

UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO - UFTM



MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL -
PROFMAT



PROFMAT

Dissertação de Mestrado

Distribuições de probabilidade: propostas de atividades para o Ensino
Médio

Patricia Facho Figueiredo

Uberaba - Minas Gerais

Abril de 2021

Distribuições de probabilidade: propostas de atividades para o Ensino Médio

Patricia Facho Figueiredo

Dissertação de Mestrado apresentada à Comissão Acadêmica Institucional do PROFMAT-UFTM, como requisito parcial para obtenção do título de Mestre Matemática.

Orientadora: Profa. Dra. Alyne Toscano Martins

Uberaba - Minas Gerais

Abril de 2021

**Catálogo na fonte: Biblioteca da Universidade Federal do
Triângulo Mineiro**

F492d Figueiredo, Patricia Facho
Distribuições de probabilidade: propostas de atividades para o ensino médio / Patricia Facho Figueiredo. -- 2021.
96 f. : il., graf., tab.

Dissertação (Mestrado Profissional em Matemática em Rede Nacional) -- Universidade Federal do Triângulo Mineiro, Uberaba, MG, 2021
Orientadora: Profa. Dra. Alyne Toscano Martins

1. Matemática – Estudo e Ensino. 2. Probabilidades. 3. Distribuição (Teoria da probabilidade). 4. Ensino Médio. I. Martins, Alyne Toscano. II. Universidade Federal do Triângulo Mineiro. III. Título.

CDU 519.2(07):373.5

Distribuições de probabilidade: propostas de atividades para o Ensino Médio

Dissertação apresentada ao Programa de Pós-Graduação em Mestrado Profissional em Matemática, área de concentração Matemática da Universidade Federal do Triângulo Mineiro como requisito parcial para obtenção do título de mestre.

Uberaba, 20 de abril de 2021.

Banca Examinadora:

Dra. Alyne Toscano Martins – Orientadora
Universidade Federal do Triângulo Mineiro

Dra. Michelli Maldonado Carretero de Oliveira
Universidade Federal do Triângulo Mineiro

Dra. Silvia Regina Viel
Centro Universitário Municipal de Franca



Documento assinado eletronicamente por **ALYNE TOSCANO MARTINS**, Professor do **Magistério Superior**, em 25/05/2021, às 23:03, conforme horário oficial de Brasília, com fundamento no art. 6º do [Decreto nº 8.539, de 8 de outubro de 2015](#) e no art. 34 da [Portaria Reitoria/UFTM nº 59, de 26 de abril de 2021](#).



Documento assinado eletronicamente por **Silvia Regina Viel, Usuário Externo**, em 27/05/2021, às 23:31, conforme horário oficial de Brasília, com fundamento no art. 6º do [Decreto nº 8.539, de 8 de outubro de 2015](#) e no art. 34 da [Portaria Reitoria/UFTM nº 59, de 26 de abril de 2021](#).



Documento assinado eletronicamente por **MICHELLI MALDONADO CARRETERO DE OLIVEIRA, Professor do Magistério Superior**, em 28/05/2021, às 09:23, conforme horário oficial de Brasília, com fundamento no art. 6º do [Decreto nº 8.539, de 8 de outubro de 2015](#) e no art. 34 da [Portaria Reitoria/UFTM nº 59, de 26 de abril de 2021](#).



A autenticidade deste documento pode ser conferida no site http://sei.uftm.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0534820** e o código CRC **1D0C1426**.

Aos meus pais.

Agradecimentos

Agradeço a meus pais Olésio e Matilde, por sempre mostrar que a educação é o melhor caminho a se percorrer. Sem eles isso não seria possível.

À minha irmã Priscila, por mesmo que distante em partes dessa fase, se fez presente, me motivando em todos os momentos dos quais eu precisei.

Ao meu esposo Caio, por me apoiar e incentivar, pela paciência e compreensão dos diversos momentos que estive ausente neste período.

Agradeço minha orientadora Profa. Dra. Alyne, pela excelente orientação, sempre presente, proporcionando todo o suporte necessário para extrair o melhor deste trabalho. Serei eternamente grata.

À Profa. Dra. Silvia, que foi de grande importância na minha graduação e pôde participar dessa minha conquista também.

E claro, aos amigos que o mestrado me proporcionou, Karina e Rhaony, que tornaram essa jornada menos árdua. Jamais esquecerei disso.

Obrigada!

“Para tudo há um tempo.”

Eclesiastes 3:1

Resumo

A busca por uma educação de qualidade no Brasil é constante na vida do professor, e um dos desafios da educação matemática é despertar o interesse do aluno em aprender os conteúdos do currículo. Ao refletir sobre motivos que geram desinteresse, conclui-se que a falta de conhecimento sobre como os conteúdos são aplicados na realidade é um deles. Assim, existe uma preocupação em associar a realidade do estudante com o conteúdo em estudo, e essa preocupação está presente inclusive na Base Nacional Comum Curricular (BNCC). Diante disto, este trabalho apresenta propostas de atividades para serem desenvolvidas no Ensino Médio, nas quais são trabalhadas as distribuições de probabilidades, especificamente as distribuições Geométrica, Binomial e Normal, com o objetivo de realizar aplicações de probabilidade em situações reais presentes na vida do aluno. São atividades, aplicáveis, adaptáveis e cumprem habilidades presentes na BNCC. Acredita-se que as mesmas possam contribuir para que professores de matemática proporcionem aos seus alunos uma aprendizagem significativa no ensino de probabilidade.

Palavras-chave: Probabilidade. Distribuições de Probabilidade. Educação Matemática. Matemática Aplicada.

Abstract

The search for quality education in Brazil is constant in the teacher's life, and one of the challenges of mathematics education is to awaken the student's interest in learning the contents of the curriculum. When reflecting on reasons that generate disinterest, it is concluded that the lack of knowledge about how the contents are applied in reality is one of them. Thus, there is a concern to associate the student's reality with the content under study, and this concern is present even in the National Common Curricular Base (BNCC). In view of this, this work presents proposals for activities to be developed in high school, in which the probability distributions are worked, specifically the Geometric, Binomial and Normal, with the objective carry out probability applications in real situations present in the student's life. They are activities, applicable, adaptable and fulfill the skills present in the BNCC. It is believed that they can contribute to mathematics teachers providing their students with meaningful learning in in probability education.

Keywords: Probability. Probability Distributions. Mathematics Education. Applied Mathematics.

Sumário

1	INTRODUÇÃO	15
2	O ENSINO DE PROBABILIDADE NO BRASIL	16
3	DISTRIBUIÇÕES DE PROBABILIDADE	21
3.1	Probabilidade	21
3.2	Variáveis aleatórias	23
3.3	Variáveis aleatórias discretas	24
3.3.1	Função de probabilidade	24
3.3.2	Função de distribuição cumulativa	25
3.3.3	Valor esperado e desvio padrão	26
3.4	Variáveis aleatórias contínuas	28
3.4.1	Função densidade de probabilidade	28
3.4.2	Função de distribuição cumulativa	29
3.4.3	Valor esperado e variância	29
3.5	Distribuição Geométrica	30
3.6	Distribuição Binomial	33
3.7	Distribuição Normal	36
3.7.1	Distribuição Normal Padrão	39
4	PROPOSTAS DE ATIVIDADES	43
4.1	Conhecimentos prévios	43
4.2	Atividade 1: A distribuição Geométrica e a Mega-Sena	45
4.3	Atividade 2: A distribuição Binomial no ENEM	57
4.4	Atividade 3: A distribuição Normal na estatura	69
4.5	Análise das habilidades da BNCC	94
5	CONSIDERAÇÕES FINAIS	96
	REFERÊNCIAS	97
A	TABELA DE DISTRIBUIÇÃO NORMAL PADRÃO	99

Lista de ilustrações

Figura 1 – Pesquisa de dissertações de probabilidade no site PROFMAT	19
Figura 2 – Representação gráfica da função de probabilidade do Exemplo 3.1	25
Figura 3 – Função de distribuição cumulativa do Exemplo 3.1	26
Figura 4 – Distribuição Geométrica do Exemplo 3.2.	33
Figura 5 – Algumas formas de obter x sucessos em n tentativas.	34
Figura 6 – Distribuições Binomiais para valores selecionados de n e p	34
Figura 7 – Distribuição Binomial do Exemplo 3.3.	36
Figura 8 – Distribuição normal com parâmetros μ e σ	37
Figura 9 – Altura de carvalhos normalmente distribuídas.	38
Figura 10 – Distribuição Normal Padrão.	39
Figura 11 – Calculando a probabilidade $P(a \leq Z \leq b)$	40
Figura 12 – Calculando a probabilidade $P(-a \leq Z \leq 0)$	40
Figura 13 – Probabilidade de um carvalho ter altura maior que 85 pés.	41
Figura 14 – Cartela de aposta.	46
Figura 15 – Cartela de números para sorteio.	46
Figura 16 – Calculando $n(S)$	47
Figura 17 – Calculando $n(S)$	48
Figura 18 – Célula com a probabilidade de sucesso.	51
Figura 19 – Criando a tabela.	51
Figura 20 – Calculando $P(1)$	52
Figura 21 – Calculando as outras probabilidades.	52
Figura 22 – Mudando o formato das células para números.	53
Figura 23 – Probabilidades em decimais.	53
Figura 24 – Copiando as células.	54
Figura 25 – Folha de resposta.	58
Figura 26 – Folha de resposta.	58
Figura 27 – Acertando duas questões.	59
Figura 28 – Possibilidades para acertar duas questões.	59
Figura 29 – Passo a.	62
Figura 30 – Passo b.	63
Figura 31 – Passo c.	64
Figura 32 – Passo c: copiando as colunas.	64
Figura 33 – Passo c: copiando as colunas.	65
Figura 34 – Passo d: construindo o gráfico.	65
Figura 35 – Passo d: editando o gráfico.	66

Figura 36 – Gráfico da distribuição Binomial.	66
Figura 37 – Exemplo de uma amostra de alturas em centímetros.	71
Figura 38 – Calculando média no Excel.	73
Figura 39 – Calculando mediana no Excel.	73
Figura 40 – Calculando moda no Excel	74
Figura 41 – Calculando o desvio padrão no Excel.	74
Figura 42 – Calculando a amplitude no Excel.	75
Figura 43 – Medidas de posição e dispersão da amostra exemplo.	75
Figura 44 – Exemplo de um histograma.	76
Figura 45 – Dados coletados na planilha do Geogebra.	78
Figura 46 – Dados coletados na planilha do Geogebra.	79
Figura 47 – Criando o histograma.	79
Figura 48 – Selecionando a opção Histograma.	80
Figura 49 – Definindo as classes manualmente, 2021.	80
Figura 50 – Histograma da amostra exemplo.	81
Figura 51 – Borboleta simétrica.	81
Figura 52 – Histograma simétrico.	82
Figura 53 – Calculando o grau de assimetria no Excel.	83
Figura 54 – Distribuição sino.	83
Figura 55 – Curva normal.	85
Figura 56 – Probabilidades na distribuição Normal.	86
Figura 57 – Janela de probabilidade do Geogebra.	88
Figura 58 – Selecionando a distribuição.	88
Figura 59 – Distribuição Normal no Geogebra.	89
Figura 60 – Comparando áreas.	90
Figura 61 – Distribuição normal padronizada.	91
Figura 62 – Lendo da tabela da variável Z	91
Figura 63 – Interpretação do segundo exemplo Z	92
Figura 64 – Simétricos em uma Distribuição Normal.	93

Lista de tabelas

Tabela 1 – Dissertações de probabilidade do PROFMAT	20
Tabela 2 – Variável aleatória X para o experimento aleatório de lançar uma moeda 3 vezes	24
Tabela 3 – Classes da tabela de frequência.	77
Tabela 4 – Tabela de frequências da amostra exemplo.	78

1 INTRODUÇÃO

Ao se falar de probabilidade é comum associá-la, a princípio, à situações que envolvam dados, moedas, baralhos e diversos outros cenários hipotéticos relacionados a jogos, principalmente. Esse pensamento também é refletido na Educação Básica, visto que muitas situações problemas abrangem esses temas. No entanto, atualmente a probabilidade não pode ser vista de forma isolada na matemática, e sim como uma área que é aplicável em outras ciências, como sociais, biológicas e econômicas, possibilitando fazer interpretações e previsões da sociedade. Nesse sentido, faz-se ainda mais importante que o aprendizado de probabilidade seja significativo na Educação Básica, em especial no Ensino Médio, nível de escolaridade que de fato o ensino de probabilidade deve ser consolidado.

Segundo a Base Nacional Comum Curricular (2018), um dos meios de tornar o ensino mais significativo é utilizando um “currículo em ação”, termo designado à um conjunto de decisões que materializam a aprendizagem. Uma dessas decisões se refere a ação de “contextualizar os conteúdos dos componentes curriculares, identificando estratégias para apresentá-los, representá-los, exemplificá-los, conectá-los e torná-los significativos, com base na realidade do lugar e do tempo nos quais as aprendizagens estão situadas” (BRASIL, 2018, p.16).

As distribuições de probabilidade aplicadas em situações que possam fazer parte do cotidiano do aluno permitem que o ensino seja mais contextualizado. Portanto, objetivo deste trabalho é propor três atividades sobre distribuições de probabilidades, direcionadas a alunos do Ensino Médio, a fim de expandir os conhecimentos, apresentando aplicações dos conceitos de probabilidades aprendidos no decorrer das aulas. Durante os processos das atividades, o aluno é instigado a fazer reflexões sobre os resultados obtidos, conhecer o caminho percorrido para a obtenção das fórmulas que são usadas, e utilizar recursos tecnológicos, reconhecendo o *software* de planilhas Excel e o Geogebra como ferramentas muito úteis na matemática.

O presente trabalho está dividido da seguinte forma: no Capítulo 2 se discorre sobre como o ensino de probabilidade é visto no Brasil e nele também apresentam-se as habilidades abrangidas nas atividades de acordo com a BNCC. No Capítulo 3 são explorados os conceitos matemáticos de distribuições de probabilidade. As atividades propostas são apresentadas no Capítulo 4, com orientações sobre como cada etapa deve ser realizada e os conceitos explorados nas mesmas. Por fim, as Conclusões são apresentadas no Capítulo 5.

2 O ensino de probabilidade no Brasil

Os registros indicam que desde antes de Cristo, os jogos de azar já faziam parte das atividades humanas, sendo eles a motivação para que matemáticos tais como Jerónimo Cardano (1501 - 1576), Blaise Pascal (1623 - 1662) e Pierre de Fermat (1601 - 1665) concretizassem os estudos da Teoria de Probabilidade como um ramo da matemática, por volta do século XVI.

Segundo Lopes e Meirelles (2005), embora o início da probabilidade tenha provindo dos jogos, atualmente essa área é de grande importância para a sociedade, principalmente quando é aliada a estatística. A probabilidade e a estatística andam de mãos dadas, e são utilizadas não só nas ciências exatas, mas também nas biológicas e sociais, tornando necessário a inserção desses conteúdos no ensino da Educação Básica. Como afirma Lopes (2008), “é imprescindível o conhecimento da probabilidade de ocorrência de acontecimentos para agilizarmos a tomada de decisão e fazermos previsões”, uma vez que o mundo está em rápida mudança.

Ao longo da Educação Básica, os alunos começam a ter contato com os conceitos de probabilidade, mas é no Ensino Médio que esse conhecimento é mais consolidado.

Analisando os documentos que norteiam o currículo educacional do Ensino Médio do país, existem três principais: os Parâmetros Curriculares Nacionais, o qual possui duas versões, o PCNEM (Parâmetros Curriculares Nacionais do Ensino Médio), elaborado no ano de 2000 e a versão de 2002, denominada de PCN+. E mais recentemente em 2018 foi criada a Base Nacional Comum Curricular (BNCC).

Os PCN's, tanto do Ensino Médio quanto do Ensino Fundamental, foram elaborados com o objetivo de dar referência para a comunidade escolar, como forma de auxiliar na elaboração da proposta curricular das escolas, e de garantir aos estudantes de todo o país o direito de “usufruir do conjunto de conhecimentos reconhecidos como necessários para o exercício da cidadania” (Silva e Figueiredo, 2019, p.3), não possuindo portanto, a função de obrigatoriedade.

No que diz respeito ao eixo temático de probabilidade do Ensino Médio, o PCN+ aborda as seguintes habilidades:

- Reconhecer o caráter aleatório de fenômenos e eventos naturais, científico-tecnológicos ou sociais, compreendendo o significado e a importância da probabilidade como meio de prever resultados.
- Quantificar e fazer previsões em situações aplicadas a diferentes áreas do conhecimento e da vida cotidiana que envolvam o pensamento probabilístico.

- Identificar em diferentes áreas científicas e outras atividades práticas modelos e problemas que fazem uso de estatísticas e probabilidades.

Diferente dos PCN's, a BNCC é obrigatória por lei e “define o conjunto orgânico e progressivo de aprendizagens essenciais que todos os alunos devem desenvolver ao longo das etapas e modalidades da Educação Básica” (BRASIL, 2018, p. 7), especificando melhor as competências e habilidades para o ensino. No entanto, mesmo com suas diferenças é possível observar que ambos documentos consideram importante a aplicação dos conteúdos à realidade, como mostra a competência da contextualização sócio-cultural proposta no PCN+ que tem o objetivo “de aproximar o aluno da realidade e fazê-lo vivenciar situações próximas que lhe permitam reconhecer a diversidade que o cerca e reconhecer-se como indivíduo capaz de ler e atuar nesta realidade” (BRASIL, 2002, p.126). Além disso, eles propõem o uso da tecnologia, como pode ser visto em diversas habilidades da BNCC, por exemplo.

É diante de todo esse embasamento que as atividades de distribuições de probabilidades presentes neste trabalho foram elaboradas, com o intuito de abordar a probabilidade em situações da realidade e trazer o uso da tecnologia aliada ao ensino. E, embora a BNCC não traga habilidades específicas de distribuições de probabilidades, o documento afirma que “no Ensino Médio o foco é a construção de uma visão integrada da Matemática, aplicada à realidade, em diferentes contextos” (BRASIL, 2018, p. 528). Portanto, distribuição de probabilidade é um conteúdo interessante para trabalhar no Ensino Médio, uma vez que abrangem habilidades sobre probabilidades e gráficos, além de ser uma aplicação da matemática à realidade.

As atividades propostas no Capítulo 4 contemplam as seguintes competências e habilidades da BNCC do Ensino Médio:

Competência específica 1: Utilizar estratégias, conceitos e procedimentos matemáticos para interpretar situações em diversos contextos, sejam atividades cotidianas, sejam fatos das Ciências da Natureza e Humanas, ou ainda questões econômicas ou tecnológicas, divulgados por diferentes meios, de modo a consolidar uma formação científica geral.

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

Competência específica 2: Propor ou participar de ações para investigar desafios do mundo contemporâneo e tomar decisões éticas e socialmente responsáveis, com base na análise de problemas sociais, como os voltados a situações de saúde, sustentabilidade, das implicações da tecnologia no mundo

do trabalho, entre outros, mobilizando e articulando conceitos, procedimentos e linguagens próprios da Matemática.

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

Competência específica 3: Utilizar estratégias, conceitos, definições e procedimentos matemáticos para interpretar, construir modelos e resolver problemas em diversos contextos, analisando a plausibilidade dos resultados e a adequação das soluções propostas, de modo a construir argumentação consistente.

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

-(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).

Competência específica 4: Compreender e utilizar, com flexibilidade e precisão, diferentes registros de representação matemáticos (algébrico, geométrico, estatístico, computacional etc.), na busca de solução e comunicação de resultados de problemas.

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

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

Competência específica 5: Investigar e estabelecer conjecturas a respeito

de diferentes conceitos e propriedades matemáticas, empregando estratégias e recursos, como observação de padrões, experimentações e diferentes tecnologias, identificando a necessidade, ou não, de uma demonstração cada vez mais formal na validação das referidas conjecturas.

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

Observa-se que são abordadas diversas habilidades, sendo a maioria voltada à probabilidade. Porém, é natural no ensino da Matemática atividades desenvolverem mais de uma, visto que faz necessário ter conhecimentos prévios para se aprender novos. No ensino público do Estado de São Paulo, por exemplo, é proposta a semana da recuperação, período no qual os professores trabalham novamente habilidades não atingidas pelas turmas, que são mostradas pelos resultados da Avaliação de Aprendizagem em Processo (AAP) realizada bimestralmente e da Avaliação Diagnóstica Complementar (ADC) realizada semestralmente.

É possível encontrar muitos trabalhos desenvolvidos no meio acadêmico com o tema de probabilidade. No site do PROFMAT, por exemplo, ao pesquisar dissertações que possuem no título a palavra “probabilidade” são encontrados 150 de um total de 5.638 registros, isto é, cerca de 2,7% no período de 2013 a 2020. Veja a Figura 1.

Figura 1 – Pesquisa de dissertações de probabilidade no site PROFMAT

Lista das Dissertações de Mestrado dos alunos do PROFMAT.

Instituição: Título: Aluno:

Busca por Título do TCC: probabilidade
Foram encontrados 150 registros.

Data de defesa	Aluno	Título da Dissertação	Instituição	Dissertação
29/10/2020	ANTONIO CARLOS RAULINO DA SILVA	PROBABILIDADE NO ENSINO MÉDIO: ELABORAÇÃO DE UM ROTEIRO DE ANÁLISE DOS LIVROS DIDÁTICOS SELECIONADOS PELAS ESCOLAS PÚBLICAS DE QUIXADÁ-CE	UECE	PDF
25/09/2020	RAFAEL LEMES DE REZENDE	ESTUDO DA TEORIA DE PROBABILIDADE ATRAVÉS DE DINÂMICAS DE JOGOS	UFG	PDF
28/08/2020	MARCO ANTÔNIO RIBEIRO	JOGOS DE POKER ATRAVÉS DE TECNOLOGIAS DIGITAIS PARA ENSINO DE PROBABILIDADE	UFSJ	PDF
07/08/2020	ADAIL BOA DE FREITAS FILHO	PROBABILIDADE: UMA PROPOSTA A LUZ DA BNCC.	UNILAB	PDF
30/07/2020	APRÍGIO DOS SANTOS VIEIRA FILHO	PROBABILIDADE E SUAS APLICAÇÕES	UNIR	PDF

Fonte: Própria autora, 2021.

Essas 150 dissertações, de forma geral, são distribuídas conforme mostra a Tabela 1.

Quanto aos trabalhos pesquisados no site do PROFMAT que também possuem como temática “distribuições de probabilidade”, verificou-se que a metodologia utilizada pelos

Tabela 1 – Dissertações de probabilidade do PROFMAT

Ano	Número de dissertações
2020	17
2019	16
2018	21
2017	14
2016	16
2015	21
2014	20
2013	25

Fonte: Própria autora, 2021.

autores difere-se das atividades que são propostas no presente trabalho. Vieira Filho (2020), por exemplo, propõe aplicações de distribuições discretas e contínuas em situações problemas, porém o autor aborda o uso direto das funções de probabilidades, sem a realização de um processo dedutivo. Da mesma forma, Rocha (2017), apresenta resoluções de exercícios com a temática do torneio de futebol “Brasileirão”, envolvendo variáveis aleatórias binomiais, que são resolvidos utilizando as aproximações da binomial por meio das distribuições normal e de Poisson. Santos (2015), apresenta questões sobre distribuição binomial dentro do contexto da medicina. Além disso, Soares (2020) propõe várias atividades de probabilidade, sendo uma delas semelhante à Atividade 2 presente no Capítulo 4 neste trabalho, no entanto, o método para resolução não trás os conceitos de distribuição de probabilidades. Gonçalves (2014) propõe uma atividade que envolve uma amostra de retiradas de um saco de laranjas a fim de resolver uma situação problema a partir da distribuição normal, porém adentrando mais na parte estatística.

Haja a vista o exposto acima, espera-se com esse trabalho, em especial, as atividades propostas no Capítulo 4, contribuir com o Ensino de Probabilidade no Ensino Médio no Brasil, à luz das exigências da BNCC.

3 Distribuições de probabilidade

Este capítulo discorrerá sobre variáveis aleatórias, desde os conceitos básicos que necessitam ser definidos para a compreensão do que é uma variável aleatória, até suas aplicações, que são as distribuições de probabilidade. Cabe ressaltar que o objetivo não é esgotar o assunto sobre, mas sim apresentar brevemente de forma objetiva os principais conceitos que serão necessários para o entendimento dos próximos capítulos pelo leitor. Para mais detalhes, indica-se a leitura de Ross (2010), Montgomery e Runger (2020), Larson e Farber (2010) e Meyer(1970), principais referências utilizadas para o desenvolvimento matemático deste trabalho.

3.1 Probabilidade

Basicamente, a probabilidade é utilizada para quantificar as chances de ocorrer um resultado em um experimento aleatório.

Definição 3.1. *Experimento aleatório é um experimento realizado sempre da mesma maneira que pode fornecer resultados diferentes a cada vez que é executado.*

Aferir pressão arterial, teste de controle de qualidade de um produto, lançamento de dados e coleta dados pluviométricos são exemplos de experimentos aleatórios. Deste modo, percebe-se que ao realizar um experimento, as possibilidades de resultados variam de acordo com cada um. No entanto, se tiver o conhecimento de todos os possíveis resultados, então sabe-se qual é o espaço amostral do experimento.

Definição 3.2. *Espaço amostral é o conjunto de todos os resultados possíveis de um experimento aleatório. É denotado por S .*

A diversidade de experimentos aleatórios permite que existam dois tipos de espaços amostrais: o discreto, quando as possibilidades de resultado do experimento podem ser representadas por um conjunto finito ou infinito contável, e o contínuo, quando esses possíveis resultados podem ser representados por um intervalo de números reais (finito ou infinito).

Ao realizar um experimento aleatório têm-se um interesse em obter alguns resultados relacionados, que são subconjuntos do espaço amostral, ou seja, um evento.

Definição 3.3. *Evento é qualquer subconjunto E do espaço amostral S .*

Em um lançamento de duas moedas, por exemplo, o espaço amostral é definido por:

$$S = \{(KK), (KC), (CK), (CC)\},$$

em que K representa sair cara, e C coroa. Além disso, pode-se considerar um evento sair o mesmo lado nos dois lançamentos, ou seja,

$$E = \{(KK), (CC)\}.$$

A definição de probabilidade pode ser feita a partir da frequência relativa de um evento, conhecida também como modelo frequentista de probabilidade.

Definição 3.4. *Seja E um evento do espaço amostral S . A probabilidade de ocorrer E é*

$$P(E) = \lim_{n \rightarrow \infty} \frac{n(E)}{n}, \quad (3.1)$$

em que $n(E)$ é o número de vezes que ocorre E em n repetições do experimento.

Deste modo, pode-se interpretar a probabilidade como sendo “o valor limite da proporção de vezes que o resultado ocorre em n repetições do experimento aleatório, a medida que n aumenta além de todos os limites” (MONTGOMERY; RUNGER, 2020, p.23).

Contudo, como afirma Ross (2010), nesta definição existe o inconveniente de não dar garantia que $n(E)/n$ convergirá para algum valor constante. Como forma de solucionar esse problema propõe-se então, que a probabilidade de ocorrência do evento E , satisfaça os três axiomas de probabilidade:

Axioma 1: $0 \leq P(E) \leq 1$.

Axioma 2: $P(S) = 1$.

Axioma 3: Para cada sequência de eventos mutuamente exclusivos E_1, E_2, \dots (isto é, eventos para os quais $E_i \cap E_j = \emptyset$ quando $i \neq j$),

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

Note que do Axioma 3 pode-se concluir que $P(\emptyset) = 0$, tomando $E_1 = S$ e $E_i = \emptyset$, para $i > 1$, visto que:

$$P(S) = \sum_{i=1}^{\infty} P(E_i) = P(S) + \sum_{i=2}^{\infty} P(\emptyset)$$

o que implica que $P(\emptyset) = 0$.

É intuitivo concluir que em muitos experimentos aleatórios com espaço amostral discreto, todos os resultados têm as mesmas chances de ocorrer, ou seja, são igualmente prováveis. Considere então que o espaço amostral S de um experimento aleatório seja $S = \{1, 2, 3, \dots, N\}$, logo neste caso tem-se que:

$$P(\{1\}) = P(\{2\}) = \dots = P(\{N\})$$

Tomando $E_i = \{i\}$, do Axioma 3 pode-se afirmar que:

$$P\left(\bigcup_{i=1}^n E_i\right) = \sum_{i=1}^n P(E_i).$$

Mas $\bigcup_{i=1}^n E_i = S$, assim pelo Axioma 2,

$$\sum_{i=1}^n P(E_i) = 1$$

$$P(\{1\}) + P(\{2\}) + \dots + P(\{N\}) = 1$$

$$nP(E) = 1$$

o que implica que $P(E) = 1/n$.

Portanto, conclui-se que para cada evento E ,

$$P(E) = \frac{n(E)}{n(S)} \quad (3.2)$$

isto é, a probabilidade de qualquer evento E é igual a proporção de resultados no espaço amostral S em que esses conjuntos estão contidos, sendo esta outra definição que se aplica para probabilidade, principalmente na Educação Básica.

3.2 Variáveis aleatórias

É intuitivo pensar que em problemas de probabilidade faz-se necessário conhecer o espaço amostral para se calcular probabilidades de eventos em um experimento, porém nem sempre é assim. Em alguns casos, é interessante associar um número a cada evento do espaço amostral e não necessariamente na descrição desses possíveis resultados. É diante dessa necessidade que insere-se o conceito de variável aleatória.

Definição 3.5. *Variável aleatória (v.a) é uma função real definida no espaço amostral de um experimento aleatório que associa um número real a cada resultado do espaço amostral. É denotada por uma letra maiúscula, como X .*

Considere o experimento aleatório a seguir:

Exemplo 3.1. *Uma moeda honesta é lançada 3 vezes e deseja-se observar a quantidade de caras.*

Neste caso, pode-se atribuir a variável aleatória X como sendo o número de caras ocorridas nos lançamentos, então X pode assumir qualquer um dos valores do conjunto $\{0, 1, 2, 3\}$. A Tabela 2 mostra a relação de X com os eventos do espaço amostral do Exemplo 3.1, em que C representa sair coroa e a letra K representa sair cara.

Tabela 2 – Variável aleatória X para o experimento aleatório de lançar uma moeda 3 vezes

Valores de X	Eventos associados
0	(CCC)
1	(CCK), (CKC), (KCC)
2	(CKK), (KCK), (KKC)
3	(KKK)

Fonte: Própria autora, 2021.

As variáveis aleatórias podem se dividir em discreta e contínuas.

3.3 Variáveis aleatórias discretas

Observe que no Exemplo 3.1 é possível enumerar a variável aleatória, de modo que pode ser representada por pontos discretos na reta real. Essa condição permite classificá-la como variável aleatória discreta.

Definição 3.6. *Uma variável aleatória é discreta quando os números de valores possíveis de X for finito ou infinito numerável, podendo ser postos em lista como $x_1, x_2, x_3, \dots, x_n$. No caso infinito numerável, a lista não possui limite superior.*

Existem muitas variáveis aleatórias que se enquadram nesse tipo, como por exemplo número de peças defeituosas de uma produção e de pessoas infectadas por um vírus.

3.3.1 Função de probabilidade

O objetivo de se utilizar variáveis aleatórias em experimentos é, principalmente, para o cálculo de distribuição de probabilidades. Essa distribuição associa uma descrição de probabilidades aos possíveis valores de X , fornecendo dados para se aplicar em problemas práticos, a fim de auxiliar em tomadas de decisões. Denomina-se essa distribuição de função de probabilidade.

Definição 3.7. *A uma variável aleatória discreta X associa-se a **Função de Probabilidade** com valores possíveis x_1, x_2, \dots, x_n uma função f que admite as seguintes condições:*

$$i) f(x_i) \geq 0$$

$$ii) \sum_{i=1}^n f(x_i) = 1$$

sendo denotada por $f(x_i) = P(X = x_i)$.

Retomando ao Exemplo 3.1, a probabilidade de sair x caras em 3 lançamentos é dada por

$$f(x) = P(X = x) = \binom{3}{x} (0,5)^3, \quad x = \{0, 1, 2, 3\}, \quad (3.3)$$

com os seguintes valores para a função de probabilidade:

$$f(0) = P(X = 0) = 0,125,$$

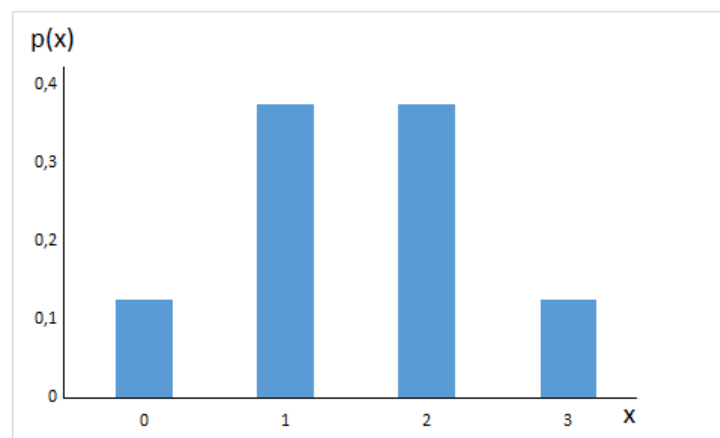
$$f(1) = P(X = 1) = 0,375,$$

$$f(2) = P(X = 2) = 0,375,$$

$$f(3) = P(X = 3) = 0,125,$$

podendo ser representada graficamente conforme a Figura 2.

Figura 2 – Representação gráfica da função de probabilidade do Exemplo 3.1



Fonte: Própria autora, 2021.

Note que a função (3.3) satisfaz as duas condições da Definição (3.7): $f(x) \geq 0$, para x_0, x_1, x_2, x_3 e o somatório das probabilidades resulta em 1. Portanto trata-se de uma função de probabilidade.

3.3.2 Função de distribuição cumulativa

Definição 3.8. A **função de distribuição cumulativa** de uma variável aleatória discreta X , denotada por $F(x)$ é definida por:

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)$$

e satisfaz as seguintes condições:

i) $0 \leq F(x) \leq 1$

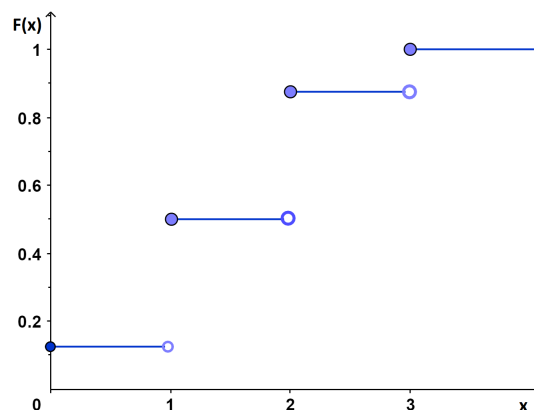
ii) Se $x \leq y$ então $F(x) \leq F(y)$

Veja a função de distribuição cumulativa do Exemplo 3.1:

$$F(x) = \begin{cases} 0,125; & x < 1 \\ 0,5; & 1 \leq x < 2 \\ 0,875; & 2 \leq x < 3 \\ 1; & 3 \leq x \end{cases}$$

cuja representação gráfica de $F(x)$ é visualizada na Figura 3.

Figura 3 – Função de distribuição cumulativa do Exemplo 3.1



Fonte: Própria autora, 2021.

É válido ressaltar que a função de distribuição cumulativa é descontínua, já que pode assumir valores não inteiros mesmo X sendo discreta. O valor de F é constante no intervalo $[x_{i-1}, x_i)$ dando um “salto” para o intervalo $[x_i, x_{i+1})$. Logo, é possível determinar $f(x_i)$ a partir da função $F(x)$, de modo que:

$$f(x_i) = P(X = x_i) = F(x_i) - \lim_{x \rightarrow x_i^-} F(x).$$

3.3.3 Valor esperado e desvio padrão

Ao se tratar de distribuição de probabilidades, o valor esperado, também chamado de esperança, é de grande importância, sendo definido como:

Definição 3.9. O valor esperado ou a esperança de uma variável aleatória discreta, representado por μ ou $E(X)$ é

$$\mu = E(X) = \sum_x xp(x).$$

Em outras palavras, é a média ponderada dos valores de X ponderando suas respectivas probabilidades.

No Exemplo 3.1 o valor esperado é:

$$E(X) = 0f(0) + 1f(1) + 2f(2) + 3f(3) = 0 \cdot 0,125 + 1 \cdot 0,375 + 2 \cdot 0,375 + 3 \cdot 0,125 = 1,5.$$

Com esse resultado, pode-se concluir que ao jogar uma moeda 3 vezes, em média, é mais provável sair 1,5 caras. No caso, o valor esperado pode fornecer um resultado decimal, mesmo tratando-se de uma variável aleatória discreta.

A média é uma medida central que em diversas situações é vista como um ponto de equilíbrio. Muitas variáveis aleatórias podem apresentar o mesmo valor esperado, porém o modo com que seus dados apresentam ou estão distantes um do outro faz ser um equívoco considerar unicamente a esperança para a interpretação de variáveis aleatórias, principalmente em distribuições que são mais dispersas. Até porque, não é uma certeza que ao realizar várias vezes um experimento serão encontrados os valores médios, mas a probabilidade de que eles apareçam é maior do que os outros eventos. Portanto, é incluído, juntamente com a média, o desvio padrão que tem como objetivo informar sobre a variabilidade/dispersão dos dados de uma *v.a.*

Para determinar o desvio padrão é necessário definir inicialmente a variância.

Definição 3.10. Se X é uma variável aleatória discreta com média μ , então a variância de X é definida por:

$$Var(X) = E(X - \mu)^2.$$

A ideia é saber quão longe os dados estão do valor central (média), por isso ela é chamada de uma medida de variabilidade. No entanto, observe que a variância é o valor esperado do quadrado da distância que a variável está de sua média, ocasionando uma mudança de unidade de medida. Mas, considerar a variância como sendo apenas $E[|X - \mu|]$ não é conveniente, já que esse valor pode tender a zero. Como forma de solucionar esse problema, para voltar à unidade de medida da variável calculá-se a raiz quadrada, que é o desvio padrão. Agora torna-se fácil definir o desvio padrão de uma *v.a.*

Definição 3.11. Atribuí-se o termo desvio padrão de uma variável aleatória como sendo:

$$\sigma = \sqrt{Var(X)}$$

Pelo fato do desvio padrão estar na mesma unidade de medida da variável, é o valor dele que será aplicado como medida de dispersão para a interpretação dos dados de uma distribuição.

Existe um resultado muito interessante que é válido para qualquer distribuição de probabilidade, o qual está associado à variância. É conhecido como Teorema de Chebyshev.

Teorema 1. *Se X é uma variável aleatória com média μ e variância σ^2 , então para qualquer valor de $K > 0$:*

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2} \quad (3.4)$$

Esse teorema estabelece limites para probabilidades quando se conhece a média e a variância de uma distribuição de probabilidade. Claro que, se a função de probabilidade da distribuição é conhecida, utilizá-la fornecerá probabilidades mais exatas. No entanto, quando isso não ocorre, o Teorema de Chebyshev é de grande utilidade.

Uma demonstração desse teorema pode ser encontrada em Ross (2010).

3.4 Variáveis aleatórias contínuas

Definição 3.12. *Define-se por variável aleatória contínua (v.a.c) a uma v.a X com um intervalo (finito ou infinito) de números reais em sua faixa.*

Há diversas variáveis que se enquadram nessa classificação, como estatura de uma população, medida de corrente elétrica e dados pluviométricos. Diferente das variáveis aleatórias discretas, não é atribuído probabilidades a valores específicos, por existir uma quantidade não enumerável. Sendo assim, a probabilidade é dada para intervalos de valores, que é obtida a partir da função densidade de probabilidade.

3.4.1 Função densidade de probabilidade

Definição 3.13. *A uma variável aleatória contínua X associa-se a **Função Densidade de Probabilidade** (fdp) que satisfaz as seguintes condições:*

i) $f(x_i) \geq 0$;

ii) $\int_{-\infty}^{\infty} f(x)dx = 1$;

iii) $P(a \leq X \leq b) = \int_a^b f(x)dx$ que é a área sob $f(x)$ de a a b para quaisquer a e b .

A partir da interpretação geométrica, conclui-se que a probabilidade é a área sob a curva da fdp. Portanto, se X é uma variável aleatória contínua, a probabilidade de

qualquer valor específico é nula ($P(X = x) = 0$), já que corresponderá a um segmento de reta. Consequentemente:

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a \leq X < b) = P(a < X < b), \quad \forall a, b.$$

3.4.2 Função de distribuição cumulativa

Definição 3.14. A **Função de distribuição cumulativa** de uma variável aleatória contínua X é:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(x)dx, \quad -\infty \leq x \leq \infty \quad (3.5)$$

onde são válidas as propriedades:

- i) $0 \leq F(x) \leq 1$;
- ii) $\lim_{x \rightarrow -\infty} F(x) = 0$;
- iii) $\lim_{x \rightarrow \infty} F(x) = 1$;
- iv) $a < b \Rightarrow F(a) \leq F(b)$.

A função de distribuição cumulativa de uma variável aleatória contínua é definida para todos os números reais, logo é uma função contínua, diferente da função cumulativa de uma *v.a.d.* Por esse motivo é possível encontrar algumas vezes *v.a.c* sendo definida através dessa característica.

Agora observe que, derivando os dois lados da igualdade da equação (3.5) obtêm-se:

$$\begin{aligned} \frac{d}{dx} F(x) &= \frac{d}{dx} \int_{-\infty}^x f(x)dx \\ \frac{d}{dx} F(x) &= f(x), \end{aligned}$$

Logo, desde que a derivada exista, é possível obter a função densidade a partir da função cumulativa.

3.4.3 Valor esperado e variância

Supondo que $f(x)$ seja a *fdp* de uma variável aleatória contínua X :

Definição 3.15. Sendo denotado por μ ou $E(X)$ então,

$$\mu = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

é a média ou valor esperado de X .

Quanto a variância,

Definição 3.16. Denotada por σ^2 ou $V(X)$ define-se

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2$$

a variância de uma variável contínua.

Consequentemente, o desvio padrão de X é $\sigma = \sqrt{\sigma^2}$.

3.5 Distribuição Geométrica

A distribuição Geométrica, que é uma distribuição de variável aleatória discreta, ocorre em experimentos que possuem uma característica específica. Essa característica é denominada de ensaios, tentativas ou experimentos de Bernoulli.

Definição 3.17. *Ensaio de Bernoulli são experimentos independentes que a cada tentativa de ocorrência apresenta apenas dois possíveis resultados, denominados de sucesso ou fracasso.*

Observa-se então que diversas situações do cotidiano se caracterizam como ensaios de Bernoulli, por exemplo, participar de um sorteio (pode ser sorteado ou não), fazer um arremesso no basquete, pois pode acertar ou não a cesta, participar de um concurso (é aprovado ou reprovado), uma peça ser defeituosa ou não, entre outras.

Geralmente, espera-se observar um resultado ao realizar um experimento, ou seja, o sucesso, o qual possui uma probabilidade p de ocorrência. Consequentemente, o fracasso terá uma probabilidade de $1 - p$, uma vez que é complementar ao sucesso. Deste modo, se em um experimento as tentativas forem realizadas até obter a ocorrência do sucesso, as probabilidades dessas tentativas definem uma distribuição Geométrica.

Definição 3.18. *Seja a variável aleatória X o número de tentativas de Bernoulli até atingir o primeiro sucesso. Então X segue uma distribuição Geométrica, cuja função de probabilidade é dada por:*

$$f(x) = (1 - p)^{x-1} p, \quad x = 1, 2, 3, \dots \quad (3.6)$$

sendo denotada por $X \sim Geo(p)$ onde p é o parâmetro.

Note que de fato a equação (3.6) é válida, pois para ocorrer o sucesso apenas na n -ésima tentativa, nas $n - 1$ primeiras devem ocorrer o fracasso, cuja probabilidade é de $(1 - p)^{n-1}$, e o sucesso na n -ésima, resultando numa probabilidade de $(1 - p)^{n-1} p$, visto que as tentativas são independentes. Ademais, trata-se de uma função de probabilidade,

pois:

i) $f(x) \geq 0$ para qualquer x ;

$$\text{ii) } \sum_{x=1}^{\infty} f(x) = \sum_{x=1}^{\infty} (1-p)^{x-1} p = p \sum_{x=1}^{\infty} (1-p)^{x-1} = p \cdot \frac{1}{1-(1-p)} = \frac{p}{1-1+p} = 1.$$

veja que $\sum_{x=1}^{\infty} (1-p)^{x-1}$ é soma infinita de uma Progressão Geométrica de razão $1-p$ com primeiro termo $a_1 = 1$.

Como visto na Seção 3.1, a toda distribuição de probabilidade pode-se atribuir o valor esperado, no caso da distribuição Geométrica, esse valor indica o número de tentativas, em média, que devem ser realizadas para ocorrer o sucesso, sendo dada por:

$$\mu = E(X) = \frac{1}{p}. \quad (3.7)$$

de fato, a partir da Definição 3.5 e da função de probabilidade dada por (3.6), têm-se:

$$E(X) = \sum_{x=1}^{\infty} xp(x) = \sum_{x=1}^{\infty} x(1-p)^{x-1}p,$$

substituindo $1-p$ por q ,

$$\sum_{x=1}^{\infty} (x-1+1)q^{x-1}p = \sum_{x=1}^{\infty} (x-1)q^{x-1}p + q^{x-1}p = \sum_{x=1}^{\infty} (x-1)q^{x-1}p + \sum_{x=1}^{\infty} q^{x-1}p.$$

Agora, note que $\sum_{x=1}^{\infty} q^{x-1}p = 1$ e fazendo $x-1 = k$ têm-se:

$$\sum_{k=0}^{\infty} kq^k p + 1 = q \sum_{k=1}^{\infty} kq^{k-1}p + 1 = q \cdot E(X) + 1$$

ou seja,

$$\begin{aligned} E(X) &= q \cdot E(X) + 1 \Rightarrow \\ E(X) - q \cdot E(X) &= 1 \Rightarrow \\ E(X)(1-q) &= 1 \Rightarrow \\ E(X) &= \frac{1}{1-q} = \frac{1}{p}. \end{aligned}$$

Quanto a variância de uma distribuição Geométrica, é definida por:

$$\sigma^2 = V(X) = \frac{1-p}{p^2}. \quad (3.8)$$

Para demonstrar o cálculo da variância usa-se os seguintes resultados:

$$V(X) = E[X^2] - (E[X])^2$$

e

$$E[X^2] = \frac{2-p}{p^2}$$

A demonstração de ambos resultados é encontrada em Ross (2010).

Agora, fica facilmente mostrar que:

$$V(X) = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.$$

Uma aplicação de distribuição Geométrica é mostrada no Exemplo 3.2.

Exemplo 3.2. *As chances de uma pessoa fazer uma cesta no basquete é $2/5$. Qual é a probabilidade dele acertar apenas no terceiro arremesso? Quantos arremessos, em média, deverá realizar para que ela acerte o arremesso?*

Observe que no Exemplo 3.2, apresenta-se uma situação de ensaios de Bernoulli, cuja probabilidade de sucesso é $p = 2/5$ e probabilidade de fracasso é $1 - p = 3/5$. Como espera-se que o jogador acerte a cesta apenas no terceiro arremesso, então deve-se ocorrer fracasso nos dois primeiros. Portanto trata-se de uma distribuição Geométrica, cuja variável aleatória é $X =$ o número de arremessos a serem realizados até fazer a cesta. Logo, a função de probabilidade é definida como:

$$f(x) = \left(\frac{3}{5}\right)^{x-1} \frac{2}{5}, \quad x = 1, 2, 3, \dots$$

Sendo assim, a probabilidade da pessoa fazer cesta no terceiro arremesso é:

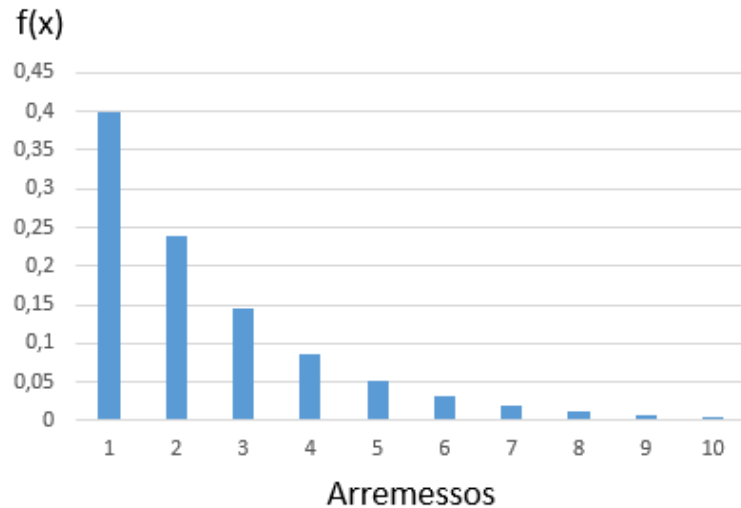
$$f(3) = \left(\frac{3}{5}\right)^2 \frac{2}{5} = \frac{9}{25} \cdot \frac{2}{5} = \frac{18}{125} = 0,144.$$

Quanto o número de arremessos que devem ser realizados para que se tenha uma probabilidade maior de acertar, será dado pelo valor esperado:

$$E(X) = \frac{1}{p} = \frac{5}{2} = 2,5 \text{ arremessos.}$$

O gráfico da Figura 4 apresenta a distribuição Geométrica do Exemplo 3.2 para os valores de $x = 1, 2, \dots, 10$. Observe que a probabilidade é quase nula para valores acima de $x > 9$, e que trata-se de um espaço amostral infinito, por não ter um limite de tentativas.

Figura 4 – Distribuição Geométrica do Exemplo 3.2.



Fonte: Própria autora, 2021.

3.6 Distribuição Binomial

Assim como na distribuição Geométrica, a distribuição Binomial surge também em ensaios de Bernoulli. No entanto, na distribuição Binomial a quantidade dessas tentativas é pré determinada, isto é, finita. Além disso, deseja-se observar x sucessos dentre todas as n tentativas, com uma probabilidade de sucesso constante e não mais o número de tentativas até o sucesso, como ocorria na Geométrica. A distribuição das probabilidades de obtenção de x sucessos em n tentativas caracteriza uma distribuição Binomial.

Definição 3.19. *Seja a variável aleatória X como sendo o número de sucessos de n tentativas de Bernoulli. Então X segue uma distribuição de probabilidade Binomial, cuja função de probabilidade é dada por:*

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n. \quad (3.9)$$

podendo ser denota por $X \sim B(n, p)$, em que n e p são os parâmetros da distribuição.

Intuitivamente, a função de probabilidade da distribuição Binomial pode ser obtida a partir da seguinte análise: de n tentativas deseja-se que ocorram x sucessos, conseqüentemente $n - x$ fracassos, resultando numa probabilidade de $p^x (1-p)^{n-x}$. Entretanto, há diversas formas de ordenar esses sucessos, conforme mostra a Figura 5. E essa quantidade é

determinada pela combinação $\binom{n}{x}$, resultando na função de probabilidade da distribuição Binomial. Além disso, de fato, a Equação (3.9) é uma distribuição de probabilidade, uma vez que, pelo Teorema Binomial têm-se:

$$\sum_{x=0}^n f(x) = \sum_{x=0}^n \binom{n}{x} p^x (1-p)^{n-x} = [p + (1-p)]^n = 1^n = 1$$

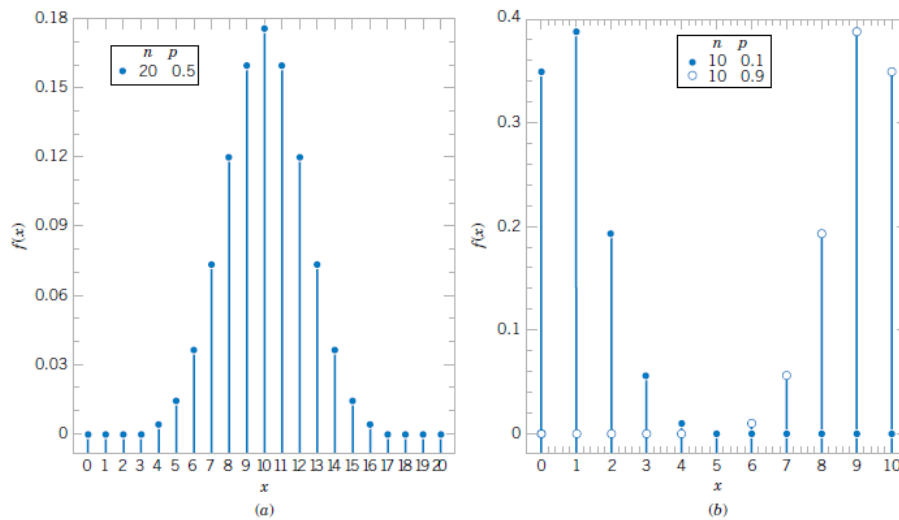
Figura 5 – Algumas formas de obter x sucessos em n tentativas.

$$\begin{array}{c} \underbrace{SSS \dots SS}_{x} \underbrace{FF \dots F}_{n-x} \\ \underbrace{FF \dots F}_{n-x} \underbrace{SSS \dots SS}_{x} \\ \underbrace{SSS \dots S}_{x-1} \underbrace{FF \dots F}_{n-x} S \end{array}$$

Fonte: Própria autora, 2021.

Veja na Figura 6 como se comporta a distribuição Binomial para alguns parâmetros.

Figura 6 – Distribuições Binomiais para valores selecionados de n e p .



Fonte: Montgomery e Runger (2002).

Quanto ao valor esperado de uma distribuição Binomial, ele indicará o número de sucessos de n tentativas que possui a maior probabilidade, sendo determinado por:

$$E(X) = np.$$

Veja a demonstração:

$$\text{Tome } E[X^K] = \sum_{i=0}^n i^k \binom{n}{i} p^i (1-p)^{n-i} = \sum_{i=1}^n i^k \binom{n}{i} p^i (1-p)^{n-i}.$$

Considerando que $i \binom{n}{i} = n \binom{n-1}{i-1}$ têm-se:

$$E[X^K] = np \sum_{i=1}^n i^{k-1} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i}.$$

Fazendo-se $j = i - 1$,

$$np \sum_{j=0}^{n-1} (j+1)^{k-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} = np E[(Y+1)^{k-1}],$$

em que $Y \sim B(n-1, p)$. Fazendo $k = 1$ conclui-se que:

$$E[X] = np E[(Y+1)^0] = np.$$

Agora, para a variância de uma distribuição Binomial, veja que,

$$\begin{aligned} V(X) &= E[X^2] - (E[X])^2 \\ &= np[(n-1)p + 1] - (np)^2 \\ &= np(1-p). \end{aligned}$$

Portanto,

$$\sigma^2 = V(X) = np(1-p).$$

Veja no Exemplo 3.3 retirado do livro Larson e Farber (2010) uma aplicação da distribuição Binomial.

Exemplo 3.3. *Cirurgias de microfraturas no joelho têm 75% de chance de sucesso em pacientes com joelhos degenerativos. Se a cirurgia for realizada em três desses pacientes, qual a probabilidade de ser um sucesso em exatamente dois pacientes?*

Neste exemplo têm-se que $X =$ número de cirurgias com sucesso. Assim, pode-se dizer que X segue uma distribuição Binomial de parâmetros $n = 3$ e $p = 3/4$. Logo,

$$f(x) = \binom{3}{x} \left(\frac{3}{4}\right)^x \left(\frac{1}{4}\right)^{3-x}, \quad x = 0, 1, 2, 3,$$

é a função de probabilidade dessa distribuição.

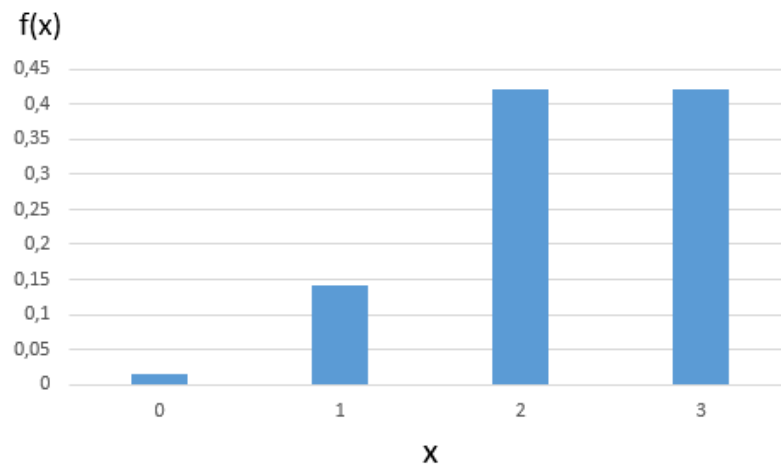
Para descobrir a probabilidade de sucesso em dois pacientes é preciso calcular $f(2)$:

$$f(2) = \binom{3}{2} \left(\frac{3}{4}\right)^2 \left(\frac{1}{4}\right) = 3 \cdot \frac{9}{16} \cdot \frac{1}{4} = \frac{27}{64} = 0,4218.$$

Portanto, a probabilidade de duas cirurgias bem sucedidas é 42,18%, aproximadamente.

A Figura 7 mostra o gráfico dessa distribuição:

Figura 7 – Distribuição Binomial do Exemplo 3.3.



Fonte: Própria autora, 2021.

Ademais, note que $E(X) = 3 \cdot 0,75 = 2,25$.

3.7 Distribuição Normal

Considere um experimento aleatório. A medida que esse experimento é replicado e a variável aleatória for igual ao resultado médio das réplicas, essa variável tenderá a ter uma distribuição normal (MONTGOMERY e RUNGER, 2020).

De acordo com Larson e Farber (2010), é possível encontrar a distribuição normal em vários estudos da natureza, indústria e negócios. Não é por acaso, que é uma das mais utilizadas ao se tratar de variável aleatória contínua.

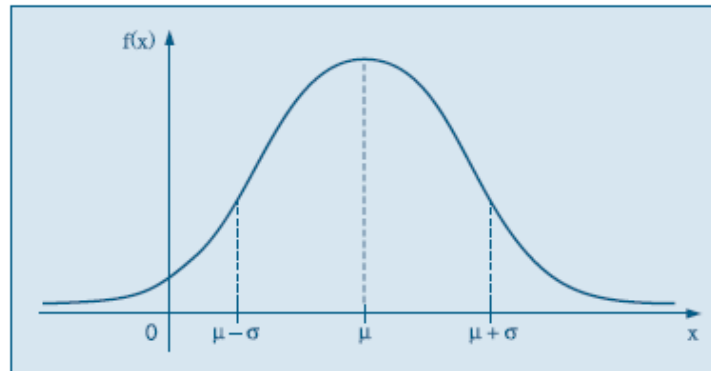
Definição 3.20. *Uma variável aleatória contínua X , cuja função densidade de probabilidade é dada por:*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (3.10)$$

define uma distribuição normal com parâmetros μ e σ^2 , com $-\infty < x < \infty$, $E(X) = \mu$ e $V(X) = \sigma^2$, podendo ser denota por $X \sim N(\mu, \sigma^2)$.

A curva de uma distribuição normal é simétrica e apresenta um formato de sino, conforme mostra a Figura 8.

Figura 8 – Distribuição normal com parâmetros μ e σ .



Fonte: Moretin e Bussab (2010).

Toda distribuição normal apresenta as seguintes propriedades:

- A média μ é ponto máximo de $f(x)$;
- A curva é simétrica em relação a reta $x = \mu$;
- A média, moda e mediana são coincidentes;
- Os pontos de inflexão ocorrem em $\mu - \sigma$ e $\mu + \sigma$;
- A curva apresenta uma assíntota em relação ao eixo x a medida que se distancia da média, isto é, $\lim_{x \rightarrow \pm\infty} f(x) = 0$.

Conforme a Definição 3.4.1, ao se tratar de variáveis aleatórias contínuas, o cálculo de probabilidades é obtido a partir da área sob a curva da função densidade de probabilidade, logo outra características da distribuição normal é:

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

A demonstração pode ser lida em Meyer (1970).

Há um outro resultando importante, o qual associa probabilidades com o desvio padrão de uma distribuição normal, sendo ele:

$$P(\mu - \sigma < X < \mu + \sigma) = 0,6827$$

$$P(\mu - 2\sigma < X < \mu + 2\sigma) = 0,9545$$

$$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0,9973$$

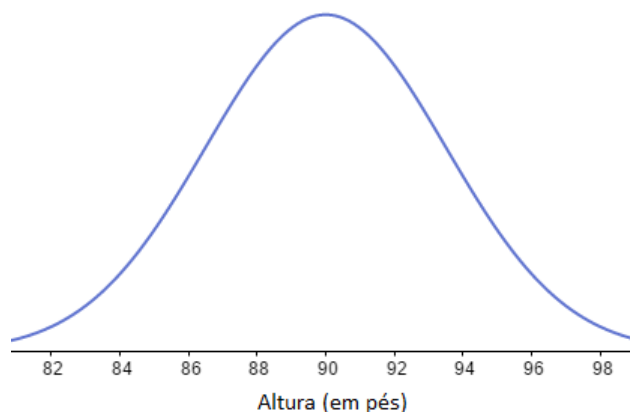
Além disso, por consequência da simetria, têm-se:

$$P(X > \mu) = P(X < \mu) = 0,5.$$

O Exemplo 3.4 é uma aplicação da distribuição Normal na biologia envolvendo alturas de árvores.

Exemplo 3.4. (LARSON e FARBER, 2010) *As alturas (em pés) de árvores de carvalho adultas são normalmente distribuídas. A curva normal representada na Figura 9 mostra essa distribuição, cujos parâmetros são $\mu = 90$ e $\sigma = 3,5$.*

Figura 9 – Altura de carvalhos normalmente distribuídas.



Fonte: Própria autora, 2021.

Fazendo uma análise das propriedades para a distribuição do Exemplo 3.4, pode-se concluir que:

a) A média das alturas dos carvalhos é de 90 pés, logo o ponto máximo de $f(x)$ ocorre para $x = 90$;

b) Há pontos de inflexão em $x = 86,5$ e $x = 93,5$. Os pontos de inflexão limitam o intervalo $(86,5; 93,5)$ que é o intervalo correspondente a $(\mu - \sigma; \mu + \sigma)$. Isso significa que, pela propriedade $P(86,5 < X < 93,5) = 0,6827$, ou seja, a probabilidade de uma árvore ter altura entre 86,5 pés e 93,5 pés é de 68,27%. Além disso, pode-se afirmar também que 95,45% equivale a probabilidade de uma árvore ter altura entre 83 e 97 pés e, 99,73% é a probabilidade de um carvalho ter altura entre 79,5 e 100,5 pés, uma vez que $P(83 < X < 97) = 0,9545$ e $P(79,5 < X < 100,5) = 0,9973$.

3.7.1 Distribuição Normal Padrão

Uma vez que existem infinitos valores de média (μ) e variância (σ) para uma variável aleatória, então há infinitas distribuições normais. Particularmente, uma distribuição normal $N(0, 1)$ é denominada de distribuição normal padrão.

Definição 3.21. *Uma distribuição normal com parâmetros $\mu = 0$ e $\sigma = 1$ é denominada de Distribuição Normal Padrão, com função densidade de probabilidade igual a:*

$$f(z) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(z)^2}{2}}, \quad (3.11)$$

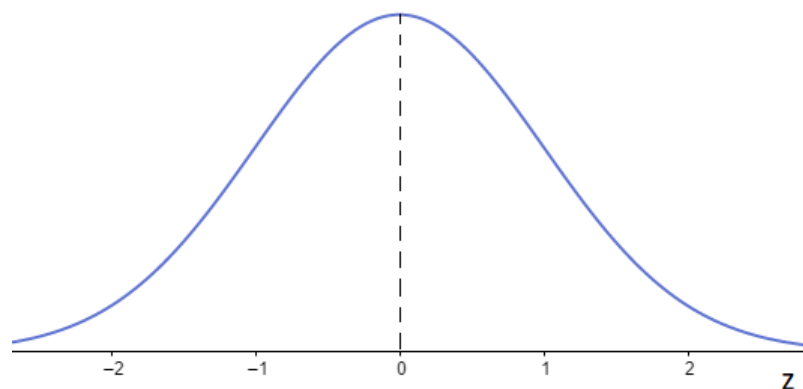
com $-\infty < x < \infty$, e função cumulativa:

$$\Phi(z) = P(Z \leq z), \quad (3.12)$$

em que Z é a notação para a variável aleatória contínua que segue uma distribuição normal padrão.

A Figura 10 mostra graficamente essa distribuição.

Figura 10 – Distribuição Normal Padrão.



Fonte: Própria autora, 2021.

Deste modo, para calcular probabilidades de uma distribuição normal padrão, deve-se calcular:

$$P(a \leq Z \leq b) = \frac{1}{\sqrt{2\pi\sigma}} \int_a^b e^{-\frac{(z)^2}{2}} dz.$$

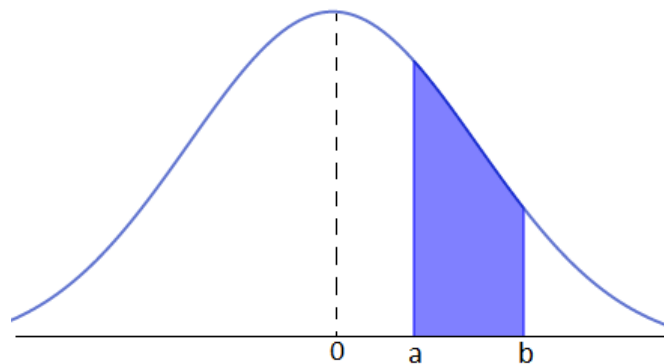
Porém, não é possível calcular essa integral a partir do Teorema Fundamental do Cálculo, visto que não existe a integral de $e^{-x^2/2}$. No entanto, existem métodos de integração numérica que permitem realizar esse cálculo, mas na prática se usa a tabela do Anexo A, uma vez que esses cálculos foram tabulados. A tabela informa o valor da

área sob a curva do intervalo $[0, z_0]$, ou seja, ao buscar o valor referente a z_0 , encontra-se a probabilidade $P(0 \leq z_0 \leq Z)$.

A princípio, entende-se que é possível determinar apenas probabilidades como $P(0 \leq z_0 \leq Z)$. Porém, a partir da interpretação geométrica verifica-se que é possível calcular probabilidades de quaisquer intervalos.

Observe a área destacada na Figura 11. É notório que a mesma refere-se a diferença entre a área compreendida do intervalo $[0, b]$ e $[0, a]$. Deste modo, ao utilizar a tabela da distribuição normal padrão, basta encontrar os valores para $z = b$ e $z = a$ e efetuar a diferença entre eles, obtendo assim, a probabilidade desejada.

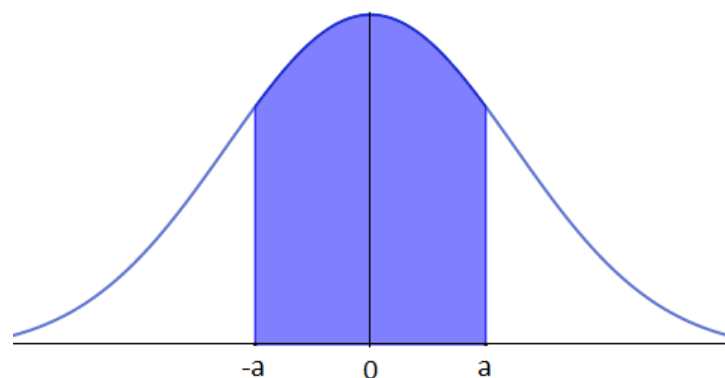
Figura 11 – Calculando a probabilidade $P(a \leq Z \leq b)$.



Fonte: Própria autora, 2021.

Agora, considere que deseja-se determinar a probabilidade de ocorrência do intervalo $[-a, 0]$, conforme Figura 12. Note que a área compreendida neste intervalo é equivalente ao $[0, a]$ devido a simetria. Portanto, independentemente do intervalo, é possível utilizar a tabela de distribuição normal padrão e obter os resultados desejados.

Figura 12 – Calculando a probabilidade $P(-a \leq Z \leq 0)$.



Fonte: Própria autora, 2021.

É válido ressaltar que a tabela de distribuição normal padrão utilizada neste trabalho não é a única existente. Alguns autores, como Montgomery e Runger (2020), trabalham com a tabela da função cumulativa, ou seja, os valores para $\Phi(z)$.

Visto que é utilizada a tabela para determinar probabilidade de uma distribuição normal padrão, como então calcular probabilidades de uma distribuição normal com parâmetros arbitrários? É nessa situação que se enquadra um dos processos mais importantes ao se lidar com uma distribuição normal: a padronização de uma variável aleatória. A Definição 3.22 mostra como pode ser padronizada uma variável aleatória normal arbitrária:

Definição 3.22. *Seja X é uma variável aleatória normal com $E(X) = \mu$ e $V(X) = \sigma^2$, a variável aleatória*

$$Z = \frac{X - \mu}{\sigma}, \quad (3.13)$$

é uma variável aleatória normal, com $E(Z) = 0$ e $V(Z) = 1$. Ou seja, Z é uma variável aleatória normal padrão.

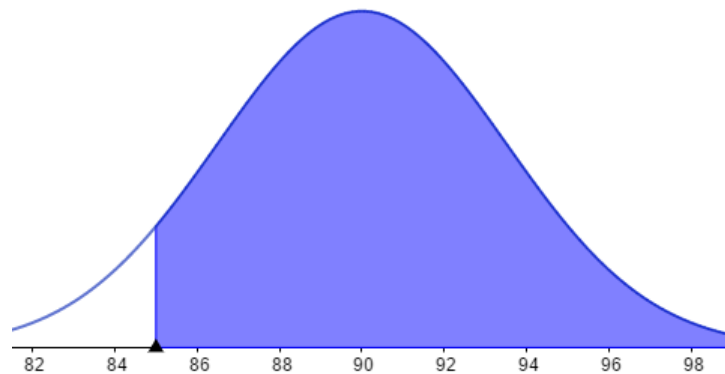
Observe que a equação (3.13) associa, a partir de uma transformação, a variável aleatória X com a variável Z , possibilitando determinar probabilidades de qualquer distribuição normal, uma vez que

$$P(X \leq x) = P\left(\frac{X - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) = P(Z \leq z).$$

Observe como é utilizado a padronização de uma *v.a.c* que segue uma distribuição normal.

Considere novamente o Exemplo 3.4. Lembrando que neste caso têm-se uma distribuição normal $N(90; 3, 5)$, suponha que deseja-se determinar a probabilidade de ocorrência de um carvalho com altura maior que 85 pés, que é equivalente a área em destaque da Figura 13.

Figura 13 – Probabilidade de um carvalho ter altura maior que 85 pés.



Fonte: Própria autora, 2021.

Para tanto, primeiramente determina-se a padronização X , onde X é a altura (em pés) da árvore carvalho. Assim, determina-se z , tal que $P(85 \leq X) = P(z \leq Z)$. Logo, é suficiente padronizar apenas o valor 85.

Como 85 está a esquerda da média, então deve-se considerar seu simétrico, que é o 95:

$$z = \frac{95 - 90}{3,5} = 1,428 \cong 1,43.$$

Buscando esse valor de $z = 1,43$ na tabela encontra-se 0,4236, o que significa que:

$$P(0 \leq Z \leq 1,43) = P(90 \leq Z \leq 95) = P(85 \leq Z \leq 90) = 0,4236.$$

Porém, note que:

$$z = \frac{85 - 90}{3,5} = -1,428,$$

ou seja, não é necessário encontrar os simétricos no momento da padronização, basta considerar o módulo de z quando o mesmo é negativo.

No entanto, a probabilidade buscada é $P(85 \leq X) = P(1,43 \leq Z)$ que é área resultante da união das áreas compreendidas dos intervalos $[1,43; 0]$ e $[0, -\infty)$, logo

$$P(85 \leq X) = P(1,43 \leq z \leq 0) + P(0 \leq z) = 0,4236 + 0,5 = 0,9236$$

Portanto, a probabilidade de ocorrência de um carvalho ter altura maior que 85 pés equivale a 92,36%, aproximadamente.

Todo o processo descrito para determinar probabilidades de uma distribuição normal se refere ao cálculo “manual”, porém, é claro que atualmente existem programas que efetuam esses cálculos, programas simples e de fácil acesso, como o *software* Geogebra, que é utilizado na Atividade 3 do Capítulo 4.

4 Propostas de atividades

Buscando aplicações de probabilidade que saem do óbvio de jogos de azar e, de acordo com os requisitos dos PCN's e da BNCC conforme discutido no Capítulo 2, neste capítulo são apresentadas propostas de atividades de distribuições de probabilidades, as quais aplicam a probabilidade em situações da vida real, além de trazer o uso das tecnologias como aliado no ensino básico.

As duas primeiras atividades trazem distribuições de variáveis discretas e a terceira, variável contínua. De acordo com os seus desenvolvimentos, as duas primeiras serão melhores aproveitadas se aplicadas na 2ª série do Ensino Médio, enquanto a terceira, pode ser tanto para a 2ª série quanto para a 3ª série do Ensino Médio, por envolver também alguns conceitos de estatística. Deste modo, professores de matemática que ministram aula nessas turmas são o público alvo para este trabalho.

4.1 Conhecimentos prévios

Pensando nas possíveis dificuldades que surgirão no desenvolvimento das atividades destacam-se: tecnologia, tempo e defasagem dos alunos sobre conceitos básicos de probabilidade.

Tecnologia: embora o uso da tecnologia na Educação Básica é indicado desde os PCN's, sabe-se que a realidade das escolas públicas do Brasil está bem distante disso, uma vez que muitas não possuem nem computadores disponíveis para o uso em aulas. De acordo com essa realidade, na descrição das etapas que abordam o uso da tecnologia são sugeridas alternativas para que possam dar sequência nas atividades.

Já nas escolas que disponibilizam esses recursos, o fator a se considerar é a falta de conhecimentos básicos dos programas, não só dos alunos como também do professor. Com intuito de facilitar, na descrição das etapas tentou-se informar o passo a passo de como utilizar os programas e além disso os arquivos das tabelas e gráficos construídos nas atividades estão disponíveis para download¹. No entanto, caso o professor deseje seus conhecimentos sobre os *softwares* sugere-se assistir aos vídeos “As 5 fórmulas básicas do Excel que você tem que aprender antes do PROCV”² e o “Geogebra por Elas: Conhecendo e Explorando o Geogebra”³, uma oficina oferecida pelo IFCE.

Cabe ressaltar que o Excel é um programa pertencente ao pacote Office da Microsoft.

¹ Disponível em: <https://drive.google.com/drive/folders/16V1yCHc12wJA0TF5rzpHZPybfo_JAQ6r?usp=sharing>

² Disponível em: <<https://www.youtube.com/watch?v=Ufmem0LjubI&list=WL&index=30>>

³ Disponível em: <<https://www.youtube.com/watch?v=3heX2rhXqYo&list=WL&index=32>>

Caso não seja viável sua instalação, é possível utilizá-lo online⁴ assim como o Geogebra, disponível gratuitamente tanto para uso online ou para download⁵. A versão do Geogebra utilizada neste trabalho é a Classico 6.

Tempo: é comum não ter tempo para desenvolver atividades extras, visto que existem redes de escolas que o professor deve cumprir um currículo. Assim, como forma de otimizar o tempo, as etapas das atividades podem ser inclusas no dia-a-dia das aulas como propostas de exercícios, desenvolvendo de acordo com os conteúdos trabalhados, pelo menos até o momento da abordagem sobre a distribuição. Outra possibilidade é considerar as atividades como um projeto, possibilitando ter mais flexibilidade para a aplicação das mesmas ou aplicar apenas uma atividade.

Defasagem na aprendizagem: acredita-se que as dificuldades que os alunos possam apresentar são consequências da defasagem do ensino de probabilidade, como por exemplo, a associação entre número decimal e porcentagem, já que em todas as propostas deste trabalho o cálculo da probabilidade são escritos apenas na forma decimal, e a noção de que em qualquer distribuição a soma das probabilidades resulta em 1. Na verdade, essas possíveis dificuldades pertencem à habilidades exigidas pela BNCC no Fundamental I e II, que são:

-(EF05MA22) Apresentar todos os possíveis resultados de um experimento aleatório, estimando se esses resultados são igualmente prováveis ou não;

-(EF05MA23) Determinar a probabilidade de ocorrência de um resultado em eventos aleatórios, quando todos os resultados possíveis têm a mesma chance de ocorrer (equiprováveis);

-(EF06MA30) Calcular a probabilidade de um evento aleatório, expressando-a por número racional (forma fracionária, decimal e percentual) e comparar esse número com a probabilidade obtida por meio de experimentos sucessivos;

-(EF08MA22) Calcular a probabilidade de eventos, com base na construção do espaço amostral, utilizando o princípio multiplicativo, e reconhecer que a soma das probabilidades de todos os elementos do espaço amostral é igual a 1.

Deste modo, caso a turma apresente defasagem dessas habilidades, sugere-se ao professor enfatizar, no decorrer do processo, a habilidade que está sendo abordada. Além disso, pode-se também realizar algumas atividades propostas por Freitas Filho (2020), tais como 3.6, 3.7, 3.8, 3.10 e 3.11 que abrangem as habilidades apresentadas acima.

Ademais, caso o professor necessite de efetivar seu conhecimento sobre o assunto, sugere-se a leitura do Capítulo 3, e como complemento, a leitura de Larson e Farber (2010), pois os autores explicam com uma linguagem simples e Ross (2010), que traz uma abordagem mais técnica e apresenta demonstrações.

⁴ Disponível em: <https://www.office.com/launch/excel?ui=pt-BR&rs=BR&auth=1>

⁵ Disponível em: <https://www.geogebra.org/download>

4.2 Atividade 1: A distribuição Geométrica e a Mega-Sena

Objetivos: Aplicação de conceitos de probabilidade em situações reais por meio da distribuição geométrica presente no cálculo da probabilidade de ganhar na Mega-Sena, revisando habilidades da BNCC e ampliando o conhecimento ao estender os conceitos para distribuições de probabilidade.

Habilidades desenvolvidas conforme texto da BNCC:

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

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

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

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

Desenvolvimento: esta atividade é construtiva, de modo que os conceitos e teorias envolvidos na distribuição geométrica sejam apresentados e discutidos com os alunos no decorrer das etapas a seguir. Cálculo de fatorial e combinação são conhecimentos prévios que os alunos devem apresentar.

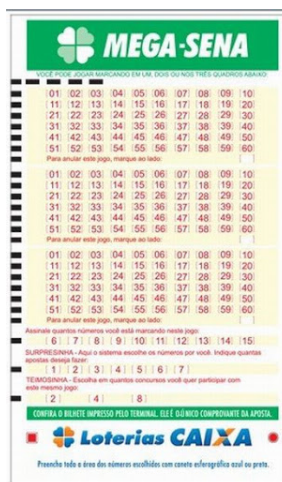
1ª Etapa: Simulando a Mega-Sena na sala de aula.

Simular o jogo da Mega-Sena com os alunos é uma boa maneira de iniciar a atividade, uma vez que o lúdico desperta uma interação e o envolvimento dos alunos. Pode-se utilizar a cartela específica do jogo conforme ilustra a Figura 14 ou até mesmo uma confeccionada pelo próprio professor.

Com a cartela em mãos cada aluno escolhe 6 números de 1 a 60, fazendo assim sua aposta para o jogo. Quanto ao sorteio dos números, o professor pode realizar como um bingo, colocando os papéis dos números em um saco e sortear seis deles. A cartela da Figura 15 pode ser utilizada. Ganhará o aluno que acertar todos os números.

⁶ Disponível em: <http://portaldoprofessor.mec.gov.br/fichaTecnicaAula.html?aula=57118>. Acesso em: 10 maio 2020.

Figura 14 – Cartela de aposta.



Fonte: Portal do professor⁶

Figura 15 – Cartela de números para sorteio.

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36
37	38	39	40	41	42
43	44	45	46	47	48
49	50	51	52	53	54
55	56	57	58	59	60

Fonte: Própria autora, 2021.

Nesse momento o professor pode usar a criatividade para deixar o sorteio descontraído. Durante o processo já trazer em pauta algumas discussões como: “Será que alguém vai ganhar?” e “Por que é tão difícil ganhar esse jogo?”

Após o sorteio, o professor pode verificar se algum aluno ganhou, quantos números cada aluno acertou e jogar mais algumas vezes, mostrando a dificuldade de ganhar esse jogo.

2ª Etapa: Calculando a probabilidade de ganhar na primeira tentativa.

Ao término do jogo, é interessante levar os alunos a refletirem sobre o porquê é tão

difícil conseguir esse prêmio, e se alguém sabe a probabilidade de ganhar fazendo apenas uma aposta em um concurso. Porém, antes de calcular a probabilidade, pode-se discutir e recordar com os alunos os conceitos envolvidos no sorteio de um número do jogo:

-Experimento aleatório não determinístico: cada vez que é sorteado um número (experimento aleatório), não sabe-se qual sairá (não determinístico).

-Espaço amostral finito: mesmo não sabendo qual número será sorteado, sabe-se quais são todas as possibilidades, que são os números de 1 a 60. Usualmente é representado pela letra S e quando é medido o tamanho desse espaço amostral, é utilizado a notação $n(S)$.

-Evento: é um subconjunto, uma parte do espaço amostral, como por exemplo, sair um dos números escolhidos. É representado por uma letra maiúscula, por exemplo evento A , e quando é medido seu tamanho, utiliza-se $n(A)$.

-Eventos equiprováveis: são eventos que tem a mesma chance (probabilidade) de ocorrer. As chances de sair o número 60 são as mesmas de sair o número 1. O mesmo se aplica para todos os 60 números que podem ser sorteados.

-Probabilidade de um evento A : é a razão entre o número de casos favoráveis (evento) e o número de todas as possibilidades (espaço amostral), ou seja,

$$P = \frac{n(A)}{n(S)}.$$

-Probabilidade de eventos equiprováveis: se todos os resultados de um experimento são equiprováveis, a probabilidade de ocorrer será $\frac{1}{60}$, dado que $n(S) = 60$.

Observe que essas reflexões são feitas considerando o sorteio de apenas um número. Porém, para calcular a probabilidade de ganhar um jogo da Mega-Sena, o espaço amostral passa a ser todas as possibilidades de agrupar 6 dos números de 1 a 60. Deste modo, considerando agora este espaço amostral, $n(S)$ será obtido através do conceito da combinação.

Uma forma de explicar para os alunos o uso da combinação é utilizando o princípio multiplicativo. O professor questiona aos alunos quantos números podem sair no primeiro sorteio, no caso 60, e considerando que não ocorre reposição, haverá 59 números para sortear o segundo e assim sucessivamente, como mostra a Figura 16.

Figura 16 – Calculando $n(S)$.

$$\frac{60}{1^\circ} \cdot \frac{59}{2^\circ} \cdot \frac{58}{3^\circ} \cdot \frac{57}{4^\circ} \cdot \frac{56}{5^\circ} \cdot \frac{55}{6^\circ} = 36.045.979.200$$

Fonte: Própria autora, 2021.

Ou seja, seriam 36.045.979.200 de apostas possíveis, no entanto, a ordem que números são sorteados não interfere. Como exemplo, o professor pode pegar uma aposta feita por

algum aluno da turma e questionar se trocar a ordem de algum número será uma aposta diferente, ou seja, suponha que o aluno escolheu os números 2, 17, 20, 34, 39 e 51. Se por acaso o sorteio sair na ordem 17, 2, 20, 34, 39 e 51, será a mesma aposta. Assim como trocar quaisquer números de posição. Portanto, para calcular todos os possíveis jogos que podem ser feitos, deve-se contar quantas vezes a escolha dos mesmos 6 números podem ser sorteadas, usando novamente o princípio multiplicativo (que pode ser tratado como permutação simples também).

Figura 17 – Calculando $n(S)$.

$$\frac{6}{1^\circ} \cdot \frac{5}{2^\circ} \cdot \frac{4}{3^\circ} \cdot \frac{3}{4^\circ} \cdot \frac{2}{5^\circ} \cdot \frac{1}{6^\circ} = 720$$

Fonte: Própria autora, 2021.

Como visto na Figura 17 os seis números podem ser ordenados em 720 maneiras distintas, sendo então necessário dividir 36.045.979.200 por 720, concluindo que há 50.063.860 maneiras de fazer uma aposta na Mega-Sena, isto é $n(S) = 50.063.860$.

Após essa explicação o professor diz que a combinação é utilizada nas situações semelhantes a essa. Para este caso, são agrupados 6 números de 60, ou seja,

$$C_{60}^6 = \binom{60}{6} = \frac{60!}{6!(60-6)!} = 50.063.860,$$

obtendo o mesmo resultado. Conseqüentemente, lembrando que a probabilidade é a razão entre $n(A)$ e $n(S)$, em que $n(A)$ é a quantidade de aposta a ser feita, a probabilidade de ganhar na Mega-Sena com uma aposta será:

$$\text{Probabilidade de sucesso} = p = \frac{1}{50063860} \cong 0,00000002.$$

Aqui o professor já pode deixar claro que irá denotar por p a probabilidade de ganhar, que também será chamada de probabilidade de sucesso. Além disso, deve ser desenvolvido o cálculo da probabilidade de não ganhar o jogo, recordando o conceito de eventos complementares.

-Eventos complementares: dois eventos são complementares quando a interseção é o conjunto vazio e a união resulta o espaço amostral.

Deste modo, a probabilidade de fracasso, sendo denotada por $1 - p$ será:

$$\text{Probabilidade de fracasso} = 1 - p = 0,99999998.$$

3ª Etapa: Calculando as probabilidades do número de apostas até conseguir ganhar.

Nesta etapa o conceito de eventos independentes deve ser retomado. Para tanto, questionar aos alunos: “considerando apenas uma aposta por jogo, qual será a probabilidade de alguém ganhar em sua segunda tentativa?” e também “Se perder no primeiro jogo, esse resultado irá interferir no segundo?” Quanto a última pergunta, espera-se que os alunos concluam que não. Sendo assim, o professor irá deixar claro que a justificativa para essa resposta é o fato desses eventos serem independentes, isto é, o resultado de um não interfere no resultado do outro. Além disso, o professor pode comentar que essa situação é também chamada de ensaios de Bernoulli independentes.

Quanto ao cálculo de probabilidade de eventos independentes ele será dado pelo produto da probabilidade de cada, ou seja, se A e B são eventos independentes, a probabilidade de ocorrer A e B é $P(A).P(B)$.

Após os alunos refletirem e darem suas respostas, sugere-se calcular as probabilidades de ganhar jogando 2, 3, 4 e 5 concursos.

- Jogando duas vezes: para que um jogador ganhe na segunda tentativa, ele deve perder a primeira. Como os eventos são independentes, então basta obter o produto da probabilidade de perder o primeiro concurso com a probabilidade de ganhar o segundo:

$$P(2) = (1 - p)p.$$

- Jogando três vezes: deve ocorrer fracasso nos dois primeiros concurso e sucesso no terceiro:

$$P(3) = (1 - p)(1 - p)p.$$

- Jogando quatro vezes: deve ocorrer fracasso nos três primeiros concursos e sucesso no quarto:

$$P(4) = (1 - p)(1 - p)(1 - p)p.$$

- Jogando cinco vezes: deve ocorrer fracasso nos quatro primeiros concursos e sucesso no quinto:

$$P(5) = (1 - p)(1 - p)(1 - p)(1 - p)p.$$

Em seguida trabalhar com os alunos a forma mais simplificada de escrever essas probabilidades, aplicando as propriedades de potenciação (se necessário rever o conceito de potência).

- Jogando duas vezes: $P(2) = (1 - p)^1 p$.

- Jogando três vezes: $P(3) = (1 - p)^2 p$.
- Jogando quatro vezes: $P(4) = (1 - p)^3 p$.
- Jogando cinco vezes: $P(5) = (1 - p)^4 p$.

Desenvolvida essa fase, o professor deve trazer uma reflexão para os alunos: vocês conseguem enxergar algum padrão nesses cálculos de probabilidades? Se sim, qual é? O que sempre se mantém e o que muda de acordo com a quantidade de tentativas?

Essas reflexões são importantes para auxiliar a dedução da função de probabilidade que deve ser feita conforme explicado a seguir:

- em 2 jogos o expoente da probabilidade de fracasso era $1 = 2 - 1$.
- em 3 jogos o expoente da probabilidade de fracasso era $2 = 3 - 1$.
- em 4 jogos o expoente da probabilidade de fracasso era $3 = 4 - 1$.
- em 5 jogos o expoente da probabilidade de fracasso era $4 = 5 - 1$.
- em todos os jogos o expoente da probabilidade de sucesso se manteve.

Em seguida, pergunta-se aos alunos: “em x jogos, como será a probabilidade? Como ficará a expressão?” E assim concluir que, para x jogos a probabilidade de ganhar será:

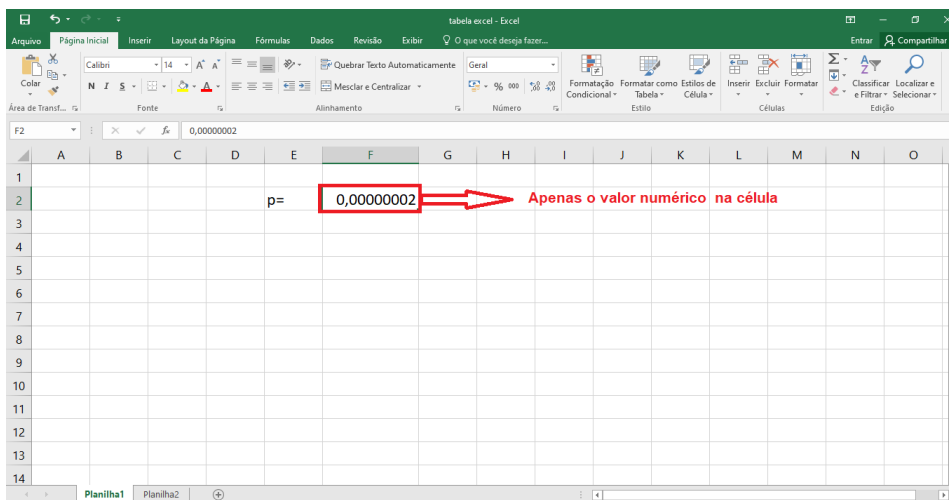
$$P(x) = (1 - p)^{x-1} p. \quad (4.1)$$

Até então, os cálculos se mantêm indicados. Como os números são muitos pequenos, quase nulos, as probabilidades podem ser calculadas utilizando o Excel. Caso os alunos não tenham acesso a computadores, se a escola possui um projetor, sugere-se ao professor projetar uma tela de um computador para que os vão acompanhando como o programa pode ser utilizado para o cálculo das probabilidades. Se a escola não possui projetor, o professor pode levar a tabela impressa para os alunos e, se possível, transmitir a ideia de como foi realizada.

Acompanhe a seguir como pode ser escrita a fórmula para calcular probabilidades da distribuição Geométrica com o parâmetro $p = 0,0000002$.

a) Em uma célula digitar a probabilidade de sucesso (aqui foi utilizado a célula F2). Veja a Figura 18.

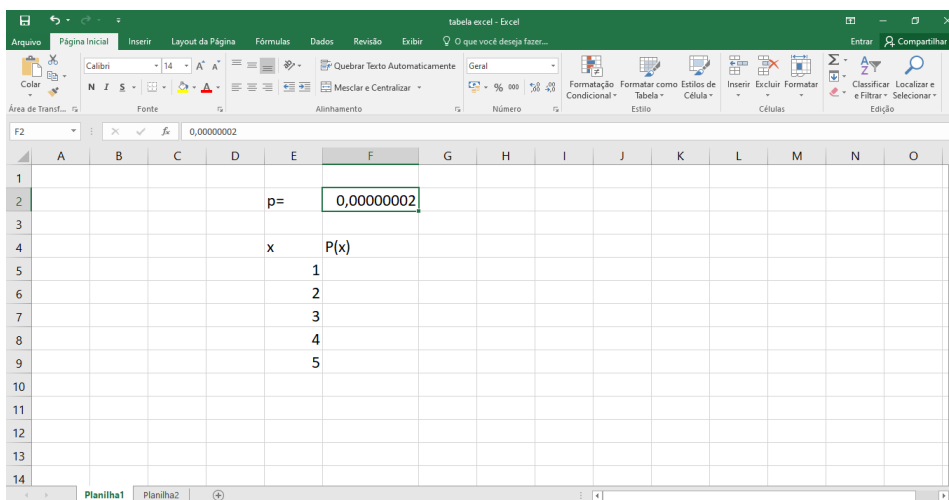
Figura 18 – Célula com a probabilidade de sucesso.



Fonte: Própria autora, 2021.

b) Construir a tabela de probabilidades, em que a primeira coluna representa os valores de x (número de tentativas para atingir o sucesso no jogo) e a segunda as respectivas probabilidades $P(x)$. Veja Figura 19.

Figura 19 – Criando a tabela.



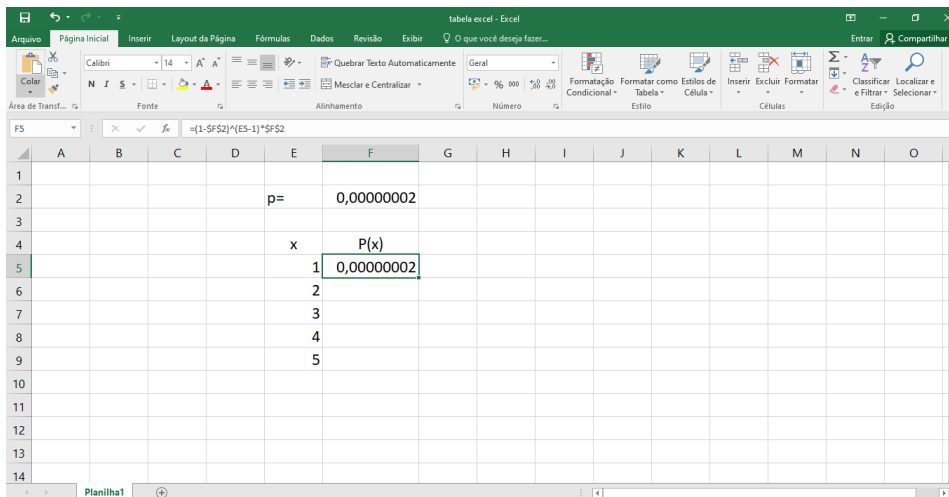
Fonte: Própria autora, 2021.

c) Calcular a probabilidade para uma tentativa ($P(1)$). Para isso, na célula respectiva digitar o comando:

$$=(1-$$$2)^(E5-1)*$$$2,$$

em que E5 é a célula de $x = 1$ e os cifrões tem a função de fixar a célula F2. Veja Figura 20.

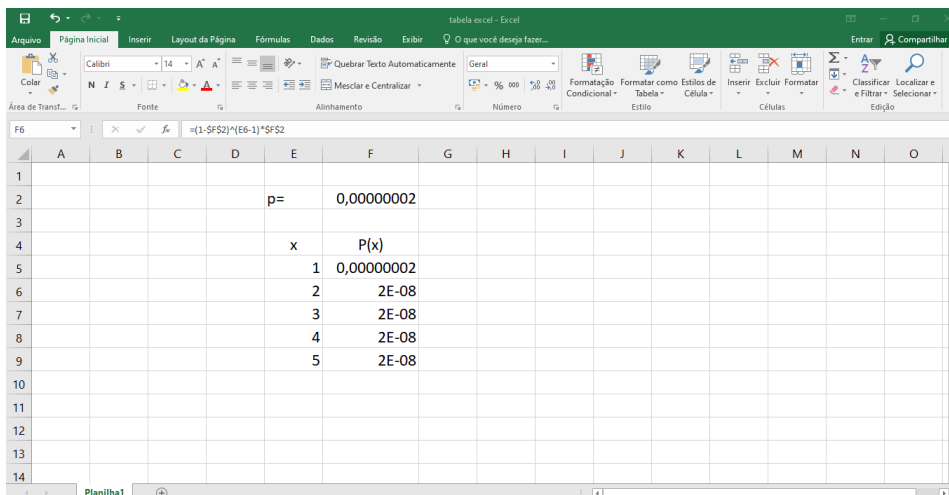
Figura 20 – Calculando P(1).



Fonte: Própria autora, 2021.

d) Para calcular P(2), P(3), P(4) e P(5), basta copiar o comando para as células. Veja Figura 21.

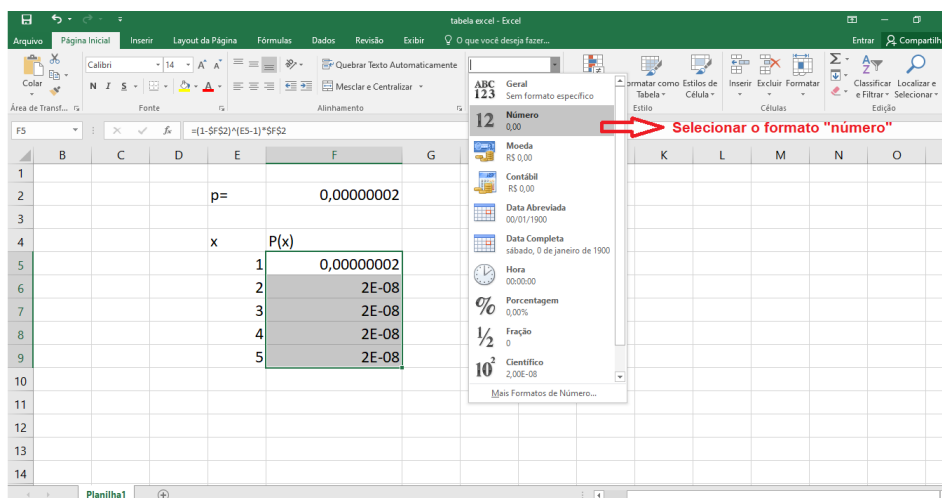
Figura 21 – Calculando as outras probabilidades.



Fonte: Própria autora, 2021.

Por serem números com muitas casas decimais, ao copiar para o restante das células, o resultado será dado em notação científica. Para transformar as probabilidades em números decimais, primeiramente coloca-se todas as células no formato de número, selecionando as células e clicando em “Número” na caixa de formato do Excel. Veja Figura 22.

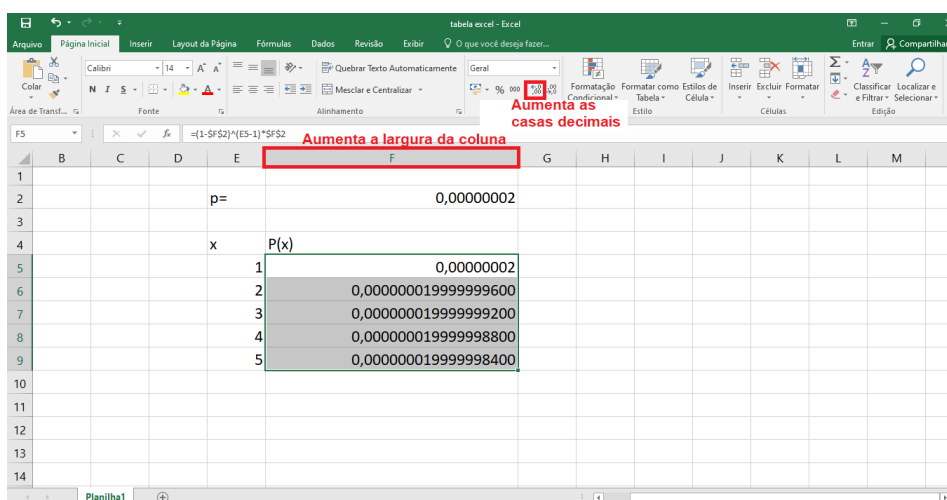
Figura 22 – Mudando o formato das células para números.



Fonte: Própria autora, 2021.

Em seguida, aumenta-se a largura da coluna e o número de casas decimais (o quanto for necessário), clicando na opção de aumentar casas decimais. Veja Figura 23.

Figura 23 – Probabilidades em decimais.

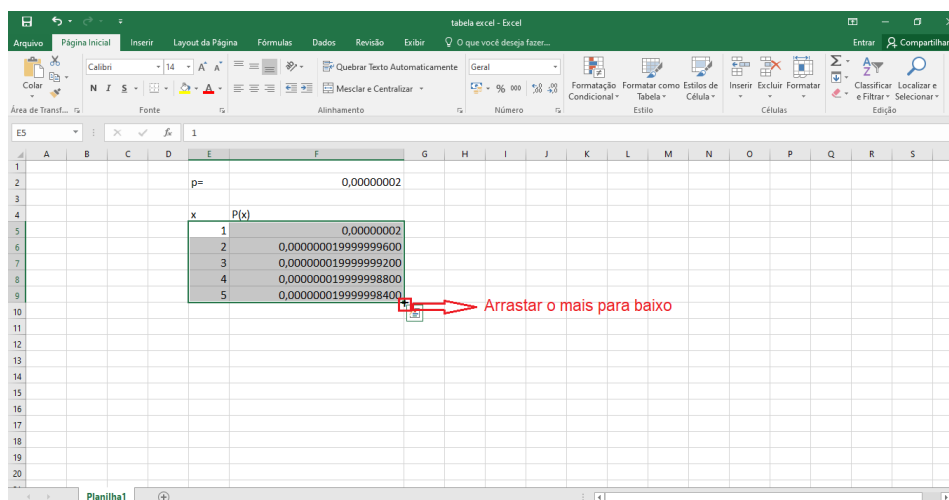


Fonte: Própria autora, 2021.

Embora a tabela tenha sido construída apenas com $x = 1, 2, 3, 4, 5$, ela pode ser estendida. Seleciona-se ambas as colunas e “arrasta” essa seleção para baixo, até obter o número desejado (veja Figura 24). Isso é interessante para realizar comparações entre as probabilidades e perceber que ao passo que as tentativas aumentam, as chances de ganhar diminuem.

4ª Etapa: A distribuição Geométrica e a Mega-Sena.

Figura 24 – Copiando as células.



Fonte: Própria autora, 2021.

Nesta etapa, será associado todo desenvolvimento da etapa anterior com a distribuição Geométrica.

O professor deve perguntar para a sala se alguém já ouviu falar de distribuição Geométrica. Provavelmente a resposta será não ou então algum aluno tentará deduzir o que possa ser. Mas, basicamente o professor responderá que o processo realizado para os cálculos das probabilidades anteriores é uma aplicação de distribuição Geométrica. Assim, a partir disto, trazer os conceitos de distribuição.

a) Variável aleatória (v.a): associa um número x a cada resultado do espaço amostral. Usualmente o seu conjunto é representado por uma letra maiúscula, como X .

Aqui o professor pode exemplificar que, no caso do jogo têm-se que a variável aleatória X é número de tentativas até atingir o sucesso, ou seja,

$$X = 1, 2, 3, \dots,$$

portanto, trata-se de uma variável discreta, já que só é possível jogar uma quantidade inteira de jogos.

Além disso, o professor pode chamar a atenção sobre o espaço amostral que pode ser finito ou infinito. No caso da probabilidade de ganhar um jogo da Mega-Sena esse espaço amostral em questão é finito, já que tem uma quantidade determinada de jogos possíveis que podem ser feitos. Porém, quando são analisadas as probabilidades do número de tentativas até ganhar, já se envolve um espaço amostral infinito, pois você só para quando ganhar, no entanto, pode passar a vida inteira jogando e nunca ganhar na Mega-Sena.

b) Função de probabilidade: é a função que informa a probabilidade para cada valor de X . A função de probabilidade da distribuição Geométrica é da forma

$$P(x) = (1 - p)^{x-1}p, \quad x = 1, 2, 3\dots$$

Observar que é a fórmula que deduziu-se em (4.1), e o que irá variar de acordo com cada situação são as probabilidades de sucesso (consequentemente de fracasso), por isso p é chamado de parâmetro e ele sempre será um valor entre 0 e 1. Além disso, a variável que está sendo analisada na situação também muda.

c) Valor esperado: informa qual é número de tentativas que, em média, deve-se realizar para conseguir maiores chances de se atingir o sucesso. Na distribuição Geométrica, seu cálculo é realizado usando a fórmula:

$$\mu = E(x) = \frac{1}{p}.$$

Para melhor compreensão desse conceito, o professor pode explicar utilizando o valor esperado do jogo da Mega-Sena. Considerando a probabilidade de sucesso sem aproximação, têm-se que $p = \frac{1}{50.063.860}$ obtendo,

$$\mu = E(x) = \frac{1}{p} = \frac{1}{\frac{1}{50.063.860}} = 50.063.860.$$

Aqui o professor vai explicar que se os alunos pudessem jogar um número infinito de vezes, essa seria a quantidade mais provável que eles teriam que jogar para ganhar, ou seja, todos os possíveis jogos.

É válido ressaltar para os alunos que nem sempre é assim que acontece, sendo este o motivo de ter pessoas que ganham o sorteio em poucas tentativas. Do mesmo modo que não é garantia obter o sucesso ao jogar esse número de vezes, pois trata-se de um experimento aleatório não determinístico.

Como curiosidade, pode ser discutido também quantas gerações de uma família seriam necessárias para manter a tradição de jogar até obter o sucesso, e dado que cada aposta custa em torno de R\$4,50, seria preciso gastar R\$225.287.370,00, aproximadamente.

5ª Etapa: Conclusão.

É importante que o professor faça com os alunos reflexões sobre todo o processo realizado, tanto em relação a interpretação dos resultados referente ao jogo, quanto à distribuição Geométrica.

No que diz respeito ao tema abordado, o professor pode discutir que ao se falar de jogos de loteria, principalmente a Mega-Sena, em geral as pessoas tem ideia de que é um

jogo difícil de se ganhar, mas nesta atividade foi mostrado, através dos números, o quanto é difícil.

Sobre a distribuição, comentar que aqui foi abordado apenas um tipo de várias que existem e que de modo geral são importantes em pesquisas, pois auxiliam, de alguma forma, a fazer a leitura da nossa sociedade.

Por fim, pensar em outras aplicações da distribuição Geométrica e calcular probabilidades através da função, uma vez que agora, espera-se que os alunos tenham compreendido o “porquê” da função de probabilidade da distribuição Geométrica.

Portanto, com esse intuito propõe-se dois exercícios que podem ser aplicados em aula. Esses exercícios são adaptações de alguns apresentados na aula de distribuição Geométrica de Liberal (UFPB).

Exercício 1: Em determinada localidade, a probabilidade de ocorrência de chuva forte em algum dia nos meses de dezembro e janeiro é igual a 0,1. Admitindo independência de um dia para o outro, responda:

a) Qual é a variável aleatória?

Resposta: $X =$ o número de dias que ocorre a primeira chuva, logo $X = 1, 2, 3, \dots, 62$

b) Qual a probabilidade da ocorrência da primeira chuva no dia 4 de dezembro?

Resposta: Pela função de probabilidade, $P(4) = (1 - 0,1)^3 \cdot 0,1 = 0,0729$.

c) E qual é o valor esperado?

Resposta: Para o valor esperado, têm-se $E(X) = \frac{1}{0,1} = 10$ dias.

Exercício 2: A probabilidade de que uma droga apresente reação positiva é 0,4. Considerando essa situação, responda:

a) Qual é a variável aleatória?

Resposta: $X =$ o número de reações até dar reação positiva, portanto $X = 1, 2, 3, \dots$

b) Qual a probabilidade de que ocorra 5 reações negativas antes da primeira positiva?

Resposta: Pela função de probabilidade, $P(6) = (1 - 0,4)^5 \cdot 0,4 = 0,0311$.

c) Qual é o valor esperado?

Resposta: $E(X) = \frac{1}{0,4} = 2,5$ ocorrências.

4.3 Atividade 2: A distribuição Binomial no ENEM

Objetivos: Aplicação de conceitos de probabilidade em situações reais por meio da distribuição Binomial para o cálculo da probabilidade de acertar x questões em uma prova do ENEM, desenvolvendo habilidades da BNCC e estendendo o conhecimento com os conceitos de distribuição de probabilidade, especificamente da Binomial.

Habilidades desenvolvidas conforme texto da BNCC:

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

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

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

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

Desenvolvimento: A atividade é construtiva, de modo que os conceitos e teorias envolvidos na distribuição binomial sejam apresentados e discutidos com os alunos no decorrer do processo. Cálculo de fatorial e combinação são conhecimentos prévios que os alunos devem apresentar.

1ª Etapa: Calculando probabilidades de acertar questões.

Partindo do princípio que uma pessoa realize a prova do ENEM sem estudar, apenas “chutando” as respostas das questões. O professor pode iniciar um diálogo com os alunos sobre qual a probabilidade de acertar uma questão do ENEM e qual seria aproximadamente o número de questões que um candidato acertaria chutando as alternativas.

É interessante coletar as respostas informadas pelos estudantes e citar que eles mesmos irão fazer os cálculos para verificar se alguma das respostas estava correta.

Iniciar os cálculos das probabilidades já considerando todas as 90 questões pode gerar dúvidas nos alunos. Deste modo, para facilitar a compreensão dos cálculos, pode-se simular o que aconteceria com o “chute” de uma prova de 5 questões. Para tornar isso mais dinâmico, o professor pode distribuir uma pequena folha de respostas para cada aluno (conforme Figura 25) onde irão marcar uma das cinco alternativas (A,B,C,D,ou E) para cada uma das questões.

Figura 25 – Folha de resposta.

Folha de resposta	
Questão	Resposta
1	
2	
3	
4	
5	

Fonte: Própria autora, 2021.

Marcadas as respostas, cada aluno irá verificar quantas questões acertou de acordo com o gabarito. O gabarito da Figura 26 são as respostas das 5 primeiras questões de matemática do ENEM 2019.

Figura 26 – Folha de resposta.

GABARITO	
Questão	Resposta
1	E
2	B
3	D
4	A
5	D

Fonte: Própria autora, 2021.

Após verificar o total de acertos, o professor pode fazer uma tabela de frequências para as quantidades de acertos possíveis (0,1,2,3,4 ou 5 questões), visualizando melhor o cenário obtido. A partir da tabela podem ser feitas algumas interpretações, como por exemplo, o número de questões que tiveram as maiores e menores frequências.

Com essas interpretações o professor traz em pauta a seguinte reflexão: “os alunos que acertaram as mesmas quantidades de questões chutaram igualmente as perguntas? Deram

as mesmas respostas?” Não existe uma única possibilidade, então podem ser mostradas duas folhas de respostas com a mesma quantidade de acertos, porém com perguntas corretas e erradas diferentes.

Para a realização dos cálculos de probabilidades, o professor pode explicar primeiro para duas questões certas, respondendo a seguinte pergunta: “qual a probabilidade de um aluno acertar duas questões?” Atribuindo S de sucesso e F de fracasso para as respostas certas e erradas, respectivamente, a Figura 27 mostra um exemplo para este cenário.

Figura 27 – Acertando duas questões.

<u>S</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>
1ª	2ª	3ª	4ª	5ª

Fonte: Própria autora, 2021.

Considerando a possibilidade representada na Figura 27, o professor vai calcular a probabilidade disso acontecer. Observando que se tratam de eventos independentes, que a probabilidade de acertar uma questão $P(S) = 1/5$ e que a probabilidade de errar uma questão $P(F) = 4/5$, acertar as duas primeiras questões terá uma probabilidade de:

$$P(S) \cdot P(S) \cdot P(F) \cdot P(F) \cdot P(F) = \frac{1}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{4}{5} \cdot \frac{4}{5}. \quad (4.2)$$

Porém, a partir do momento que se tem interesse nas chances de ocorrer o acerto de duas questões, como visto pelas respostas dos próprios alunos, não existe essa única possibilidade, o aluno pode acertar a primeira e quarta pergunta ou a segunda e terceira. Com base nisso, o professor vai propor para a turma verificar de quantas maneiras um aluno pode acertar duas questões. A Figura 28 mostra todas as possibilidades de ocorrerem dois sucessos.

Figura 28 – Possibilidades para acertar duas questões.

<u>S</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>S</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>
<u>S</u>	<u>F</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>S</u>
<u>S</u>	<u>F</u>	<u>F</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>S</u>	<u>S</u>	<u>F</u>
<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>S</u>	<u>F</u>	<u>S</u>
<u>F</u>	<u>S</u>	<u>S</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>F</u>	<u>S</u>	<u>S</u>	<u>S</u>

Fonte: Própria autora, 2021.

Ao verificar todas as possibilidades, os alunos vão concluir que são dez maneiras, deste modo, a probabilidade de acertar duas questões será o resultado obtido em (4.2) multiplicado por 10:

$$P(2) = 10 \cdot \left(\frac{1}{5} \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{4}{5} \cdot \frac{4}{5} \right). \quad (4.3)$$

Alguns pontos são importantes para discutir com os alunos. O primeiro deles é que a quantidade de possibilidades verificadas pode ser calculada a partir da combinação, como foi desenvolvido na Atividade 1. Sendo assim, o professor pode recordar que trata-se de uma combinação, uma vez que a ideia é “encaixar” dois sucessos “S” em cinco posições diferentes. Logo basta fazer $C_5^2 = \binom{5}{2} = 10$. O segundo ponto a ser levantado é que no produto das probabilidades de sucesso e fracasso, o que realmente importa é a quantidade de cada um. A ordem não irá alterar o resultado porque já foi consirada ao realizar o produto por 10. Portanto o resultado obtido em (4.3) pode ser reescrito utilizando o conceito de potência:

$$P(2) = \binom{5}{2} \left(\frac{1}{5} \right)^2 \left(\frac{4}{5} \right)^3. \quad (4.4)$$

Acredita-se que neste momento os estudantes tenham embasamento para compreender os cálculos realizados, podendo considerar agora a prova do ENEM que possui 90 questões. O processo descrito a seguir desenvolve os cálculos das probabilidades de acertar de 0 a 3 questões dentre as 90, para que através disso, os alunos consigam observar os padrões nos cálculos e deduzir a função de probabilidade da distribuição binomial. A calculadora científica do celular pode ser utilizada como recurso.

- Acertar 0 questões: deve-se errar todas

$$P(0) = \left(\frac{4}{5} \right)^{90} = 0,0000000019.$$

-Acertar 1 questão: pode ser 1 dentre as 90 e errar 89

$$P(1) = \binom{90}{1} \left(\frac{1}{5} \right)^1 \left(\frac{4}{5} \right)^{89} = 0,0000000427.$$

-Acertar 2 questões: podem ser 2 dentre as 90 e errar 88

$$P(2) = \binom{90}{2} \left(\frac{1}{5} \right)^2 \left(\frac{4}{5} \right)^{88} = 0,0000004749.$$

-Acertar 3 questões: pode ser 3 dentre as 90 e errar 87

$$P(3) = \binom{90}{3} \left(\frac{1}{5} \right)^3 \left(\frac{4}{5} \right)^{87} = 0,0000034824.$$

2ª Etapa: Deduzindo a função de probabilidade.

Com os cálculos realizados na primeira etapa, o professor pode perguntar aos alunos se eles conseguem observar alguns padrões e, se através deles, é possível calcular a probabilidade de acertar qualquer quantidade de questões na prova do Enem.

Reconhecer os padrões é importante para deduzir e obter a distribuição Binomial. Logo, caso necessário, o professor pode ir comparando os cálculos, dando ênfase nos números de questões, nos expoentes, até obter a função:

$$P(x) = \binom{90}{x} \left(\frac{1}{5}\right)^x \left(\frac{4}{5}\right)^{90-x}. \quad (4.5)$$

Se a atividade anterior foi desenvolvida com a mesma turma, possivelmente os alunos vão perceber que se trata de uma distribuição de probabilidade. Cabe ao professor instruí-los a generalizar para probabilidade de sucesso p e concluir que a equação (4.6) é a função de probabilidade de uma distribuição Binomial.

$$P(x) = \binom{n}{x} (p)^x (1-p)^{n-x}, \quad x = 0, 1, 2, \dots, n. \quad (4.6)$$

É importante mostrar que os mesmos conceitos básicos discutidos na atividade da Mega-Sena também se aplicam nessa distribuição, porém a função apresenta características diferentes.

a) Variável aleatória (v.a): associa um número x a cada resultado do espaço amostral. No caso da prova do ENEM temos que a variável aleatória X e é o número de questões corretas “chutando” as respostas da prova do ENEM, logo

$$X = 0, 1, 2, 3, \dots, 90.$$

É uma variável aleatória discreta, já que só é possível acertar uma quantidade inteira de questões. Além disso, trata-se de um espaço amostral finito, uma vez que existe um limite para o número de questões que o aluno pode acertar, que são de 0 a 90.

b) Função de probabilidade: é a função que informa a probabilidade para cada valor de X . A função de probabilidade da distribuição binomial é a obtida em (4.6), onde n e p são chamados de parâmetros.

c) Valor esperado: na distribuição Binomial o valor esperado é dado por:

$$E(x) = n \cdot p.$$

Logo, considerando os parâmetros $n = 90$ e $p = 0,2$, têm-se que:

$$E(x) = 90 \cdot 0,2 = 18,$$

o que significa que, em média, os alunos irão acertar 18 questões ao chutar as respostas das 90 questões da prova.

Cabe lembrar que o valor esperado é o que tem mais chance de acontecer, e não necessariamente o que irá acontecer.

3ª Etapa: Construindo o gráfico da distribuição Binomial.

O uso da tecnologia como aliada do ensino é interessante para mostrar ao aluno como a matemática pode ser desenvolvida na prática, além de apresentar recursos dos quais muitos ainda nem conhecem, mesmo utilizando frequentemente computadores e celulares.

Até o momento, os cálculos de probabilidade foram realizados na calculadora, porém o seu uso é limitado já que estamos lidando com números de diversas casas decimais. O programa de planilhas Excel oferece muitos recursos se tratando de matemática, inclusive na área de probabilidade. Sendo assim, os alunos irão construir um gráfico dessa distribuição Binomial no Excel. Caso a escola não apresente esse recurso, uma alternativa é o professor executar essa etapa, junto com os alunos, utilizando um projetor ou levar o gráfico da Figura 36 já impresso para que seja possível a interpretação do mesmo.

A seguir apresenta-se os passos para calcular no Excel as probabilidades de uma distribuição Binomial.

a) Primeiramente deve-se construir uma tabela de probabilidades com os valores de $P(x)$. Para isso, em células distintas escreve-se os valores dos parâmetros p e n , seguido das colunas ‘Número de questões (x)’ e ‘Probabilidade de acertar ($P(x)$)’. Veja Figura 29.

Figura 29 – Passo a.

	A	B	C	D	E	F	G	H	I	J
1										
2		p=	0,2							
3		n=	90							
4										
5		Número de questões (x)	Probabilidade de acertar (P(x))							
6		0								
7		1								
8		2								
9		3								
10		4								
11		5								
12										
13										

Fonte: Própria autora, 2021.

b) Calcular as probabilidades na coluna $P(x)$ utilizando o comando da distribuição Binomial do Excel. Para isso, digitar na célula =DISTR.BINOM(B6;\$C\$3;\$C\$2;FALSO), em que:

B6: é a célula referente a $x = 0$;

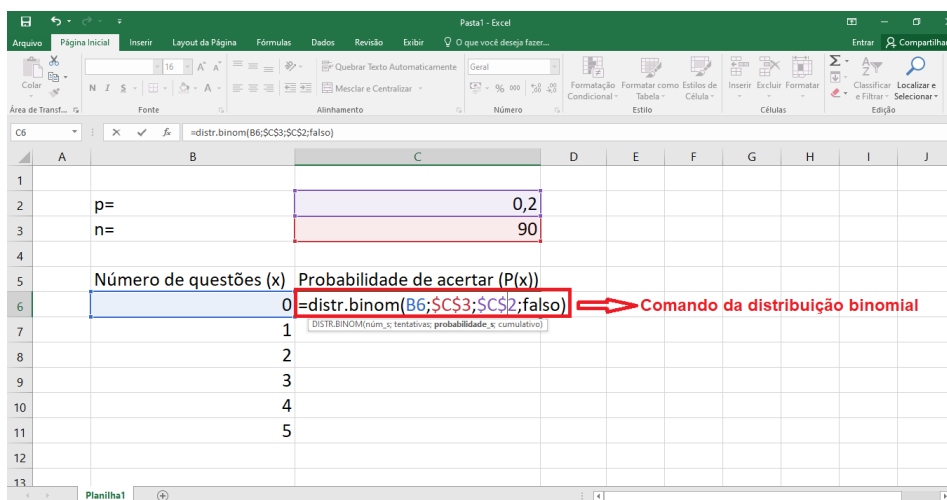
C3: célula do parâmetro n ;

C2: célula do parâmetro p ;

FALSO: indica que não é a probabilidade cumulativa.

As células C2 e C3 estão entre o cifrão para deixá-las fixas quando esse comando for copiado para outras células. Veja Figura 30.

Figura 30 – Passo b.

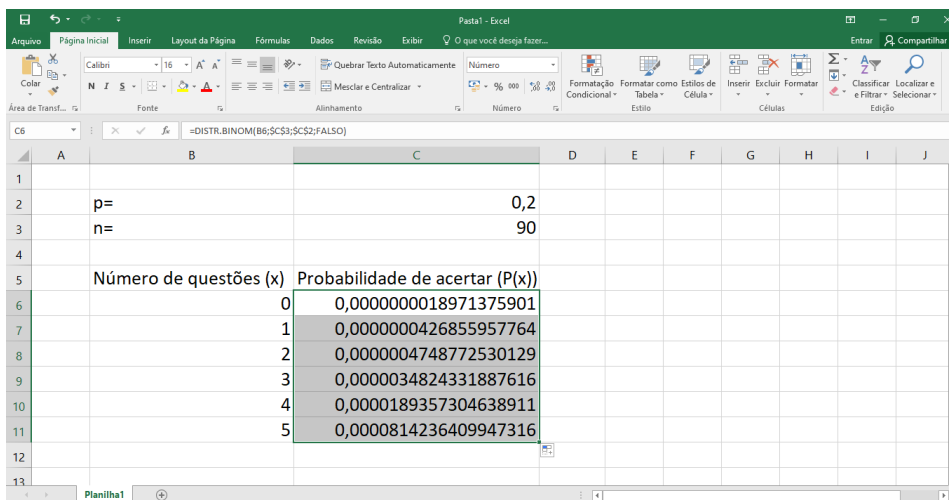


Fonte: Própria autora, 2021.

Se o resultado ficar em notação científica, faz-se necessário deixar a célula no formato de número e aumentar as casas decimais, como mostram as Figuras 22 e 23.

c) Com a probabilidade em decimal, copiar para as células referentes a $P(x)$, $x = 1, 2, 3, 4$.
Veja Figura 31.

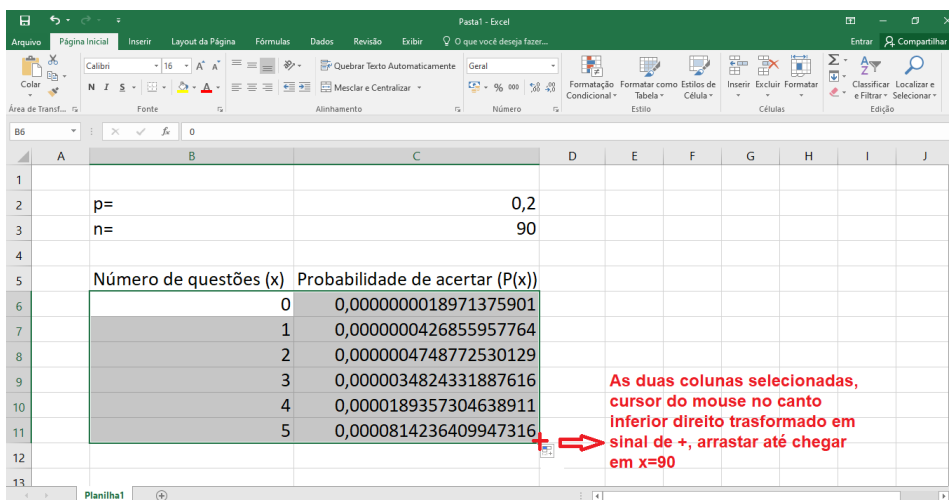
Figura 31 – Passo c.



Fonte: Própria autora, 2021.

Para em seguida, copiar ambas as colunas até chegar em $x = 90$. Veja Figura 32.

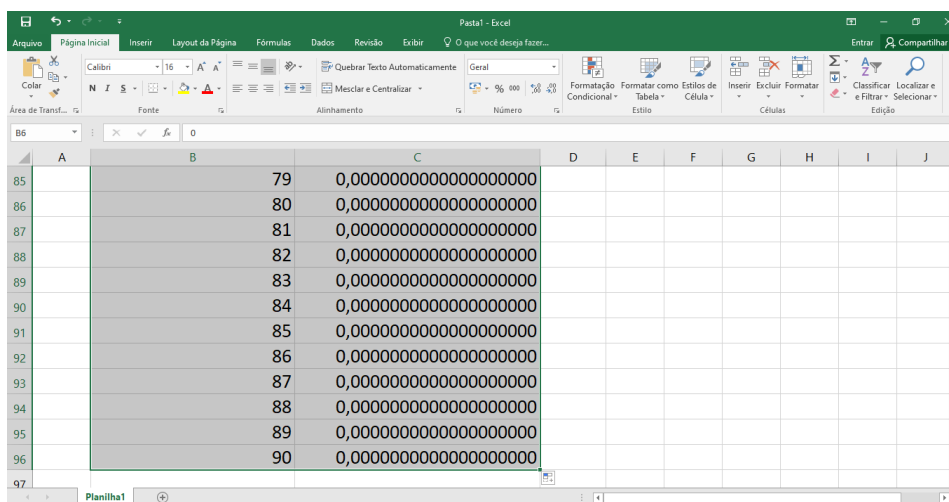
Figura 32 – Passo c: copiando as colunas.



Fonte: Própria autora, 2021.

Após selecionar, arrastar até chegar em 90, conforme mostra a Figura 33.

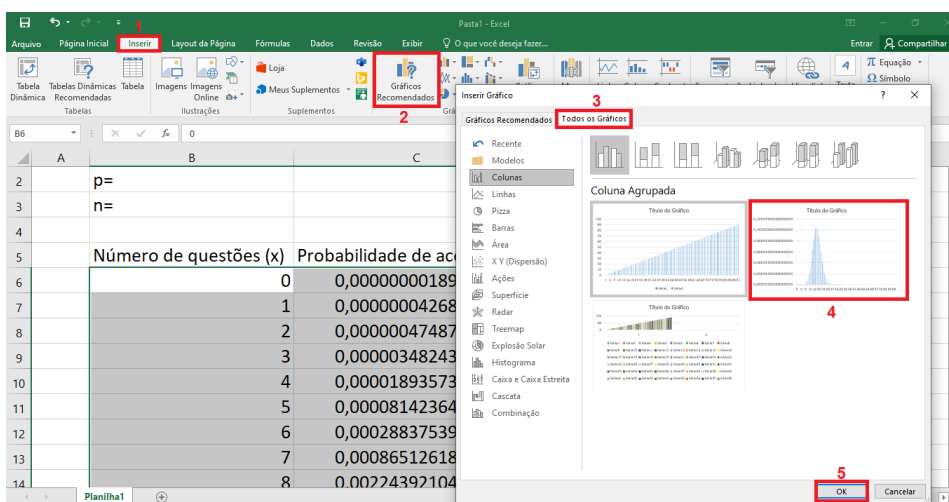
Figura 33 – Passo c: copiando as colunas.



Fonte: Própria autora, 2021.

d) Com a tabela pronta, seleciona-se todas as linhas e colunas. Em seguida pode-se construir o gráfico da distribuição, como mostra a Figura 34.

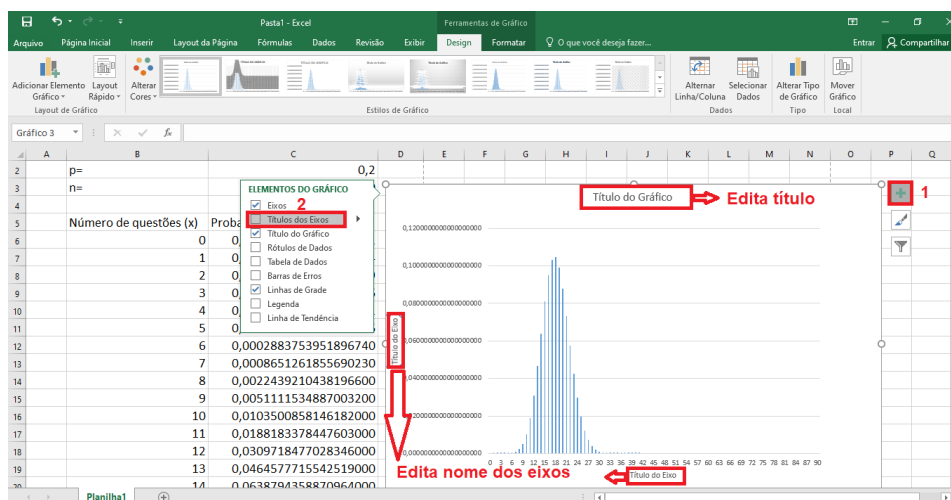
Figura 34 – Passo d: construindo o gráfico.



Fonte: Própria autora, 2021.

Aqui foi escolhido o gráfico de barras para representar melhor essa distribuição. Além disso, algumas melhorias podem ser realizadas no gráfico, como editar o título, inserir nome nos eixos e entre outros itens quando se trata de gráfico. Veja Figura 35.

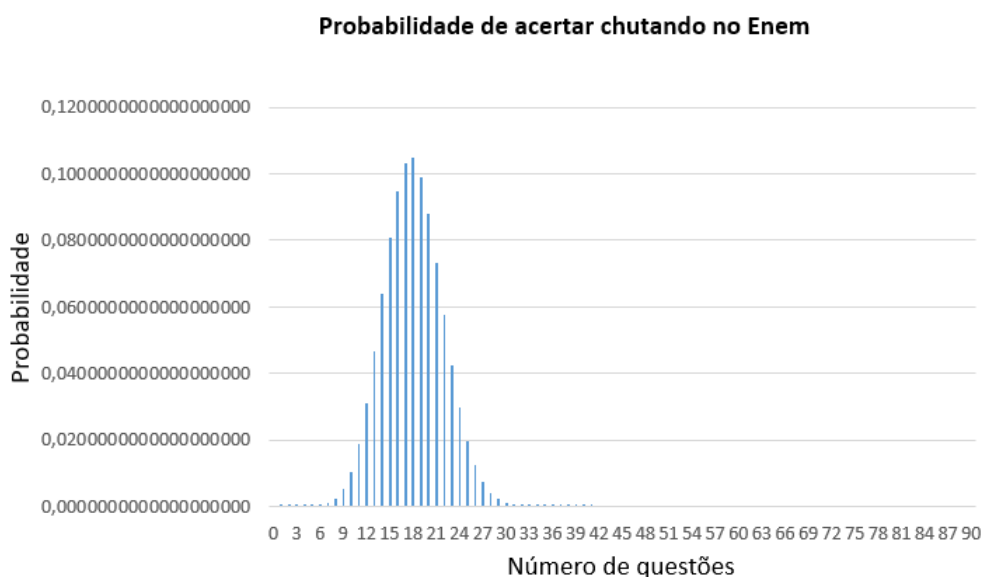
Figura 35 – Passo d: editando o gráfico.



Fonte: Própria autora, 2021.

Após as algumas dessas edições, finalmente obtêm-se o gráfico conforme a Figura 36.

Figura 36 – Gráfico da distribuição Binomial.



Fonte: Própria autora, 2021.

4ª Etapa: Conclusão.

Após todo esse processo de construção chega o momento de fazer reflexões e obter conclusões sobre o tema abordado. Claro que no decorrer deste caminho os próprios alunos já terão observado e até exposto algumas dessas conclusões, porém, é importante garantir que o objetivo desta atividade seja alcançado. Para tanto, o professor deve certificar-se da

compreensão dos alunos sobre os seguintes tópicos:

- Quantas questões é mais provável o estudante acertar ao realizar a prova sem estudar? É possível observar que as maiores probabilidades ocorrem entre 15 e 20 questões, e que a maior é 18 questões $P(18) = 10,46\%$, sendo este o número de questões obtidas no valor esperado, isto é, na distribuição Binomial é possível visualizar esse resultado no gráfico, como sendo o que apresenta a maior probabilidade.

-Retomar a pergunta realizada no início da atividade e as que surgiram no decorrer, respondendo agora usando os resultados obtidos ao aplicar a distribuição Binomial.

- Discutir que não é porque as probabilidades são baixas que sempre vai ser assim (como mostram a tabela e o gráfico). Pode ser que alguém acerte mais ou menos questões, mas que na maioria dos casos é esse cenário que vai ocorrer.

- Reforçar a habilidade EM13MAT106, que diz sobre as escolhas realizadas pelos alunos que levam-se riscos probabilísticos, possibilitando associar a uma conversa de incentivo aos estudos e que as chances da sorte não estar ao nosso lado são maiores. Além disso, outro ponto a se destacar quanto a nota do ENEM é que ela não é definida só pela quantidade de acertos, mas também pela coerência entre as respostas. Por exemplo, dois candidatos que tenham acertado o mesmo número de questões podem ter notas diferentes, aquele que apresentou mais respostas certas de nível difícil do que nível fácil, terá uma nota menor daquele que acertou na maior parte as de níveis fácil a médio. Isso é uma metodologia de avaliação chamada de TRI (teoria de resposta ao item). A dissertação de Soares (2020) apresenta melhores informações sobre isso.

- Questionar aos alunos outras situações que envolvam uma distribuição Binomial e propor a resolução de algumas questões. Como sugestão, podem ser aplicados os dois exercícios a seguir. O exercício 1 é uma adaptação de um apresentado no trabalho de Santos (2015) e o exercício 2 é uma questão do ENEM 2017. A resolução dessas questões podem ser realizadas utilizando a calculadora.

Exercício 1: O Traço Falciforme nada mais é que uma alteração genética herdada de um dos pais que não é suficiente para se manifestar como doença, e é bastante comum. Um casal descobre que cada um deles é portador do traço falsiforme. A probabilidade de que qualquer criança nascida deles tenha a anemia falciforme AF é 0,25. Considerando que o casal tenha 4 filhos, responda as perguntas:

a) Descreva a variável aleatória para esse problema. Porque trata-se de uma distribuição

binomial?

Resposta: $X =$ o número de filhos que tem AF em 4 nascimentos ($X = 0, 1, 2, 3, 4$). É uma distribuição binomial pois em cada nascimento o filho terá ou não AF e são eventos independentes. Além disso, a probabilidade de ter AF é constante para cada nascimento. Logo trata-se de ensaios de Bernoulli, dentro de um espaço amostral finito.

b) Qual é a probabilidade de 4 filhos do casal ter AF?

Resposta: deve-se calcular $P(4) = 0,00390625$

c) Qual a quantidade de filhos que é mais provável ter AF?

Resposta: É o valor esperado, $E(X) = 4 \cdot 0,25 = 1$ filho.

Exercício 2: (ENEM 2017) Numa avenida existem 10 semáforos. Por causa de uma pane no sistema, os semáforos ficaram sem controle durante uma hora, e fixaram suas luzes unicamente em verde ou vermelho. Os semáforos funcionam de forma independente; a probabilidade de acusar a cor verde é de $2/3$ e de acusar a cor vermelha é de $1/3$. Uma pessoa percorreu a pé toda essa avenida durante o período da pane, observando a cor da luz de cada um desses semáforos. Qual a probabilidade de que esta pessoa tenha observado exatamente um sinal na cor verde?

a) $\frac{10 \times 2}{3^{10}}$

b) $\frac{10 \times 2^9}{3^{10}}$

c) $\frac{2^{10}}{3^{100}}$

d) $\frac{2^{90}}{3^{100}}$

e) $\frac{2}{3^{100}}$

Resposta: Como deseja-se observar a cor verde, a variável aleatória é $X =$ o número de cor verde nos 10 semáforos, sendo assim, os parâmetros são $p = 2/3$ (chance de acusar verde) e $n = 10$. Logo têm-se a função de probabilidade:

$$P(x) = \binom{10}{x} \left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{10-x}$$

E, observar um sinal na cor verde, significa que $x = 1$. Substituindo têm-se:

$$P(x) = \binom{10}{1} \left(\frac{2}{3}\right)^1 \left(\frac{1}{3}\right)^{10-1} = 10 \cdot \left(\frac{2}{3}\right)^1 \left(\frac{1}{3}\right)^9 = \frac{10 \times 2}{3^{10}}$$

Logo a alternativa correta é a letra a.

4.4 Atividade 3: A distribuição Normal na estatura

Objetivos: Apresentar a distribuição Normal a partir do tratamento de uma amostra das alturas dos alunos da turma, desenvolvendo conceitos de tabela de frequências, histogramas, simetrias em histogramas e distribuições. Além disso, utilizar ferramentas computacionais, como o programa de planilhas Excel e o *software* Geogebra.

Habilidades desenvolvidas conforme texto da BNCC:

-(EM13MAT202): Planejar e executar pesquisa amostral usando dados coletados ou de diferentes fontes sobre questões relevantes atuais, incluindo ou não, apoio de recursos tecnológicos;

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

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

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

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

-(EM13MAT511): Reconhecer a existência de diferentes tipos de espaços amostrais, discretos ou não, de eventos equiprováveis ou não, e investigar as implicações no cálculo de probabilidades.

Desenvolvimento: É uma atividade construtiva, de modo que, os conceitos e teorias envolvidos sejam apresentados e discutidos com os alunos no decorrer do processo.

1ª Etapa: Coletando dados.

A conversa inicial que o professor pode ter com os alunos é que eles conhecerão, na prática, alguns dos processos realizados em pesquisas, uma vez que existe um procedimento a se cumprir para torná-las confiáveis.

O primeiro requisito é decidir o assunto a ser pesquisado, no caso desta pesquisa, o tema será a estatura dos alunos do Ensino Médio da escola.

Vale ressaltar aqui que, a escolha desse tema não é aleatória, pois nem toda amostra segue uma distribuição Normal, porém a altura de uma população possui essa propriedade, de acordo com Ross (2010). O professor pode chamar atenção para isso, mas não falar ainda da distribuição Normal e sim informar aos alunos que existe um critério para esse tema ser escolhido e que no final vão compreender qual é.

Seguido disso, sugere-se ao professor dizer que, com o tema definido, os pesquisadores analisam qual é a população e a amostra, explicando os conceitos:

-População: é o conjunto total que se enquadra nas características do tema pesquisado. Na pesquisa da turma, a população será os alunos do Ensino Médio matriculados na escola.

-Amostra: é um subconjunto da população no qual são coletados os dados. Na pesquisa deles, a ideia é de que todos os alunos da turma sejam a amostra, porém farão parte apenas aqueles presentes no dia da coleta.

É importante deixar claro para os alunos que em uma pesquisa a intenção é, a partir da análise da amostra, obter conclusões sobre a população. O professor pode mencionar que mais a frente dessa atividade os dados dessa amostra de altura serão utilizados para saber como estão distribuídas as alturas dos alunos do Ensino Médio e, a partir disso, permitir que probabilidades envolvendo essas alturas possam ser calculadas. Para complementar essa fala, pode-se abordar também o conceito de censo, uma pesquisa cuja amostra é igual a população, como por exemplo, as pesquisas do IBGE (Instituto Brasileiro de Geografia e Estatística).

Neste momento, é hora coletar os dados. A altura de cada aluno precisa ser medida com uma fita métrica e, para que ocorra uma interação, a turma pode ser separada em grupos os quais ficam responsáveis por medir a altura de cada integrante. Lembrando que é necessário utilizar a mesma unidade de medida, metro ou centímetro.

Após cada grupo realizar a coleta eles devem informar as medidas obtidas enquanto o professor anota na lousa, assim todos os alunos terão acesso a amostra completa. Para facilitar as etapas posteriores já organizar os dados em Rol (em ordem crescente ou decrescente), abordando também esse conceito com os alunos.

Figura 37 – Exemplo de uma amostra de alturas em centímetros.

147,5	150,9	152,4	155,0	155,4
155,8	157,8	160,2	160,6	161,9
162,3	162,8	163,5	165,0	165,0
165,2	165,8	165,8	166,5	166,9
167,2	168,2	168,3	168,5	168,6
169,7	169,9	170,3	170,8	171,9
175,0	175,5	176,8	179,8	183,8

Fonte: Amostra gerada pelo Action Stat.

A Figura 37 representa uma amostra normal gerada pelo *software* Action Stat (disponibilizado pela plataforma Portal Action) com média de 165,4cm e desvio padrão de 8cm, supondo alturas de 35 alunos. A média foi definida a partir da média das alturas medianas de adolescentes de 15 a 17 anos, do sexo feminino e masculino, da “Tabela 2645 - Estimativas populacionais das medianas de altura e peso de crianças, adolescentes e adultos, por sexo, situação do domicílio e idade - Brasil e Grandes Regiões”⁷ do Banco de Tabelas Estatísticas SIDRA, do IBGE. Essa amostra será utilizada para exemplificar a execução das etapas desta atividade, que no decorrer do texto será referida como ‘amostra exemplo’.

É válido ressaltar que quando se referir a resultados obtidos da amostra exemplo, o professor deve considerar os correspondentes resultados da amostra da turma.

2ª Etapa: Medidas de posição e dispersão.

Nesta etapa o professor irá explorar os conceitos de medidas de posição e dispersão. Ressalta-se que o intuito é abordar mais o conceito e não como realizar as contas, um dos motivos dos quais propõe-se o uso do Excel para isso.

Talvez os alunos saibam o que são as medidas, uma vez que a habilidade EF08MA25 aborda esses assunto no 8º ano. Mas, é interessante o professor comentar com a turma que esses valores são significativos para realizar inferências de uma pesquisa e que basicamente, as medidas de posição; média, mediana e moda; são usadas para compreender a posição em que os dados estão, como o próprio nome diz. Enquanto que as medidas de dispersão; amplitude e desvio padrão; informam o quão distante os dados estão uns dos outros.

Após essa pequena introdução, abordar os conceitos individuais:

-Média (\bar{x}): pode ser vista como “o equilíbrio dos dados”. O professor pode explicar, por exemplo, que se juntar todas as alturas da amostra (somatório) e distribuir igualmente

⁷ Disponível em: <<https://sidra.ibge.gov.br/tabela/2645>>. Acesso em: 10 out. 2020.

esse total para cada aluno (a divisão pela quantidade de elementos somados), a altura que cada um “receberia” é a média da amostra.

Uma abordagem com exemplo tornará mais fácil a compreensão desses conceitos. A média de notas também pode ser citada, já que é algo bastante presente na vida escolar dos alunos.

-Moda (M_o): é o elemento que mais se repetiu. Essa “repetição” é o que chama-se de frequência.

-Mediana (M_e): com os dados ordenados em ordem crescente, é número que está na posição central, dividindo a amostra ou a população em duas partes iguais.

Como exemplo mais prático, o professor pode mostrar que a mediana da série 1, 3, 3, 4, 5, 6 e 6 é 4.

-Amplitude total: é a diferença entre o maior e o menor elemento da amostra.

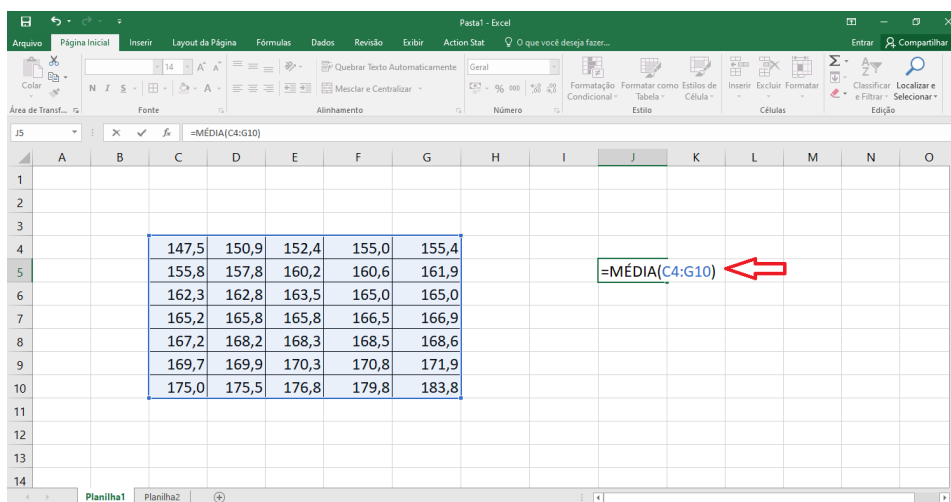
-Desvio Padrão (s): mede o “quanto distante” os valores da amostra estão da média. Quanto maior o desvio padrão, mais dispersos os dados estão.

Definidos os conceitos, é o momento de calcular as medidas de posição e dispersão da amostra no Excel (caso não seja possível utilizá-lo, os cálculos podem ser realizados manualmente). É interessante utilizar esse programa pelo fato de ser uma planilha mais acessível, uma vez que está presente na maioria dos computadores e por ser comum utilizá-la na área administrativa, proporcionando ao aluno um conhecimento que talvez utilize no futuro como profissional.

Em primeiro lugar, é necessário inserir a amostra na planilha, digitando uma altura por célula. Após realizar isso, seguir os passos:

a) Para calcular a média (\bar{x}), em uma célula qualquer digitar o comando =MÉDIA(), onde dentro dos parênteses deve conter o intervalo das células dos dados da amostra. Após isso, digitar *enter* para efetuar o resultado. A Figura 38 representa esse processo.

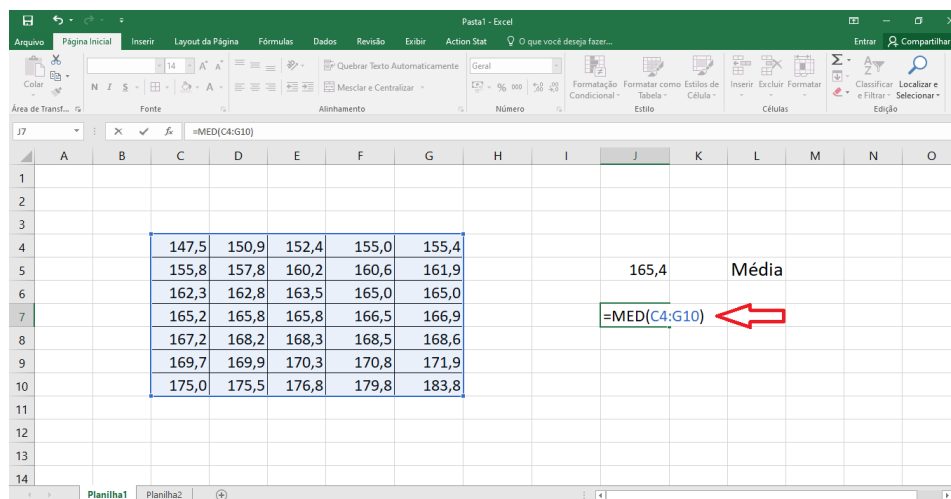
Figura 38 – Calculando média no Excel.



Fonte: Própria autora, 2021.

b) Para calcular a mediana (Me): usa-se o comando `=MED()`. Assim como a média, inserir dentro dos parênteses o intervalo das células da amostra e digitar *enter*. Observe a Figura 39.

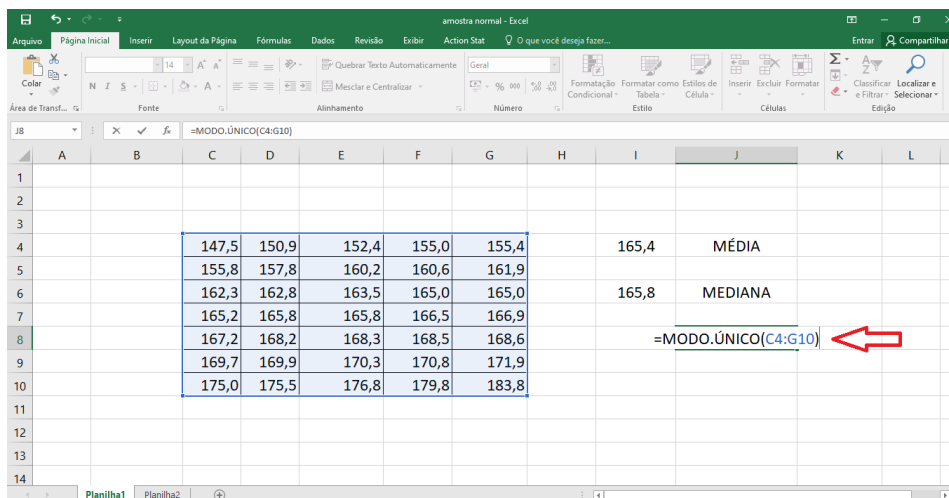
Figura 39 – Calculando mediana no Excel.



Fonte: Própria autora, 2021.

c) Para calcular a moda (Mo): usa-se o comando `=MODO.ÚNICO()`, procedendo como nos itens anteriores. Veja Figura 40.

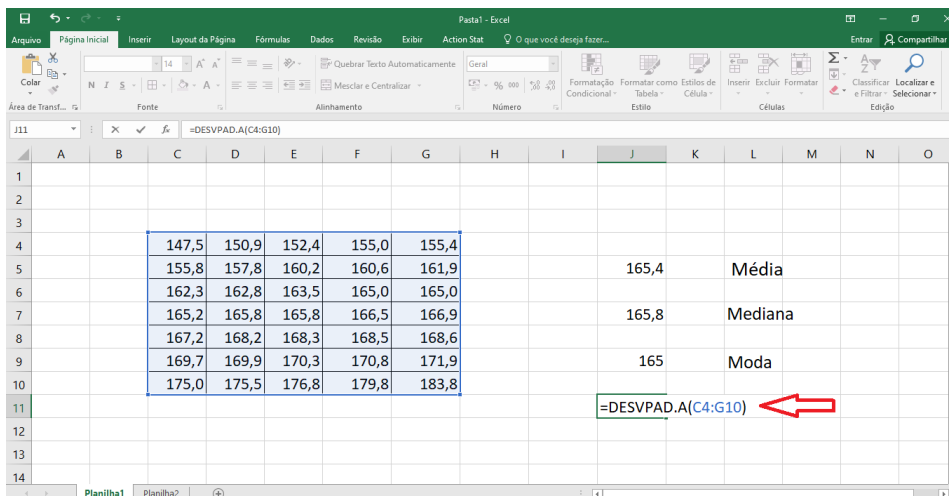
Figura 40 – Calculando moda no Excel



Fonte: Própria autora, 2021.

d) Para o desvio padrão utiliza-se o comando =DESPAD.A(). Conforme a Figura 41.

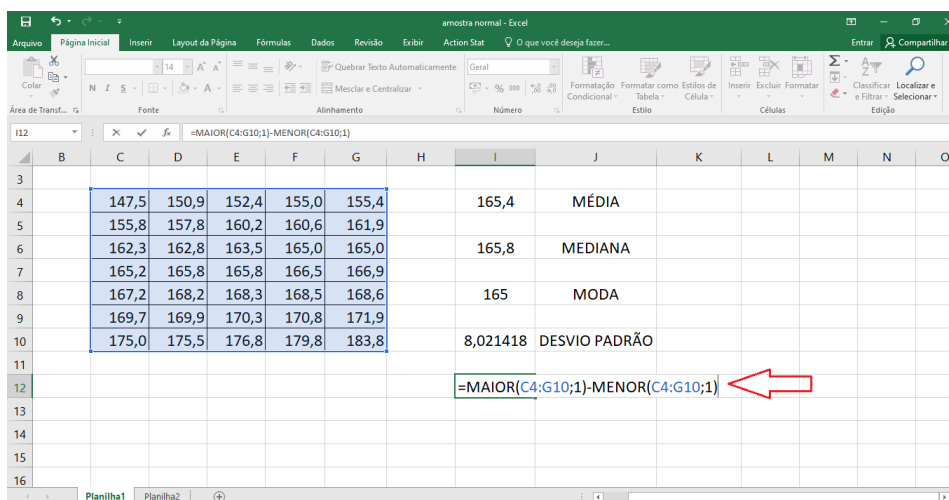
Figura 41 – Calculando o desvio padrão no Excel.



Fonte: Própria autora, 2021.

e) A amplitude é facilmente calculada subtraindo a altura menor da maior, que no Excel pode ser feita usando o comando =MAIOR(C4:G10;1)-MENOR(C4:G10;1), em que C4:G10 é o intervalo da amostra, e o número 1 indica que deve ser o primeiro maior/menor. Veja a Figura 42.

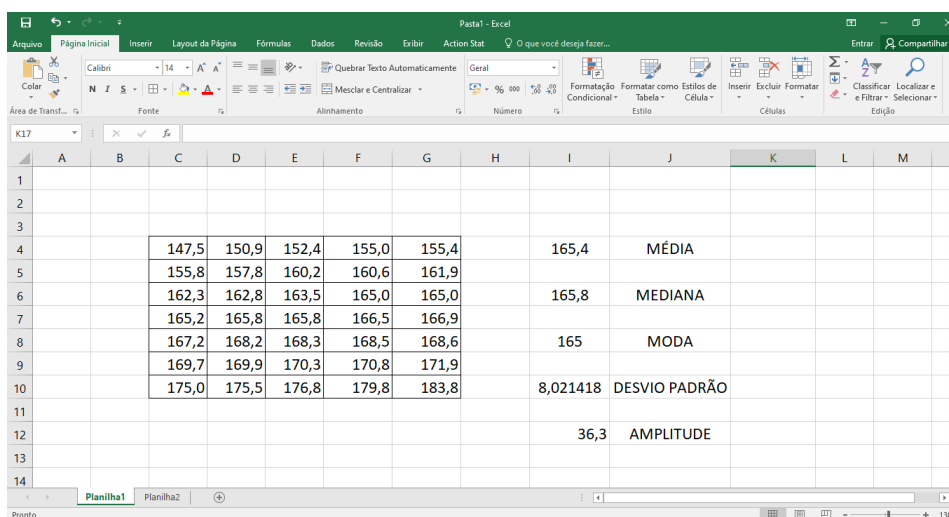
Figura 42 – Calculando a amplitude no Excel.



Fonte: Própria autora, 2021.

Os resultados da média, mediana, moda, desvio padrão e amplitude da amostra exemplo foram, respectivamente, 165,4cm; 165,8cm; 165cm; 8cm; 36,3cm como pode ser visto na imagem 43.

Figura 43 – Medidas de posição e dispersão da amostra exemplo.



Fonte: Própria autora, 2021.

Ao término desta etapa o professor irá retomar os conceitos, principalmente da média e desvio padrão, uma vez que agora sabe-se essas medidas. Pode ser dito novamente, por exemplo, que se fosse distribuir igualmente as alturas entre os alunos cada um deveria ter aproximadamente a altura resultante da média. Porém, percebe-se que muitos alunos tem alturas distantes dessa média, por isso a importância do desvio padrão, uma vez que ele vai indicar o quanto disperso esse conjunto de dados está. Por fim, deixar claro que

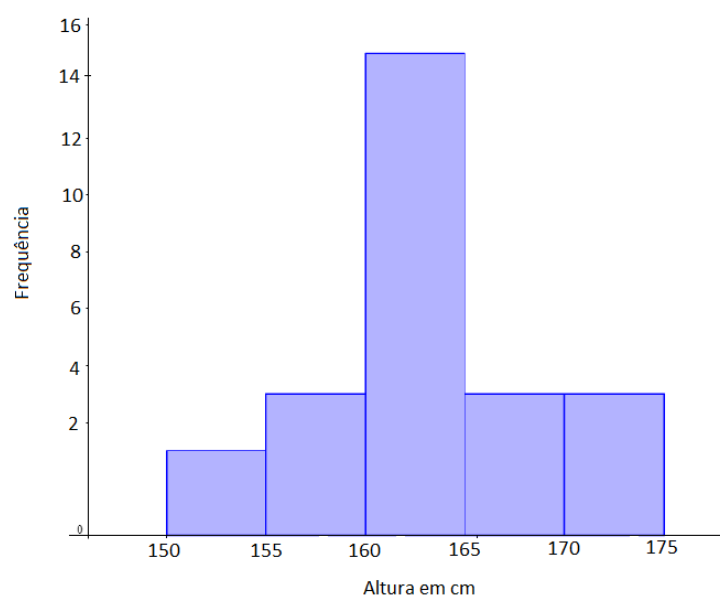
quanto menor o desvio padrão, mais homogênea é a amostra.

3ª Etapa: Construindo o histograma.

Os resultados de uma pesquisa usualmente são mostrados em gráficos. Pensando nisso, perguntar para a turma qual tipo de gráfico seria ideal para representar a amostra das alturas. Possivelmente algum aluno comentará sobre o gráfico de colunas (caso não, indicar essa ideia) e considerando essa opção perguntar quantas colunas seriam necessárias ter. Espera-se que os alunos percebam que seriam diversas devido aos números decimais, proporcionando o momento ideal para o professor comentar que as alturas são variáveis contínuas, pois abrange a reta numérica real (por serem números reais) e que, para essa situação, o gráfico ideal é o histograma.

O histograma é um gráfico utilizado para representar a distribuição de frequências de variáveis contínuas. A Figura 44 é um exemplo de um histograma, que inclusive deve ser mostrada para a sala.

Figura 44 – Exemplo de um histograma.



Fonte: Própria autora, 2021.

Os alunos podem ter a noção do histograma, uma vez que sua interpretação é abordada inclusive no 7º ano na disciplina de Geografia (habilidade EF07GE10), porém cabe orientar e reforçar com alunos a observarem que neste histograma as colunas representam as classes, indicando os intervalos. Essas colunas são justapostas para indicar exatamente a continuidade da variável analisada. A altura de cada coluna, por sua vez, informa a respectiva frequência (ocorrência de elementos que pertencem àquele intervalo).

Conhecendo agora o que é um histograma, propor a construção de um para a amostra das alturas.

O primeiro passo a se realizar é definir a quantidade (k) de colunas, que na verdade são chamadas de classes. É empírico pensar que quanto maior a amostra mais classes terão, ou seja, depende do tamanho da amostra. E embora não exista uma regra, usualmente determina-se k a partir da raiz quadrada de n , onde n é a quantidade de elementos da amostra.

O segundo passo é calcular a amplitude da classe, que se refere a “largura” de cada intervalo, cujo valor é obtido da divisão da amplitude total da amostra pelo número de classes k .

Antes de construir o histograma, sugere-se construir a tabela de frequência, manualmente mesmo, pois assim os alunos terão uma visão melhor sobre como ficará o histograma.

Para fazer a tabela os passos são semelhantes ao do histograma, visto que apresentam as mesmas informações, mas de maneiras diferentes. Os passos podem ser explicados na prática, aplicando-se para a amostra das alturas. Neste trabalho, isso será realizado com a amostra exemplo.

Primeiramente, calcula-se a quantidade de classes. Como $n = 35$, então $k = \sqrt{35} = 5,9$, que pela aproximação, poderiam ser consideradas 6 classes. Porém, como visa um histograma mais simétrico possível (discussão feita na próxima etapa), uma quantidade ímpar de classes, a priori, facilita a análise da simetria. Portanto, a tabela da amostra exemplo terá 5 classes.

Agora será calculada a amplitude das classes para definir os intervalos. A amplitude total da amostra exemplo é 36,3, dividindo pelo número de classes têm-se $36,3 \div 5 = 7,26$. Para facilitar, será tomado a aproximação 7,5. Veja que se considerar a amplitude aproximada para 7 seria necessário inserir mais um intervalo de classe para abranger todas as alturas da amostra, logo é preferível fazer a aproximação para mais.

O próximo passo é definir as classes. Na amostra exemplo, a altura mínima é 147,5cm, que pode ser usada como valor inicial, mas, por opção, será usado 147cm. Organizando de 7,5cm em 7,5cm, têm-se as classes:

Tabela 3 – Classes da tabela de frequência.

Altura (cm)
147 – 154,5
154,5 – 162
162 – 169,5
169,5 – 177
177 – 184,5

Fonte: Própria autora, 2021.

Nesse momento deve-se lembrar de explicar para a sala sobre a notação dos intervalos, que escrever $a \text{ – } b$, significa que estão incluídos todos os elementos maiores ou iguais a a e menores que b . Isso é importante considerar na hora de fazer a contagem das alturas de

cada classe, ou seja, identificar as frequências, sendo este o último passo para construir a tabela. Analisando a quantidade de alturas pertencentes a cada intervalo, obtêm-se a Tabela 4 da amostra exemplo.

Tabela 4 – Tabela de frequências da amostra exemplo.

Altura (cm)	Frequência
147 † 154,5	3
154,5 † 162	7
162 † 169,5	15
169,5 † 177	8
177 † 184,5	2

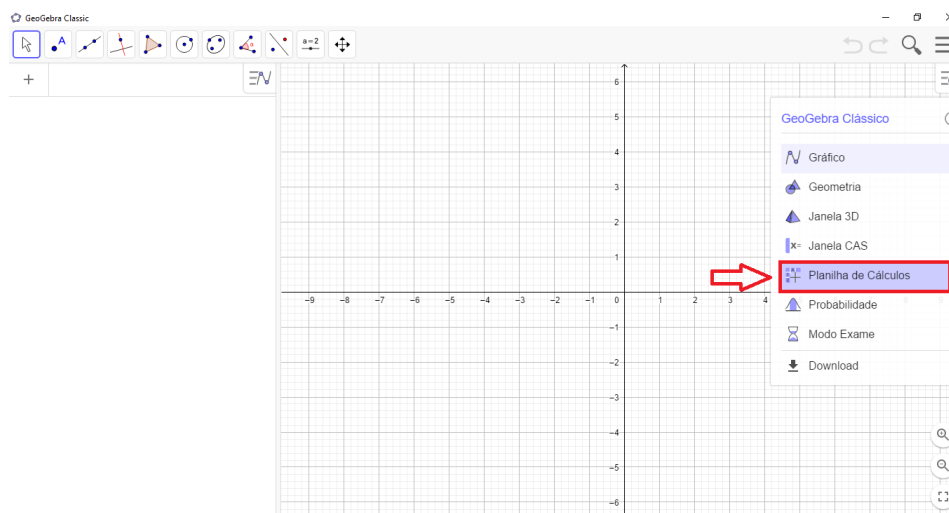
Fonte: Própria autora, 2021.

Da mesma maneira os alunos devem proceder com a amostra da turma.

Para a construção do histograma, os alunos podem fazer manualmente no caderno ou utilizar o *software* Geogebra, seguindo os passos descritos a seguir.

a) Certifique-se que esteja aberta a janela ‘Planilha de Cálculos’ conforme Figura 45 .

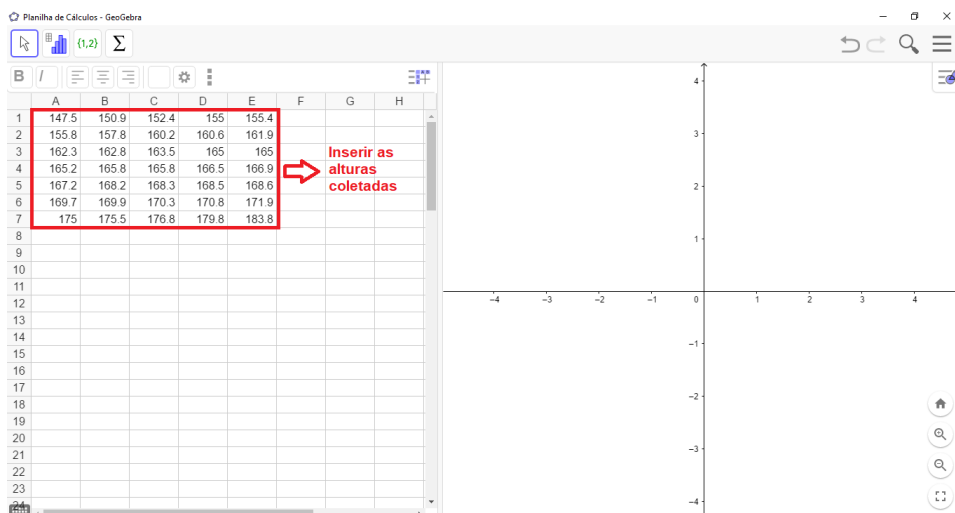
Figura 45 – Dados coletados na planilha do Geogebra.



Fonte: Própria autora, 2021.

b) Inserir as alturas coletadas na planilha do Geogebra. Veja Figura 46.

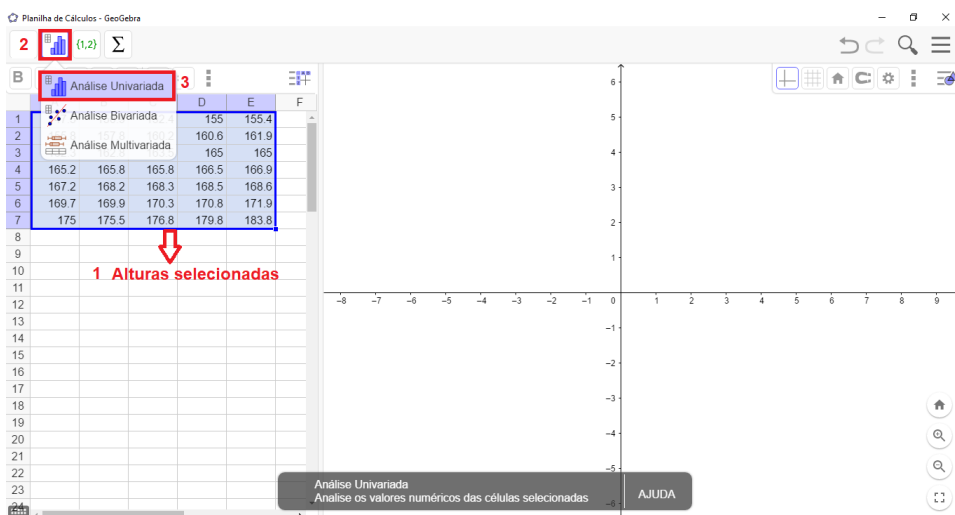
Figura 46 – Dados coletados na planilha do Geogebra.



Fonte: Própria autora, 2021.

c) Criar o histograma selecionando a opção ‘Análise Univariada’ após a seleção dos dados. Veja Figura 47.

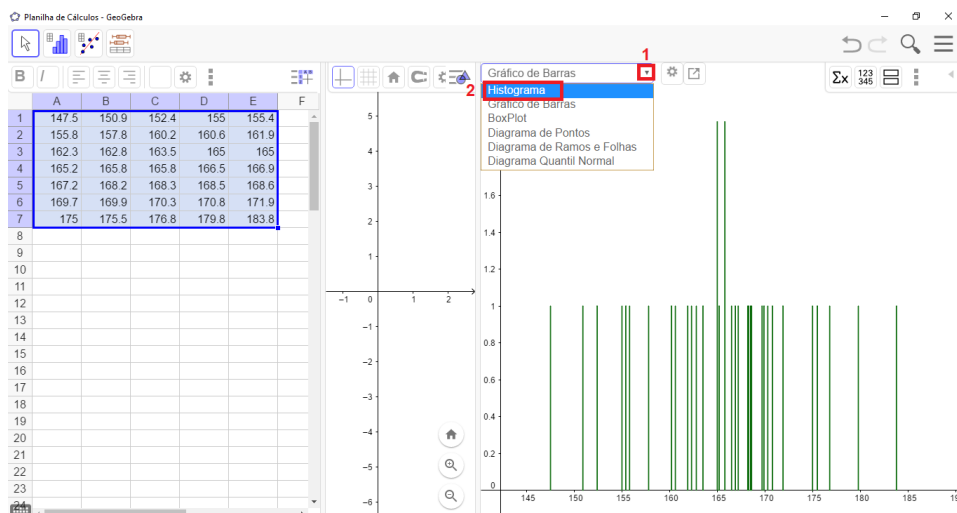
Figura 47 – Criando o histograma.



Fonte: Própria autora, 2021.

Caso seja plotado outro tipo de gráfico, basta selecionar ‘Histograma’ na caixa de opções, como mostra a Figura 48.

Figura 48 – Selecionando a opção Histograma.

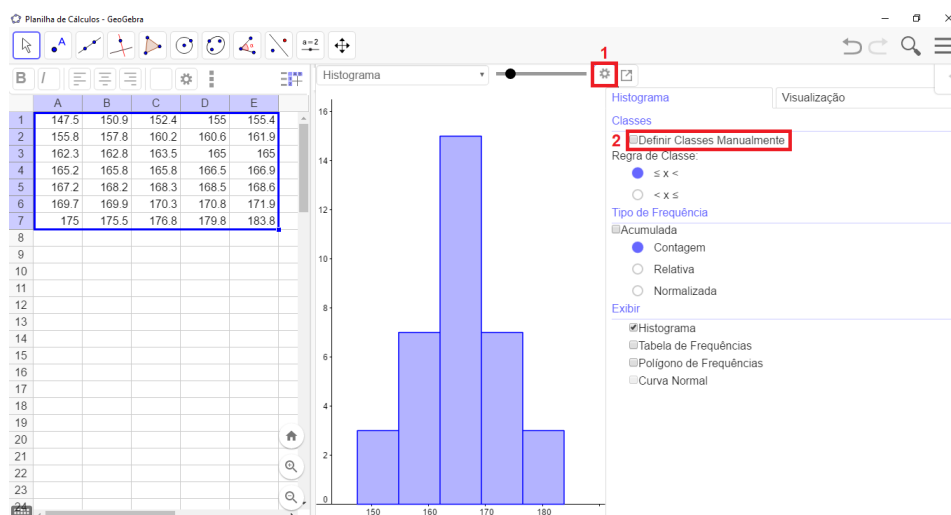


Fonte: Própria autora, 2021.

Pode ser que o histograma não esteja com os mesmos intervalos da tabela de frequência. Porém, é possível modificar as classes de acordo com as definidas pelos alunos, conforme explica o passo c.

d) Na caixa de opções, clicar em ‘Definir classes manualmente’. Veja Figura 49.

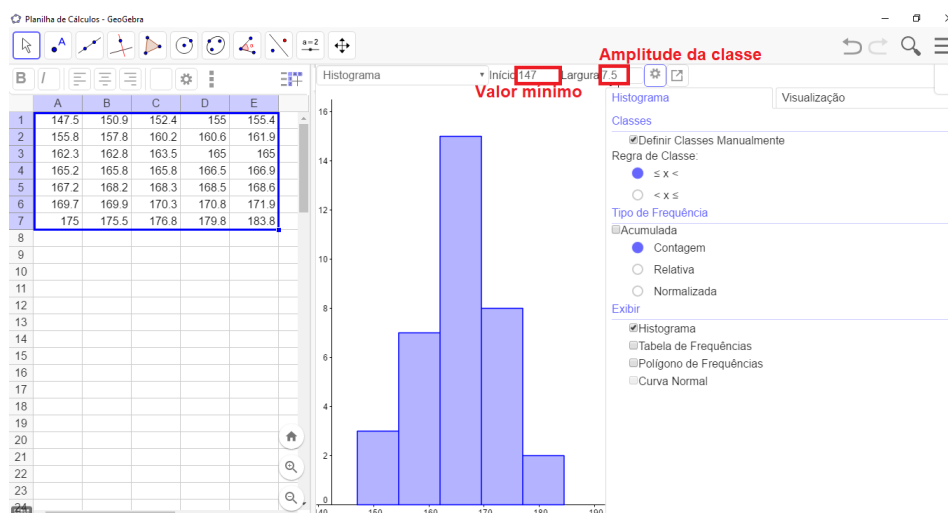
Figura 49 – Definindo as classes manualmente, 2021.



Fonte: Própria autora, 2021.

Isso mostrará dois campos, o ‘Início’, que se refere ao valor mínimo da primeira classe, e o campo ‘Largura’, que indica a amplitude das classes. Digitando os valores estabelecidos anteriormente têm-se o histograma desejado. A Figura 50 mostra o histograma da amostra exemplo após a alteração das classes.

Figura 50 – Histograma da amostra exemplo.



Fonte: Própria autora, 2021.

4ª Etapa: Simetria e a distribuição em forma de sino.

Como dito anteriormente as simetrias/assimetrias em histogramas serão utilizadas para associar a distribuição Normal, o que torna esta uma das mais importantes etapas da atividade, pois é a ‘ponte’ que liga a amostra das estaturas com a distribuição Normal. Portanto, precisa ser construída de forma efetiva com os alunos.

Sugere-se ao professor questionar primeiro os alunos sobre o que é simetria e se eles conseguem pensar em algo que é simétrico. Após serem exploradas suas respostas, reforçar então que um elemento é simétrico quando dividido ao meio resulta em duas partes iguais. Uma imagem simétrica precisa ser mostrada na aula, como por exemplo a Figura 51, para reforçar a compreensão.

Figura 51 – Borboleta simétrica.

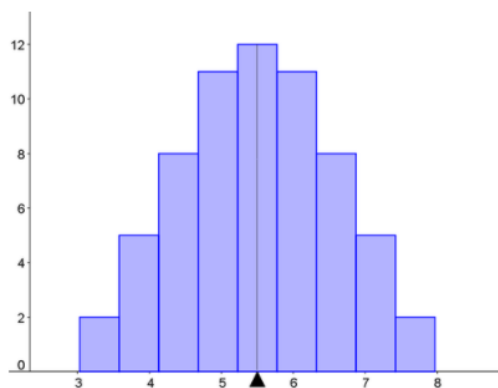


Fonte: Freepik.⁸

⁸ Disponível em https://br.freepik.com/vetores-premium/colecao-de-borboletas_3902529.htm#page=4&query=borboletas&position=14 Acesso em: 28 nov. 2020.

Seguindo essa ideia, comentar com os alunos que assim como em figuras, existe simetria em histogramas. Com isso, apresentar um histograma simétrico. Uma sugestão é a Figura 52.

Figura 52 – Histograma simétrico.



Fonte: Um livro aberto.⁹

Após essa discussão pedir para os alunos analisarem se o histograma das alturas da sala é simétrico ou não. Possivelmente não será, concluindo que se trata de uma assimetria. Desta maneira, o professor irá explicar que existe um cálculo que mede o grau de assimetria de uma amostra, chamado de Coeficiente de Assimetria de Pearson (LOESCH, 2012), cuja fórmula é:

$$T = \frac{(\bar{x} - Mo)}{s} \quad (4.7)$$

onde,

se $|T| \leq 0,15$ a distribuição é praticamente simétrica;

se $0,15 < |T| \leq 1$ é uma assimetria moderada;

se $|T| > 1$ é uma forte assimetria.

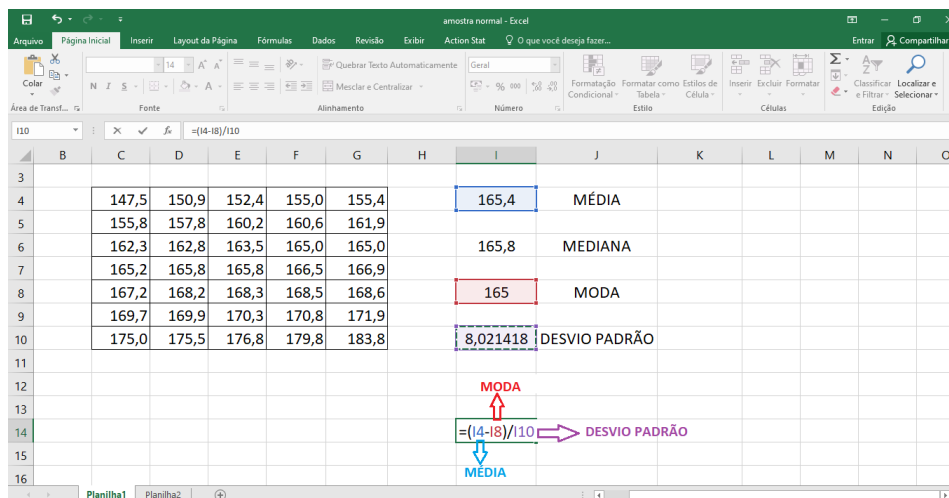
Considerando essa fórmula, propor para os alunos calcularem o coeficiente de assimetria da amostra e verificar em qual grau se enquadra. Esse cálculo pode ser realizado facilmente com uma calculadora, manualmente ou até mesmo no Excel, utilizando os valores da média, moda e desvio padrão determinados na 2ª etapa. A Figura 53 mostra como pode ser feito no Excel usando o comando:

$$= (l4 - l8)/l10$$

onde $l4$, $l8$ e $l10$ são as células com os valores da média, moda e desvio padrão, respectivamente.

⁹ Disponível em: https://www.umlivroaberto.org/BookCloud/Volume_1/master/view/PE104-1.html
Acesso em: 28 nov. 2020.

Figura 53 – Calculando o grau de assimetria no Excel.

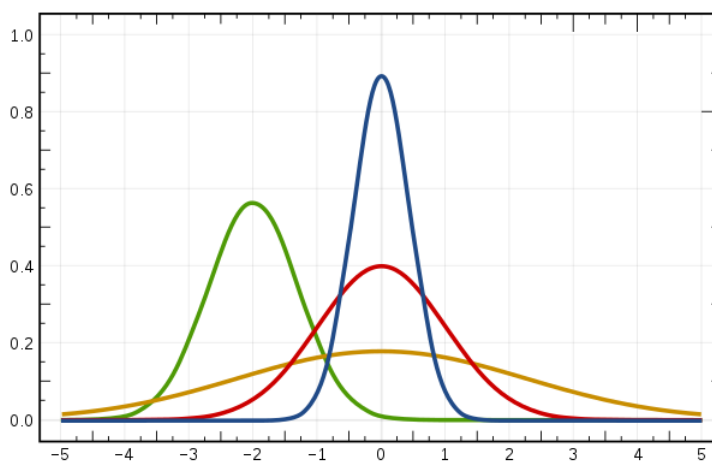


Fonte: Própria autora, 2021.

Para a amostra exemplo, $T = 0,05$. Logo é praticamente simétrica.

Discutido sobre simetrias, o professor comentará com os alunos que existem algumas variáveis contínuas que seguem uma distribuição simétrica e, além disso, apresentam formato semelhante a um sino. Veja que a Figura 54 pode ser mostrada como um exemplo e possui várias curvas para os alunos visualizarem.

Figura 54 – Distribuição sino.



Fonte: Wikipedia ¹⁰

Pode-se dizer também que, uma característica importante desse tipo de distribuição é que a média, mediana e moda têm o mesmo valor e finalizar comentando que a essa distribuição em formato de sino, denomina-se distribuição Normal.

¹⁰ Disponível em: https://pt.wikipedia.org/wiki/Distribui%C3%A7%C3%A3o_normal#/media/Ficheiro:Normal_Distribution.PDF.svg Acesso em: 28 nov. 2020.

Tendo agora a ideia do que é uma distribuição Normal, chegou um dos principais momentos desta atividade: a partir do histograma e do cálculo de assimetria de Pearson para as alturas da amostra, será que pode-se concluir que as alturas dos alunos do Ensino Médio (população) seguem uma distribuição Normal?

Uma vez que essa atividade é totalmente empírica, as amostras serão complementemente diferentes em cada sala de aula que esta for aplicada. Assim, existem dois possíveis cenários: a amostra tem uma assimetria de grau leve a moderado; ou não. Caso o primeiro tenha ocorrido, os alunos poderão responder que a amostra segue uma distribuição Normal. Deste modo, o professor deve orientá-los explicando que não é possível dar certeza da normalidade dos dados apenas pela análise da assimetria e complementar dizendo que em outro momento oportuno, quando eles tiverem mais maturidade, poderão aprender “Teste de Hipóteses” para checar essa normalidade de forma mais eficiente. Porém, por hora, existem indícios sobre (no caso a simetria observada no histograma, e o coeficiente de Pearson) e que esses indícios podem ser utilizados para ilustrar a situação.

Caso o segundo cenário tenha acontecido, deve-se concluir que a distribuição não é Normal, proporcionando um momento interessante para comentar que na vida real é assim que acontece, as vezes não se obtém aquilo que era esperado. Além disso, explicar que, talvez não se obteve por ser uma amostra pequena e se fosse realizado um censo seria melhor observado a normalidade dos dados.

Portanto, uma sugestão caso o segundo cenário aconteça é realizar o censo. Outra possibilidade é utilizar a amostra exemplo desta atividade permitindo dar continuidade neste trabalho, uma vez que a mesma segue uma distribuição Normal.

Comparada com as atividades de distribuições discreta pode-se perceber que a abordagem é realizada de maneira diferente, isso deve-se a complexidade presente na função densidade da distribuição Normal. É por esse motivo que decidiu-se utilizar o método empírico descrito nesta etapa, para na seguinte formalizar os conceitos sobre a distribuição Normal.

5ª Etapa: A distribuição Normal.

Por se tratar do cálculo de probabilidades de variáveis contínuas, os conceitos básicos apresentados nas duas primeiras atividades são semelhantes, tais como variável aleatória e o valor esperado. Se a turma não realizou nenhuma delas, sugere-se ao professor explicar detalhadamente esses conceitos, conforme exposto na 4ª Etapa da Atividade 1. Claro que deve-se ressaltar as diferenças que existem entre variáveis aleatórias discretas e contínuas, bem como a diferença de função de probabilidade para a função densidade. Caso a turma tenha realizado alguma delas, a explicação dos próximos itens será mais fácil, podendo fazer uma retomada ou comparações com os conceitos dos dois tipos de variáveis.

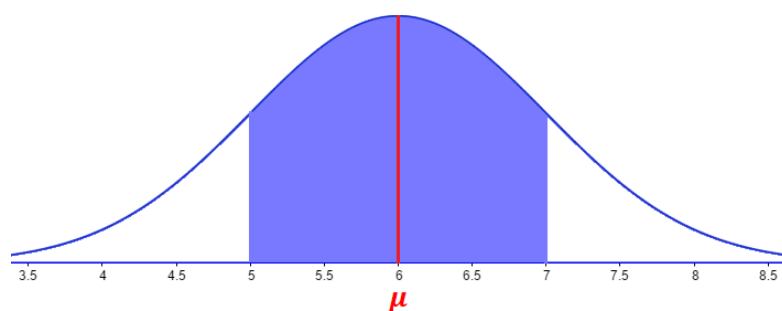
- a) Variável aleatória contínua: possui um intervalo de números reais, incluindo assim

os números decimais. No caso dessa pesquisa, $X =$ alturas dos alunos.

O professor pode solicitar para os alunos falarem outros exemplos de variáveis contínuas. Além disso, deixar claro que, nem todas variáveis contínuas seguem uma distribuição Normal, mas já é comprovado que a estatura de uma população é, por isso que esse tema foi escolhido, retomando a fala do início da primeira etapa.

b) Curva normal: possui como característica a simetria e o formato de sino, como visto na Figura 55. O ponto máximo corresponde a média, moda e mediana, pois são coincidentes.

Figura 55 – Curva normal.



Fonte: Própria autora, 2021.

Outro ponto importante é o professor falar que no caso de variáveis aleatórias contínuas, a probabilidade é dada pela área sob a curva da função densidade. Na Figura 55, por exemplo, a área em azul indica a probabilidade do intervalo $[5, 7]$. Sendo assim, é possível calcular probabilidades apenas de intervalos, já que um ponto projeta apenas um segmento de reta, como o segmento vermelho traçado indicando a média da distribuição da figura em questão.

Além disso concluir que a área total sob a curva é 1, uma vez que a probabilidade de todo o espaço amostral (1) é distribuída entre os possíveis intervalos da amostra.

c) Função densidade de probabilidade da distribuição Normal é dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

onde:

$-\infty < x < \infty$;

μ é a média;

σ é o desvio padrão.

Aqui o professor irá ressaltar que a curva da distribuição Normal é a função densidade de probabilidade e que a probabilidade de uma variável aleatória contínua não é obtida diretamente por essa função, mas sim pela área sob ela.

Dizer também que pode ser usado também a notação $X \sim N(\mu, \sigma)$ para representar uma distribuição Normal, uma vez que a média e o desvio padrão são os parâmetros. Retomando a Figura 54, aquelas curvas apresentavam formas diferentes justamente por ter médias e desvios padrões distintos.

Se surgir dúvida quanto a notação, cabe explicar que, quando está trabalhando com população a notação da média e desvio padrão são, respectivamente, μ e σ ; quando se é amostra, utiliza-se \bar{x} e s .

d) Valor esperado: para uma distribuição Normal, o valor esperado é a própria média.

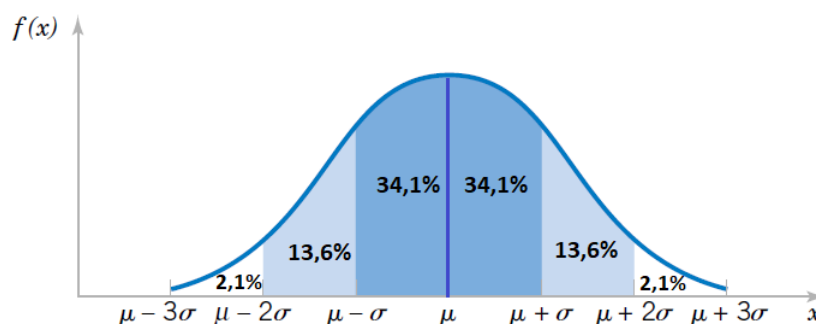
6ª Etapa: Calculando probabilidades na distribuição Normal.

A introdução desta etapa que o professor fará com a turma é explicando que assim como nas distribuições de variáveis discretas, é possível calcular probabilidades de variáveis contínuas a partir função de densidade. Porém, para a distribuição Normal, sua complexidade faz com que esse meio não seja utilizado.

Diante disto, eles aprenderão três maneiras de calcular probabilidades desse tipo de distribuição: por meio da propriedade relacionada aos desvios padrões; utilizando recursos computacionais, no caso, o Geogebra e por fim, pela padronização.

De Moivre (1667-1754) mostrou que em todas as distribuições normais 68% dos dados pertencem ao intervalo $(\mu - \sigma, \mu + \sigma)$, 95% estão no intervalo $(\mu - 2\sigma, \mu + 2\sigma)$ e 99% em $(\mu - 3\sigma, \mu + 3\sigma)$ e, mais especificamente, como a distribuição é simétrica têm-se as probabilidades na Figura 56.

Figura 56 – Probabilidades na distribuição Normal.



Fonte: Montgomery e Runger (2002, adaptada).

Neste momento, propor algumas perguntas para os alunos que podem ser respondidas utilizando as propriedades.

A princípio, o professor pode usar a média e o desvio-padrão da amostra como sendo o da população, ou seja, a turma está fazendo uma estimativa pontual para os valores reais de μ e σ , que até então são desconhecidos.

Essa observação pode ser feita para os alunos e complementar dizendo que existem formas mais sofisticadas (mais eficientes) de estimar esses valores, como o uso de Intervalos de Confiança. Mas, é um assunto que exige um nível maior de conhecimento e é geralmente abordado em cursos de graduação.

Logo, para fins ilustrativos dessa atividade, sem perda de generalidade, será usado $\mu = 165,4$ e $\sigma = 8$.

Segue algumas sugestões de perguntas:

1) Escolhendo um aluno do Ensino Médio, qual a probabilidade dele ter altura acima da média?

Resposta: Pela propriedade da distribuição normal, a probabilidade é 50%.

2) Qual a probabilidade de um aluno ter altura entre 165,4cm e 173,4cm?*

*Note: esse intervalo poderá alterar conforme a amostra, mas garantir que é o intervalo equivalente ao $(\mu, \mu + \sigma)$ da amostra utilizada.

Resposta: Espera-se que os alunos concluam que é 34,1%.

3) Qual a probabilidade de um aluno ter altura entre 159cm e 175cm? (equivalente ao intervalo $(\mu - \sigma, \mu + \sigma)$).

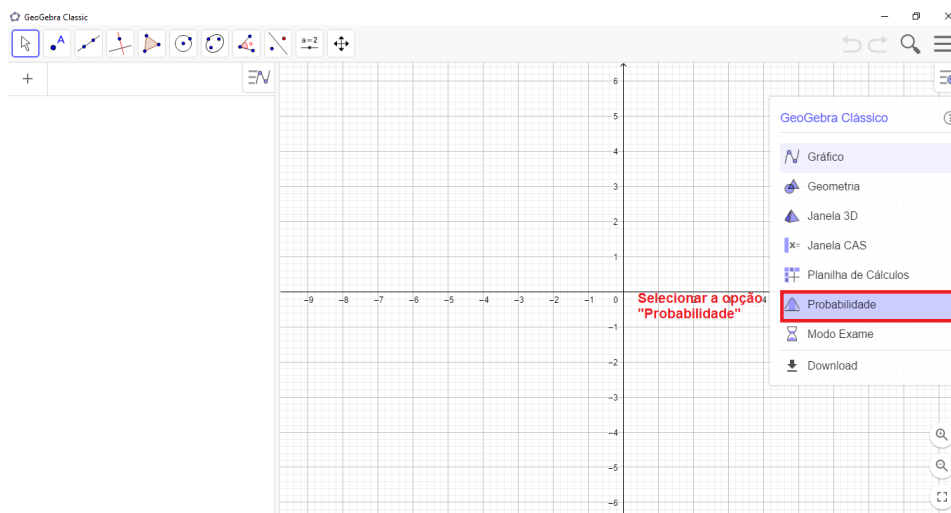
Resposta: Os alunos devem concluir que é 68,2% .

Para complementar, utilizar o Geogebra para calcular outras probabilidades, pois é simples e dinâmico o uso *software* para esta modalidade. Deste modo, os alunos podem calcular facilmente as probabilidades, incluindo as que responderam anteriormente utilizando as propriedades, comparando os resultados obtidos. Novamente, se a escola não disponibiliza acesso a um computador, pode utilizar um projetor e mostrar aos alunos como utilizar o Geogebra para o cálculo de probabilidades de uma distribuição Normal. Se não for viável o uso do projetor, pode-se dar um panorama geral sobre como o programa funciona e avançar para a 7ª Etapa.

Segue um tutorial de como utilizar o Geogebra para a distribuição Normal:

a) Assim que o programa for aberto, selecionar a janela de “Probabilidade”. Veja Figura 57.

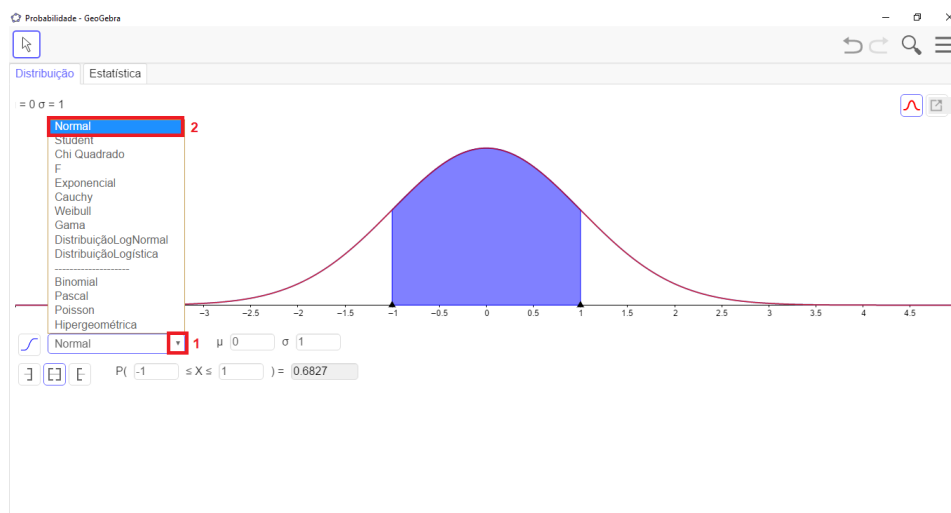
Figura 57 – Janela de probabilidade do Geogebra.



Fonte: Própria autora, 2021.

A janela de distribuição será aberta. Se por acaso não estiver selecionada a distribuição Normal, basta clicar na janela de opções e alterar para a desejada. Figura 58.

Figura 58 – Selecionando a distribuição.

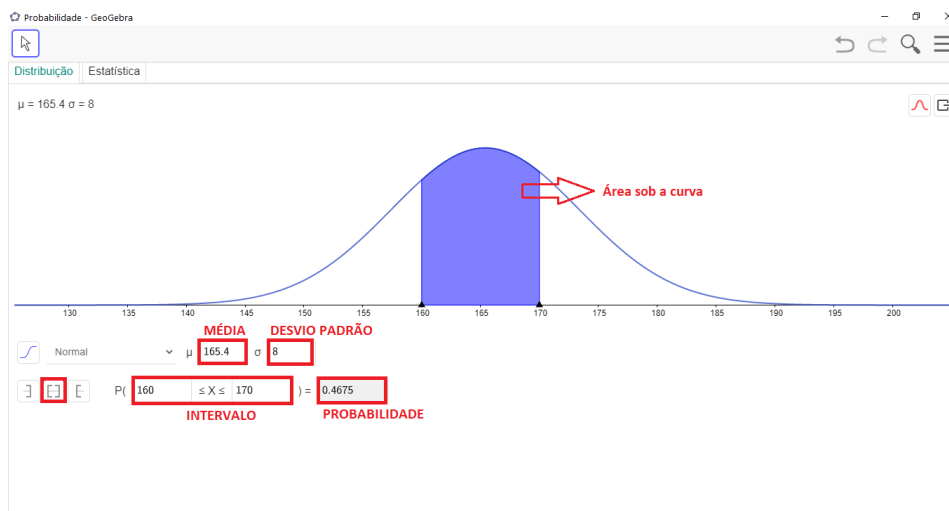


Fonte: Própria autora, 2021.

b) Os valores da média e do desvio padrão podem ser alterados. Para calcular probabilidades, pode mudar o intervalo digitando ou arrastando as setas do eixo do gráfico. Figura 59.

Quanto as probabilidades que serão calculadas no Geogebra, propõe-se que também sejam calculadas probabilidades de intervalos que não pertencem à propriedade, tais como: 160cm a 170cm e 175cm a 180cm.

Figura 59 – Distribuição Normal no Geogebra.



Fonte: Própria autora, 2021.

Enquanto os alunos exploram o Geogebra, é um momento para o professor colocar em pauta a importância do uso da tecnologia nas pesquisas, por facilitarem todo o processo. Além disso, orientá-los que embora o programa realize as operações, isso não invalida a necessidade de saber matemática, pois eles viram que é preciso compreender os valores que devem ser inseridos e claro, saber interpretar os resultados obtidos.

Por fim, levantar a dúvida: “E sem esses programas, como podem ser calculadas as probabilidades de uma distribuição normal para qualquer intervalo?”

É neste momento que entra a padronização.

7ª Etapa: Distribuição Normal Padrão.

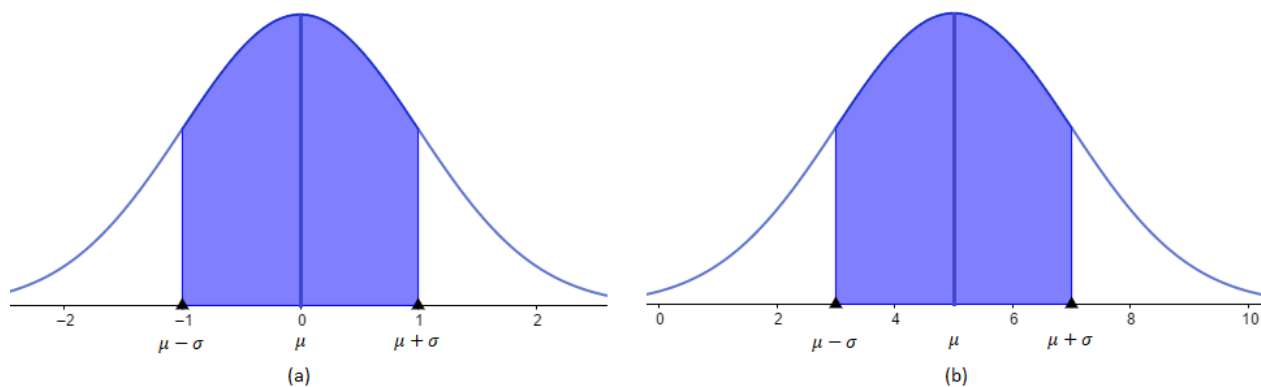
O professor pode expor para os alunos que na etapa anterior eles utilizaram a propriedade como um meio de resolver os exercícios propostos. Porém, observa-se que se fica limitado em calcular probabilidades apenas para os intervalos dos desvios padrões. Deste modo, para trabalhar com qualquer intervalo, sem uso da tecnologia, é utilizada a padronização.

Com o intuito de explicar a padronização de modo mais simples, sugere-se ao professor comentar primeiramente da distribuição Normal Padrão, que é uma distribuição Normal com média $\mu = 0$ e desvio padrão $\sigma = 1$ e, que existe uma tabela já com todas as probabilidades calculadas para todos os valores possíveis dessa distribuição. Além disso, falar que estudos mostraram que a área sob a curva de qualquer distribuição Normal, é equivalente a área da distribuição Normal Padrão, de acordo com os desvios padrões.

A Figura 60, pode ser mostrada para a turma, deixando claro que a da esquerda é a representação da distribuição Normal Padrão e a da direita é uma distribuição Normal com média $\mu = 5$ e desvio padrão $\sigma = 2$. Ademais enfatizar que, mesmo sendo valores

de parâmetros diferentes, as áreas em destaque são equivalentes. O mesmo ocorre para quaisquer outros intervalos e parâmetros.

Figura 60 – Comparando áreas.



Fonte: Própria autora, 2021.

Assim, deve ser dito que a “ligação” entre qualquer distribuição normal com a padrão, é a fórmula da variável aleatória normal padrão Z , a qual transforma intervalos de distribuições Normais de parâmetros $N(\mu, \sigma)$ para intervalos da distribuição Normal Padrão.

Variável aleatória normal padrão Z : define-se como $z = \frac{x - \mu}{\sigma}$ o valor de z obtido pela padronização de x .

Um exemplo será feito a seguir para uma explicação mais simples e direta, sendo a mesma maneira que o professor prosseguirá com a turma. Além disso, é importante que os alunos tenham em mãos a tabela da Distribuição Normal Padronizada do Anexo A.

Suponha-se que queira determinar a probabilidade de um aluno ter altura entre 165,4cm a 170cm, onde 165,4cm é média da amostra exemplo (a qual o professor substituirá para a média da amostra da turma).

A primeira parte da resolução é determinar a representação desses valores do intervalo na variável Z . Como dito, a média é 0, mas pode-se calcular para ela também, mostrando que, de fato, ela corresponde ao 0:

$$z_1 = \frac{165 - 165}{8} = 0$$

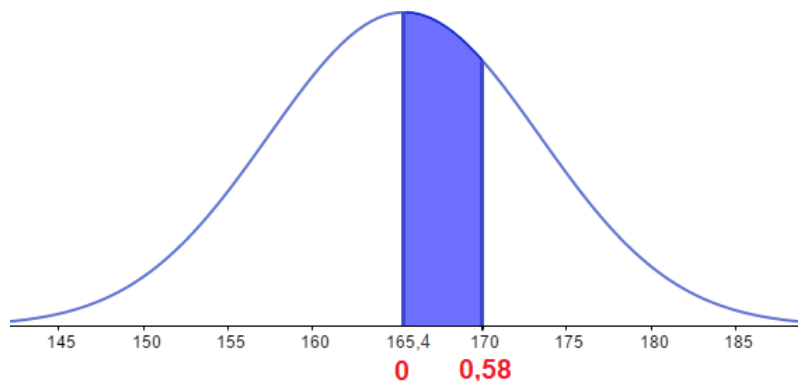
$$z_2 = \frac{170 - 165,4}{8} = 0,575 \cong 0,58,$$

Os subíndices 1 e 2 são utilizados apenas para diferenciar as variáveis.

É importante que os alunos compreendam o que esses valores significam, tornando a representação gráfica fundamental neste processo. Para tanto, o professor irá desenhar uma distribuição que contenha o intervalo citado no exemplo e irá associar aos resultados de z_1

e z_2 , como mostra a Figura 61. Após isso deve deixar claro que, determinar a probabilidade entre 165,4cm e 170cm é equivalente a do intervalo de 0 a 0,58 na distribuição padronizada.

Figura 61 – Distribuição normal padronizada.



Fonte: Própria autora, 2021.

Neste momento, a tabela será usada para determinar a probabilidade. Para utilizá-la, o primeiro ponto a explicar para a turma é que a mesma fornece a probabilidade do intervalo $(0, z)$, o que significa que ao procurar 0,58 por exemplo, o número correspondente indica a probabilidade de ocorrer um valor entre 0 e 0,58, ou seja, $P(0 < Z < 0,58)$.

O outro ponto é como localizar z . O professor precisa falar que a primeira coluna representa o valor do inteiro e o primeiro decimal de z , enquanto a primeira linha corresponde ao segundo decimal. Por exemplo, o número $0,58 = 0,5 + 0,08$, possui 0,5 como parte inteira e o primeiro (que será buscado na primeira coluna) e o 8 é o segundo decimal (representado na primeira linha), a interseção dessa linha e coluna informam a probabilidade do intervalo desejado.

Sugere-se ao professor destacar esses resultados na tabela, conforme a Figura 62.

Figura 62 – Lendo da tabela da variável Z .

z_0	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830

Fonte: Centro de informática - UFPE ¹¹

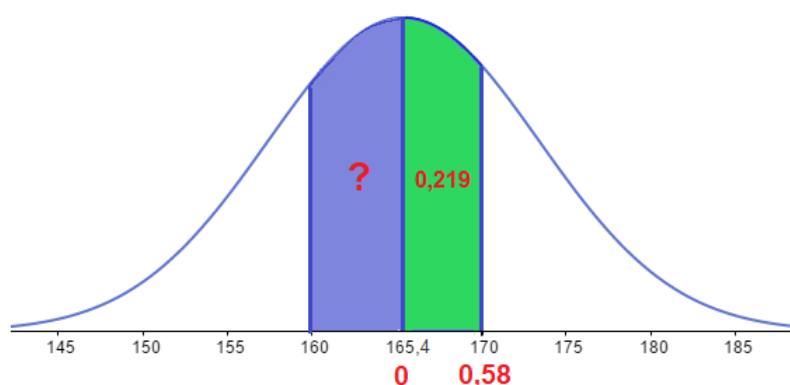
¹¹ Disponível em: <https://cin.ufpe.br/~et586cc/download/Normal.pdf>. Acesso em: 03 nov. 2020.

Agora pode-se responder o exemplo, concluindo que a probabilidade de um aluno ter altura entre 165,4cm e 170cm é de 0,2190, aproximadamente.

Depois de explicar a ideia básica do uso da tabela, questionar aos alunos: “E se quisessem determinar a probabilidade de um aluno ter altura entre 160cm e 170cm?” Intervalo o qual eles calcularam usando o Geogebra.

Para responder essa pergunta, auxiliar os alunos a fazerem a representação gráfica. Caso os alunos não tenham destacado a média nessa distribuição, solicitar para eles fazerem e, além disso, inserir as informações encontradas no exemplo anterior.

Figura 63 – Interpretação do segundo exemplo Z .



Fonte: Própria autora, 2021.

Com uma representação semelhante a Figura 63, é possível perceber que a probabilidade de ocorrer um valor entre 160cm a 170cm, é a união das probabilidades de 160cm a 165,4cm e 165,4cm a 170cm, ou seja, a área azul e verde juntas. A área verde já foi determinada no exemplo anterior, basta agora determinar a azul. Para isso, o professor pode sugerir aos alunos calcular z para 160. Considerando a média e desvio padrão da amostra exemplo, têm-se:

$$z = \frac{160 - 165,4}{8} = -0,675 \cong -0,68.$$

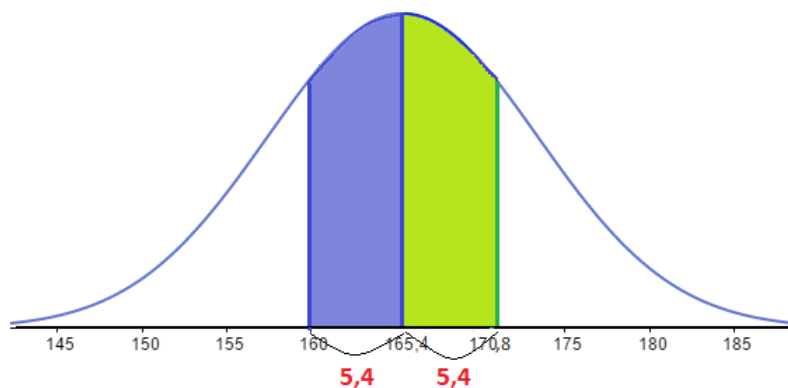
Veja que o valor deu negativo, assim o professor irá explicar que é pelo fato de ser uma altura menor que a média. A tabela não possui valores negativos, porém deve ser lembrado que a distribuição é simétrica, e o simétrico de 160cm para essa amostra exemplo é 170,8, pois são equidistantes (mesma distância) da média. Veja a Figura 64

Deste modo, concluir que calcular a probabilidade de 160cm a 165,4cm é equivalente ao do intervalo (165,4; 170,8). Determinando a variável z para 170,8 obtém:

$$z = \frac{170,8 - 165,4}{8} = 0,675 \cong 0,68,$$

ou seja, o simétrico de $-0,68$ na distribuição padronizada, logo, na prática basta desconsiderar o negativo no momento de buscar o z na tabela.

Figura 64 – Simétricos em uma Distribuição Normal.



Fonte: Própria autora, 2021.

Finalmente pode-se calcular a probabilidade de ocorrer uma altura entre 160cm a 170cm, somando-se 0,2517 e 0,2190, onde 0,2517 é probabilidade do intervalo $(-0,68, 0)$ correspondente, a área azul da Figura 63.

Portanto, a resposta que se busca é $0,2517 + 0,2190 = 0,4707$.

8ª Etapa: Conclusão.

Depois de todo esse processo, os alunos já têm a noção do que é uma distribuição Normal e sobre o cálculo de probabilidades. Aqui pode ser proposto que os alunos realizem a mesma atividade porém considerando como variável os pesos da turma. Além disso, o professor pode reforçar a ideia de que, quanto maior a amostra, mais se aproximará da distribuição Normal.

Para finalizar, os alunos podem confeccionar um cartaz com imagens do histograma, da distribuição normal, contendo informações sobre probabilidades da estatura dos alunos e expor no mural da escola. Quanto as informações podem ser, por exemplo, que 50% dos alunos do Ensino Médio tem altura menor que 165,4cm; 47% possuem a estatura entre 160cm e 170cm e etc. Além disso, pode ser proposto mais dois exercícios de distribuição Normal. Segue algumas sugestões:

Exercício 1) (LARSON e FARBER, 2010) Suponha que os níveis de colesterol dos homens nos Estados Unidos sejam normalmente distribuídos, com uma média de 215 miligramas pro decilitro e um desvio padrão de 25 miligramas por decilitro. Você seleciona aleatoriamente um homem dos Estados Unidos. Qual é a probabilidade de que o nível de colesterol dele seja menor que 175? Resolva utilizando a padronização e confirme o resultado no Geogebra.

Resposta: $P(Z \leq 175) = 0,0548 = 5,48\%$

Exercício 2) (LARSON e FARBER, 2010) Uma pesquisa foi conduzida para medir o número de horas que adultos nos Estados Unidos passam em frente ao computador todas as semanas. Na pesquisa, o número de horas era normalmente distribuído, com uma média de 7 horas e um desvio padrão de 1 hora. Um participante da pesquisa é escolhido aleatoriamente. Encontre a probabilidade desse participante passar entre 5,5 e 9,5 horas por semana em frente ao computador.

Resposta: $P(5,5 \leq Z \leq 9,5) = 0,927 = 92,7\%$

4.5 Análise das habilidades da BNCC

Tendo em vista a descrição das atividades é possível, agora, identificar os momentos em que as habilidades citadas se desenvolvem.

As Atividades 1 e 2 são de metodologia semelhantes por se tratarem de distribuições de probabilidade discretas. Nelas, pode-se perceber que a habilidade EM13MAT106 está presente no momento de reflexões: quando é analisado se alguém vai ganhar o jogo da Mega-Sena, o porquê é difícil ganhar e também sobre as chances de um candidato acertar determinado número de questões na prova Enem apenas chutando as respostas. A EM13MAT311 se aplica no momento de calcular as probabilidades de ganhar apenas um jogo da Mega-Sena ou a probabilidade de acertar uma questão da prova. Ao calcular as várias probabilidades para se deduzir as funções trabalha-se a EM13MAT312, pois envolve experimentos aleatórios sucessivos. E por fim, a EM13MAT511, uma vez que na distribuição geométrica é trazido o conceito de espaço amostral infinito, assim como o espaço amostral finito e discreto em ambas atividades.

Ao analisar mais a fundo, não são apenas essas habilidades que são abrangidas. Embora a EM13MAT406 indica a construção e interpretação de gráficos e tabelas com base em dados obtidos em pesquisas por amostragens, de certo modo, parte dela é trabalhada nas etapas que se constroem as tabelas e o gráfico no Excel. Além disso, o processo realizado para deduzir as funções de probabilidade, desenvolve uma habilidade que se classificaria na competência específica 5, uma vez que condiz a investigar e estabelecer conjecturas de diferentes conceitos e propriedades matemáticas ao empregar a observação de padrões.

Quanto a Atividade 3 que é de variável aleatória contínua, as habilidades EM13MAT202 e EM13MAT316 são exploradas ao determinar e interpretar as medidas de tendência central e de dispersão usando recursos tecnológicos. A EM13MAT406 no momento em que se constroe a tabela de frequências da amostra e a EM13MAT407 na construção do histograma, inclusive ao reconhecer que ele é o mais eficiente por estar lidando com variável contínua. Em relação a parte de probabilidade são desenvolvidas as habilidades EM13MAT311 e EM13MAT511, pois propõe o cálculo de probabilidades após assumir que se trata de uma distribuição Normal e apresenta um espaço amostral de variável contínua.

Vale ressaltar que o professor tem total autonomia para adaptar as atividades de acordo com a turma, visto que cada uma possui sua particularidade. É por esse motivo que não estimou-se tempo de duração para a aplicação.

5 CONSIDERAÇÕES FINAIS

O papel do professor tem mudado frequentemente com o decorrer dos anos, e hoje, é primordial buscar ferramentas e atividades que relacionam os conteúdos do currículo aplicados a situações da realidade, e as distribuições são excelentes aplicações da matemática na área de probabilidade.

Relacionar os conteúdos da sala de aula à realidade do aluno se faz necessário, pois torna-se a aprendizagem mais significativa. Esse tipo de aprendizagem faz com que os alunos realmente compreendam a importância da probabilidade para a formação de cidadãos mais críticos, uma vez que ela auxilia na leitura e interpretação de diversas pesquisas que são realizadas para compreender a sociedade de modo geral.

No Capítulo 4 pode-se perceber que o objetivo proposto é cumprido. No decorrer das atividades o aluno é estimulado a pensar matematicamente e perceber que muitos resultados na matemática são uma leitura do mundo, quebrando a ideia de que matemática “não serve para nada”, frase muito comum dita por alunos nas aulas de matemática. Esse cumprimento do objetivo pode ser medido por uma avaliação contínua, isto é, pela percepção do progresso que os alunos apresentaram durante as atividades e outra maneira de avaliar é através da desenvoltura da resolução dos exercícios sugeridos para aplicação ao final de cada atividade.

Infelizmente, não foi possível colocar as atividades em prática, devido a pandemia presente no momento em que este trabalho foi escrito. Sendo assim, como trabalhos futuros pretende-se aplicá-las em uma sala de aula para verificar o cumprimento das habilidades da BNCC na prática, além de relatar as experiências proporcionadas pelas mesmas.

Ademais, espera-se que este trabalho auxilie professores a desenvolverem uma aprendizagem de modo mais significativo com alunos do Ensino Médio, seja pela aplicação das atividades em sua totalidade ou adaptações, e até ideias para novas contribuições na educação matemática.

Referências

- BRASIL. Ministério da Educação. PCN+ Ensino Médio: Ciências da Natureza, Matemática e suas Tecnologias. Brasília, 2002. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/ciencian.pdf>. Acesso em: 20 jan. 2020.
- BRASIL. Ministério da Educação. Base Nacional Comum Curricular. Brasília, 2018. Disponível em: <http://basenacionalcomum.mec.gov.br/abase/>. Acesso em: 20 jan. 2020.
- FREITAS FILHO, Adail Boa de. Probabilidade: Uma proposta à luz da bncc. Dissertação de Mestrado (PROFMAT) - Universidade da Integração Internacional da Lusofonia Afro-Brasileira, Redenção - CE, 2020. Disponível em: https://sca.profmtat-sbm.org.br/sca_v2/get_tcc3.php?cpf=05422270371&d=20210318175638&h=fc722136cd436f2c9bfd7e315b78d4e06f3f31d6. Acesso em: 01 nov. 2020.
- GONÇALVES, Paulo Henrique Rodrigues. Uma abordagem da distribuição normal através da resolução de uma situação problema com a utilização do software geogebra. Dissertação de Mestrado (PROFMAT) - Universidade Federal de Goiás, Jataí - GO, 2014. Disponível em: https://sca.profmtat-sbm.org.br/sca_v2/get_tcc3.php?cpf=86614746120&d=20210318165828&h=8c5c589938ea45bc785df35c36b9326e6bb261c1. Acesso em: 11 maio 2020.
- LARSON, Ron; FARBER, Betsy. *Estatística aplicada*. 4. ed. São Paulo: Pearson, 2010.
- LIBERAL, Tarciana. Probabilidade 1. 16 slides. Disponível em: <http://www.de.ufpb.br/~tarciana/Probabilidade/Aula14.pdf>. Acesso em: 11 maio. 2020.
- LOESCH, Cláudio. *Probabilidade e estatística*. 1. ed. Rio de Janeiro: LTC, 2015.
- LOPES, Celi Espasandin. O ensino da estatística e da probabilidade na educação básica e a formação de professores. Cadernos CEDES, Campinas, 2008.
- LOPES, Celi Espasandin and MEIRELLES, Elaine. Estocástica nas séries iniciais. XVIII Encontro Regional de Professores de Matemática, Campinas, 2005.
- MEYER, Paul L. *Introductory probability and statistical applications*. 2. ed. United States of America: Addison-Wesley Publishing Company, 1970.
- MONTGOMERY, Douglas C.; RUNGER, George C. *Applied Statistics and Probability for Engineers*. 3. ed. United States of America: John Wiley e Sons, 2003.
- MONTGOMERY, Douglas C.; RUNGER, George C. *Estatística aplicada e probabilidade para engenheiros*. 6. ed. Rio de Janeiro: LTC, 2020.
- MORETTIN, Pedro A.; BUSSAB, Wilton de O. *Estatística Básica*. 6. ed. São Paulo: Saraiva, 2010.
- ROCHA, Samy Marques. Distribuição Binomial e aplicações. Dissertação de Mestrado (PROFMAT) - Universidade Federal do Maranhão, São Luís - MA, 2017. Disponível em: https://sca.profmtat-sbm.org.br/sca_v2/get_tcc3.php?cpf=76086860300&d=

20210318164649&h=480008bc95c60cf672b8c55f97db20c28ee8874b}). Acesso em: 29 jan. 2020.

ROSS, Sheldon. *Probabilidade: Um curso moderno com aplicações*. 8. ed. Porto Alegre: Bookman, 2010.

SANTOS, Fábio Nascimento dos. Algumas aplicações da distribuição Binomial. Dissertação de Mestrado (PROFMAT) - Universidade Federal Fluminense, Niterói - RJ, 2015. Disponível em: https://sca.proformat-sbm.org.br/sca_v2/get_tcc3.php?cpf=02951122713&d=20210318231634&h=3208efc24a00d0debf5d116c52f637ae7151d750). Acesso em: 11 maio 2020.

SILVA, Nelson Antonio; FIGUEIREDO, Helenara R. Sampaio. A educação estatística na educação básica do Brasil, Estados Unidos, França e Espanha segundo os documentos curriculares. REVEMAT, Florianópolis - SC, v. 14, 2019. Disponível em: <https://periodicos.ufsc.br/index.php/revemat/article/view/1981-1322.2019.e62813>). Acesso em: 20 jan. 2020.

SOARES, Jonas Dutra. Probabilidade: uma proposta didática para se trabalhar no ensino médio. Dissertação de Mestrado (PROFMAT) - Universidade Federal de Lavras, Lavras - MG, 2020. Disponível em: https://sca.proformat-sbm.org.br/sca_v2/get_tcc3.php?cpf=05404042680&d=20210318165527&h=ee7772994cbcb4d460b239928ecf4184dbb9065). Acesso em: 01 nov. 2020.

VIEIRA FILHO, Aprígio dos Santos. Probabilidade e suas aplicações. Dissertação de Mestrado (PROFMAT) - Fundação Universidade Federal de Rondônia, Porto Velho - RO, 2020. Disponível em: https://sca.proformat-sbm.org.br/sca_v2/get_tcc3.php?cpf=48622516415&d=20210318163543&h=1743a6769783162070afb1b8b62f1d3fe6a8866a). Acesso em: 10 set. 2020.

