar na aba <Propriedades>, em seguida na aba <Programação> e digitar o seguinte comando na aba <Ao Clicar>:  

$$\text{DefinirValor}[\text{Dados}, \text{Remover}[\text{Dados}, \{a\}]]$$
8. Configurar o botão *btlimpar* para excluir todos os valores existentes no conjunto *Dados*.  
 Para isso, clicar com o botão direito do *mouse* sobre o botão *btlimpar*, clicar na aba <Propriedades>, em seguida na aba <Programação> e digitar o seguinte comando na aba <Ao Clicar>:  

$$\text{DefinirValor}[\text{Dados}, \{\}]$$
9. Criar um rol com os elementos do conjunto *Dados* digitando a seguinte sintaxe na caixa de entrada e clicando <Enter>:  

$$\text{Rol} = \text{Ordenar}(\text{Dados})$$
10. Clicar sobre a representação algébrica do *Rol*, na janela de álgebra, e arrastar para a janela de visualização para criar uma caixa de texto (*texto1*) contendo o conjunto de dados que será analisado;
11. Construir uma caixa de texto (*texto2*) para exibir o número de elementos no conjunto de dados em análise.  
 Para isso, clicar na ferramenta Texto e no local onde se deseja posicionar a caixa de texto, escrever no campo <Editar> da janela de diálogo  $n = b$  (o *b* deve ser selecionado no *menu* suspenso <Objetos>);
12. Calcular a média (*Me*) do conjunto de dados *Rol*, digitando a seguinte sintaxe na barra de entrada e clicando <Enter>:  

$$\text{Me} = \text{média}(\text{Rol})$$

13. Criar uma caixa de texto (*texto3*) para exibir o valor da média aritmética dos elementos do conjunto analisado;

14. Calcular a mediana (*Md*) do conjunto de dados *Rol*, digitando a seguinte sintaxe na barra de entrada e clicando <Enter>:

$$Md = Mediana(Rol)$$

15. Criar uma caixa de texto (*texto4*) para exibir o valor da mediana do conjunto em análise;

16. Calcular o desvio padrão (*DP*) do conjunto de dados *Rol*, digitando a seguinte sintaxe na barra de entrada e clicando <Enter>:

$$DP = dp(Rol)$$

17. Criar uma caixa de texto (*texto5*) para exibir o valor do desvio padrão calculado para o conjunto de dados em análise;

18. Construir um gráfico de dispersão unidimensional (*pontos*) para os elementos do conjunto de dados *Rol*:

$$pontos = DiagramaDePontos(Rol, false, 1)$$

19. Criar um segmento de reta (*c*) que representa graficamente a média dos valores do conjunto de dados *Rol*:

$$c = Segmento((Me, 0), (Me, Me))$$

Obs: o comando *Segmento* constrói um segmento de reta entre dois pontos especificados na sintaxe

$$Segmento(< Ponto >, < Ponto >)$$

Este comando não está descrito no corpo da dissertação por fugir ao tema do trabalho.

20. Construir um segmento de reta (*d*) que representa graficamente a mediana do conjunto de dados em análise:

$$d = Segmento((Md, 0), (Md, Md))$$

21. Construir um *Boxplot* (*box*) representativo dos dados do conjunto em análise:

$$box = BoxPlot(-2, 0.5, Rol, true)$$

22. Construir um controle deslizante (*i*), variando de 2 a 9, inclusive, com incremento 1;

23. Criar um histograma (*histo*) dos dados do conjunto, tendo a quantidade de intervalos de classe definida pelo controle deslizante *i*:

$$histo = Histograma(Clases(Rol, i), Rol)$$

24. Criar um ponto ( $C$ ) que indicará a média do conjunto de dados em análise:

$$C = (Me, 0)$$

25. Construir uma lista ( $Desvio$ ) cujos elementos são pontos de ordenada zero e abscissa igual a cada um dos valores do conjunto de dados  $Rol$ . Essa será a representação gráfica do Desvio Padrão:

$$Desvio = Sequência((Elemento(Rol, i), 0), i, 1, Comprimento(Rol))$$

26. Construir uma caixa para exibir/esconder objetos ( $f$ ), sem legenda e sem objeto relacionado.

27. Construir uma tabela com a distribuição de frequência absoluta ( $texto6$ ) dos dados do conjunto em análise:

$$TabelaDeFrequências(Rol, 1)$$

28. Configurar o  $texto6$  para ser exibido quando a caixa para exibir/esconder objetos  $f$  assumir valor verdadeiro ( $true$ ).

Para isso, clicar com o botão direito do *mouse* sobre a tabela de distribuição de frequência absoluta, clicar na aba <Propriedades>, em seguida na aba <Avançado>, digitar o seguinte comando no campo <Condição para Exibir Objeto(s)>:

$$f = true$$

29. Criar uma tabela com a distribuição dos dados em intervalos de classe ( $texto7$ ), com o número de classes definidas pelo controle deslizante  $i$ :

$$TabelaDeFrequências(Classes(Rol, i), Rol)$$

30. Configurar o  $texto7$  para ser exibido quando a caixa para exibir/esconder objetos  $f$  assumir valor falso ( $false$ ).

Para isso, clicar com o botão direito do *mouse* sobre a tabela de distribuição dos dados em intervalos de classe, clicar na aba <Propriedades>, em seguida clicar na aba <Avançado>, digitar o seguinte comando no campo <Condição para Exibir Objeto(s)>:

$$f = false$$

31. Construir um botão ( $btdados$ ) para alternar a exibição tabular dos dados entre a tabela de distribuição de frequência absoluta ( $texto6$ ) e a tabela de distribuição em intervalos de classe ( $texto7$ ).

32. Configurar o botão  $btdados$  para, ao ser clicado, alternar a exibição do  $texto6$  e  $texto7$ .

Para isso, clicar com o botão direito do *mouse* sobre o botão  $btdados$ , depois clicar na aba <Propriedades>, na aba <Programação>, depois na aba <Ao Clicar> e digitar o seguinte código:

$$Se[f==true, DefinirValor[f, false], DefinirValor[f, true]]$$



Se[ $f == true$ , DefinirLegenda[btdados, "Dados Agrupados"], DefinirLegenda[btdados, "Dados Não Agrupados"]]

Obs: A primeira linha do código é um comando condicional que, a cada clique no botão *btdados*, alterna o valor assumido pela caixa para exibir/esconder objetos *f* entre *true* e *false*;

A segunda linha do código é um comando condicional que, a cada clique no botão *btdados*, alterna a legenda exibida por esse botão entre "Dados Agrupados", quando *f* assumir o valor *true* e "Dados Não Agrupados", quando *f* assumir o valor *false*.

33. Construir três caixas para exibir/esconder objetos (*hist*, *disp*, *desv*) para controlar a exibição das seguintes representações gráficas do conjunto de dados: histograma, gráfico de dispersão unidimensional e desvio padrão, respectivamente.

34. Configurar a caixa para exibir/esconder objetos *hist* para que, ao ser selecionada, exiba a representação gráfica do histograma (*histo*) e oculte as caixas para exibir/esconder objetos *disp* e *desv*.

Para isso, clicar com o botão direito do *mouse* sobre a representação gráfica do histograma na janela de visualização, ou na sua representação algébrica na janela de álgebra, em seguida clicar na aba <Propriedades>, na aba <Avançado> e digitar o seguinte código no campo <Condição para Exibir Objeto(s)>:

$hist = true$

Ainda, clicar com o botão direito do *mouse* sobre a caixa para exibir/esconder objetos *hist*, em seguida na aba <Propriedades>, depois em <Avançado>, digitar o seguinte código no campo <Condição para Exibir Objeto(s)>:

$(\neg disp) \wedge (\neg desv)$

Obs: ao clicar no campo <Condição para Exibir Objeto(s)>, O GeoGebra disponibiliza um ícone ( $\alpha$ ) para uma tabela que auxilia na inserção de caracteres especiais.

35. Configurar a caixa para exibir/esconder objetos *disp* para que, ao ser selecionada, exiba o gráfico de dispersão unidimensional (*pontos*) e oculte as caixas para exibir/esconder objetos *hist* e *desv*.

Para isso, clicar com o botão direito do *mouse* sobre a representação gráfica do gráfico de dispersão unidimensional na janela de visualização, ou na sua representação algébrica na janela de álgebra, em seguida clicar na aba <Propriedades>, na aba <Avançado> e digitar o seguinte código no campo <Condição para Exibir Objeto(s)>:

$disp = true$

Ainda, clicar com o botão direito do *mouse* sobre a caixa para exibir/esconder objetos *disp*, em seguida na aba <Propriedades>, depois em <Avançado>, digitar o seguinte código no campo <Condição para Exibir Objeto(s)>:

$(\neg hist) \wedge (\neg desv)$

36. Configurar a caixa para exibir/esconder objetos *desv* para que, ao ser selecionada, exiba a representação gráfica do desvio padrão (composta pelo ponto *C* e pela lista *Desvio*), oculte as caixas para exibir/esconder objetos *hist* e *disp*.

Para isso, clicar com o botão direito do *mouse* sobre a representação gráfica de cada um dos objetos que compõem o desvio padrão na janela de visualização, ou na representação algébrica na janela de álgebra, em seguida clicar na aba <Propriedades>, na aba <Avançado> e digitar o seguinte código no campo <Condição para Exibir Objeto(s)>:

*desv = true*

Ainda, clicar com o botão direito do *mouse* sobre a caixa para exibir/esconder objetos *disp*, em seguida na aba <Propriedades>, depois em <Avançado>, digitar o seguinte código no campo <Condição para Exibir Objeto(s)>:

$(\neg hist) \wedge (\neg disp)$

37. Construir três caixas para exibir/esconder objetos (*plot*, *media*, *mediana*) para exibir a representação gráfica do *Boxplot*, da média e da mediana, respectivamente.

38. Configurar a caixa para exibir/esconder objetos *plot* para exibir a representação gráfica do *Boxplot* (*box*) e ser ocultada quando a caixa *desv* for selecionada.

Para isso, clicar com o botão direito do *mouse* na representação gráfica do *Boxplot* na janela de visualização, ou na representação algébrica na janela de álgebra, clicar na aba <Propriedades>, na aba <Avançado>, digitar o seguinte código no campo <Condição para Exibir Objeto(s)>:

*plot = true*

Em seguida, clicar com o botão direito do *mouse* sobre a caixa *plot*, em seguida na aba <Propriedades>, na aba <Avançado> e digitar o seguinte comando no campo <Condição para Exibir Objeto(s)>:

$\neg desv$

Depois, no campo <Condição para Exibir Objeto(s)> da caixa para exibir/esconder objetos *desv*, alterar o código existente para:

$(\neg hist) \wedge (\neg disp) \wedge (\neg plot)$

39. Configurar a caixa para exibir/esconder objetos *mediana* a fim de exibir a representação gráfica da mediana (segmento de reta *d*) dos valores do conjunto de dados em análise e ser ocultada quando alguma das caixas para exibir/esconder objetos *plot* ou *desv* forem selecionadas.

Para isso, clicar com o botão direito do *mouse* sobre a representação gráfica do segmento de reta *d* na janela de visualização, ou da sua representação algébrica na janela de álgebra, em seguida na aba <Propriedades>, depois na aba <Avançado> e digitar o seguinte comando no campo <Condição para Exibir Objeto(s)>:

*mediana = true*

Depois, digitar o seguinte comando no campo <Condição para Exibir Objeto(s)>

da caixa para exibir/esconder objetos *mediana*:

$(\neg plot) \wedge (\neg desv)$

40. Configurar a caixa para exibir/esconder objetos *media* a fim de exibir a representação gráfica da média (segmento de reta *c*) dos valores do conjunto de dados em análise e ser ocultada quando alguma das a caixas para exibir/esconder objetos *plot* ou *desv* forem selecionadas.

Para isso, clicar com o botão direito do *mouse* sobre a representação gráfica do segmento de reta *c* na janela de visualização, ou da sua representação algébrica na janela de álgebra, em seguida na aba <Propriedades>, depois na aba <Avançado> e digitar o seguinte comando no campo <Condição para Exibir Objeto(s)>:

$media = true$

Depois, digitar o seguinte comando no campo <Condição para Exibir Objeto(s)> da caixa para exibir/esconder objetos *media*:

$(\neg plot) \wedge (\neg desv)$

Em seguida, no campo <Condição para Exibir Objeto(s)> da caixa para exibir/esconder objetos *plot*, alterar o código existente para:

$(\neg desv) \wedge (\neg media) \wedge (\neg mediana)$

Finalmente, no campo <Condição para Exibir Objeto(s)> da caixa para exibir/esconder objetos *desv*, alterar o código existente para:

$(\neg hist) \wedge (\neg disp) \wedge (\neg plot) \wedge (\neg media) \wedge (\neg mediana)$

## 5.10 ATIVIDADE

Supondo que os dados coletados pela equipe de coleta, consultando os elementos selecionados aleatoriamente nas turmas de 1º, 2º e 3º ano do Ensino Médio da escola pesquisada, foram organizados e cada grupo realizou suas análises com base nos dados descritos na Tabela 15.

Tabela 15: Dados coletados de uma amostra de 30 alunos e alunas da escola pública de tempo integral pesquisada sobre sexo, altura, massa e conceito dado ao almoço servido na cantina.

Nº	Sexo	Altura	Massa	Conceito
1	F	1,59	48	R
2	F	1,61	59	B
3	F	1,61	68	O
4	F	1,62	61	B
5	M	1,66	70	O
6	F	1,67	65	B
7	F	1,67	68	B
8	F	1,67	72	O
9	F	1,67	74	O
10	M	1,68	64	R
11	F	1,69	70	O
12	F	1,70	65	R
13	M	1,72	79	B
14	F	1,73	70	B
15	M	1,73	86	O
16	M	1,73	90	O
17	M	1,74	74	O
18	M	1,74	89	B
19	M	1,74	99	O
20	M	1,75	62	R
21	M	1,75	68	R
22	M	1,75	79	O
23	M	1,75	97	O
24	F	1,76	84	O
25	M	1,76	91	B
26	M	1,78	85	O
27	M	1,80	84	B
28	M	1,81	89	B
29	M	1,82	95	O
30	M	1,88	102	B

Fonte: Próprio autor

O IMC é obtido pela divisão da massa do indivíduo pelo quadrado da sua altura, cujo resultado, ao ser comparado aos níveis indicados na Figura 26, permite classificar a condição corporal do aluno em um dos intervalos descritos na ilustração.

O IMC é dado por:

$$IMC = \frac{m}{a^2} \quad (25)$$

Sendo a massa  $m$  em quilograma e a altura  $a$  em metro.

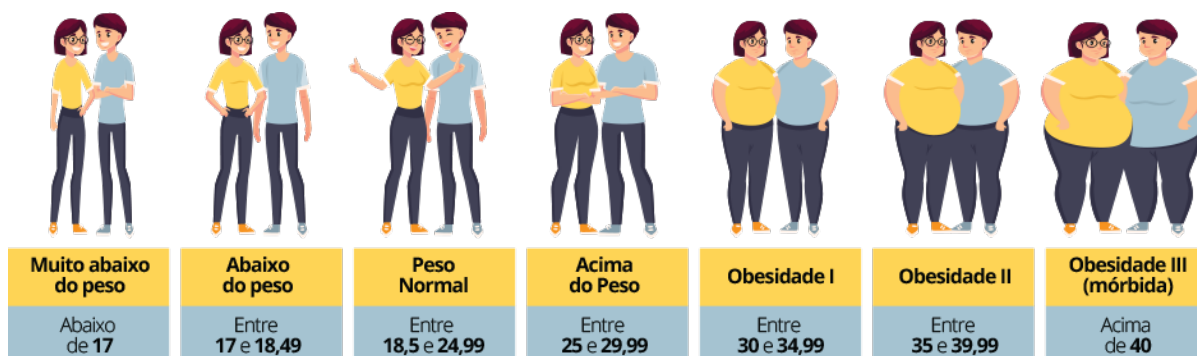


Figura 26: Ilustração dos intervalos de classificação da massa do indivíduo, segundo o índice de massa corporal.

Fonte: [www.sigagermed.com.br/calculo-imc](http://www.sigagermed.com.br/calculo-imc)

No problema 1, solicita-se que seja verificada a existência de relação entre os conceitos atribuídos ao almoço servido na escola e o IMC dos estudantes do Ensino Médio; trata-se de uma associação entre uma variável qualitativa e uma variável quantitativa. Nesse caso, deve-se analisar o que ocorre com a variável quantitativa dentro de cada variável qualitativa, o que pode ser feito por intermédio de medidas resumo, histogramas ou *boxplot* [20]. Para isso, calcula-se o IMC de cada estudante, faz-se uma distribuição das frequências do IMC e constrói-se uma tabela contendo as medidas resumo dessa nova variável, segundo a variável qualitativa (Tabela 16).

Tabela 16: Medidas resumo para a variável IMC dos alunos e alunas da amostra, segundo a variável conceito, na escola pesquisada.

Conceito	$n$	$\bar{x}$	$S^2$	$S$	$x_1$	$q_1$	$q_2$	$q_3$	$x_n$
Ruim	5	21,32	2,12	1,46	18,99	19,62	22,20	22,59	22,68
Bom	11	25,80	6,52	2,55	22,76	23,35	26,32	29,12	29,40
Ótimo	14	27,47	6,11	2,47	24,44	25,60	26,53	29,40	32,70
<b>Amostra</b>	<b>30</b>	<b>25,93</b>	<b>10,45</b>	<b>3,23</b>	<b>18,99</b>	<b>23,31</b>	<b>25,88</b>	<b>28,73</b>	<b>32,70</b>

Fonte: Tabela 15

Os valores apresentados na Tabela 16 sugerem uma dependência entre o conceito atribuído ao almoço servido na escola e o IMC dos estudantes. A variância da variável IMC para toda a amostra mede a dispersão dos dados de forma geral; o valor da variância da variável IMC dentro de cada categoria qualitativa é menor do que a da amostra, isso significa que existe uma relação entre as variáveis [20]. A dependência entre Conceito e IMC pode ser observada pela representação gráfica do *boxplot* da Figura 27, construída no GeoGebra.

Os problemas 2 e 3 buscam informações para a escolha de uma alimentação adequada a ser servida pela cantina da escola e para a definição do tipo de exercício físico a ser

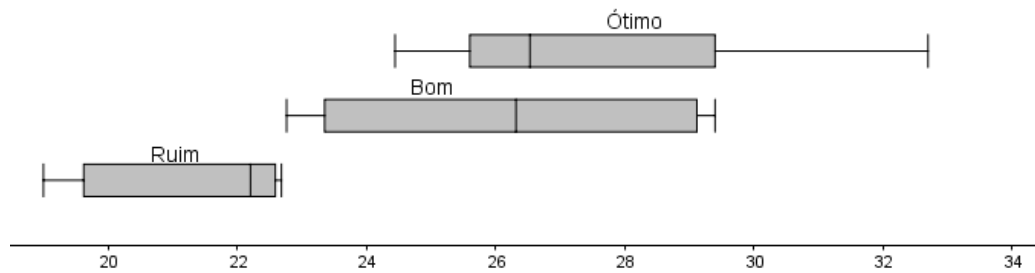


Figura 27: *Boxplot* da variável IMC, segundo a variável conceito.  
Fonte: Tabela 15

praticado nas aulas de educação física, respectivamente, considerando-se, em ambos os casos, o IMC dos estudantes avaliados.

A média pode não ser uma medida adequada para representar um conjunto de dados, pois é afetada, exageradamente, por valores extremos; porém, torna-se ótima medida de representação se a distribuição dos dados aproximar-se de uma distribuição normal [20]. Ao passo que a mediana é uma medida de posição robusta, mais resistente a valores extremos que a média.

Caso a diferença  $q_2 - x_{(1)}$  aproxime-se do valor de  $x_{(n)} - q_2$ , a distribuição dos dados é aproximadamente simétrica [20]. Isso autoriza o uso da média para representar o conjunto de dados.

Considerando a amostra coletada entre os alunos e alunas do Ensino Médio da escola pesquisada, cujos dados estão apresentados na Tabela 15 e as medidas resumo dispostas na Tabela 16, tem-se:

$$q_2 - x_{(1)} = x_{(n)} - q_2$$

$$25,88 - 18,99 \approx 32,70 - 25,88$$

$$6,89 \approx 6,82$$

Portanto, os valores são aproximados, o que indica uma distribuição aproximadamente simétrica; logo, a média amostral ( $\bar{x} = 25,86$ ) representa o IMC médio dos alunos do Ensino Médio da escola pesquisada. Comparando esse valor com os dados ilustrados na Figura 26, observa-se que os alunos do Ensino Médio da escola pesquisada podem ser classificados como Acima do Peso e, portanto, justifica que a nutricionista adote uma alimentação menos calórica no almoço servido na cantina da escola aos alunos do Ensino Médio e autoriza o professor de Educação Física a realizar atividades com exercícios aeróbicos para as turmas do Ensino Médio, durante suas aulas.

O último problema sugerido nesta proposta didática trata de identificar que grupo apresenta maior regularidade na distribuição da variável IMC.

O desvio padrão de um conjunto de dados deve ser interpretado como uma medida que indica o quanto, em média, os valores desse conjunto se desviam da média do conjunto.

Quanto mais heterogêneo, mais espalhados estão os valores que compõem o conjunto de dados, maior será o seu desvio padrão [16].

Portanto, a Tabela 17 apresenta as medidas resumo da variável IMC, segundo a variável Sexo, da amostra dos estudantes do Ensino Médio da escola pesquisada.

Tabela 17: Medidas resumo para a variável IMC dos alunos e alunas da amostra, segundo a variável sexo, na escola pesquisada.

Sexo	$n$	$\bar{x}$	$S^2$	$S$	$x_1$	$q_1$	$q_2$	$q_3$	$x_n$
Feminino	12	24,06	4,59	2,14	18,99	23,00	23,89	26,03	29,38
Masculino	18	27,05	10,20	3,19	20,24	25,40	27,00	29,38	32,70
<b>Amostra</b>	<b>30</b>	<b>25,93</b>	<b>10,45</b>	<b>3,23</b>	<b>18,99</b>	<b>23,31</b>	<b>25,88</b>	<b>28,73</b>	<b>32,70</b>

Fonte: Tabela 15

Ao comparar os valores do desvio padrão da variável IMC, segundo a variável Sexo separada em feminino e masculino, percebe-se que o conjunto de dados do IMC do grupo formado pelas alunas do Ensino Médio da escola pesquisada é mais regular, mais homogêneo, dado que possui menor variação em torno da média, indicado pelo desvio padrão menor em comparação com o desvio padrão do grupo formado pelos alunos.

## 5.11 AVALIAÇÃO

A avaliação da tarefa sugerida na proposta didática está voltada para o cumprimento de metas e o desenvolvimento de competências pelos aprendentes, é realizada de modo constante e constituída por uma avaliação do ensino, uma avaliação do aprendizado e uma avaliação da atividade, considerando os seguintes tópicos:

### *Avaliação do Ensino*

- O tempo destinado para a revisão teórica do conteúdo de estatística é suficiente?
  - A ferramenta de apoio apresentada atende às necessidades dos estudantes para análise dos dados?
  - O professor acompanhou e orientou a coleta de dados realizada pela equipe de coleta de dados?
  - O professor foi consultado, sugeriu ou orientou a análise dos dados realizada pelos grupos?

### *Avaliação do Aprendizado*

- Os estudantes realizaram a coleta dos dados dividindo os estudantes do Ensino Médio em estratos, conforme descrito na Metodologia?
  - Os dados foram organizados em tabelas para facilitar a compreensão?

- Os integrantes dos grupos se encontraram em algum momento do período destinado para análise dos dados e discussão das hipóteses levantadas em busca de solução aos problemas?
- Os resultados foram exibidos em tabelas e gráficos e indicadas as justificativas estatísticas para a solução apresentada?
- Durante a apresentação dos resultados, os estudantes mostraram compreensão do assunto de estatística e domínio do processo de busca da solução ao problema sugerido?
- Durante a apresentação, os alunos da assistência realizaram perguntas sobre o que estava sendo exposto, apresentaram críticas ou sugestões?

#### *Avaliação da Atividade*

- O tempo para coleta de dados foi suficiente?
  - O tamanho da equipe de coleta de dados (oito estudantes) é ideal para a realização dessa parte da pesquisa?
  - A quantidade de balança e fita métrica foi suficiente para a coleta dos dados dos estudantes da amostra?
  - Os problemas sugeridos possuem significado, contextualização, podem ser identificados no cotidiano do estudante e conduzem à compreensão dos conceitos estatísticos envolvidos?
  - São necessárias aulas específicas sobre o GeoGebra para que os estudantes aprendam a utilizar o programa?



## CONSIDERAÇÕES FINAIS

---

No princípio desta dissertação citou-se o uso de estatísticas enviesadas que podem conduzir o leitor a conclusões equivocadas, seja pelo desejo de apresentar informações de modo tendencioso ou por que seus analistas não possuam conhecimento técnico para fazê-lo de forma clara. Ainda, observou-se que a formação para uma interpretação correta de dados e, a partir dela, a construção de argumentação consistente, atitudes que devem ser comuns aos cidadãos e cidadãs, podem ser iniciadas no ambiente escolar.

É possível de se compreender conceitos estatísticos com o auxílio de programas didáticos que inter-relacionem álgebra, estatística e geometria e, para isso, escolheu-se pesquisar sobre o GeoGebra por razões descritas neste documento; no entanto, constatou-se a escassez de material instrutivo e gratuito relacionando o aplicativo ao estudo de probabilidade e estatística descritiva, motivo que conduziu à elaboração desta dissertação e da qual espera-se que possa levar formação complementar a estudantes e professores, com reflexos no desenvolvimento do aprendizado e no trabalho docente realizado em sala de aula. Este trabalho tornou-se fonte de consulta constante para o seu autor.

As ferramentas e comandos do GeoGebra, apresentados no capítulo 4, são suficientes para que docentes possam construir um aplicativo destinado a análise estatística de um conjunto de dados. Para análises mais técnicas de grandes coleções de valores, em cursos de graduação em estatística ou de ciências de dados, orienta-se utilizar programas técnicos estatísticos. O uso do GeoGebra é recomendado para análises didáticas, com a vantagem de poder se construir diferentes representações gráficas de um mesmo conjunto de dados, ou uma mesma representação gráfica de diversos conjuntos de dados, exibidos na mesma janela de visualização e utilizando um único sistema cartesiano ortogonal, facilitando comparações entre os conjuntos analisados.

Foram utilizados alguns comandos de programação na criação da ferramenta de apoio apresentada na proposta didática, cujo protocolo de construção está descrito no Apêndice. Tais comandos são imprescindíveis, foram utilizados apenas para organizar a informação gráfica exibida na janela de visualização e tornar a ferramenta mais atrativa.

O ensino a distância, imposto pelas medidas de controle sanitário para conter a propagação do coronavírus, tornou-se realidade no país, escolas passaram a adotar tecnologias para a realização de aulas remotas, oportunidade em que se pôde empregar o GeoGebra nas aulas de matemática; portanto, mostra-se que o conteúdo desta dissertação é aplicável, inclusive na modalidade não presencial de estudo.

Esta pesquisa não esgota as possibilidades de se abordar o uso do GeoGebra no estudo de probabilidade e estatística descritiva em trabalhos acadêmicos, sugere-se que ela pode ser complementada com investigações acerca das funções e comandos internos avançados,

construções utilizando linguagens de programação ou integradas a outros *softwares*, o uso embutido em programas de apresentação em *slide*, formas de realizar construções colaborativas *on-line* ou aprofundar o estudo da representação gráfica do desvio padrão apresentada na proposta didática, inclusive na janela de visualização 3D.

Contudo, é necessário lembrar que o uso de recursos tecnológicos é um complemento das metodologias utilizadas, ao tempo que o professor precisa participar de forma ativa do processo de construção do conhecimento dos alunos e alunas, sendo mediador, motivador e orientador da aprendizagem [21]; para isso, o grande desafio é desenvolver um programa dinâmico, apresentando a ciência relacionada aos problemas atuais, indo ao encontro dos interesses dos estudantes, com o uso de tecnologias, pois de outra maneira distancia-se da realidade vivida por eles [11]; porém, apesar de necessárias, é preciso ter cautela no uso de tecnologias em sala de aula, existem componentes que podem tirar o foco da construção de conceitos e, a utilização inadequada, tornar-se um óbice à conquista dos objetivos educacionais [18].

Brilhantemente, o excelsior patrono da educação brasileira registrou que "o homem concreto deve se instrumentar com o recurso da ciência e da tecnologia para melhor lutar pela causa de sua humanização e de sua libertação", a educação [12].



## REFERÊNCIAS BIBLIOGRÁFICAS

---

- [1] AMERICAN STATISTICAL ASSOCIATION (ASA). Guidelines College Report ASA Revision Committee (GAISE), “Guidelines for Assessment and Instruction in Statistics Education College Report 2005.”
- [www.amstat.org/asa/files/pdfs/GAISE/2005GaiseCollege\\_Full.pdf](http://www.amstat.org/asa/files/pdfs/GAISE/2005GaiseCollege_Full.pdf)
- [2] \_\_\_\_\_. Guidelines College Report Revision Committee (GAISE), “Guidelines for Assessment and Instruction in Statistics Education College Report 2016.”
- [www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege\\_Full.pdf](http://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf)
- [3] BÍBLIA, A. T. 2 Samuel. *In*: Bíblia de Estudo NTLH. São Paulo: Sociedade Bíblica do Brasil, 2000. p. 307.
- [4] BOLFARINE, Heleno; BUSSAB, Wilton O. Elementos de Amostragem. São Paulo: Blucher, 2005.
- [5] BORTOLOSSI, Humberto José. **Criando conteúdo educacionais digitais interativos em matemática e estatística com o uso integrado de tecnologias: GeoGebra, JavaView, HTML, CSS, MathML e JavaScript.** **Revista do Instituto GeoGebra Internacional de São Paulo**, São Paulo, v.1 n. 1, 2012.
- <https://revistas.pucsp.br/IGISP/article/view/8823/6595>.
- [6] \_\_\_\_\_. O uso do *software* GeoGebra no ensino e na aprendizagem de estatística e probabilidade. **Vidya Revista Eletrônica**, v.36, n.2, p.429-440, 2016.
- <https://periodicos.ufn.edu.br/index.php/VIDYA/article/view/1804/1749>.
- [7] BRASIL. Ministério da Educação. Base Nacional Comum Curricular. Brasília: MEC, 2018.
- [http://basenacionalcomum.mec.gov.br/images/historico/BNCC\\_EnsinoMedio\\_embaixa\\_site\\_110518.pdf](http://basenacionalcomum.mec.gov.br/images/historico/BNCC_EnsinoMedio_embaixa_site_110518.pdf)
- [8] \_\_\_\_\_. Ministério da Educação. Parâmetros Curriculares Nacionais. Ciências da Natureza, Matemática e suas Tecnologias. Ensino Médio. Brasília: MEC, 1997.
- <http://portal.mec.gov.br/seb/arquivos/pdf/ciencian.pdf>
- [9] CARNEIRO, Reginaldo Fernando; PASSOS, Carmém Lúcia Brancaglioni. A utilização das Tecnologias da Informação e Comunicações nas aulas de Matemática: limites e possibilidades. **Revista Eletrônica da Educação**, São Carlos, v.8 n. 2, 2014.

- <http://www.reveduc.ufscar.br/index.php/reveduc/article/view/729/328>.
- [10] CRESPO, Antônio Arnot. *Estística Fácil*. São Paulo: Ed Saraiva, 1999.
- [11] D'AMBRÓSIO, Ubiratan. **Educação Matemática**: Da teoria à prática. 23. ed. Campinas: Papirus, 2012.
- [12] FREIRE, Paulo. *A Educação na Cidade*. 5. ed. São Paulo: Cortez, 2001.
- [13] \_\_\_\_\_. **Pedagogia da autonomia**: saberes necessários à prática educativa. 50. ed. Rio de Janeiro: Paz e Terra, 2015.
- [14] GEOGEBRA - Manual oficial, 2020. Acesso em: 6 abr. 2020. Disponível em:  
[https://wiki.geogebra.org/en/Category:Manual\\_\(official\)](https://wiki.geogebra.org/en/Category:Manual_(official))
- [15] IBGE. Fundação Instituto Brasileiro de Geografia e Estatística. Normas de Apresentação Tabular - Rio de Janeiro, 1979.  
<https://biblioteca.ibge.gov.br/visualizacao/livros/liv82498.pdf>
- [16] LARSON, Ron; FARBER, Betsy. *ESTATÍSTICA APLICADA*; Tradução: José Fernando Pereira Gonçalves. 6. ed. São Paulo: Pearson, 2015.
- [17] LOPES, Celi Espasandin; TAVARES, Fernando Gonzales. Mapeamento do uso do GeoGebra no ensino de estatística. **Revista Eletrônica de Educação Matemática**, Florianópolis, v.14, Edição Especial Estatística, p.1-20, 2019.  
<https://periodicos.ufsc.br/index.php/revemat/article/view/1981-1322.2019.e62800>.
- [18] CORTELLA, Mario Sergio. **Educação, escola e docência**: novos tempos, novas atitudes; São Paulo: Cortez Editora, 2014.
- [19] MLODINOW, Leonard. **O andar do bêbado**: como o acaso determina nossas vidas; Tradução: Diego Alfaro. Rio de Janeiro: Zahar, 2009.
- [20] MORETTIN, Pedro A.; BUSSAB, Wilton O. *ESTATÍSTICA BÁSICA*. 9. ed. São Paulo: Saraiva, 2017.
- [21] SOUZA, Leandro Coqueiro. **A TIC na educação**: uma grande aliada no aumento da aprendizagem no Brasil. **Revista EIXO**, Brasília, n. 1, jan.-jun. 2016. <http://revistaeixo.ifb.edu.br/index.php/RevistaEixo/article/view/315/197>.
- [22] TRIOLA, Mario F. *Introdução à Estatística*; Tradução: Vera Regina Lima de Farias e Flores. 10. ed. Rio de Janeiro: LTC, 2008.
- [23] WHEELAN, Charles. **ESTATÍSTICA**: o que é, para que serve, como funciona; Tradução: George Schlesinger. 1. ed. Rio de Janeiro: Zahar, 2016.