

**UNIVERSIDADE DE SÃO PAULO**  
Instituto de Ciências Matemáticas e de Computação

**Abordando a curva normal no ensino médio**

**Ligia Maria Ferreira de Paula Silva**

Dissertação de Mestrado do Programa de Mestrado Profissional em  
Matemática em Rede Nacional (PROFMAT)



SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: \_\_\_\_\_

**Ligia Maria Ferreira de Paula Silva**

## Abordando a curva normal no ensino médio

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestra em Ciências – Mestrado Profissional em Matemática em Rede Nacional. *VERSÃO REVISADA*

Área de Concentração: Mestrado Profissional em Matemática em Rede Nacional

Orientadora: Profa. Dra. Geraldine Góes Bosco

**USP – São Carlos**  
**Julho de 2021**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi  
e Seção Técnica de Informática, ICMC/USP,  
com os dados inseridos pelo(a) autor(a)

L586a Ligia Maria Ferreira de Paula Silva, Silva  
Abordando a curva normal no ensino médio. /  
Silva Ligia Maria Ferreira de Paula Silva;  
orientador Geraldine Góes Bosco. -- São Carlos,  
2021.  
92 p.

Dissertação (Mestrado - Programa de Pós-Graduação  
em Mestrado Profissional em Matemática em Rede  
Nacional) -- Instituto de Ciências Matemáticas e de  
Computação, Universidade de São Paulo, 2021.

1. Probabilidade. 2. Variável Aleatória. 3. Curva  
Normal. 4. Aplicações. I. Bosco, Geraldine Góes,  
orient. II. Título.

**Ligia Maria Ferreira de Paula Silva**

## High school normal curve addressing

Dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP – in accordance with the requirements of the Professional Master's Program in Mathematics in National Network, for the degree of Master in Science. *FINAL VERSION*

Concentration Area: Professional Master Degree Program in Mathematics in National Network

Advisor: Profa. Dra. Geraldine Góes Bosco

**USP – São Carlos**  
**July 2021**



# AGRADECIMENTOS

---

---

Agradeço a todos os que foram meus docentes na USP (Ribeirão Preto), em especial, a minha orientadora Geraldine, que sempre me atendia, mesmo aos sábados, domingos e feriados. Agradeço a ela por ter sido exigente comigo, e muitas vezes ter falado coisas que eu não gostaria de ouvir. Foi a atitude dela que me fez crescer. Aprendi a correr atrás e não desistir diante das dificuldades. Minha gratidão aos organizadores do programa Profmat por escolher professores tão competentes para a nossa formação.

A todos os meus companheiros de PROFMAT. Tornaram-se amigos muito queridos que sempre estavam prontos a colaborar comigo e, mesmo a distância, continuam me ajudando. São pessoas muito competentes e que morarão no meu coração.

Agradeço ao meu filho e minha mãe, Pedro e Ana Rosa, pela paciência em lidar com minha ausência nos finais de semana, pois eu tinha sempre que estudar ou escrever o trabalho. Ao meu marido, Sinésio, pela força e palavras de incentivo, que sempre me motivava.

Minha eterna gratidão ao meu querido irmão Marcos e admirada cunhada Talita, sua esposa. Com toda a paciência e sabedoria, eles leram o meu trabalho nos seus tempos livres e mostraram modos de melhorá-lo. São pessoas iluminadas.

Ao meu pai amado, que não está mais comigo, dedico esse trabalho pois ele sempre acreditou em mim. Quanta falta ele me faz.

Agradeço aos amigos e parentes pela paciência que tiveram comigo e que me ajudaram quando mais precisei.

Agradeço a Deus, acima de tudo, por me permitir chegar até aqui.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.





*“Feliz aquele que transfere o que sabe e aprende o que ensina.”*

*(Cora Coralina)*



# RESUMO

SILVA, L. M. F. DE P. **Abordando a curva normal no ensino médio**. 2021. 92 p. Dissertação (Mestrado em Ciências – Mestrado Profissional em Matemática em Rede Nacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2021.

Este trabalho apresenta os passos teóricos básicos para o professor de Matemática do ensino médio entender a Curva Normal e algumas de suas aplicações. Apresentamos uma trilha de conhecimentos, saindo da definição usual da Probabilidade, passando pela definição axiomática da Probabilidade, que é pouco conhecida pelos docentes do ensino básico; passando, pelo conceito de variáveis aleatórias simples como a Bernoulli e a Binomial, que se relacionam com a curva Normal, através de um teorema chamado Teorema Central do Limite. Apresentamos também conceitos introdutórios de Estatística, como estimadores, estimação, intervalos de confiança e testes de hipóteses. As aplicações da Curva Normal que propusemos usam esses conceitos estatísticos para analisar características de uma população a partir de uma amostra. Dependendo da população escolhida, podemos analisar características como peso e altura de jovens, ou a vida útil de uma lâmpada ou aparelho. A base teórica que apresentamos dá autonomia para o professor de Matemática criar roteiros de aulas de acordo com os interesses de suas turmas. E pensando nos alunos do ensino básico, não há como deixar de propor o uso de aplicativos e planilhas eletrônicas quando falamos da Curva Normal e suas aplicações. Nesse sentido, este trabalho tenta contribuir para fazer uma ponte entre a teoria apresentada e sua contrapartida digital. Por isso, para auxiliar o professor apresentamos um exemplo de roteiro de aula que envolve a Curva da Normal e o uso eletrônicas e gráficas.

**Palavras-chave:** Probabilidade, Curva Normal, Aplicações, Resolução de Problemas.



# ABSTRACT

SILVA, L. M. F. DE P. **High school normal curve addressing**. 2021. 92 p. Dissertação (Mestrado em Ciências – Mestrado Profissional em Matemática em Rede Nacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2021.

This work presents the basic theoretical steps for the high school Mathematics teacher to understand the Normal Curve and some of its applications. We present a path of knowledge, leaving from the usual definition of Probability, passing through the axiomatic definition of Probability, which is little known by teachers of basic education; passing through the concept of simple random variables such as Bernoulli and Binomial, which are related to the Normal curve, through a theorem called Central Limit Theorem. We also introduce introductory Statistics concepts such as estimators, estimation, confidence intervals, and hypothesis testing. The Normal Curve applications we proposed use these statistical concepts to analyze characteristics of a population from a sample. Depending on the population chosen, we can analyze characteristics such as weight and height of young people, or the life of a lamp or appliance. The theoretical basis that we present gives the Mathematics teacher autonomy to create lesson scripts according to the interests of their classes. And thinking about elementary school students, it is impossible not to propose the use of applications and electronic spreadsheets when we talk about the Normal Curve and its applications. In this sense, this work tries to contribute to make a bridge between the presented theory and its digital counterpart. Therefore, to help the teacher, we present an example of a class script that involves the Normal Curve and the use of electronics and graphics.

**Keywords:** Probability, Normal Curve, Applications, Problems Solving.



# LISTA DE ILUSTRAÇÕES

---

---

Figura 1	– Curva Normal tomada para estipular os tamanhos das saias (PP, P, M, G ou GG) confeccionadas por alunos de nonos anos . . . . .	23
Figura 2	– Variável aleatória $X$ levando os valores de $\Omega$ até $\mathbb{R}$ e $Y$ levando os valores $-1, 0$ e $1$ em $0$ e $1$ . . . . .	36
Figura 3	– Representação gráfica de $F(x)$ , do exemplo 2.2.8, em que $X$ é uma variável aleatória binomial com $n = 5$ e $p = \frac{1}{2}$ . . . . .	43
Figura 4	– Representação gráfica de $f_X(x)$ e da probabilidade de $X$ assumir valores entre $a$ e $b$ , dada pela área da figura hachurada . . . . .	52
Figura 5	– Representação gráfica da probabilidade de $X$ estar entre os números $8$ e $10$ . . . . .	54
Figura 6	– Representação gráfica da probabilidade com que a variável aleatória $X$ pertença ao intervalo $[8, 12]$ . . . . .	55
Figura 7	– Na tabela normal, os números estão na $1^a$ coluna representam valores reais maiores ou iguais a zero e sua primeira casa decimal. Os números que aparecem na $1^a$ linha indicam a segunda casa decimal de $x$ . Os números que estão nas demais linhas e colunas, representam $\mathbb{P}(Z \leq x)$ . No texto, usamos $z$ ao invés de $x$ . . . . .	56
Figura 8	– Representação gráfica do exemplo 3.2.4 mostrando a probabilidade de $X$ estar entre $124$ e $132$ de uma população com média $\mu = 128$ e desvio padrão $\sigma = 16$ . . . . .	60
Figura 9	– Tabela com Índice de Massa Corporal de adolescentes dos sexos feminino e masculino, retirada do site (EFDEPORTES.COM, 2010). . . . .	69
Figura 10	– tabela com todos os resultados possíveis da soma dos números que estão nas faces superiores dos dados. . . . .	76
Figura 11	– tabela com todos os resultados possíveis da soma que aparecem na tabela 10, com a frequência com que elas aparecem . . . . .	76
Figura 12	– Gráfico de dispersão relacionando a soma dos valores das faces superiores de dois dados e a frequência com que elas aparecem. . . . .	77
Figura 13	– Frequência das somas dos valores obtidos ao lançar três dados . . . . .	78
Figura 14	– Gráfico de dispersão relacionando a soma dos valores das faces superiores de três dados e a frequência com que elas aparecem . . . . .	78
Figura 15	– Tabela e gráfico de linha relacionando a soma dos valores das faces superiores de quatro dados e a frequência com que elas aparecem . . . . .	79

Figura 16 – Representação gráfica de uma distribuição normal com média 14 e desvio-padrão 0,77 . . . . .	79
Figura 17 – Representação gráfica de uma distribuição normal para o estudo de suas características . . . . .	80
Figura 18 – Representação gráfica de três curvas normais com mesmas médias e desvios-padrão diferentes: 0,25, 0,2 e 0,15 . . . . .	81
Figura 19 – Representação gráfica de três curvas normais com mesmos desvios-padrão e médias diferentes: -3, 0 e 3 . . . . .	81
Figura 20 – A área entre a curva e a reta $x$ , delimitada pelas retas $x = a$ e $x = b$ , é igual a $\mathbb{P}(a \leq X \leq b)$ , em que $X$ equivale a variável que descreve a característica de interesse. . . . .	82
Figura 21 – Representação gráfica para o cálculo da probabilidade pedida no exemplo . . . . .	83
Figura 22 – Representação gráfica da probabilidade com que um jogador de futebol profissional tenha menos que 80 kg . . . . .	83
Figura 23 – Representação gráfica das distribuições de $X$ e $Z$ , em que $X$ tem média 74 kg e desvio-padrão 4 kg e $Z$ tem média 0 e desvio-padrão 1 . . . . .	84
Figura 24 – Uso da tabela normal para o encontro da $\mathbb{P}(Z < 1.5)$ . . . . .	85
Figura 25 – Uso da tabela normal para o encontro da $\mathbb{P}(Z < -1.5)$ . . . . .	86



# LISTA DE TABELAS

---

---

Tabela 1 – Os valores apresentados na tabela correspondem à soma dos valores ( $i$ e $j$ ) das faces superiores de dois dados, obtidos após o lançamento de ambos . . .	32
Tabela 2 – A tabela apresenta os valores assumidos pela variável aleatória $X$ , que assume os numeros de faces caras em 3 lançamentos de uma moeda, e suas respectivas probabilidades com que $X$ assume os seus valores . . . . .	34
Tabela 3 – Distribuição conjunta de $X$ e $Y$ , definidas no exemplo 2.3.1 . . . . .	44
Tabela 4 – Distribuição conjunta de $X$ e $Y$ . . . . .	46
Tabela 5 – Tipos de erros em um teste de hipóteses . . . . .	68



# SUMÁRIO

---

---

1	INTRODUÇÃO . . . . .	19
1.1	Estatística e BNCC . . . . .	20
1.2	Estatística no Currículo do Estado de São Paulo . . . . .	21
1.3	Contribuições para este trabalho . . . . .	22
1.4	Descrição do Trabalho . . . . .	24
2	EM BUSCA DA NORMAL - PARTE 1: UMA BREVE INTRODUÇÃO À PROBABILIDADE E VARIÁVEIS ALEATÓRIAS . . . . .	27
2.1	Probabilidade . . . . .	27
2.1.1	<i>Experimento aleatório, espaço amostral e eventos</i> . . . . .	27
2.1.2	<i>Relações entre Eventos</i> . . . . .	28
2.1.3	<i>Definição Axiomática de Probabilidade</i> . . . . .	28
2.1.4	<i>Abordagem Clássica de Probabilidade</i> . . . . .	29
2.1.5	<i>Probabilidade condicional e eventos Independentes</i> . . . . .	30
2.2	Variável Aleatória . . . . .	31
2.2.1	<i>Esperança de uma variável aleatória</i> . . . . .	33
2.2.2	<i>Variância e Desvio Padrão de uma variável aleatória</i> . . . . .	36
2.2.3	<i>Variável aleatória discreta com distribuição de probabilidade de Bernoulli</i> . . . . .	39
2.2.4	<i>Variável Aleatória com Distribuição de Probabilidade Binomial</i> . . . . .	40
2.2.5	<i>Função de Distribuição</i> . . . . .	41
2.3	Distribuição Conjunta de Probabilidade para duas variáveis aleatórias discretas . . . . .	43
2.3.1	<i>Distribuições Marginais de duas variáveis aleatórias discretas</i> . . . . .	44
2.3.2	<i>Variáveis Aleatórias Independentes</i> . . . . .	45
2.3.3	<i>Esperança e Variância de uma função de duas ou mais variáveis aleatórias discretas</i> . . . . .	47
2.3.4	<i>Esperança e Variância da Variável Aleatória Binomial</i> . . . . .	49
3	EM BUSCA DA NORMAL - PARTE 2: UMA BREVE INTRODUÇÃO À ESTATÍSTICA . . . . .	51
3.1	Variável Aleatória Contínua . . . . .	51
3.1.1	<i>Esperança e Variância de uma variável aleatória contínua</i> . . . . .	52

3.1.2	<i>Variável Aleatória com Distribuição Normal</i>	52
3.2	Inferência Estatística	57
3.2.1	<i>Estimação Pontual</i>	57
3.2.2	<i>Teorema Central do Limite</i>	59
3.2.3	<i>Estimação Intervalar</i>	61
3.2.3.1	<i>Intervalos de Confiança para média populacional</i>	61
3.2.3.2	<i>Intervalo de confiança para <math>p</math></i>	63
3.3	Teste de Hipóteses	67
3.3.1	<i>Teste de Hipóteses para a média, com variância conhecida</i>	67
3.3.2	<i>Teste de Hipóteses para a proporção</i>	70
4	ENCONTRANDO A NORMAL NA SALA DE AULA: EXEMPLO DE ROTEIRO DE AULA	73
4.1	Roteiro de Ensino	73
4.2	Introduzindo a Distribuição Normal	77
4.2.1	<i>Características de uma Distribuição Normal</i>	80
4.2.2	<i>Calculando Probabilidades</i>	81
5	CONSIDERAÇÕES FINAIS	87
	REFERÊNCIAS	89
	APÊNDICE A DEMONSTRAÇÕES	91

---

## INTRODUÇÃO

---

A curva normal aparece em várias áreas da ciência, e modela, por exemplo, a altura de homens, mulheres ou jovens, a pressão sanguínea e a frequência cardíaca de esportistas, o peso de bebês recém-nascidos. Em todos esses exemplos o interesse é saber com que probabilidade a característica de interesse (altura, pressão, frequência, peso) assume valores em um certo intervalo. Por exemplo, qual a probabilidade de um bebê recém nascido pesar entre 1 e 2 kg.

Na área de Probabilidade, como veremos no capítulo 3, a curva normal corresponde a uma densidade de probabilidade de uma variável aleatória. Nessa área usa-se a expressão "distribuição normal de probabilidade", ou só, "distribuição normal" para se referir à curva normal. Também usaremos esses termos nesse trabalho.

O gráfico que descreve uma distribuição de probabilidade normal tem a forma de sino e também recebe o nome de Curva Gaussiana em homenagem ao matemático Carl Friederich Gauss (1777 – 1855), pois foi ele, juntamente com Robert Adrian (1775 – 1843), em 1808, segundo (BITTENCOURT; VIALI, 2006), que desenvolveram a equação da distribuição normal:

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

A distribuição normal também é muito importante na área de Inferência Estatística, que, em palavras, analisa características de uma população a partir de uma amostra. Assim, se estivermos interessados em saber qual será o político mais votado numa eleição para presidente, não precisamos e nem é viável que entrevistemos todos os eleitores.

Um dos resultados mais importantes para a Inferência Estatística é o Teorema Central do Limite. Em palavras, esse teorema nos diz que a média dos valores coletados da amostra poderá ser modelada por uma curva normal, quando o tamanho da amostra for suficientemente grande. Como veremos, permite que sejam testados os resultados provenientes da amostragem.

Conhecendo um pouco mais sobre a Distribuição de probabilidade normal e toda a sua importância dentro da Estatística, resolvi escrever esse trabalho que tem como objetivo me

aprofundar no assunto e colaborar para a melhoria do ensino da Estatística.

## 1.1 Estatística e BNCC

A Base Nacional Comum Curricular (BNCC) (MEC, 2017) é um documento federal estruturado para definir quais competências e habilidades serão desenvolvidas ao longo de toda a educação básica, de modo a garantir os direitos de aprendizagem, superar as desigualdades e o desenvolvimento integral de todos os estudantes. Tal documento, organiza os conteúdos a serem desenvolvidos na Educação Básica na União, nos Estados e Municípios. A fim de se nortear o trabalho docente, e melhorar a qualidade da Educação nas escolas públicas e privadas, este documento define competências e habilidades pautadas no desenvolvimento global de cada estudante.

Em relação aos estudantes do Ensino Médio, busca-se oferecer a eles uma escola cuja formação permite que eles definam os seus projetos de vida. Para isso, a escola precisa garantir a eles a leitura da realidade, a preparação básica para o trabalho, o aprimoramento do educando como pessoa humana, a busca de sentido das aprendizagens, autonomia intelectual, pensamento crítico, protagonismo em suas aprendizagens, desenvolvimento da abstração, da capacidade de trabalhar em equipe, promoção do diálogo não violento para solucionar conflitos, etc.

Com a homologação da BNCC, o Ensino Médio vem sofrendo mudanças, de modo a garantir a formação integral do estudante, aumentando para 1800 aulas para incorporar uma parte flexível, chamada de Itinerários Formativos. O termo flexível significa que tais itinerários serão escolhidos pelos estudantes, que buscarão as áreas e os cursos de formação técnica que se identificam.

Em relação à Matemática, a BNCC espera que os estudantes do Ensino Médio se aprofundem nos conhecimentos adquiridos no Ensino Fundamental. "Ganha importância a competência de comunicar; não só por meio de símbolos matemáticos, como também por meio da língua materna"(MEC, 2017). A comunicação mencionada na BNCC está relacionada à capacidade de argumentar, apresentar justificativa.

São 5 as competências gerais que devem ser desenvolvidas nos estudantes de Ensino Médio, na área de Matemática e suas Tecnologias. Cada competência engloba várias habilidades. As habilidades envolvendo os assuntos relacionados a Estatística são citados em todas as competências. Por exemplo:

### Competência 3

Utilizar estratégias, conceitos e procedimentos matemáticos, em seus campos - Aritmética, Álgebra, Grandezas e medidas, Geometria, Probabilidade e Estatística - para interpretar, construir modelos e resolver problemas em diversos contextos, analisando a plausibilidade dos re-

sultados e a adequação das soluções propostas, de modo a construir argumentação consistente. (MEC, 2017)

Tal competência engloba a seguinte habilidade:

"(EM13MAT316) Resolver e elaborar problemas, em diferentes contextos que envolvem cálculo e interpretação das medidas de tendência central (média, moda e mediana) e das de dispersão (amplitude, variância e desvio padrão)"(MEC, 2017, p.528).

O código EM13MAT316, significa que a habilidade é desenvolvida da 1<sup>a</sup> a 3<sup>a</sup> série do Ensino Médio, no componente curricular Matemática, sendo a 16<sup>a</sup> habilidade referente à 3<sup>a</sup> competência.

O presente trabalho está alinhado com a BNCC. Uma das habilidades propostas nela e desenvolvidas no trabalho recomenda o uso da tecnologia para a realização de cálculos e construção de gráficos estatísticos.

"(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"(MEC, 2017).

Tal habilidade faz parte da seguinte competência:

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. (MEC, 2017)

## 1.2 Estatística no Currículo do Estado de São Paulo

No Estado de São Paulo, os conteúdos estão organizados até o nono ano do Ensino Fundamental, no Currículo Paulista (SEE, 08/2019). O currículo do Ensino Médio está sendo modificado gradativamente. No ano de 2021, a primeira série do Ensino Médio já faz parte do chamado Novo Ensino Médio. No ano de 2023, pretende-se que todas as três séries já façam parte dessa nova modalidade.

Em relação aos conteúdos que são pré-requisitos para o entendimento da Distribuição Normal, Probabilidade e Estatística aparecem em todos os anos do Ensino Fundamental. No 6<sup>o</sup> ano, trabalham-se os seguintes assuntos: leitura e interpretação de tabelas e construção de gráficos, probabilidade como a razão entre o número de casos favoráveis e o número de casos possíveis e por meio de repetição. No 7<sup>o</sup> ano desenvolve-se os assuntos: experimentos aleatórios, planejamento de pesquisa, gráfico de setores, média e amplitude dos dados. No 8<sup>o</sup> os alunos aprendem que a soma das probabilidades de todos os elementos de um espaço amostral é igual a 1. Desenvolverão a habilidade de fazer um planejamento de pesquisa, organizar os dados de uma

variável contínua em classes e saber qual o tipo de gráfico deve ser usado para representá-los. Trabalha-se também, as medidas de tendência central e dispersão. No 9º ano são desenvolvidos os conteúdos: leitura, interpretação e representação de dados de pesquisa em tabelas e gráficos (colunas simples, agrupadas, gráfico de barras, de setores e pictóricos), análise de gráficos divulgados pela mídia e elementos que podem induzir a erros de leitura, escrita de relatório e cálculo de probabilidades de ocorrer eventos dependentes e independentes.

Ao terminar o Ensino Fundamental, espera-se que os alunos já estejam familiarizados com aspectos básicos da Probabilidade e Estatística, para que os conteúdos trabalhados no Ensino Médio aconteçam de forma mais significativa.

Até o ano de 2019, no Ensino Médio, os assuntos que embasam o processo de ensino-aprendizagem de Probabilidade e Estatística começavam a serem desenvolvidos no 3º bimestre da 2ª série, sendo eles: número de possibilidades de ocorrência de um evento usando os princípios aditivos e multiplicativos, arranjos, combinações, cálculo de probabilidades em diferentes situações-problema, Probabilidades Condicionais, Binômio de Newton e Distribuição Binomial de Probabilidade, (COPED, 2020).

A Distribuição Normal de Probabilidade era desenvolvida no 4º bimestre da 3ª série do Ensino Médio, iniciando com interpretação de tabelas e gráficos a partir de dados obtidos por amostras, medidas de tendência central (média, mediana e moda), medidas de dispersão (desvio padrão). Por meio da curva normal, é desenvolvida a habilidade de "reconhecer as características de conjuntos de dados distribuídos normalmente, utilizar a curva normal em estimativas pontuais e intervalares."(COPED, 2020, p. 89)

Analisando os livros didáticos mais atuais, tais como, (CEVADA, 2020), (BONJORNO, 2020) e (ANDRADE, 2020), (SMOLE, 2020) que estão colocados para consulta dos professores de Ensino Básico para o PNLD (Programa Nacional do Livro e do Material Didático), pode-se notar que somente o livro (SMOLE, 2020), mostra a Distribuição Normal. Os demais textos consultados desenvolvem as ideias de Estatística Descritiva e um pouco sobre Estatística Inferencial, definindo amostra, técnicas de amostragem e população.

A quantidade pequena de livros didáticos de Ensino Médio que abordam a Distribuição Normal, foi a motivação para desenvolver este trabalho.

### 1.3 Contribuições para este trabalho

Nesta seção será feita uma breve explanação do que alguns autores já publicaram em suas dissertações para o Profmat sobre a Curva Normal e metodologia baseada em resolução de problemas. As ideias desenvolvidas nelas contribuíram para a escrita deste trabalho.

A autora Cássia Cristina Marcomini, no desenvolvimento da sua dissertação intitulada como A Matemática, a Estatística e o Corte e Costura (MARCUMINI, 2020), desenvolveu uma



atividade prática relacionando Estatística e Geometria, que tem como produto final uma saia godê confeccionada pelos alunos de nono ano, da escola em que ela é docente. Os conceitos, tanto de Geometria como de Estatística, envolvidos são: comprimento da circunferência, área do círculo, diâmetro, população, amostra, média, desvio-padrão e gráficos.

Para decidir sobre o tamanho das saias, mediu-se a cintura das alunas e, a partir desses valores, construíram um histograma junto com a curva normal. Os tamanhos PP, P, M, G e GG foram estipulados, a partir da média, conforme a figura 1.

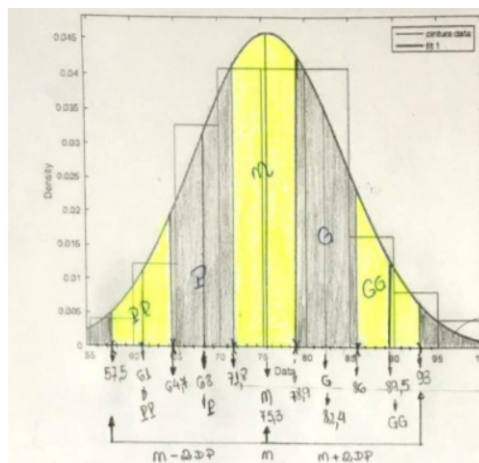


Figura 1 – Curva Normal tomada para estipular os tamanhos das saias (PP, P, M, G ou GG) confeccionadas por alunos de nonos anos

O autor Maurício Gonçalo de Carvalho (CARVALHO, 2018) usou a planilha de Cálculos EXCEL para simular o lançamento de uma moeda 50 vezes e calculou as probabilidades com que a face cara seja voltada para cima 0, 1, 2, ..., 50 vezes. Repetiu esse procedimento 200 e 1000 vezes, fez a representação gráfica e mostrou a aproximação de uma distribuição binomial para a normal.

Assim como Maurício (CARVALHO, 2018), João Luís Rodrigues Freire (FREIRE, 2017), abordou a curva normal por meio de simulações. No seu trabalho, Introdução à Inferência Estatística, o autor propõe uma sequência didática, a partir do lançamento de dois dados e discute o comportamento da variável que assume o valor da soma dos números das faces superiores de dois dados.

A sequência didática utilizada por (FREIRE, 2017), a partir do lançamento de dois dados foi o ponto de partida para o roteiro de ensino descrito no Capítulo 4. Foram seguidas todas as etapas do ensino através de resolução de problemas, propostas por (BINOTTO, 2019) que são:

- Preparação do problema;
- Leitura do problema individual e em grupo;
- Resolução do problema (em grupo);
- Professor: observa e incentiva;

- Registro das resoluções na lousa;
- Plenária;
- Busca de Consenso;
- Formalização do Conteúdo.

A resolução de problemas faz parte do que chamamos de Metodologias Ativas. O objetivo dessas metodologias é modificar o modo como os alunos aprendem, levando-os a participar mais ativamente das aulas. Segundo (EDUCAÇÃO, 2021), a implantação de metodologias ativas no ensino traz benefícios para os alunos, tais como:

- Melhoria na qualidade do aprendizado;
- Aumento da autonomia;
- Aprofundamento do censo crítico;
- Maior colaboração com os colegas;
- Desenvolvimento do censo de responsabilidade;
- Compreensão da importância da participação na sociedade

Ainda, segundo (EDUCAÇÃO, 2021), quando existem discussões e debates, o aluno absorve mais que 50% do que é ensinado. No roteiro de aula sugerido neste trabalho, essas discussões e debates acontecem quando os alunos estão buscando a resolução do problema nos grupos e quando são convidados a fazerem os registros na lousa, a partir das discussões.

## 1.4 Descrição do Trabalho

A fundamentação teórica dos assuntos presentes no trabalho está no Capítulo 2, que está dividido em seções. A seção 2.1 introduz os conceitos básicos de **Probabilidade**. Inicialmente são apresentadas as definições de experimento aleatório, espaço amostral e eventos. São dois os tipos de abordagem apresentados no capítulo sobre a Probabilidade: a axiomática e a clássica. A definição axiomática apresenta a Probabilidade como uma função e a clássica é a fórmula que geralmente é usada no Ensino Básico: razão entre número de casos favoráveis e o número de casos possíveis.

O assunto desenvolvido na seção 2.2 é **Variáveis Aleatórias**, que são divididas em discretas e contínuas. Inicialmente, são discutidos conceitos voltados para variáveis aleatórias discretas, tais como: distribuição de probabilidade, esperança, variância, desvio padrão e função de distribuição acumulada. Serão 2 os tipos de variáveis aleatórias discretas abordados no texto: a variável aleatória de Bernoulli e a Binomial.

Na seção 2.3 é abordada a distribuição de probabilidade conjunta para 2 variáveis aleatórias. Os conceitos são estendidos, dentro das limitações possíveis, para  $n$  variáveis aleatórias.

A Distribuição Normal será estudada no capítulo 3, usando a idéia de probabilidade como a área sob a curva de uma função. A **variável aleatória contínua com distribuição normal**, ganha um destaque nesse capítulo, sendo pré-requisito para o estudo da **Inferência Estatística**, assunto tratado na subseção 3.1.1.

Por meio de dados obtidos de uma amostra, é possível concluir, com grande probabilidade de acerto, o que ocorre com a população, já que o estudo da população demanda tempo e altos custos. Assuntos como inferência pontual, intervalar (muito usada pela mídia ao mostrar resultados de uma pesquisa eleitoral) e testes de hipóteses são desenvolvidos na seção 3.2. O Teorema que embasa a Inferência Estatística é chamado de **Teorema Central do Limite**, e fala da convergência de distribuição de probabilidade da soma de variáveis aleatórias para a distribuição de probabilidade de uma variável aleatória normal.

No capítulo 4, é apresentado um Roteiro de Ensino que segue todas as etapas da Resolução de Problemas, descritas no trabalho de (BINOTTO, 2019). Como ponto de partida, é sugerida uma avaliação diagnóstica que norteará o trabalho do professor, fazendo com ele verifique quais são os alunos que mais necessitam de uma atenção especial ao longo da atividade proposta.



---

# EM BUSCA DA NORMAL - PARTE 1: UMA BREVE INTRODUÇÃO À PROBABILIDADE E VARIÁVEIS ALEATÓRIAS

---

Os assuntos trabalhados neste capítulo são pré-requisitos para a compreensão da Distribuição Normal de Probabilidades. Inicialmente vamos abordar conceitos teóricos básicos relacionados à Probabilidade, e seguindo uma trilha, passando por variáveis aleatórias discretas, dando ênfase em Distribuição de Probabilidade de Bernoulli, bem como a Distribuição Binomial com suas esperanças e variâncias, busca-se chegar na Distribuição Normal de Probabilidades e suas aplicações, com os conceitos básicos já apreendidos.

Ao escrever os capítulos 2 e 3, pensando na atualização dos professores do Ensino Médio em assuntos referentes à Probabilidade e Estatística, pretende-se fornecer uma base teórica, para que estes possam aplicá-la em sala de aula, por meio do exemplo proposto no capítulo 4, ou ainda, criar suas próprias aulas.

## 2.1 Probabilidade

### 2.1.1 *Experimento aleatório, espaço amostral e eventos*

Experimentos aleatórios são mecanismos que geram resultados de maneira aleatória. O lançamento de um dado, por exemplo, é um experimento aleatório.

O conjunto de todos os resultados de um experimento aleatório recebe o nome de **espaço amostral** e é representado pela letra grega  $\Omega$ . Um elemento do espaço amostral é representado pela letra minúscula  $\omega$ . Um subconjunto do espaço amostral recebe o nome de **evento**.

**Exemplo 2.1.1.** No experimento aleatório, lançamento de um dado, o espaço amostral é dado por:  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .

Se considerarmos o evento:  $A = \text{"ocorrer um número par"}$ , temos:

$$A = \{2, 4, 6\}$$

### 2.1.2 Relações entre Eventos

Vamos relembrar algumas operações com conjuntos já os considerando como eventos, ou seja, subconjuntos do espaço amostral.

Sejam  $A$  e  $B$  dois eventos de um espaço amostral  $\Omega$ . Então:

- $A \cup B$ : evento  $A$  ou evento  $B$  ocorrem, ou ainda ambos. Assim:

$$A \cup B = \{\omega \in \Omega : \omega \in A \text{ ou } \omega \in B\}$$

- $A \cap B$ : eventos  $A$  e  $B$  ocorrem de forma simultânea.

$$A \cap B = \{\omega \in \Omega : \omega \in A \text{ e } \omega \in B\}$$

Se  $A \cap B = \emptyset$ , então  $A$  e  $B$  são chamados de eventos mutuamente excludentes ou disjuntos.

- O evento complementar de um conjunto  $A$ , é dado por:

$$A^c = \{\omega \in \Omega : \omega \notin A\}$$

Em palavras, se o evento  $A^c$  ocorre, então não ocorre o evento  $A$ .

### 2.1.3 Definição Axiomática de Probabilidade

O estudo da Teoria das Probabilidades começou no sec. *XVI* e tem suas raízes em jogos e apostas. Jerônimo Cardano (1501 – 1576) foi o primeiro a fazer observações sobre probabilidades a partir do lançamento de um dado honesto.

Segundo ([WIKIPÉDIA, 2019](#)), foi a partir dos Axiomas de Kolmogorov (1903 – 1987) que se iniciou um estudo mais rigoroso da teoria da Probabilidade.

Nesse contexto, vamos apresentar a definição rigorosa da probabilidade, conhecida como Definição Axiomática da Probabilidade.

**Definição 2.1.1.** A probabilidade é uma função que possui, como domínio, o conjunto formado por todos os eventos de  $\Omega$  e contradomínio o intervalo  $[0, 1]$ , satisfazendo os seguintes axiomas:

1.  $\mathbb{P}(A) \geq 0, \forall A \subset \Omega$
2.  $\mathbb{P}(\Omega) = 1$

3. Se  $A_1, A_2, \dots$  forem eventos disjuntos de  $\Omega$ , então:

$$\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

### 2.1.4 Abordagem Clássica de Probabilidade

O primeiro contato que temos com a probabilidade, no ensino médio, como alunos, vem muito associado às técnicas de combinatória. Por isso, a probabilidade ficou muito relacionada ao número de elementos de um conjunto. Para diferenciar da definição axiomática, essa "definição" passou a ser chamada de "definição" clássica da probabilidade. Neste trabalho usaremos o termo abordagem, pois não é de fato uma definição.

Dado  $\Omega$  o espaço amostral, com uma quantidade finita de elementos, e  $A$  um evento de  $\Omega$ , a probabilidade de ocorrer  $A$  é dada por:

$$\mathbb{P}(A) = \frac{n(A)}{n(\Omega)}, \quad (2.1.1)$$

onde  $n(A)$  é o número de elementos de  $A$  e  $n(\Omega)$  é o número de elementos de  $\Omega$ .

**Exemplo 2.1.2.** No lançamento de um dado honesto, vamos calcular a probabilidade de sair um número maior ou igual a 5. Vamos chamar de  $A$  o evento "sair resultado maior ou igual a 5". Assim, temos que:

$\Omega = \{1, 2, 3, 4, 5, 6\}$  e  $A = \{5, 6\}$  Então:

$$\mathbb{P}(A) = \frac{n(A)}{n(\Omega)} = \frac{2}{6} = \frac{1}{3}$$

Vamos mostrar que a razão 2.1.1 obedece os axiomas definidos na subseção 2.1.3. A demonstração se baseia no fato do número de elementos de um conjunto ser um número maior ou igual a zero, e também atuar da mesma forma, que a probabilidade, nos conjuntos disjuntos.

1.  $\mathbb{P}(\Omega) = \frac{n(\Omega)}{n(\Omega)} = 1$  com  $n(\Omega) > 0$
2.  $0 \leq \mathbb{P}(A) \leq 1$ , visto que:  $0 \leq n(A) \leq n(\Omega)$ .

Dividindo todos os membros da desigualdade anterior por  $n(\Omega)$ , obteremos:

$$\frac{0}{n(\Omega)} \leq \frac{n(A)}{n(\Omega)} \leq \frac{n(\Omega)}{n(\Omega)}, n(\Omega) \neq 0 \Rightarrow 0 \leq \mathbb{P}(A) \leq 1$$

3. Dados dois eventos disjuntos  $A$  e  $B$ , então:

$$\mathbb{P}(A \cup B) = \frac{n(A \cup B)}{n(\Omega)} = \frac{n(A) + n(B)}{n(\Omega)} = \frac{n(A)}{n(\Omega)} + \frac{n(B)}{n(\Omega)},$$

com  $n(\Omega) > 0$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$$

### 2.1.5 Probabilidade condicional e eventos Independentes

No lançamento de dois dados honestos, jogados sequencialmente, vamos considerar o evento "o resultado do lançamento do segundo dado é múltiplo ou divisor do resultado do lançamento do primeiro dado". Vamos chamar esse evento de  $A$ . Por exemplo: sai face 6 em um dado e face 2 no outro.

O espaço amostral ( $\Omega$ ) é formado por todos os pares ordenados  $(i, j)$ , tal que  $i = 1, \dots, 6$  e  $j = 1, \dots, 6$ . O número de elementos de  $\Omega$  é 36. E o evento  $A$  pode ser escrito como:

$$A = \{(1, 1), \dots, (1, 6), (2, 1), (2, 2), (2, 4), (2, 6), (3, 1), \dots, (5, 1), (5, 5), (6, 1), (6, 2), (6, 3), (6, 6)\}.$$

Como os dois dados são honestos, temos que:

$$\mathbb{P}(A) = \frac{22}{36}.$$

Supondo que saibamos que saiu 2 no lançamento do primeiro dado, a probabilidade do evento  $A$  fica alterada, pois houve a adição de nova informação. Seja  $B$  o evento "sair 2 no lançamento do primeiro dado".

$$B = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)\}$$

O cálculo da probabilidade de  $A$ , agora, pode ser feito tomando como novo espaço amostral, o conjunto  $B$ . Como em  $B$ , os pares ordenados em que uma componente é múltipla ou divisor da outra são:  $(2, 1), (2, 2), (2, 4), (2, 6)$ , a probabilidade de ocorrer o evento  $A$  é  $\frac{4}{6}$ , sendo 6 o número de elementos de  $B$ .

Vamos agora mostrar uma outra forma de se calcular a probabilidade de  $A$  quando temos uma informação adicional. A probabilidade  $\mathbb{P}(A|B)$ , chamada probabilidade condicional de  $A$  dado  $B$ , é definida como:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}, \mathbb{P}(B) > 0 \quad (2.1.2)$$

Retomando o exemplo, vamos calcular, agora,  $\mathbb{P}(A|B)$ , usando 2.1.2. Primeiro vamos calcular a probabilidade do numerador.

$$A \cap B = \{(2, 1), (2, 2), (2, 4), (2, 6)\} \Rightarrow \mathbb{P}(A \cap B) = \frac{4}{36}.$$

Como  $\mathbb{P}(B) = \frac{6}{36}$ , temos:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\frac{4}{36}}{\frac{6}{36}} = \frac{4}{6} = \frac{2}{3},$$

o mesmo resultado que encontramos anteriormente.

O próximo exemplo nos ajudará a apresentar o conceito de eventos independentes.



**Exemplo 2.1.3.** Uma urna contém 4 bolas pretas e 6 bolas brancas. Retiramos, ao acaso, a 1ª bola e colocamos novamente na urna. Depois retiramos a 2ª. Definimos os eventos:

$$A : \text{"a primeira bola retirada é branca"} \Rightarrow \mathbb{P}(A) = \frac{6}{10}.$$

$$B : \text{"a segunda bola retirada é preta"} \Rightarrow \mathbb{P}(B) = \frac{4}{10}$$

Como houve reposição, a probabilidade de que a segunda bola seja preta, dado que a primeira bola foi branca é dada por:

$$\mathbb{P}(B|A) = \mathbb{P}(B) = \frac{4}{10}$$

Ou seja, a ocorrência de  $A$  não modificou a probabilidade do evento  $B$  ocorrer.

Dois eventos  $A$  e  $B$  são independentes, se:

$$\mathbb{P}(B|A) = \mathbb{P}(B), \mathbb{P}(A) > 0, \quad (2.1.3)$$

ou seja, se a informação trazida por  $A$  não influencia a probabilidade de  $B$ .

De 2.1.3, e usando o fato de  $A$  e  $B$  serem independentes, temos que:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \Rightarrow \mathbb{P}(B) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} \Rightarrow \mathbb{P}(B \cap A) = \mathbb{P}(B) \cdot \mathbb{P}(A)$$

Ou seja, se dois eventos são independentes, a probabilidade da intersecção é igual ao produto de suas probabilidades.

## 2.2 Variável Aleatória

Uma variável aleatória é uma função que, a cada elemento de  $\Omega$ , atribui um valor real.

$$X : \Omega \longrightarrow \mathbb{R}$$

**Exemplo 2.2.1.** (BOSCO, 2010) Consideremos o experimento aleatório envolvendo o lançamento de dois dados honestos. Nesse caso, o espaço amostral é:

$$\Omega = \{(1, 1), (1, 2), \dots, (1, 6), (2, 1), (2, 2), \dots, (2, 6), \dots, (6, 1), \dots, (6, 6)\}$$

Por exemplo, vamos denotar um elemento do espaço amostral por  $\omega_{ij} = (i, j)$ , e definir a variável aleatória  $X(\omega_{ij}) = i + j$ . Por exemplo:  $\omega_{11} = (1, 1)$ , então  $X(\omega_{11}) = 2$ .

A tabela 1 apresenta todos os valores de  $X$ .

Como  $X$  pode assumir valores entre 2 e 12, então o conjunto imagem de  $X$  é:

$$I_x = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$	$i = 6$
$j = 1$	2	3	4	5	6	7
$j = 2$	3	4	5	6	7	8
$j = 3$	4	5	6	7	8	9
$j = 4$	5	6	7	8	9	10
$j = 5$	6	7	8	9	10	11
$j = 6$	7	8	9	10	11	12

Tabela 1 – Os valores apresentados na tabela correspondem à soma dos valores ( $i$  e  $j$ ) das faces superiores de dois dados, obtidos após o lançamento de ambos

Uma variável aleatória é discreta se o seu conjunto imagem é finito ou infinito enumerável.

Vamos tomar, agora, um evento  $\{X = x\}$  como sendo o conjunto de todos os elementos de  $\Omega$  que são levados no valor  $x$  por meio da função  $X$ . No exemplo anterior,

$$\{X = 3\} = \{\omega \in \Omega : X(\omega) = 3\} = \{(1, 2), (2, 1)\}$$

É natural nos perguntarmos com quais probabilidades uma variável aleatória  $X$ , assume seus valores. Voltando no exemplo anterior, sabemos que  $X$  assume os valores  $2, 3, 4, \dots, 12$ , portanto a sua distribuição de massa de probabilidade é formada pelas probabilidades com que  $X$  assume esses valores.

Usaremos a notação  $\mathbb{P}(X = x)$  para a probabilidade de ocorrer o evento  $\{X = x\}$ , sendo  $X$  o símbolo usado para a variável aleatória e  $x$  um valor específico assumido por ela. Então, a distribuição de probabilidade da variável aleatória  $X$ , do exemplo, é:

$$\mathbb{P}(X = 2) = \mathbb{P}(\{X = 2\}) = \mathbb{P}(\{\omega_{11}\}) = \frac{1}{36}$$

$$\mathbb{P}(X = 3) = \mathbb{P}(\{\omega_{12}, \omega_{21}\}) = \frac{1}{36} + \frac{1}{36} = \frac{2}{36}$$

$$\mathbb{P}(X = 4) = \mathbb{P}(\{\omega_{13}, \omega_{22}, \omega_{31}\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{3}{36}$$

$$\mathbb{P}(X = 5) = \mathbb{P}(\{\omega_{14}, \omega_{23}, \omega_{32}, \omega_{41}\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{4}{36}$$

$$\mathbb{P}(X = 6) = \mathbb{P}(\{\omega_{15}, \omega_{24}, \omega_{33}, \omega_{42}, \omega_{51}\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{5}{36}$$

$$\mathbb{P}(X = 7) = \mathbb{P}(\{\omega_{16}, \omega_{25}, \omega_{34}, \omega_{43}, \omega_{52}, \omega_{61}\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{6}{36}$$

$$\mathbb{P}(X = 8) = \mathbb{P}(\{\omega_{26}, \omega_{35}, \omega_{44}, \omega_{53}, \omega_{62}\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{5}{36}$$

$$\mathbb{P}(X = 9) = \mathbb{P}(\{\omega_{36}, \omega_{45}, \omega_{54}, \omega_{63}\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{4}{36}$$

$$\mathbb{P}(X = 10) = \mathbb{P}(\{\omega_{46}, \omega_{55}, \omega_{64}\}) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{3}{36}$$

$$\mathbb{P}(X = 11) = \mathbb{P}(\{\omega_{56}, \omega_{65}\}) = \frac{1}{36} + \frac{1}{36} = \frac{2}{36}$$

$$\mathbb{P}(X = 12) = \mathbb{P}(\{\omega_{66}\}) = \frac{1}{36}$$

### 2.2.1 Esperança de uma variável aleatória

Embora nem todos os professores do Ensino Básico façam o cálculo da nota bimestral da mesma forma, vários ainda usam o cálculo da Média Aritmética para essa finalidade. Supondo que um professor aplique 3 avaliações no bimestre, todas com o mesmo peso, a média ( $M$ ) é calculada assim:

$$M = \frac{N_1 + N_2 + N_3}{3} = N_1 \cdot \frac{1}{3} + N_2 \cdot \frac{1}{3} + N_3 \cdot \frac{1}{3}$$

em que  $N_1$ ,  $N_2$  e  $N_3$  são as notas das avaliações 1, 2 e 3. Nos cursos de graduação, alguns professores usam pesos para cada avaliação. Tomemos um exemplo em que é atribuído peso 1 para  $N_1$ , 2 para  $N_2$  e 3 para  $N_3$ . O cálculo da Média Ponderada é feito assim:

$$M = \frac{1 \cdot N_1 + 2 \cdot N_2 + 3 \cdot N_3}{6} = N_1 \cdot \frac{1}{6} + N_2 \cdot \frac{2}{6} + N_3 \cdot \frac{3}{6} \quad (2.2.1)$$

Usamos esses exemplos para dar, ao leitor, alguma intuição sobre o conceito de Esperança de uma variável aleatória  $X$ . A esperança de  $X$  é a média ponderada de  $X$ .

**Definição 2.2.1.** Dada uma variável aleatória  $X$ , e sua distribuição de probabilidade  $P(X = x)$ , definimos a esperança (ou valor esperado, ou valor médio) de  $X$ , como:

$$\mathbb{E}(X) = \sum_x x \cdot \mathbb{P}(X = x) \quad (2.2.2)$$

**Exemplo 2.2.2.** Considerando 3 lançamentos de uma moeda, vamos definir uma variável aleatória  $X$ , como sendo o número de caras nos lançamentos e calcular  $\mathbb{E}(X)$ .

Os valores assumidos por  $X$  são: 0, 1, 2 e 3. Para calcularmos a  $\mathbb{E}(X)$ , precisamos saber os valores das probabilidades com que  $X$  assume os seus valores. Vamos então observar a tabela abaixo. Representaremos a face cara por  $H$  e coroa por  $T$ .

"faces"	$X$	$\mathbb{P}(X = x)$
<i>HHH</i>	3	$\frac{1}{8}$
<i>HHT</i>	2	$\frac{1}{8}$
<i>HTH</i>	2	$\frac{1}{8}$
<i>THH</i>	2	$\frac{1}{8}$
<i>TTT</i>	0	$\frac{1}{8}$
<i>TTH</i>	1	$\frac{1}{8}$
<i>THT</i>	1	$\frac{1}{8}$
<i>HTT</i>	1	$\frac{1}{8}$

Tabela 2 – A tabela apresenta os valores assumidos pela variável aleatória  $X$ , que assume os números de faces caras em 3 lançamentos de uma moeda, e suas respectivas probabilidades com que  $X$  assume os seus valores

A esperança de  $X$  é dada por:

$$\mathbb{E}(X) = \sum_{i=0}^3 x_i \mathbb{P}(X = x_i) = 0 \cdot \mathbb{P}(X = 0) + 1 \cdot \mathbb{P}(X = 1) + 2 \cdot \mathbb{P}(X = 2) + 3 \cdot \mathbb{P}(X = 3) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8}$$

$$= \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = \frac{12}{8} = 1,5$$

**Propriedade da Esperança de  $X$ :**

Sejam  $a$  e  $b$  constantes,

$$\begin{aligned}\mathbb{E}(aX + b) &= \sum_x (ax + b) \cdot \mathbb{P}(X = x) = \sum_x (ax \cdot \mathbb{P}(X = x) + b \cdot \mathbb{P}(X = x)) \\ &= \sum_x ax \cdot \mathbb{P}(X = x) + \sum_x b \cdot \mathbb{P}(X = x) = a \sum_x x \cdot \mathbb{P}(X = x) + b \cdot \sum_x \mathbb{P}(X = x) = a\mathbb{E}(X) + b.\end{aligned}$$

A definição abaixo será necessária para o cálculo da variância de uma variável aleatória que será apresentado na sequência.

Vamos definir, agora, uma variável aleatória  $Y$ , tal que  $Y = g(X)$ . A esperança de  $Y$  é dada por:

$$\mathbb{E}(Y) = \sum_{\{x|y=g(x)\}} g(x) \cdot \mathbb{P}(X = x) \quad (2.2.3)$$

Segue um exemplo de aplicação da definição 2.2.3.

**Exemplo 2.2.3.** Considerando o espaço amostral

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

e a variável aleatória  $X$ , tal que:

$$X(1) = -1$$

$$X(2) = X(3) = 0$$

$$X(4) = X(5) = X(6) = 1$$

Vamos encontrar o conjunto imagem de  $Y = g(X) = X^2$  e sua esperança.

Como  $Y = X^2$ , e  $X$  assume os valores  $-1, 0$  e  $1$ , então:

$$X = -1 \Rightarrow Y = (-1)^2 = 1$$

$$X = 0 \Rightarrow Y = 0^2 = 0$$

$$X = 1 \Rightarrow Y = 1^2 = 1$$

Portanto, o conjunto imagem de  $Y$  é:

$$I_y = \{0, 1\}$$

Antes de calcularmos a esperança de  $Y$ , vamos representar o exemplo em uma figura.

A figura 2 mostra a variável aleatória  $X$  levando os elementos de  $\Omega$  aos números  $-1, 0$  e  $1$ . Vamos tomar então, como exemplo, os elementos  $4, 5$  e  $6$  de  $\Omega$ . Estes elementos são levados em  $1$ , por  $X$  que, por sua vez, é levado em  $1$ , pela variável aleatória  $Y$ . Outro elemento que

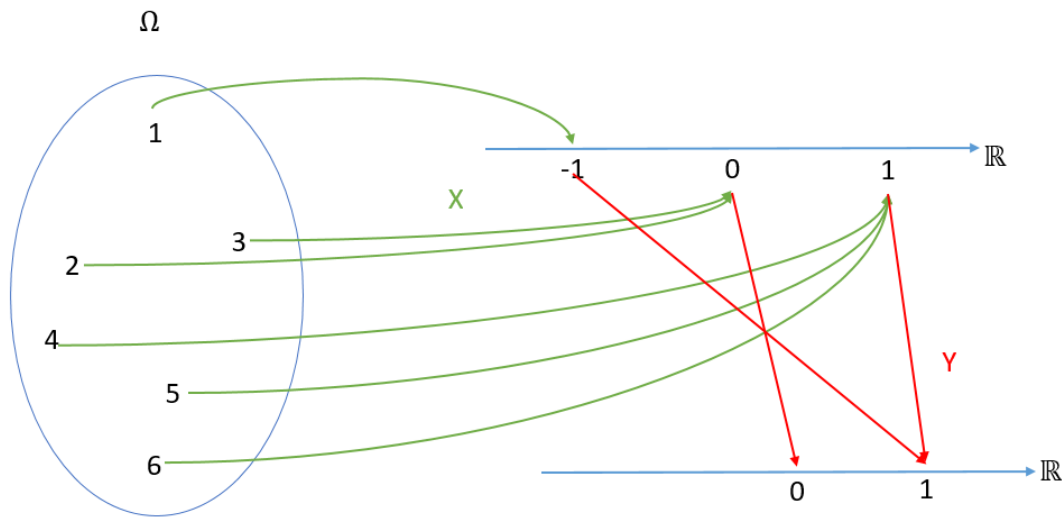


Figura 2 – Variável aleatória  $X$  levando os valores de  $\Omega$  até  $\mathbb{R}$  e  $Y$  levando os valores  $-1, 0$  e  $1$  em  $0$  e  $1$ .

também é levado ao valor  $1$  por  $Y$  é  $-1$ , que é imagem do elemento  $1$ , de  $\Omega$ . Vamos, então, calcular com que probabilidade  $Y$  assume o valor  $1$ .

$$\mathbb{P}(Y = 1) = \mathbb{P}(X = -1) + \mathbb{P}(X = 1) = \frac{1}{6} + \frac{3}{6} = \frac{4}{6}$$

De modo análogo, vamos calcular  $\mathbb{P}(Y = 0)$ .

$$\mathbb{P}(Y = 0) = \mathbb{P}(X = 0) = \frac{2}{6}$$

E agora a  $\mathbb{E}(Y)$  pode ser calculada, como segue:

$$\begin{aligned} \mathbb{E}(Y) &= \sum_y y \cdot \mathbb{P}(Y = y) = 0 \cdot \mathbb{P}(Y = 0) + 1 \cdot \mathbb{P}(Y = 1) \\ &= 0 \cdot \frac{2}{6} + 1 \cdot \frac{4}{6} = 0 \cdot \frac{2}{6} + 1 \cdot \left( \frac{1}{6} + \frac{3}{6} \right) = 0 \cdot \frac{2}{6} + 1 \cdot \frac{1}{6} + 1 \cdot \frac{3}{6} = 1 \cdot \frac{1}{6} + 0 \cdot \frac{2}{6} + 1 \cdot \frac{3}{6} \\ &= (-1)^2 \mathbb{P}(X = -1) + 0^2 \cdot \mathbb{P}(X = 0) + 1^2 \cdot \mathbb{P}(X = 1) \\ &= g(-1) \cdot \mathbb{P}(X = -1) + g(0) \cdot \mathbb{P}(X = 0) + g(1) \cdot \mathbb{P}(X = 1) \\ \mathbb{E}(Y) &= \sum_{\{x|y=g(x)\}} g(x) \cdot \mathbb{P}(X = x) \end{aligned}$$

## 2.2.2 Variância e Desvio Padrão de uma variável aleatória

Vamos usar um exemplo adaptado do livro de (ROSS, 2010), para termos uma noção intuitiva a respeito de como os valores assumidos pela variável aleatória se distribuem em torno da média.

**Exemplo 2.2.4.** Sejam  $Y$  e  $Z$ , variáveis aleatórias discretas e suas seguintes distribuições de probabilidade.

$$Y = \begin{cases} -1, & \text{com probabilidade } \frac{1}{2} \\ 1, & \text{com probabilidade } \frac{1}{2} \end{cases}$$

$$Z = \begin{cases} -100, & \text{com probabilidade } \frac{1}{2} \\ 100, & \text{com probabilidade } \frac{1}{2} \end{cases}$$

Vamos, agora, calcular  $\mathbb{E}(Y)$  e  $\mathbb{E}(Z)$ .

$$\mathbb{E}(Y) = \sum_y y \cdot \mathbb{P}(Y = y) = -1 \cdot \mathbb{P}(Y = -1) + 1 \cdot \mathbb{P}(Y = 1) = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = -\frac{1}{2} + \frac{1}{2} = 0$$

$$\begin{aligned} \mathbb{E}(Z) &= \sum_z z \cdot \mathbb{P}(Z = z) = -100 \cdot \mathbb{P}(Z = -100) + 100 \cdot \mathbb{P}(Z = 100) = -100 \cdot \frac{1}{2} + 100 \cdot \frac{1}{2} \\ &= -50 + 50 = 0 \end{aligned}$$

Vamos calcular a diferença entre os valores que  $Y$  e  $Z$  assumem e suas respectivas esperanças.

- $Y - \mathbb{E}(Y)$

$$-1 - 0 = -1$$

$$1 - 0 = 1$$

- $Z - \mathbb{E}(Z)$

$$-100 - 0 = -100$$

$$100 - 0 = 100$$

Para que tenhamos uma noção intuitiva de como os valores de  $Y$  e  $Z$  se distribuem em torno de suas médias, vamos calcular  $\mathbb{E}(Y - \mathbb{E}(Y))$  e  $\mathbb{E}(Z - \mathbb{E}(Z))$ .

- $\mathbb{E}(Y - \mathbb{E}(Y)) = \sum_y (y - \mathbb{E}(Y)) \mathbb{P}(Y = y) = -1 \cdot \mathbb{P}(Y = -1) + 1 \cdot \mathbb{P}(Y = 1) = -1 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = -\frac{1}{2} + \frac{1}{2} = 0.$

- $\mathbb{E}(Z - \mathbb{E}(Z)) = \sum_z (z - \mathbb{E}(Z)) \mathbb{P}(Z = z) = -100 \cdot \mathbb{P}(Z = -1) + 100 \cdot \mathbb{P}(Z = 1) = -100 \cdot \frac{1}{2} + 100 \cdot \frac{1}{2} = -50 + 50 = 0.$

Observando a situação acima, podemos verificar que os valores de  $Y$  não se distanciam tanto em relação à média (0), quanto os valores de  $Z$ . Mesmo assim, a esperança da distância entre os valores assumidos por  $Y$  e  $Z$  e suas respectivas médias resulta em 0. Logo, nada podemos

concluir a respeito da distribuição dos valores de  $Y$  e  $Z$  a partir dos cálculos realizados acima. O cálculo de  $|Y - \mathbb{E}(Y)|$  e  $|Z - \mathbb{E}(Z)|$ , nos ajudaria a ter uma melhor visualização de como os valores de  $Y$  e  $Z$  se distribuem em torno da média. Sendo a potência uma alternativa melhor tratável analiticamente, fazemos  $(Y - \mathbb{E}(Y))^2$  e  $(Z - \mathbb{E}(Z))^2$  e calculamos a esperança dos resultados, nos levando, naturalmente, a definição do que se chama de Variância de uma variável aleatória, que segue abaixo:

$$\text{VAR}(X) = \mathbb{E}[(X - \mathbb{E}(X))^2] \quad (2.2.4)$$

O problema do uso de 2.2.4 é que não podemos relacionar o seu resultado com a  $\mathbb{E}(X)$ , pois ambos estão em unidades diferentes. Enquanto  $X$  e  $\mathbb{E}(X)$ , por exemplo, são dadas em  $cm$ , a variância será dada em  $cm^2$ . A solução é extrair a raiz quadrada de  $\text{VAR}(X)$ , encontrando o desvio padrão, que representaremos pela letra  $\sigma$ . Então:

$$\sigma = \sqrt{\text{VAR}(X)}$$

Usando a propriedade da esperança de  $X$  e 2.2.3, desenvolvemos uma forma mais fácil de se calcular a  $\text{VAR}(X)$ :

$$\text{VAR}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (2.2.5)$$

A demonstração da fórmula alternativa para o cálculo da variância pode ser encontrada em A.

**Exemplo 2.2.5.** Voltando no exemplo 2.2.4, vamos calcular a  $\text{VAR}(Y)$  e  $\text{VAR}(Z)$ , usando A.0.2. Temos que, inicialmente, calcular  $\mathbb{E}(Y^2)$  e  $\mathbb{E}(Z^2)$ .

$$\mathbb{E}(Y^2) = \sum_y y^2 \cdot \mathbb{P}(Y = y) = (-1)^2 \cdot \mathbb{P}(y = -1) + 1^2 \cdot \mathbb{P}(y = 1) = \frac{1}{2} + \frac{1}{2} = 1$$

$$\begin{aligned} \mathbb{E}(Z^2) &= \sum_z z^2 \cdot \mathbb{P}(Z = z) = (-100)^2 \cdot \mathbb{P}(z = -100) + 100^2 \cdot \mathbb{P}(z = 100) \\ &= \frac{10000}{2} + \frac{10000}{2} = 5000 + 5000 = 10000 \end{aligned}$$

$$\mathbb{E}(Y) = \mathbb{E}(Z) = 0 \Rightarrow (\mathbb{E}(Y))^2 = (\mathbb{E}(Z))^2 = 0$$

Vamos, agora, calcular a  $\text{VAR}(Y)$  e  $\text{VAR}(Z)$ .

$$\text{VAR}(Y) = \mathbb{E}(Y^2) - (\mathbb{E}(Y))^2 = 1 - 0 = 1$$

$$\text{VAR}(Z) = \mathbb{E}(Z^2) - (\mathbb{E}(Z))^2 = 10000 - 0 = 10000$$

O desvio padrão de  $Y$  é  $\sqrt{\text{VAR}(Y)}$  e, portanto igual a 1, e o desvio padrão de  $Z$  é  $\sqrt{10000} = 100$ .



### Propriedade da variância de X

Sejam  $a$  e  $b$  "pertencentes a  $\mathbb{R}$ ". Então temos a seguinte propriedade da variância:

$$\begin{aligned} \text{VAR}(aX + b) &= \mathbb{E} \left[ (aX + b)^2 \right] - [\mathbb{E}(aX + b)]^2 = \mathbb{E} (a^2X^2 + 2aXb + b^2) - [a\mathbb{E}(X) + b]^2 \\ &= \mathbb{E}(a^2X^2) + 2ab\mathbb{E}(X) + \mathbb{E}(b^2) - [a^2\mathbb{E}(X)^2 + 2a\mathbb{E}(X)b + b^2] \\ &= a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - a^2\mathbb{E}(X)^2 - 2ab\mathbb{E}(X) - b^2 = a^2 [\mathbb{E}(X^2) - (\mathbb{E}(X))^2] = a^2\text{VAR}(X). \end{aligned}$$

Portanto:

$$\text{VAR}(aX + b) = a^2\text{VAR}(X) \quad (2.2.6)$$

### 2.2.3 Variável aleatória discreta com distribuição de probabilidade de Bernoulli

A variável aleatória  $X$  com distribuição de probabilidade de Bernoulli aparece em experimentos em que ocorrem apenas dois resultados, os quais chamamos de sucesso e fracasso. Desta forma,  $X$  assume apenas 2 valores. Geralmente usa-se 1 para representar sucesso, que ocorre com probabilidade  $p$ , e 0 para representar fracasso, que ocorre com probabilidade  $1 - p$ . Usamos a notação  $X \sim \text{Bernoulli}(p)$ , para dizer que  $X$  tem distribuição de Bernoulli com probabilidade  $p$  de sucesso.

A variável aleatória  $X$  com distribuição de Bernoulli é usada para modelar situações com duas respostas. Tais situações recebem o nome de Experimentos de Bernoulli. Por exemplo, numa pesquisa de opinião, em que um aluno pesquisado responde apenas se concorda ou não com uma determinada mudança na escola. Nesse caso, podemos atribuir a  $X$ , o valor 0 caso a resposta do aluno entrevistado seja "não concordo" e 1, se a resposta for "concordo".

A probabilidade com que  $X$  assume os valores 1 e 0, fica resumida na seguinte expressão:

$$\mathbb{P}(X = x) = p^x \cdot (1 - p)^{1-x}, \quad (2.2.7)$$

onde  $x$  assume 0 ou 1 apenas.

A esperança de  $X$  é, portanto:

$$\mathbb{E}(X) = \sum_x x \cdot \mathbb{P}(X = x) = 0 \cdot (1 - p) + 1 \cdot p = p$$

Para calcularmos a  $\text{VAR}(X)$ , vamos primeiro calcular a  $\mathbb{E}(X^2)$

$$\mathbb{E}(X^2) = \sum_x x^2 \cdot \mathbb{P}(X = x) = 0^2(1 - p) + 1^2 \cdot p = 0 \cdot (1 - p) + 1 \cdot p = p$$

Logo,

$$\text{VAR}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = p - p^2 = p \cdot (1 - p)$$

**Exemplo 2.2.6.** Um certo jogo se inicia somente se a face superior do dado for igual a 6.

Consideremos o lançamento de um dado honesto. Definiremos como  $X$  a variável aleatória que assume o valor 1 quando a face voltada para cima for 6 e 0 caso contrário. Nesse caso,  $X \sim \text{Bernoulli}\left(\frac{1}{6}\right)$ . Vamos calcular a esperança e a variância de  $X$

O espaço amostral é dado por:  $\Omega = \{1, 2, 3, 4, 5, 6\}$

Em termos de  $X$  nos interessam os seguintes eventos:

$$\{X = 1\} = \{\omega \in \Omega : X(\omega) = 1\} = \{6\} \Rightarrow \mathbb{P}(X = 1) = p = \mathbb{P}(\{6\}) = \frac{1}{6}$$

$$\{X = 0\} = \{\omega \in \Omega : X(\omega) = 0\} = \{1, 2, 3, 4, 5\} \Rightarrow \mathbb{P}(X = 0) = \mathbb{P}(\{1, 2, 3, 4, 5\}) = 1 - p = \frac{5}{6}$$

A esperança e a variância de  $X$  são dadas por:

$$\mathbb{E}(X) = p = \frac{1}{6}$$

$$\text{VAR}(X) = p(1 - p) = \frac{1}{6} \cdot \frac{5}{6} = \frac{5}{36}$$

## 2.2.4 Variável Aleatória com Distribuição de Probabilidade Binomial

Vamos partir da situação citada no início da subseção anterior e supor que pesquisemos 100 alunos para saber se eles concordam ou não com uma mudança proposta na escola. Definiremos, então, uma variável aleatória  $X_i$ , que assume o valor 1 ou 0 se a resposta for favorável ou não favorável, respectivamente. Dessa forma, estamos repetindo o experimento de Bernoulli 100 vezes. Definimos a variável aleatória  $Y$  que assume o número de respostas favoráveis. Essa variável aleatória  $Y$ , apresenta Distribuição de Probabilidade Binomial com parâmetros  $n$  e  $p$  e é representada por:

$$Y \sim \text{Binomial}(n, p)$$

A probabilidade com que  $Y$  assume  $k$  sucessos,  $\{k \in \mathbb{N} : 0 \leq k \leq n\}$  em  $n$  experimentos de Bernoulli, é dada por:

$$\mathbb{P}(Y = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \quad (2.2.8)$$

A esperança e a variância de uma variável aleatória com distribuição binomial de probabilidade são dadas por:

$$\mathbb{E}(Y) = np \quad (2.2.9)$$

e

$$\text{VAR}(Y) = np(1 - p) \quad (2.2.10)$$

A demonstração de 2.2.9 e 2.2.10 será feita na seção 2.3.4.

**Exemplo 2.2.7.** Definindo uma variável aleatória  $X$  que assume o número de faces cara quando uma moeda é lançada 100 vezes, vamos calcular a esperança e a variância de  $X$ , usando 2.2.9 e 2.2.10.

$$\mathbb{E}(X) = np = 100 \cdot \frac{1}{2} = 50$$

$$\text{VAR}(X) = np(1-p) = 100 \cdot \frac{1}{2} \cdot \frac{1}{2} = 25$$

### 2.2.5 Função de Distribuição

Seja  $X$  uma variável aleatória que assume os valores  $x_1, x_2, \dots, x_n$ . A Função de Distribuição ou Função de Distribuição Acumulada ( $F(x)$ ), com domínio em  $\mathbb{R}$  e contradomínio no intervalo  $[0, 1]$ , é dada por:

$$F(x) = \mathbb{P}(X \leq x)$$

A função de Distribuição tem as seguintes características:

- $\lim_{x \rightarrow -\infty} F(x) = 0$  e  $\lim_{x \rightarrow \infty} F(x) = 1$ .
- $F(x)$  é contínua à direita, ou seja, para um certo valor  $a \geq 0$ ,
  - Existe  $F(a)$ ,
  - Existe  $\lim_{x \rightarrow a} F(x)$ ,
  - $\lim_{x \rightarrow a} F(x) = F(a)$
- $F(x)$  é não decrescente.

O exemplo abaixo foi adaptado do livro de (BUSSAB, 2002):

**Exemplo 2.2.8.** Seja  $X$  uma variável aleatória com distribuição binomial com parâmetros  $n = 5$  e  $p = \frac{1}{2}$ . Vamos calcular a sua função de distribuição e exibir sua representação gráfica. Inicialmente, vamos calcular com que probabilidades  $X$  assume cada um de seus valores: 0, 1, 2, 3, 4, 5. Usando 2.2.8:

$$\mathbb{P}(X = 0) = \binom{5}{0} \cdot \left(\frac{1}{2}\right)^0 \cdot \left(\frac{1}{2}\right)^5 = \frac{5!}{0! \cdot 5!} \cdot \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

$$\mathbb{P}(X = 1) = \binom{5}{1} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^4 = \frac{5!}{1! \cdot 4!} \cdot \left(\frac{1}{2}\right)^5 = 5 \cdot \frac{1}{32} = \frac{5}{32}$$

$$\mathbb{P}(X = 2) = \binom{5}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^3 = \frac{5!}{2! \cdot 3!} \cdot \left(\frac{1}{2}\right)^5 = 10 \cdot \frac{1}{32} = \frac{10}{32}$$

$$\mathbb{P}(X = 3) = \binom{5}{3} \cdot \left(\frac{1}{2}\right)^1 \cdot \left(\frac{1}{2}\right)^4 = \frac{5!}{3! \cdot 2!} \cdot \left(\frac{1}{2}\right)^5 = 10 \cdot \frac{1}{32} = \frac{10}{32}$$

$$\mathbb{P}(X = 4) = \binom{5}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^1 = \frac{5!}{4! \cdot 1!} \cdot \left(\frac{1}{2}\right)^5 = 5 \cdot \frac{1}{32} = \frac{5}{32}$$

$$\mathbb{P}(X = 5) = \binom{5}{5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^0 = \frac{5!}{5! \cdot 0!} \cdot \left(\frac{1}{2}\right)^5 = \frac{1}{32}$$

Calcularemos agora  $F(x)$  para todo  $x \in \mathbb{R}$ :

- $-\infty < x < 0$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\emptyset) = 0$$

- $0 \leq x < 1$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X = 0) = \frac{1}{32}$$

- $1 \leq x < 2$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) = \frac{1}{32} + \frac{5}{32} = \frac{6}{32}$$

- $2 \leq x < 3$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = \frac{6}{32} + \frac{10}{32} = \frac{16}{32}$$

- $3 \leq x < 4$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) = \frac{16}{32} + \frac{10}{32} = \frac{26}{32}$$

- $4 \leq x < 5$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4) = \frac{26}{32} + \frac{5}{32} = \frac{31}{32}$$

- $x \geq 5$

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) + \mathbb{P}(X = 4) + \mathbb{P}(X = 5) = 1$$

A função de distribuição de  $X$  é dada por:

$$F(x) = \begin{cases} 0, & x \in ]-\infty, 0[ \\ \frac{1}{32}, & x \in [0, 1[ \\ \frac{6}{32}, & x \in [1, 2[ \\ \frac{16}{32}, & x \in [2, 3[ \\ \frac{26}{32}, & x \in [3, 4[ \\ \frac{31}{32}, & x \in [4, 5[ \\ 1, & x \in [5, \infty[ \end{cases}$$

A figura 3 mostra a representação gráfica de  $F(x)$ :

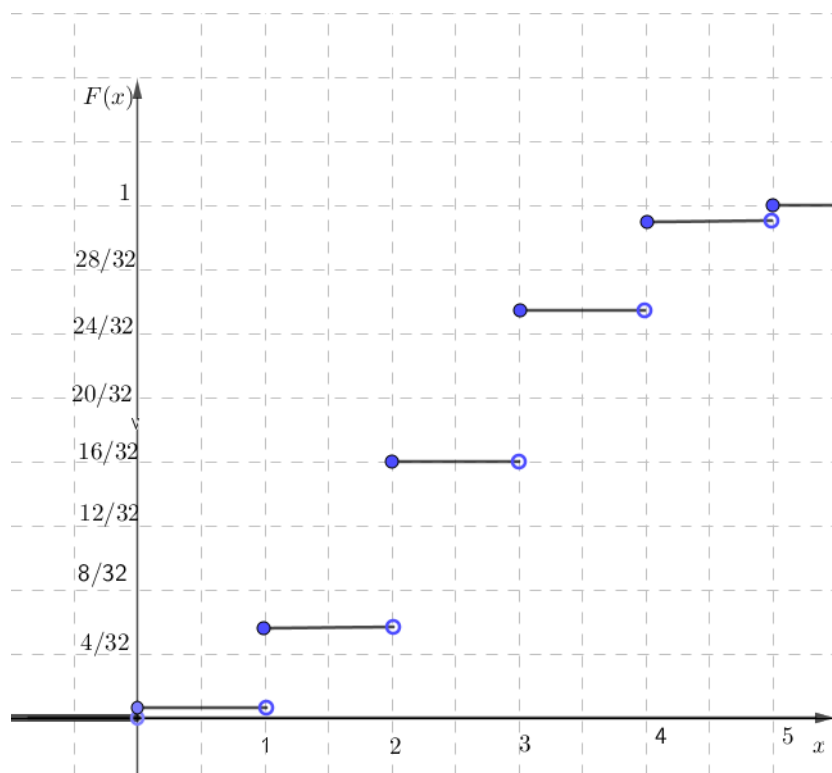


Figura 3 – Representação gráfica de  $F(x)$ , do exemplo 2.2.8, em que  $X$  é uma variável aleatória binomial com  $n = 5$  e  $p = \frac{1}{2}$

## 2.3 Distribuição Conjunta de Probabilidade para duas variáveis aleatórias discretas

Aplicações interessantes deste trabalho vieram do conceito de sequência finita de variáveis aleatórias  $X_1, X_2, \dots, X_n$ . Para entendermos melhor os conceitos, vamos apresentar o caso particular de duas variáveis aleatórias discretas.

Assim como no caso unidimensional, também vamos definir distribuição de probabilidade para o caso de duas variáveis aleatórias.

**Definição 2.3.1.** Sejam  $X$  e  $Y$  duas variáveis aleatórias discretas, assumindo os valores  $x_1, x_2, \dots, x_m$  e  $y_1, y_2, \dots, y_n$ . A distribuição de probabilidade conjunta de  $X$  e  $Y$ , é a probabilidade com que  $X$  e  $Y$  assumem os valores  $x_i$  e  $y_j$ , com  $i = 1, \dots, m$  e  $j = 1, \dots, n$ , de forma simultânea, ou seja, a probabilidade dos eventos  $\{X = x_i\} \cup \{Y = y_j\}$ , para qual usaremos a notação  $\mathbb{P}(X = x_i, Y = y_j)$ , com a vírgula representando a intersecção.

**Exemplo 2.3.1.** Consideremos o lançamento de um dado e de uma moeda. Chamaremos de  $X$  que assume os seguintes valores:

$$X(1) = X(2) = 1$$

$$X(3) = X(4) = X(5) = 2$$

$$X(6) = 3$$

Vamos definir também a variável aleatória  $Y$  que assume os valores 1 e 0,

$$Y = \begin{cases} 1, & \text{se a face superior da moeda for cara} \\ 0, & \text{se a face superior da moeda for coroa} \end{cases}$$

A tabela ?? apresenta as probabilidades associadas aos eventos  $\{X = x, Y = y\}$ , com  $x = 1, 2, e 3$  e  $y = 0, 1$ .

Usaremos a seguinte notação:

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(x, y)$$

A distribuição conjunta de probabilidade de  $X$  e  $Y$  é usualmente representada de acordo com a tabela 3.

	$X = 1$	$X = 2$	$X = 3$
$Y = 0$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{1}{12}$
$Y = 1$	$\frac{2}{12}$	$\frac{3}{12}$	$\frac{1}{12}$

Tabela 3 – Distribuição conjunta de  $X$  e  $Y$ , definidas no exemplo 2.3.1

### 2.3.1 Distribuições Marginais de duas variáveis aleatórias discretas

Sejam duas variáveis aleatórias  $X$  e  $Y$ , assumindo os valores  $x_1, x_2, \dots, x_m$  e  $y_1, y_2, \dots, y_n$  respectivamente. Vamos supor que saibamos a probabilidade conjunta de  $X$  e  $Y$  e precisemos encontrar as distribuições de probabilidade de  $X$  e  $Y$ , separadamente, que são chamadas de distribuições marginais de  $X$  e  $Y$ . A distribuição marginal de  $X$ , é calculada, como segue abaixo, a partir da distribuição conjunta de  $X$  e  $Y$ :

$$\mathbb{P}(X = x_i) = \mathbb{P}(X = x_i, Y = y_1) + \mathbb{P}(X = x_i, Y = y_2) + \dots + \mathbb{P}(X = x_i, Y = y_n)$$

$$\Rightarrow \mathbb{P}(X = x_i) = \sum_{j=1}^n \mathbb{P}(X = x_i, Y = y_j)$$

De forma análoga, calculamos a distribuição marginal de  $Y$ .

**Exemplo 2.3.2.** Vamos calcular as distribuições marginais de  $X$  e  $Y$ , definidas no exemplo 2.3.1.

Distribuição marginal de  $X$ :

$$\mathbb{P}(X = 1) = \mathbb{P}(X = 1, Y = 0) + \mathbb{P}(X = 1, Y = 1) = \frac{2}{12} + \frac{2}{12} = \frac{4}{12}$$

$$\mathbb{P}(X = 2) = \mathbb{P}(X = 2, Y = 0) + \mathbb{P}(X = 2, Y = 1) = \frac{3}{12} + \frac{3}{12} = \frac{6}{12}$$

$$\mathbb{P}(X = 3) = \mathbb{P}(X = 3, Y = 0) + \mathbb{P}(X = 3, Y = 1) = \frac{1}{12} + \frac{1}{12} = \frac{2}{12}$$

Distribuição marginal de  $Y$ :

$$\mathbb{P}(Y = 0) = \mathbb{P}(X = 1, Y = 0) + \mathbb{P}(X = 2, Y = 0) + \mathbb{P}(X = 3, Y = 0) = \frac{2}{12} + \frac{3}{12} + \frac{1}{12} = \frac{6}{12}$$

$$\mathbb{P}(Y = 1) = \mathbb{P}(X = 1, Y = 1) + \mathbb{P}(X = 2, Y = 1) + \mathbb{P}(X = 3, Y = 1) = \frac{2}{12} + \frac{3}{12} + \frac{1}{12} = \frac{6}{12}$$

### 2.3.2 Variáveis Aleatórias Independentes

Dadas as variáveis aleatórias  $X$  e  $Y$ , podemos calcular a distribuição de probabilidade condicional com que  $X$  assume os seus valores, dado que ocorreu  $\{Y = y\}$ , usando o conceito de probabilidade condicional, vista para eventos anteriormente:

$$\mathbb{P}(\{X = x\} | \{Y = y\}) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}, \quad (2.3.1)$$

com  $\mathbb{P}(Y = y) > 0$ . O cálculo de 2.3.1 vale todos os valores de  $X$  e  $Y$ , exceto para os valores que  $Y$  assume tal que  $\mathbb{P}(Y = y) = 0$  (quando  $\mathbb{P}(Y = y) = 0$ , não há uma distribuição condicional para  $X$  dado  $\{Y = y\}$ ).

Se  $X$  e  $Y$  são variáveis aleatórias independentes, a ocorrência de um determinado valor de  $X$  não é influenciada pela ocorrência de um determinado valor de  $Y$ , ou seja, a probabilidade de  $X$  assumir quaisquer um de seus valores não é alterada pela ocorrência de  $\{Y = y\}$ . O contrário também é válido. Sendo assim:

$$\mathbb{P}(\{X = x\} | \{Y = y\}) = \mathbb{P}(X = x),$$

para quaisquer valores de  $X$  ou  $Y$ . De 2.3.1

$$\mathbb{P}(X = x) = \frac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}$$

$$\Rightarrow \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y), \quad (2.3.2)$$

para todo os valores que  $X$  e  $Y$  assumem.

Ou seja,  $X$  e  $Y$  são variáveis aleatórias independentes se e somente se a probabilidade com que  $X$  e  $Y$  assumem valores de forma conjunta é igual ao produto das probabilidades com que  $X$  e  $Y$  assumem os seus valores de forma individual.

**Exemplo 2.3.3.** Dada a tabela 4, que apresenta a distribuição de probabilidade conjunta de  $X$  (que assume os valores 1, 2 e 3) e  $Y$  (assumindo os valores 0, 1 e 2), vamos verificar se  $X$  e  $Y$  são independentes.

	$X = 1$	$X = 2$	$X = 3$
$Y = 0$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{10}$
$Y = 1$	$\frac{2}{10}$	0	$\frac{3}{10}$
$Y = 2$	0	$\frac{1}{10}$	$\frac{1}{10}$

Tabela 4 – Distribuição conjunta de  $X$  e  $Y$

Inicialmente, vamos calcular as probabilidades com que  $X$  e  $Y$  assumem os valores 3 e 2, respectivamente:

$$\mathbb{P}(X = 3) = \frac{1}{10} + \frac{3}{10} + \frac{1}{10} = \frac{5}{10}$$

$$\mathbb{P}(X = 2) = 0 + \frac{1}{10} + \frac{1}{10} = \frac{2}{10}.$$

Multiplicando os resultados obtidos anteriormente, temos:

$$\mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 2) = \frac{5}{10} \cdot \frac{2}{10} = \frac{10}{100} = \frac{1}{10}.$$

$$\mathbb{P}(X = 3, Y = 2) = \frac{1}{10}$$

$$\mathbb{P}(X = 3, Y = 2) = \mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 2) \tag{2.3.3}$$

Não podemos concluir, a partir de 2.3.3, que as variáveis aleatórias  $X$  e  $Y$  são independentes. Temos, então que verificar se 2.3.2 ocorre para outros valores de  $X$  e  $Y$ .

Vamos tomar, agora, os eventos:  $\{X = 2\}$  e  $\{Y = 1\}$ , e verificar com que probabilidade as variáveis  $X$  e  $Y$  assumem esses valores.

$$\mathbb{P}(X = 2) = \frac{1}{10} + 0 + \frac{1}{10} = \frac{2}{10}$$

e

$$\mathbb{P}(Y = 1) = \frac{2}{10} + 0 + \frac{3}{10} = \frac{5}{10}$$



Então:

$$\mathbb{P}(X = 2) \cdot \mathbb{P}(Y = 1) = \frac{2}{10} \cdot \frac{5}{10} = \frac{1}{10}$$

Observando a tabela 4, veremos que  $\mathbb{P}(X = 2, Y = 1) = 0$ .

Como  $\mathbb{P}(X = 2, Y = 1) \neq \mathbb{P}(X = 2) \cdot \mathbb{P}(Y = 1)$ , podemos afirmar, então, que as variáveis aleatórias  $X$  e  $Y$  não são independentes.

### 2.3.3 Esperança e Variância de uma função de duas ou mais variáveis aleatórias discretas

Na subsecção 2.2.4, dissemos sem provar, que se  $X$  for uma variável aleatória com distribuição binomial, então  $\mathbb{E}(X) = n \cdot p$  e  $\text{VAR}(X) = n \cdot p \cdot (1 - p)$ . Essa conclusão vem do fato de que  $X$  é uma soma de variáveis aleatórias  $X_1, X_2, \dots, X_n$  com Distribuição de Bernoulli. Demonstramos, na subsecção 2.2.3, que  $\mathbb{E}(X_i) = p$  e  $\text{VAR}(X_i) = p \cdot (1 - p)$ .

Demonstraremos, nesta subsecção e na seguinte, as afirmações 2.2.9 e 2.2.10.

Sejam  $X$  e  $Y$  duas variáveis aleatórias discretas, que assumem os valores  $x_1, x_2, \dots, x_m$  e  $y_1, y_2, \dots, y_n$  com distribuição conjunta de probabilidades  $\mathbb{P}(X = x_i, Y = y_j)$  e  $g(X, Y)$  uma função que atua nos valores assumidos por elas. Então:

$$\mathbb{E}(g(X, Y)) = \sum_{i=1}^m \sum_{j=1}^n g(x_i, y_j) \mathbb{P}(X = x_i, Y = y_j)$$

Vamos calcular  $\mathbb{E}(g(X, Y))$  em que  $g(X, Y) = X + Y$ .

$$\begin{aligned} \mathbb{E}(X + Y) &= \sum_{i=1}^m \sum_{j=1}^n (x_i + y_j) \mathbb{P}(X = x_i, Y = y_j) \\ &= \sum_{i=1}^m \sum_{j=1}^n x_i \cdot \mathbb{P}(X = x_i, Y = y_j) + \sum_{i=1}^m \sum_{j=1}^n y_j \cdot \mathbb{P}(X = x_i, Y = y_j). \end{aligned}$$

Já vimos anteriormente que, se é dada a distribuição conjunta de  $X$  e  $Y$ , podemos calcular as suas distribuições marginais:

$$\mathbb{P}(X = x_i) = \sum_{j=1}^n \mathbb{P}(X = x_i, Y = y_j)$$

e

$$\mathbb{P}(Y = y_j) = \sum_{i=1}^m \mathbb{P}(X = x_i, Y = y_j)$$

Logo,  $\mathbb{E}(X + Y)$  é dada por:

$$\sum_{i=1}^m x_i \sum_{j=1}^n \mathbb{P}(X = x_i, Y = y_j) + \sum_{j=1}^n y_j \sum_{i=1}^m \mathbb{P}(X = x_i, Y = y_j) = \sum_{i=1}^m x_i \mathbb{P}(X = x_i) + \sum_{j=1}^n y_j \mathbb{P}(Y = y_j)$$

$$= \mathbb{E}(X) + \mathbb{E}(Y)$$

Então, podemos concluir que dadas duas variáveis aleatórias  $X$  e  $Y$ , a esperança da soma delas é a soma de suas esperanças, ou seja:

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

Esse resultado pode ser estendido para a soma finita de mais de duas variáveis.

$$\mathbb{E}(X_1 + X_2 + \dots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n) \quad (2.3.4)$$

**Teorema 2.3.1.** Para duas variáveis aleatórias discretas independentes, temos:

$$\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y) \quad (2.3.5)$$

Este teorema pode ser demonstrado conforme a seguir:

$$\mathbb{E}(XY) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j \mathbb{P}(X = x_i, Y = y_j) \quad (2.3.6)$$

Mas, como as duas variáveis são independentes, então:

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \cdot \mathbb{P}(Y = y) \quad (2.3.7)$$

Logo, substituindo a equação 2.3.7 em 2.3.6, obtemos:

$$\begin{aligned} \mathbb{E}(XY) &= \sum_{i=1}^n \sum_{j=1}^m x_i y_j \cdot \mathbb{P}(X = x_i, Y = y_j) = \sum_{i=1}^n \sum_{j=1}^m x_i \cdot y_j \cdot \mathbb{P}(X = x_i) \cdot \mathbb{P}(Y = y_j) \\ &= \sum_{i=1}^n \sum_{j=1}^m x_i \cdot \mathbb{P}(X = x_i) \cdot y_j \cdot \mathbb{P}(Y = y_j) = \sum_{i=1}^n x_i \cdot \mathbb{P}(X = x_i) \cdot \sum_{j=1}^m y_j \cdot \mathbb{P}(Y = y_j) = \mathbb{E}(X) \cdot \mathbb{E}(Y) \end{aligned}$$

Vamos agora calcular  $\text{VAR}(X + Y)$ , usando o fato de que  $X$  e  $Y$  são variáveis aleatórias independentes e a propriedade da esperança:

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b$$

Pela definição de variância:

$$\begin{aligned} \text{VAR}(X + Y) &= \mathbb{E}((X + Y)^2) - (\mathbb{E}(X + Y))^2 = \mathbb{E}(X^2 + 2 \cdot X \cdot Y + Y^2) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\ &= \mathbb{E}(X^2) + \mathbb{E}(2 \cdot X \cdot Y) + \mathbb{E}(Y^2) - (\mathbb{E}(X))^2 - 2 \cdot \mathbb{E}(X) \cdot \mathbb{E}(Y) - (\mathbb{E}(Y))^2 \end{aligned}$$

Como  $X$  e  $Y$  são variáveis aleatórias independentes, pelo teorema 2.3.1:

$$\begin{aligned} \text{VAR}(X + Y) &= \mathbb{E}(X^2) + 2 \cdot \mathbb{E}(X) \cdot \mathbb{E}(Y) + \mathbb{E}(Y^2) - (\mathbb{E}(X))^2 - 2 \cdot \mathbb{E}(X) \cdot \mathbb{E}(Y) - (\mathbb{E}(Y))^2 \\ &= [\mathbb{E}(X^2) - (\mathbb{E}(X))^2] + [\mathbb{E}(Y^2) - (\mathbb{E}(Y))^2] = \text{VAR}(X) + \text{VAR}(Y) \end{aligned}$$

Podemos estender esse resultado para a Variância da soma de  $n$  variáveis aleatórias independentes  $X_1, X_2, \dots, X_n$ .

$$\text{VAR}(X_1 + X_2 + \dots + X_n) = \text{VAR}(X_1) + \text{VAR}(X_2) + \dots + \text{VAR}(X_n)$$

### 2.3.4 Esperança e Variância da Variável Aleatória Binomial

Já vimos na seção anterior que a esperança da soma de  $n$  variáveis aleatórias independentes é igual à soma de suas esperanças e o mesmo ocorre com a variância. Concluiremos agora o que foi proposto na subseção 2.2.4: demonstrar 2.2.9 e 2.2.10.

Seja  $X_i$ , ( $i = 1, \dots, n$ ), uma variável aleatória com distribuição de probabilidade de Bernoulli, assumindo os valores 0 e 1, com probabilidades  $p$  e  $p(1 - p)$ , e  $X$  uma variável aleatória com distribuição binomial de probabilidades, tal que  $X = X_1 + X_2 + \dots + X_n$ , então:

$$\begin{aligned} \text{VAR}(X) &= \text{VAR}(X_1 + X_2 + \dots + X_n) = \text{VAR}(X_1) + \text{VAR}(X_2) + \dots + \text{VAR}(X_n) \\ &= p(1 - p) + p(1 - p) + \dots + p(1 - p) = np(1 - p). \end{aligned}$$

$$\mathbb{E}(X) = \mathbb{E}(X_1 + X_2 + \dots + X_n) = \mathbb{E}(X_1) + \mathbb{E}(X_2) + \dots + \mathbb{E}(X_n) = p + p + \dots + p = np$$



---

## EM BUSCA DA NORMAL - PARTE 2: UMA BREVE INTRODUÇÃO À ESTATÍSTICA

---

Após serem explanados os assuntos que são pré-requisitos para o entendimento da Variável Aleatória Normal, busca-se agora levar os professores a aprofundar sobre ela.

Convém que os professores do Ensino Básico se atualizem sobre assuntos que são tratados no Ensino Superior, tais como derivadas, integral, função contínua para melhor aproveitamento do capítulo.

### 3.1 Variável Aleatória Contínua

Uma variável,  $X$ , com  $\Omega = \mathbb{R}$  ou um intervalo de  $\mathbb{R}$  é chamada de contínua se tiver a ela associada uma função densidade de probabilidade, que denotaremos por  $f_X(x)$ . Essa função diz como varia a densidade de probabilidade ao longo da reta real, que é o caso particular que estudamos neste trabalho. A função  $f_X(x)$  tem as seguintes propriedades:

**a)**  $f_X(x) \geq 0$

**b)**  $\int_{-\infty}^{\infty} f_X(x) dx = 1$

Para termos uma intuição a respeito, podemos pensar que  $f_X(x)$  tem como unidade, probabilidade por unidade de comprimento. Então, se tivermos um "pedaço de  $\mathbb{R}$ " (intervalo de  $\mathbb{R}$ ) de comprimento  $\Delta x$ , então a "quantidade de probabilidade nesse pedaço" será  $f_X(x) \cdot \Delta x$ .

A probabilidade da variável aleatória  $X$  assumir valores em um intervalo  $[a, b]$  é dada por:

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) dx,$$

e corresponde à área sob o gráfico da  $f_X(x)$ , para  $x$  variando entre  $a$  e  $b$ , conforme mostra a figura 4.

Há uma ligação direta entre  $F(x)$  e  $f_X(x)$ ,

$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f_X(x) dx.$$

Sem entrar em maiores detalhes, porque fugiria ao escopo deste trabalho, temos pelo Teorema Fundamental do Cálculo que  $F(x)$  é a primitiva de  $f_X(x)$ , ou seja,  $f_X(x)$  é a derivada de  $F(x)$ .

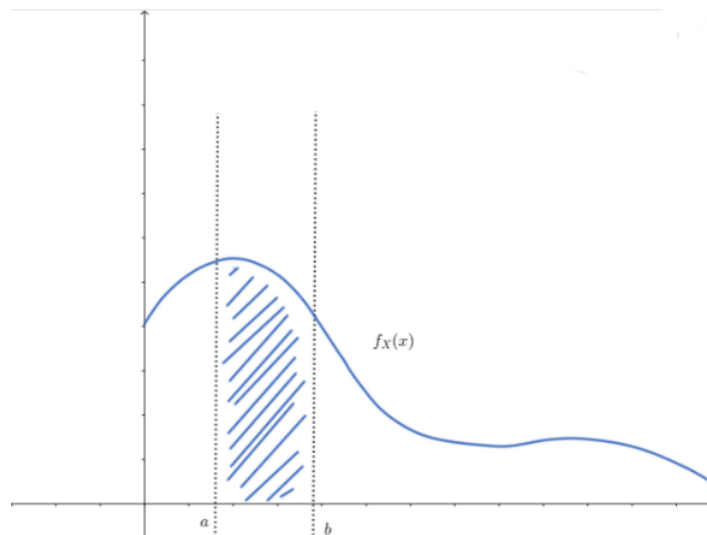


Figura 4 – Representação gráfica de  $f_X(x)$  e da probabilidade de  $X$  assumir valores entre  $a$  e  $b$ , dada pela área da figura hachurada

### 3.1.1 Esperança e Variância de uma variável aleatória contínua

A esperança de uma variável aleatória contínua  $X$  é calculada usando a função  $f_X(x)$ :

$$\mathbb{E}(X) = \int_a^b x f_X(x) dx$$

A **variância de  $X$**  pode ser calculada usando a mesma fórmula usada para o cálculo da variância de uma variável aleatória discreta, ou seja,  $\text{VAR}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ , sendo a  $\mathbb{E}(X^2)$  calculada por meio da resolução da integral:

$$\mathbb{E}(X^2) = \int_a^b x^2 f_X(x) dx$$

### 3.1.2 Variável Aleatória com Distribuição Normal

Uma variável aleatória contínua  $X$  possui Distribuição Normal de Probabilidade, se tiver a seguinte função densidade de probabilidade:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\mu-x)^2}{2\sigma^2}} \quad (3.1.1)$$

onde  $x$  pertence ao conjunto  $\mathbb{R}$ ,  $\mu$  é a esperança de  $X$  e  $\sigma^2$ , a sua variância.

Usamos a notação  $X \sim N(\mu, \sigma^2)$ , para nos referirmos a uma variável aleatória  $X$  com distribuição normal de probabilidade de esperança  $\mu$  e variância  $\sigma^2$ .

A função densidade de probabilidade  $f_X(x)$ , dada por 3.1.1, possui as seguintes propriedades.

- $f_X(x)$  é simétrica em relação à média  $\mu$ .
- Seu gráfico tem a forma de um sino.
- A área compreendida entre o gráfico e o eixo das abscissas é igual a 1.
- $\lim_{x \rightarrow \pm\infty} f_X(x) = 0$ .
- O maior valor de  $f_X(x)$  acontece quando  $x = \mu$ .

Para o cálculo de probabilidades de variáveis aleatórias contínuas, podemos usar:

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f_X(x) dx \quad (3.1.2)$$

No caso de uma variável aleatória Normal usamos  $f_X(x)$  dada por 3.1.1, e nesse caso, a integral só pode ser resolvida por meio de métodos numéricos, por isso, faz-se o uso de tabelas com valores previamente calculados. As tabelas usadas para o cálculo de probabilidades são padronizadas com média  $\mu = 0$  e variância  $\sigma^2 = 1$ . Por isso, se  $X$  for uma variável aleatória com distribuição normal com parâmetros  $\mu \neq 0$  e  $VAR(X) \neq 1$ , fazemos a transformação abaixo, definindo uma variável aleatória  $Z$ , tal que:

$$Z = \frac{X - \mu}{\sigma}, \quad (3.1.3)$$

Assim, encontramos uma variável aleatória  $Z$ , com  $\mathbb{E}(Z) = 0$  e  $VAR(Z) = 1$ .

Usando as propriedades da esperança e da variância, verificamos a afirmação acima:

$$\mathbb{E}(Z) = \mathbb{E}\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma} \mathbb{E}(X - \mu) = \frac{\mathbb{E}(X) - \mathbb{E}(\mu)}{\sigma} = \frac{\mu - \mu}{\sigma} = 0$$

$$VAR(Z) = VAR\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma^2} (VAR(X) - VAR(\mu)) = \frac{1}{\sigma^2} (\sigma^2 - 0) = 1$$

O exemplo seguinte, adaptado do livro de (BUSSAB, 2002), mostra como calculamos a probabilidade com que uma variável aleatória  $X$ , com distribuição normal, assume valores entre  $x_1$  e  $x_2$ .

**Exemplo 3.1.1.** Sendo  $X \sim N(10, 4)$ , vamos encontrar as seguintes probabilidades:

- $\mathbb{P}(8 \leq X \leq 10) = \mathbb{P}\left(\frac{8 - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{10 - \mu}{\sigma}\right)$

Como  $\mu = 10$  e  $\sigma = \sqrt{4} = 2$ ,

$$\begin{aligned} \mathbb{P}(8 \leq X \leq 10) &= \mathbb{P}\left(\frac{8 - 10}{2} \leq Z \leq \frac{10 - 10}{2}\right) \\ &= \mathbb{P}(-1 \leq Z \leq 0) \end{aligned} \tag{3.1.4}$$

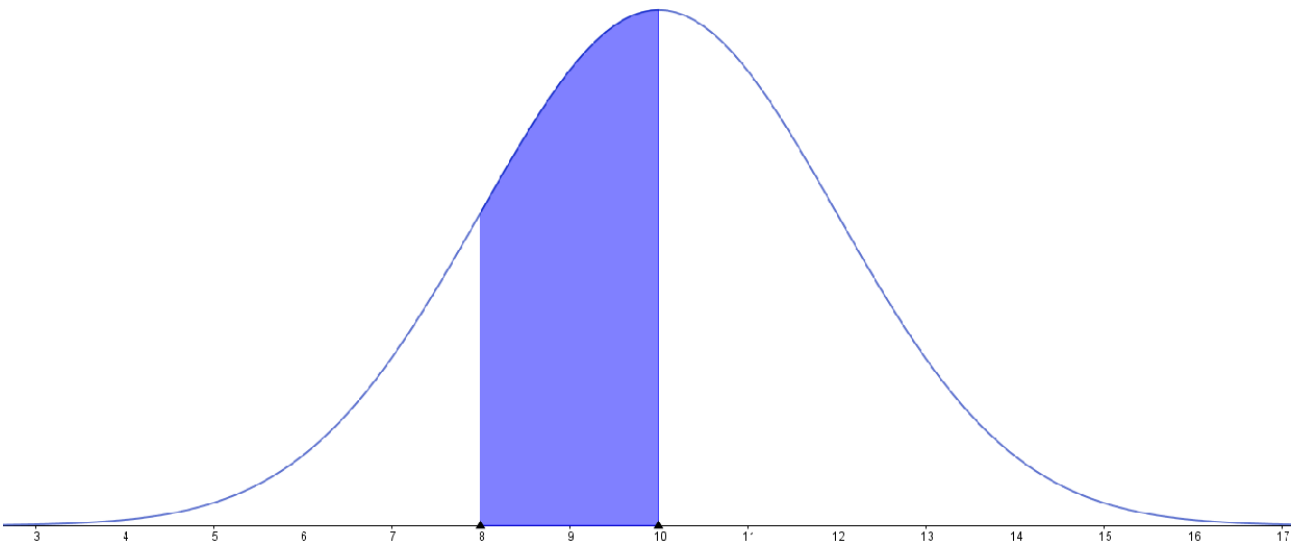


Figura 5 – Representação gráfica da probabilidade de  $X$  estar entre os números 8 e 10.

Como a tabela da figura 7 mostra a  $\mathbb{P}(Z \leq z)$ , precisamos desenvolver a expressão 3.1.4.

$$\mathbb{P}(-1 \leq Z \leq 0) = \mathbb{P}(Z \leq 0) - \mathbb{P}(Z \leq -1).$$

Não temos, na tabela da figura 7,  $\mathbb{P}(Z \leq -1)$ , então fazemos:

$$\mathbb{P}(Z \leq -1) = \mathbb{P}(Z \geq 1) = 1 - \mathbb{P}(Z \leq 1) \tag{3.1.5}$$

Logo:

$$\Rightarrow \mathbb{P}(Z \leq 0) - \mathbb{P}(Z \leq -1) = \mathbb{P}(Z \leq 0) - [1 - \mathbb{P}(Z \leq 1)]$$

$$\Rightarrow \mathbb{P}(-1 \leq Z \leq 0) = \mathbb{P}(Z \leq 0) + \mathbb{P}(Z \leq 1) - 1 \tag{3.1.6}$$

Na tabela da figura 7, os números que estão na 1ª coluna representam valores reais maiores ou iguais a zero e sua primeira casa decimal. E os valores da 1ª linha indicam a segunda casa decimal de  $x$ . Os valores que estão nas demais linhas e colunas, representam a



probabilidade de  $Z$  assumir valores entre 0 e 3,49. Por exemplo, se quisermos encontrar na tabela o resultado de  $\mathbb{P}(Z \leq 1,21)$ , buscamos na 1ª coluna o número 1,2 e na 1ª coluna, o número 0,01. No cruzamento da linha horizontal (que tem o 1,2) e vertical (do número 0,01), está  $\mathbb{P}(Z \leq 1,21) = 0,8869$ .

No exemplo 3.1.4,

$$\mathbb{P}(Z \leq 0) = 0,5$$

e

$$\mathbb{P}(Z \leq 1) = 0,8413.$$

E assim podemos, finalmente, calcular o resultado procurado.

$$\mathbb{P}(8 \leq X \leq 10) = 0,5 + 0,8413 - 1 = 0,3413$$

A área sob o gráfico da figura 5 equivale a probabilidade de, aproximadamente, 34% de  $X$  assumir valores no intervalo  $[8, 10]$

Para a mesma variável aleatória  $X \sim N(10, 4)$ , vamos calcular:

- $\mathbb{P}(8 \leq X \leq 12)$

$$\begin{aligned} \mathbb{P}(8 \leq X \leq 12) &= \mathbb{P}\left(\frac{8 - \mu}{\sigma} \leq \frac{X - \mu}{\sigma} \leq \frac{12 - \mu}{\sigma}\right) = \mathbb{P}\left(\frac{8 - 10}{2} \leq Z \leq \frac{12 - 10}{2}\right) \\ &= \mathbb{P}(-1 \leq Z \leq 1). \end{aligned}$$

De 3.1.5,

$$\begin{aligned} \mathbb{P}(8 \leq X \leq 12) &= \mathbb{P}(Z \leq 1) - \mathbb{P}(Z \leq -1) = \mathbb{P}(Z \leq 1) - [1 - \mathbb{P}(Z \leq 1)] \\ &= 2 \cdot \mathbb{P}(Z \leq 1) - 1 = 2 \cdot 0,8413 - 1 = 0,6826 \end{aligned}$$

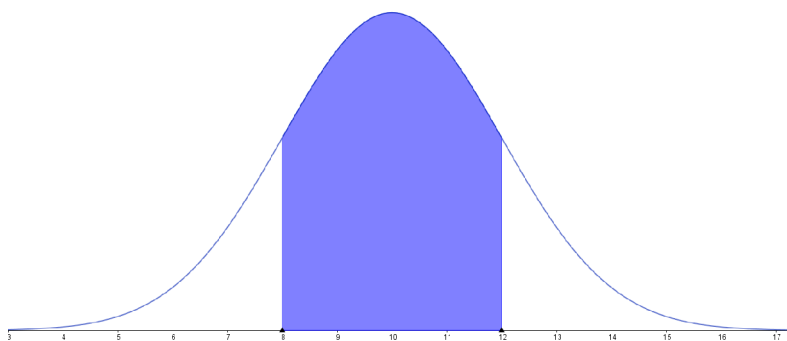
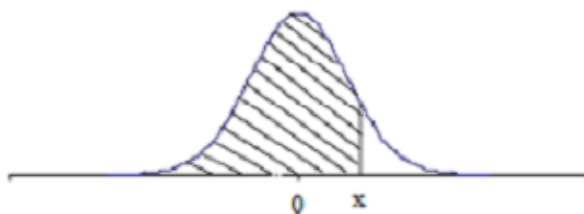


Figura 6 – Representação gráfica da probabilidade com que a variável aleatória  $X$  pertença ao intervalo  $[8, 12]$ .

### Distribuição Normal Padrão

Os valores tabelados correspondem à área abaixo representada:  $P(Z \leq x)$



x	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

Figura 7 – Na tabela normal, os números estão na 1ª coluna representam valores reais maiores ou iguais a zero e sua primeira casa decimal. Os números que aparecem na 1ª linha indicam a segunda casa decimal de  $x$ . Os números que estão nas demais linhas e colunas, representam  $\mathbb{P}(Z \leq x)$ . No texto, usamos  $z$  ao invés de  $x$ .

## 3.2 Inferência Estatística

O estudo de uma característica populacional, esbarra em questões como tempo e recursos financeiros para realizar uma pesquisa ampla. Por isso, o estudo é realizado a partir de uma amostra representativa da população. Esse processo de estudo é chamado de **inferência estatística**.

Segundo (BUSSAB, 2002), **população** é o conjunto de indivíduos ou objetos cujas características queremos estudar e **amostra** é um subconjunto da população.

Uma amostra aleatória simples (A. A. S.) de tamanho  $n$  de uma variável aleatória  $X$ , com dada distribuição, é um conjunto de  $n$  variáveis aleatórias independentes e identicamente distribuídas (i.i.d.),  $X_1, X_2, \dots, X_n$ , cada uma com a mesma distribuição de  $X$ .

### 3.2.1 Estimação Pontual

Vamos supor que estudando uma determinada característica de uma população, associemos a ela uma variável aleatória  $X$ . Sabemos a distribuição de probabilidade de  $X$ , mas não os valores de seus parâmetros. Usaremos estimadores, que são funções de  $X_1, X_2, \dots, X_n$  para os parâmetros desconhecidos. Representamos por  $\hat{\theta}$  o estimador de um parâmetro  $\theta$ .

Para que  $\hat{\theta}$  seja um bom estimador de  $\theta$ , ele deve ter as seguintes características:

- 1ª : Ser não viciado ou não viesado.

Um estimador é não viesado ou não viciado se:

$$\mathbb{E}(\hat{\theta}) = \theta$$

**Exemplo 3.2.1.** Consideremos a amostra aleatória simples,  $X_1, X_2, \dots, X_n$ , onde  $X_i$  é uma variável aleatória com distribuição normal com média  $\mu$  e variância  $\sigma^2$ , ambos desconhecidos. Usaremos estimadores para a média e variância de  $X$ , e vamos verificar se são viesados ou não.

Afirmamos que  $\bar{X}$ , dado por

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

é um estimador não viesado para  $\mu$ , pois:

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\frac{\sum_{i=1}^n X_i}{n}\right) = \frac{1}{n} \mathbb{E}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot n\mu = \mu$$

Seria natural que o estimador de  $\sigma^2$ , chamado de  $S^2$ , fosse dado pela fórmula:

$$S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \quad (3.2.1)$$

No apêndice A, será demonstrado que  $S^2$  é um estimador viesado para  $\sigma^2$ .

Contudo, se multiplicarmos  $S^2$  por  $\frac{n}{n-1}$ , e calcularmos a sua esperança, temos:

$$\mathbb{E}\left(\frac{n}{n-1}S^2\right) = \frac{n}{n-1}\mathbb{E}(S^2) = \frac{n}{n-1} \cdot \frac{n-1}{n}\sigma^2 = \sigma^2$$

Concluimos que um estimador não viesado de  $\sigma^2$  seria:

$$\frac{n}{n-1}S^2 = \frac{n}{n-1} \cdot \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Portanto, um estimador não viesado de  $\sigma^2$  pode ser:

$$\hat{\sigma}^2 = \frac{1}{n-1} \cdot \sum_{i=1}^n (X_i - \bar{X})^2$$

- 2ª : Dizemos que  $\hat{\theta}$  é um estimador consistente para  $\theta$ , se:

$$\lim_{n \rightarrow \infty} E(\hat{\theta}) = \theta \quad (3.2.2)$$

e

$$\lim_{n \rightarrow \infty} VAR(\hat{\theta}) = 0 \quad (3.2.3)$$

**Exemplo 3.2.2.** O estimador  $\bar{X}$  é um estimador consistente de  $\mu$ , pois satisfaz 3.2.2 e 3.2.3, conforme pode ser verificado a seguir:

$$\lim_{n \rightarrow \infty} \mathbb{E}(\bar{X}) = \lim_{n \rightarrow \infty} \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \lim_{n \rightarrow \infty} \frac{1}{n} n\mu = \mu$$

Usando a propriedade da variância, vista em 2.2.2

$$\lim_{n \rightarrow \infty} VAR(\bar{X}) = \lim_{n \rightarrow \infty} VAR\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \lim_{n \rightarrow \infty} \frac{1}{n^2} \sum_{i=1}^n VAR(X_i) = \lim_{n \rightarrow \infty} \frac{1}{n^2} n\sigma^2 = \lim_{n \rightarrow \infty} \frac{1}{n}\sigma^2 = 0$$

- 3ª : Dados dois estimadores, o estimador mais eficiente é aquele que possui menor variância.

O exemplo a seguir foi retirado do texto de (BOSCO, 2010).

**Exemplo 3.2.3.** Suponha uma Amostra Aleatória Simples (A. A. S.)  $X_1, X_2, \dots, X_n$  de uma v.a.  $X$ , tem distribuição de Bernoulli de parâmetro  $p$  e assume os valores 1, para sucesso, e 0 para fracasso. Seja  $Y = X_1 + X_2 + \dots + X_n$  o número de sucessos e considere os estimadores para  $p$ .

1.  $\hat{p}_1 = \frac{Y}{n}$
2.  $\hat{p}_2 = X_1$

Vamos verificar qual dos dois estimadores é o mais eficiente.

$$\begin{aligned} - \text{VAR}(\hat{p}_1) &= \text{VAR}\left(\frac{Y}{n}\right) = \frac{1}{n^2} \text{VAR}(Y) = \frac{1}{n^2} np(1-p) = \frac{p(1-p)}{n} \\ - \text{VAR}(\hat{p}_2) &= \text{VAR}(X_1) = p(1-p) \end{aligned}$$

Como  $\text{VAR}(\hat{p}_1) < \text{VAR}(\hat{p}_2)$ , concluímos que  $\hat{p}_1$  é mais eficiente que  $\hat{p}_2$ .

### 3.2.2 Teorema Central do Limite

**Teorema 3.2.1.** Seja  $X_1, X_2, \dots, X_n$ , uma sequência de um número  $n$ , finito, de variáveis aleatórias i.i.d., com  $E(X_i) = \mu$  e  $\text{VAR}(X_i) = \sigma^2$ . O Teorema Central do Limite (T. C. L.)<sup>1</sup> afirma que, para um valor de  $n$  suficientemente grande, a distribuição de probabilidade de  $Z_n$ , dada por:

$$Z_n = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{X_1 + X_2 + \dots + X_n - n \cdot \mu}{\sqrt{n} \cdot \sigma}.$$

converge para uma distribuição normal com parâmetros  $\mu = 0$  e  $\sigma^2 = 1$ .

Como resultado do Teorema central do Limite,  $\mathbb{P}(Z_n \leq z)$  vai convergir para

$$F(z) = \frac{1}{2 \cdot \pi} \cdot \int_{-\infty}^z e^{-\frac{x^2}{2}} dx,$$

onde  $F(z) = \mathbb{P}(Z_n \leq z)$ ,  $Z_n \sim N(0, 1)$

O exemplo a seguir foi retirado do texto de (BOSCO, 2010).

**Exemplo 3.2.4.** Considere uma amostra de uma população, para a qual se atribui uma variável aleatória  $X$  a uma característica de interesse. Sabe-se que  $X$  tem média 128 e desvio padrão igual a 16. Se uma amostra de tamanho 100 é selecionada, vamos usar o Teorema Central do Limite para calcular, aproximadamente, a probabilidade da média amostral estar entre 124 e 132.

A Variável Aleatória  $X$  tem parâmetros  $\mu = 128$ ,  $\sigma = 16$ .

Vamos, agora, calcular  $\mathbb{E}(\bar{X})$  e  $\text{VAR}(\bar{X})$ , para uma amostra  $X_1, X_2, \dots, X_n$  i.i.d.

$$\begin{aligned} \bullet \mathbb{E}(\bar{X}) &= \mathbb{E}\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot \mathbb{E}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n} \cdot \sum_{i=1}^n \mathbb{E}(X_i) = \frac{1}{n} \cdot n\mu = \mu = 128 \\ \bullet \text{VAR}(\bar{X}) &= \text{VAR}\left(\frac{1}{n} \cdot \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \cdot \text{VAR}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \text{VAR}(X_i) = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n}. \end{aligned}$$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{16}{\sqrt{100}} = \frac{16}{10}$$

<sup>1</sup> Não demonstraremos o Teorema Central do Limite pois foge do escopo do trabalho

$$\mathbb{P}(124 < \bar{X} < 132) = \mathbb{P}\left(\frac{124 - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{132 - \mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

O Teorema Central do Limite, afirma que, para valores grandes de  $n$ , distribuição de probabilidade da variável aleatória  $Z$ , tal que:

$$Z_n = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}},$$

converge para uma distribuição normal de parâmetros  $\mu = 0$  e  $\sigma^2 = 1$ . Usaremos a tabela 7, para encontramos a probabilidade pedida.

$$\mathbb{P}(124 < \bar{X} < 132) = \mathbb{P}\left(\frac{124 - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{132 - \mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= \mathbb{P}\left(\frac{124 - 128}{\frac{16}{10}} < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{132 - 128}{\frac{16}{10}}\right) = \mathbb{P}\left(\frac{-4}{1,6} < Z_n < \frac{4}{1,6}\right) = \mathbb{P}(-2,5 < Z < 2,5)$$

$$= \mathbb{P}(Z_n < 2,5) - \mathbb{P}(Z_n < -2,5) = \mathbb{P}(Z_n \leq 2,5) - \mathbb{P}(Z_n > 2,5) = \mathbb{P}(Z_n < 2,5) - (1 - \mathbb{P}(Z_n < 2,5))$$

$$= 2 \cdot \mathbb{P}(Z_n < 2,5) - 1 \Rightarrow \mathbb{P}(124 < \bar{X} < 132) = 2 \cdot 0,9938 - 1 = 0,9876$$

A figura 8 ilustra a probabilidade procurada do exemplo 2.4.1.

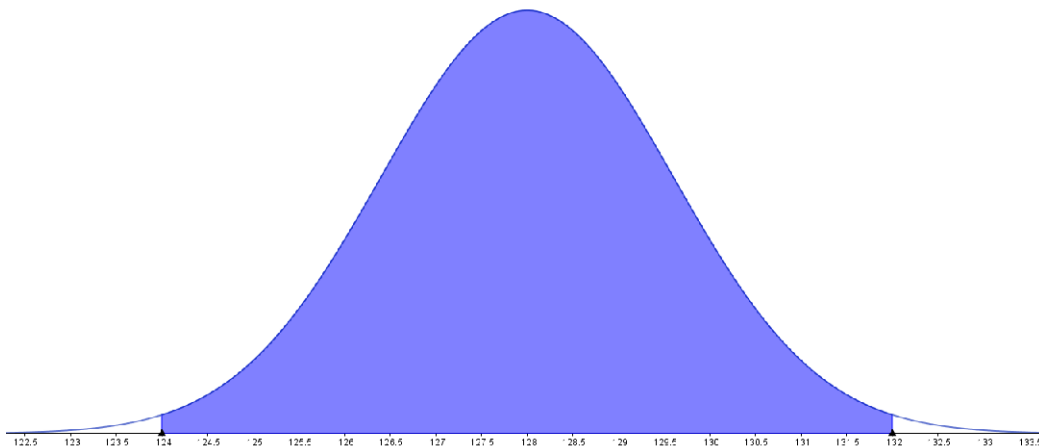


Figura 8 – Representação gráfica do exemplo 3.2.4 mostrando a probabilidade de  $X$  estar entre 124 e 132 de uma população com média  $\mu = 128$  e desvio padrão  $\sigma = 16$

### 3.2.3 Estimação Intervalar

Na presente subsecção, mostraremos como construir intervalos aleatórios, através dos dados da amostra, que contenham o verdadeiro valor do parâmetro populacional, com um índice alto de probabilidade ( $\gamma$ ).

#### 3.2.3.1 Intervalos de Confiança para média populacional

Vamos supor, como exemplo, que tenhamos uma população formada por alunos concluintes do ensino básico nas escolas públicas paulistas e pretendemos saber qual a média de altura ( $\mu$ ) desses alunos. Como não é possível consultarmos todos os alunos, selecionamos uma amostra e, a cada aluno da amostra, associamos uma variável aleatória  $X_i$ ,  $i = 1, \dots, n$  que assume o valor da altura do estudante, sendo  $X_i \sim N(\mu, \sigma^2)$ . A altura média dos estudantes que fazem parte da amostra é dada por:

$$\bar{X} = \frac{S_n}{n}$$

em que

$$S_n = \sum_{i=1}^n X_i.$$

Usaremos o TCL para fazer os cálculos e construiremos, a partir da amostra, um intervalo de confiança que tem probabilidade  $\gamma$  de conter o verdadeiro valor de  $\mu$ .

Para um certo valor de  $VAR(X)$ <sup>2</sup>, vamos construir Intervalos de Confiança para  $\mu$ , encontrando o valor  $c$ ,  $c > 0$  que satisfaça:

$$\mathbb{P}(|\bar{X} - \mu| \leq c) = \gamma$$

em que  $0 < \gamma < 100\%$ . Geralmente, usa-se  $\gamma$  com valores acima de 90%.

$$\mathbb{P}(-c \leq \bar{X} - \mu \leq c) = \gamma \quad (3.2.4)$$

Vamos, agora, dividir ambos os membros por  $\frac{\sigma}{\sqrt{n}}$ .

$$\mathbb{P}\left(\frac{-c}{\frac{\sigma}{\sqrt{n}}} \leq \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{c}{\frac{\sigma}{\sqrt{n}}}\right) = \gamma$$

Como:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (3.2.5)$$

<sup>2</sup> o valor da variância já deve ser conhecido, pois a construção de Intervalos de Confiança foge do escopo deste trabalho, caso não saibamos o valor da variância.

tem distribuição Normal com parâmetros  $\mu = 0$  e  $\frac{\sigma^2}{n} = 1$ ,

$$\mathbb{P}\left(\frac{-c}{\frac{\sigma}{\sqrt{n}}} \leq Z_n \leq \frac{c}{\frac{\sigma}{\sqrt{n}}}\right) = \gamma \quad (3.2.6)$$

Para usarmos a tabela da Normal Padrão (7), deixaremos 3.2.6 na forma de uma desigualdade do tipo  $\mathbb{P}(Z_n \leq x)$ .

$$\mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right) - \mathbb{P}\left(Z_n \leq \frac{-c \cdot \sqrt{n}}{\sigma}\right) = \gamma \quad (3.2.7)$$

Como

$$\mathbb{P}\left(Z_n \leq \frac{-c \cdot \sqrt{n}}{\sigma}\right) = \mathbb{P}\left(Z_n \geq \frac{c \cdot \sqrt{n}}{\sigma}\right) = 1 - \mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right),$$

substituiremos esse resultado em 3.2.7 para obtermos:

$$\begin{aligned} \mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right) - \left(1 - \mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right)\right) &= \gamma = \mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right) - 1 + \mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right) \\ \Rightarrow 2 \cdot \mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right) &= \gamma + 1 \Rightarrow \mathbb{P}\left(Z_n \leq \frac{c \cdot \sqrt{n}}{\sigma}\right) = \frac{\gamma + 1}{2} \end{aligned}$$

Buscaremos agora, na tabela 7, o valor

$$z = \frac{c \cdot \sqrt{n}}{\sigma},$$

que satisfaça

$$\mathbb{P}(Z_n \leq z) = \frac{\gamma + 1}{2}, \forall \gamma$$

Usando  $\gamma = 95\%$ , como no exemplo, temos que:

$$\begin{aligned} \mathbb{P}(Z_n \leq z) &= \frac{0,95 + 1}{2} = 0,975 \\ z &= 1,96 \end{aligned}$$

Podemos, agora, encontrar uma expressão para  $c$ .

$$\frac{c \cdot \sqrt{n}}{\sigma} = 1,96 \Rightarrow c \cdot \sqrt{n} = 1,96 \cdot \sigma \Rightarrow c = \frac{1,96 \cdot \sigma}{\sqrt{n}}.$$

Voltaremos então, em 3.2.4 e vamos substituir o  $c$  por  $\frac{1,96 \cdot \sigma}{\sqrt{n}}$

$$\begin{aligned} \mathbb{P}\left(-\frac{1,96 \cdot \sigma}{\sqrt{n}} \leq \bar{X} - \mu \leq \frac{1,96 \cdot \sigma}{\sqrt{n}}\right) &= 0,95 \\ \Rightarrow \mathbb{P}\left(-\bar{X} - \frac{1,96 \cdot \sigma}{\sqrt{n}} \leq -\mu \leq -\bar{X} + \frac{1,96 \cdot \sigma}{\sqrt{n}}\right) &= 0,95 \\ \Rightarrow \mathbb{P}\left(\bar{X} - \frac{1,96 \cdot \sigma}{\sqrt{n}} \leq \mu \leq \bar{X} + \frac{1,96 \cdot \sigma}{\sqrt{n}}\right) &= 0,95 \end{aligned}$$

Logo, o intervalo de confiança para  $\mu$ , dado  $\gamma = 95\%$  fica descrito da seguinte forma:

$$IC(\mu; 95\%) = \left[ \bar{X} - 1,96 \cdot \frac{\sigma}{\sqrt{n}}; \bar{X} + 1,96 \cdot \frac{\sigma}{\sqrt{n}} \right] \quad (3.2.8)$$



**Exemplo 3.2.5.** Vamos, agora, voltar na situação descrita no início da seção, no qual queríamos estimar a altura de todos os estudantes concluintes do ensino básico. Ao tomar uma amostra aleatória simples com 100 alunos, é obtido, para  $\bar{X}$ , o valor 163 cm. Supondo que  $X$  tenha distribuição normal  $N(\mu, \sigma^2)$ . Conhecendo-se o valor de  $\sigma = 5\text{cm}$ , vamos construir um intervalo de confiança para  $\mu$  com  $\gamma = 95\%$

$$\begin{aligned} IC(\mu, 95\%) &= \left[ 163 - 1,96 \cdot \frac{5}{\sqrt{100}}; 163 + 1,96 \cdot \frac{5}{\sqrt{100}} \right] = [163 - 0,98; 163 + 0,98] \\ &= [162,02; 163,98]. \end{aligned}$$

Concluimos que, a probabilidade de que o valor verdadeiro de  $\mu$  pertença ao intervalo  $[162,02; 163,98]$  é de 95%.

### 3.2.3.2 Intervalo de confiança para $p$

Tomemos uma amostra  $X_1, X_2, \dots, X_n$ , i.i.d., em que cada variável aleatória  $X_i$  tem distribuição de probabilidade de Bernoulli, com parâmetro  $p$ . Retomando os conceitos desenvolvidos na subseção 2.2.4, já vimos que a variável aleatória  $S_n$  dada por:

$$S_n = \sum_{i=1}^n X_i$$

tem distribuição de probabilidade binomial com parâmetros  $n$  e  $p$ .

Seja  $\hat{p}$ , estimador para  $p$ , dado por:

$$\hat{p} = \frac{S_n}{n}.$$

Nesta subseção iremos construir intervalos de confiança para  $p$ , com um nível de confiança de 95%, encontrando um valor  $c$ , ( $c > 0$ ), que satisfaça:

$$\mathbb{P}(|\hat{p} - p| \leq c) = 95\% = \mathbb{P}(-c \leq \hat{p} - p \leq c)$$

Sendo  $\hat{p}$ , estimador para  $p$ , dado por:

$$\hat{p} = \frac{S_n}{n}.$$

Dividindo ambos os membros da desigualdade pelo desvio-padrão amostral, dado por

$$\sqrt{\frac{pq}{n}},$$

com  $q = 1 - p$ . obteremos:

$$\mathbb{P}(|\hat{p} - p| \leq c) = \mathbb{P}(-c \leq \hat{p} - p \leq c) = \mathbb{P}\left(\frac{-c}{\sqrt{\frac{pq}{n}}} \leq \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \leq \frac{c}{\sqrt{\frac{pq}{n}}}\right) = 0,95 \quad (3.2.9)$$

Se  $X_i, i = 1, \dots, n$ , uma variável aleatória de Bernoulli, tal que  $\mathbb{E}(X_i) = p$  e  $\text{VAR}(X_i) = p \cdot (1 - p) = pq$  e  $S_n = X_1 + X_2 + \dots + X_n$ . o TCL afirma que para um número  $n$  finito de Variáveis Aleatórias, a distribuição de probabilidade de  $Z$ , tal que:

$$Z_n = \frac{X_1 + X_2 + \dots + X_n - n \cdot \mu}{\sqrt{n} \cdot \sigma} = \frac{S_n - n \cdot p}{\sqrt{n} \cdot \sqrt{pq}} = \frac{\frac{S_n}{n} - \frac{np}{n}}{\frac{\sqrt{n} \cdot \sqrt{pq}}{n}} = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}}$$

converge para uma distribuição normal com parâmetros  $\mu = 0$  e  $\sigma^2 = 1$ .

Vamos, então, fazer a seguinte substituição, em 3.2.9:

$$Z_n = \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}},$$

$$\mathbb{P}\left(\frac{-c}{\sqrt{\frac{pq}{n}}} \leq Z_n \leq \frac{c}{\sqrt{\frac{pq}{n}}}\right) = 0,95 \quad (3.2.10)$$

Assim como foi feito na subseção anterior, para usarmos a tabela da figura 7, precisamos obter uma desigualdade no formato  $\mathbb{P}(Z_n \leq z)$ . Elaborando a desigualdade 3.2.10, obtemos o seguinte resultado:

$$\mathbb{P}\left(Z_n \leq \frac{c}{\sqrt{\frac{pq}{n}}}\right) = \frac{0,95 + 1}{2} = \frac{1,95}{2} = 0,9750$$

Na tabela da figura 7, podemos ver que o valor de  $z$  que satisfaz

$$\mathbb{P}(Z_n \leq z) = 0,9750$$

é 1,96. Logo:

$$\frac{c}{\sqrt{\frac{pq}{n}}} = 1,96 \Rightarrow c = 1,96 \cdot \sqrt{\frac{pq}{n}}$$

Voltaremos agora em 3.2.9 e substituiremos  $\frac{c}{\sqrt{\frac{pq}{n}}}$  por 1,96, e chegaremos a um intervalo que contenha o verdadeiro valor de  $p$  com 95% de chance.

$$\mathbb{P}\left(-1,96 \leq \frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \leq 1,96\right) = \mathbb{P}\left(-1,96 \cdot \sqrt{\frac{pq}{n}} \leq \hat{p} - p \leq 1,96 \cdot \sqrt{\frac{pq}{n}}\right)$$

$$= \mathbb{P} \left( -\hat{p} - 1,96 \cdot \sqrt{\frac{pq}{n}} \leq -p \leq -\hat{p} + 1,96 \cdot \sqrt{\frac{pq}{n}} \right) = \mathbb{P} \left( \hat{p} - 1,96 \cdot \sqrt{\frac{pq}{n}} \leq p \leq \hat{p} + 1,96 \cdot \sqrt{\frac{pq}{n}} \right)$$

Logo, o intervalo de confiança para  $p$  com índice de confiança  $\gamma = 95\%$  é:

$$IC(p; 0,95) = \left[ \hat{p} - 1,96 \cdot \sqrt{\frac{pq}{n}}; \hat{p} + 1,96 \cdot \sqrt{\frac{pq}{n}} \right] \quad (3.2.11)$$

Podemos observar que  $p$  aparece nos extremos do intervalo 3.2.11, pois a variância da Binomial depende de  $p$ . Não conhecemos seu valor, e justamente, queremos um intervalo de confiança para esse parâmetro. Há duas maneiras de contornar esse problema. A primeira seria trocar  $p$  pela estimativa de  $\hat{p}$  fornecida, que segue abaixo. a outra opção seria usar o valor máximo da função  $p(1-p)$ , descrita abaixo também.

1. Trocaremos  $p$  pelo seu estimador  $\hat{p}$  e  $q$  por  $\hat{q} = 1 - \hat{p}$ , obtendo o seguinte intervalo:

$$IC(p; 0,95) = \left[ \hat{p} - 1,96 \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}}; \hat{p} + 1,96 \cdot \sqrt{\frac{\hat{p}\hat{q}}{n}} \right]$$

2. Vamos substituir  $pq$  pelo seu valor máximo.

O valor máximo que  $pq = p(1-p) = p - p^2$  atinge é calculado encontrando a sua derivada e igualando o resultado a zero.

$$\frac{d}{dp}(p - p^2) = 1 - 2p = 0 \Rightarrow 1 = 2p \Rightarrow p = \frac{1}{2} \Rightarrow q = 1 - p = \frac{1}{2}$$

Logo,  $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ . Então:

$$\begin{aligned} pq \leq \frac{1}{4} &\Rightarrow \frac{pq}{n} \leq \frac{1}{4n} \Rightarrow \sqrt{\frac{pq}{n}} \leq \sqrt{\frac{1}{4n}} \Rightarrow -1,96 \cdot \sqrt{\frac{pq}{n}} \geq -1,96 \cdot \sqrt{\frac{1}{4n}} \\ &\Rightarrow \hat{p} - 1,96 \cdot \sqrt{\frac{pq}{n}} \geq \hat{p} - 1,96 \cdot \sqrt{\frac{1}{4n}} \end{aligned}$$

Logo:

$$\hat{p} - 1,96 \cdot \sqrt{\frac{1}{4n}} \leq \hat{p} - 1,96 \cdot \sqrt{\frac{pq}{n}} \quad (3.2.12)$$

Por outro lado:

$$\sqrt{\frac{pq}{4n}} \leq \sqrt{\frac{1}{4n}} \Rightarrow 1,96 \cdot \sqrt{\frac{pq}{n}} \leq 1,96 \cdot \sqrt{\frac{1}{4n}}$$

$$\hat{p} + 1,96 \cdot \sqrt{\frac{pq}{n}} \leq \hat{p} + 1,96 \cdot \sqrt{\frac{1}{4n}} \quad (3.2.13)$$

De 3.2.12 e 3.2.13, chegaremos ao seguinte intervalo:

$$\hat{p} - 1,96 \cdot \sqrt{\frac{1}{4n}} \leq \hat{p} - 1,96 \cdot \sqrt{\frac{pq}{n}} \leq Z_n \leq \hat{p} + 1,96 \cdot \sqrt{\frac{pq}{n}} \leq \hat{p} + 1,96 \cdot \sqrt{\frac{1}{4n}}$$

$$\Rightarrow \hat{p} - 1,96 \cdot \sqrt{\frac{1}{4n}} \leq Z_n \leq \hat{p} + 1,96 \cdot \sqrt{\frac{1}{4n}}$$

$$\Rightarrow IC(p; 0,95) = \left[ \hat{p} - 1,96 \cdot \sqrt{\frac{1}{4n}}, \hat{p} + 1,96 \cdot \sqrt{\frac{1}{4n}} \right] \quad (3.2.14)$$

Este intervalo é chamado conservador, e contém o intervalo de confiança dado por 3.2.11. Então, concluímos que se  $p$  pertence ao intervalo dado em 3.2.11 com 95% de chance, tb pertencerá ao intervalo conservador 3.2.14. O Exemplo abaixo foram usados dados fictícios, com o objetivo de fazer simulações.

**Exemplo 3.2.6.** Vamos supor que, no ano de 2020, com a paralisação do ensino presencial, a equipe responsável pela educação no Estado de São Paulo, preocupada com os problemas psicológicos que o ensino online pode acarretar, resolveu fazer uma pesquisa com pais e professores das escolas públicas estaduais, perguntando a opinião sobre a volta às aulas, desde que tomados todos os cuidados para evitar a contaminação pelo COVID-19. Numa determinada escola, de todos os 5000 pesquisados que responderam, apenas 2000 concordaram com a volta as aulas. Alguns, por algum motivo, não responderam.

Vamos encontrar um intervalo de confiança para  $p$ , proporção de pais e professores que receberam o formulário de pesquisa, usando os dois métodos já vistos, para a situação descrita, em que 2000 pessoas responderam que concordam com a volta as aulas, para um nível de confiança igual a 95%.

O valor de  $\hat{p}$  é dado por:

$$\hat{p} = \frac{2000}{5000} = 0,40 \Rightarrow \hat{q} = 0,60$$

- Trocando  $p$  por  $\hat{p}$  e  $q$  por  $\hat{q}$ : Temos o seguinte intervalo de confiança:

$$IC(p, 0,95) = \left[ 0,40 - 1,96 \cdot \sqrt{\frac{0,4 \cdot 0,6}{5000}}; 0,40 + 1,96 \cdot \sqrt{\frac{0,4 \cdot 0,6}{5000}} \right]$$

$$IC(p, 0,95) = [0,40 - 1,96 \cdot 0,0069; 0,40 + 1,96 \cdot 0,0069] = [0,40 - 0,0135; 0,40 + 0,0135]$$

$$\therefore IC(p, 0,95) = [0,3865; 0,4135] = [38,65\%; 41,35\%]$$

- como fica o intervalo conservador:

$$IC(p, 0,95) = \left[ 0,40 - 1,96 \cdot \sqrt{\frac{1}{4 \cdot 5000}}; 0,40 + 1,96 \cdot \sqrt{\frac{1}{4 \cdot 5000}} \right]$$

$$= [0,40 - 1,96 \cdot 0,007; 0,40 + 1,96 \cdot 0,007] = [0,40 - 0,0137; 0,40 + 0,0137] = [0,3863; 0,4137]$$

$$\therefore IC(p; 0,95) = [38,63\%; 41,37\%]$$

Mesmo usando métodos diferentes para construirmos os intervalos de confiança, chegamos em resultados muito próximos. No segundo caso, houve um aumento na amplitude do intervalo, sendo portanto um intervalo mais conservador, ou seja, é menos provável que o valor de  $p$  fique fora do intervalo.

Podemos concluir que o verdadeiro valor de  $p$  pertence ao intervalo  $[38,63;41,37]$ .

## 3.3 Teste de Hipóteses

### 3.3.1 Teste de Hipóteses para a média, com variância conhecida

Na subseção 3.2.3, construímos intervalos de confiança para inferirmos sobre um parâmetro populacional. Testes de hipóteses também podem ser usados para fazer inferência a respeito de parâmetros populacionais, porém partimos de uma afirmação sobre o seu valor, e verificamos se é correta, através de cálculos com valores obtidos da amostra. Tal afirmação recebe o nome de hipótese.

Seja uma variável aleatória  $X$ , com média  $\mu$  e desvio padrão  $\sigma$ . Nessa subseção, faremos teste para o parâmetro  $\mu$ , que é desconhecido. Chamamos esse valor de  $\mu_0$ . Serão testadas duas hipóteses, que chamamos de Hipótese Nula ( $H_0$ ) e Hipótese Alternativa ( $H_1$ ). tal que:

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu > \mu_0$$

ou

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu < \mu_0,$$

ou ainda

$$H_0 : \mu = \mu_0$$

$$H_1 : \mu \neq \mu_0.$$

O dois primeiros pares de hipóteses são o que se chama de testes unilaterais (teste unilateral à direita, no primeiro par e, à esquerda, no segundo) e o terceiro par de hipóteses é chamado teste bilateral.

Tomemos uma amostra  $X_1, X_2, \dots, X_n$ . A variável aleatória  $\bar{X}$ , tal que:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

é chamada de **Teste Estatístico**. Teste estatístico, é uma função da amostra, portanto uma variável aleatória, cuja estatística é usada para decidirmos se aceitamos ou rejeitamos a hipótese nula.

Em testes de hipóteses podem ocorrer os seguintes erros:

- Erro do tipo *I*: Acontece quando rejeitamos a hipótese nula, estando ela correta. A probabilidade de ocorrência do erro do tipo *I* recebe o nome de nível de significância e é representado pela letra  $\alpha$ .
- Erro do tipo *II*: Acontece quando aceitamos a hipótese nula, sendo falsa. A probabilidade de ocorrência do erro do tipo *II* é representada pela letra  $\beta$ .

A tabela abaixo facilita a compreensão sobre os tipos de erros que ocorrem em um teste de hipótese.

	$H_0$ é verdadeira	$H_0$ é falsa
Rejeitamos $H_0$	Erro tipo <i>I</i>	Decisão correta
Não rejeitamos $H_0$	Decisão correta	Erro tipo <i>II</i>

Tabela 5 – Tipos de erros em um teste de hipóteses

Construiremos um intervalo, chamado de **Região Crítica**, ou região de rejeição de  $H_0$ , que dependerá se o teste for unilateral à esquerda, unilateral à direita ou bilateral. Assim:

$$RC = ]-\infty, z_c]$$

ou

$$RC = [z_c, \infty[$$

ou ainda,

$$RC = ]-\infty, z_1] \cup [z_2, \infty[$$

Nos dois primeiros intervalos, a Hipótese Nula é rejeitada quando a estimativa do Teste Estatístico for maior ou menor que  $z_c$ , respectivamente. No terceiro caso,  $H_0$  é rejeitada quando a estimativa estiver em um dos intervalos indicados. Por meio do exemplo 3.3.1, usando dados fictícios, vamos trabalhar todos os conceitos vistos nessa subsecção.

**Exemplo 3.3.1.** Supomos que uma professora de 7<sup>o</sup> ano estava desenvolvendo com seus alunos o conteúdo "Razões Especiais- Índice de Massa Corporal (IMC)". Então, pediu que seus 100 alunos dos sétimos anos A, B e C trouxessem anotados seus pesos e alturas.

Como a professora já havia visto a tabela da figura 9, (EFDEPORTES.COM, 2010), ela resolveu fazer a média dos IMC dos meninos e das meninas e comparar com os valores da tabela 9. O que chamou a sua atenção foi que a média das meninas era bem superior à média da tabela, fazendo com que ela achasse que 18,3, que é o valor encontrado na tabela para as alunas de 12 anos (idade com que os alunos estão no 7<sup>o</sup> ano), muito baixo, pois seus cálculos mostravam que o IMC médio das alunas era de 22,1.

Tabela 3. Valores médios e desvio padrão do Índice de Massa Corporal para a idade

Idades	Masculino	Feminino
10	17,5±3,74	18,0±3,35
11	17,8±3,25	17,9±3,09
12	19,0±3,68	18,3±3,49
13	18,7±3,11	19,3±3,60
14	18,5±2,55	19,6±3,31

Figura 9 – Tabela com Índice de Massa Corporal de adolescentes dos sexos feminino e masculino, retirada do site ([EFDEPORTES.COM](http://EFDEPORTES.COM), 2010).

Vamos modelar o problema e descobrir se, a partir dos dados da amostra, a professora tem razão em duvidar que o IMC médio das meninas seja de 18,3. Consideremos, então, as duas hipóteses:

$$H_0 : \mu = 18,3$$

$$H_1 : \mu > 18,3$$

Vamos então encontrar o valor de  $z_c$ , com  $\alpha = 5\%$ . Como  $\alpha$  é a probabilidade de rejeitar  $H_0$ , sendo ela verdadeira, então:

$$\alpha = \mathbb{P}(\bar{X} > t_c; \mu = 18,3)$$

$$0,05 = \mathbb{P}\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{t_c - \mu}{\frac{\sigma}{\sqrt{n}}}; \mu = 18,3\right) \quad (3.3.1)$$

Já sabemos, pelo Teorema Central do Limite, que para  $n$  suficientemente grande, a distribuição da variável aleatória:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}},$$

converge para a distribuição de probabilidade de uma Variável Aleatória  $Z_n$  Normal com parâmetros  $\mu = 0$  e  $\sigma = 1$ . No site ([EFDEPORTES.COM](http://EFDEPORTES.COM), 2010), na descrição da tabela 9, podemos ver que  $\sigma = 3,49$ . Vamos, agora, voltar em 3.3.1 e encontrar o valor de  $t_c$ , substituindo o valor de  $\mu$  por 18,3 e  $n$  por 100, que é o número de alunos da amostra,

$$0,05 = \mathbb{P}\left(Z > \frac{z_c - 18,3}{\frac{3,49}{10}}; \mu = 18,3\right)$$

$$\Rightarrow 0,05 = \mathbb{P}\left(Z_n > \frac{z_c - 18,3}{0,349}; \mu = 18,3\right).$$

$$\mathbb{P}\left(Z_n < \frac{z_c - 18,3}{0,349}\right) = 0,95 = 1 - \mathbb{P}\left(Z_n \leq \frac{z_c - 18,3}{0,349}; \mu = 18,3\right) = 0,05$$

$$\Rightarrow \mathbb{P}\left(Z \leq \frac{z_c - 18,3}{0,349}\right) = 0,95$$

temos:

$$\frac{z_c - 18,3}{3,56} = 1,645 \Rightarrow z_c - 18,3 = 1,645 \cdot 0,349 = 0,574$$

$$\Rightarrow z_c = 0,574 + 18,3 = 18,874$$

Concluimos que a região de rejeição de  $H_0$  é dada por:

$$RC = [18,874, \infty[,$$

e como o valor do IMC médio feminino encontrado pela professora foi 22,1, que pertence a  $RC$  rejeitamos  $H_0$ .

### 3.3.2 Teste de Hipóteses para a proporção

Neste tipo de teste a A.A.S. é formada por  $X_i \sim \text{Bernoulli}(p)$ . As hipóteses  $H_0$  e  $H_1$  sobre os valores de  $p$  são mostradas abaixo:

$$H_0 : p = p_0$$

$$H_1 : p > p_0,$$

ou

$$H_0 : p = p_0$$

$$H_1 : p < p_0,$$

ou ainda,

$$H_0 : p = p_0$$

$$H_1 : p \neq p_0.$$

Como na subseção anterior, os dois primeiros pares de hipóteses são de testes unilaterais e o terceiro, de um teste bilateral. Construiremos o intervalo  $RC$ , calculando um valor para  $p_c$  (chamado  $p$  crítico e  $p_c \neq p_0$ ), tal que:

$$RC : [p_c, \infty[,$$

ou

$$RC : ]-\infty, p_c],$$

ou ainda

$$RC : ]-\infty, p_1] \cup [p_2, \infty[.$$

O exemplo seguinte foi formulado pela autora e os valores são fictícios.



**Exemplo 3.3.2.** Suponha que alguns funcionários de uma determinada escola tem observado que os alunos, durante o intervalo, tem feito sua refeição com alimentos de baixo valor nutricional e, conversando com a equipe escolar, decidiram fazer uma pesquisa com alunos, pais e equipe escolar sobre os alimentos vendidos na cantina. Elaboraram, então, um formulário com a seguinte questão: você concorda que a cantina passe a vender apenas alimentos mais saudáveis, como suco e lanches naturais? A equipe escolar resolveu tomar a atitude de vender alimentos saudáveis se 60% das pessoas entrevistadas concordassem com a troca. Enviaram o formulário para 1000 pessoas escolhidas aleatoriamente. No entanto, apenas 830 responderam e, dentre eles, 450 apoiaram a mudança.

Tomemos então uma amostra aleatória  $X_1, X_2, \dots, X_n$ , com  $n = 830$ , que é o número de pessoas que responderam ao formulário. Cada  $X_i$ , assume 1 ou 0, caso a resposta seja favorável ou não à mudança, respectivamente. Logo as  $X_{is}$  são Bernoullis de parâmetro  $p$  desconhecido, que corresponde à proporção de professores e pais que concordariam com a mudança no método de ensino. Para saber se ainda haveria chance de não rejeitar  $H_0$ , vamos testar a hipótese de  $p$  ser menor que 0,6 ao nível de significância de 5%

$$H_0 : p = 0,6$$

$$H_1 : p < 0,6$$

Assim, como foi feito na subseção 3.3.1, encontraremos agora o valor de  $p_c$ , que satisfaça:

$$\alpha = \mathbb{P}(\hat{p} < p_c; p = 0,6) = \mathbb{P}\left(\frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} < \frac{p_c - p}{\sqrt{\frac{pq}{n}}}; p = 0,6\right)$$

O Teorema Central do Limite afirma que, para valores suficientemente grandes de  $n$ , a distribuição de probabilidade de:

$$\frac{\hat{p} - p}{\sqrt{\frac{pq}{n}}} \tag{3.3.2}$$

converge para a distribuição de uma variável aleatória  $Z_n \sim N(0, 1)$ . Assim, temos que:

$$\begin{aligned} \Rightarrow 0,05 &= \mathbb{P}\left(Z_n < \frac{p_c - 0,6}{\sqrt{\frac{0,6 \cdot 0,4}{830}}}; p = 0,6\right) = \mathbb{P}\left(Z_n < \frac{p_c - 0,6}{\sqrt{\frac{0,24}{830}}}; p = 0,6\right) \\ &= \mathbb{P}\left(Z_n < \frac{p_c - 0,6}{0,016}; p = 0,5\right) = \mathbb{P}\left(Z_n > -\left(\frac{p_c - 0,6}{0,017}\right); p = 0,6\right) \\ &\Rightarrow \mathbb{P}\left(Z_n < -\left(\frac{p_c - 0,6}{0,017}\right); p = 0,6\right) = 0,95 \\ &\Rightarrow -\left(\frac{p_c - 0,6}{0,017}\right) = 1,645 \Rightarrow -p_c + 0,6 = 0,029 \Rightarrow -p_c = -0,571 \end{aligned}$$

$$\Rightarrow p_c = 0,571.$$

A Região crítica é escrita como:

$$RC = ]-\infty; 0,571]$$

Como  $\hat{p} = \frac{450}{830} = 0,54$  pertence a região crítica, há evidências que podemos rejeitar  $H_0$ , com 5% de probabilidade de estarmos rejeitando uma hipótese verdadeira. Ou seja, há indícios de que não se deve modificar o que já é vendido na cantina, segundo a equipe escolar.

---

# ENCONTRANDO A NORMAL NA SALA DE AULA: EXEMPLO DE ROTEIRO DE AULA

---

Neste capítulo, será apresentado um roteiro de aula, especificando quais são os objetivos, competências, habilidades, atividades desenvolvidas e avaliação que podem ser aplicadas na 3ª série do Ensino Médio.

## 4.1 Roteiro de Ensino

1. **Área do conhecimento:** Matemática e suas tecnologias - Probabilidade e Estatística
2. **Título/tema:** Curva Normal
3. **Série:** 3ª série do Ensino Médio
4. **Competências desenvolvidas da BNCC nas aulas:**

- **Competência 3**

Utilizar estratégias, conceitos e procedimentos matemáticos, em seus campos - Aritmética, Álgebra, Grandezas e medidas, Geometria, Probabilidade e Estatística - 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. (MEC, 2017)

- **Competência 4**

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. (MEC, 2017)

## 5. Habilidades:

- "(EM13MAT316) Resolver e elaborar problemas, em diferentes contextos que envolvem cálculo e interpretação das medidas de tendência central (média, moda e mediana) e das de dispersão (amplitude, variância e desvio padrão)"(MEC, 2017, p.528).
- "(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"(MEC, 2017, p.539).

6. **Local da Aula:** Sala equipada com computador para uso do professor, com a ferramenta EXCEL e um projetor.

7. **Duração da atividade:** 5 aulas aproximadamente.

- **Objetivo Geral:**

O aluno deve ser capaz de calcular probabilidades em uma população normalmente distribuída.

- **Objetivos Específicos**

- Construir a curva normal;
- Visualizar a média e desvio-padrão na curva-normal;
- Mostrar as características da curva normal.

## DADOS DA AULA

### 1. Conhecimentos prévios trabalhados:

Representação gráfica, cálculo de probabilidades, média, variância, desvio padrão.

É importante que o professor faça uma avaliação Diagnóstica, buscando atividades em livros didáticos sobre os conteúdos elencados acima. Caso seja necessário, o professor deve fazer uma retomada dos assuntos antes de iniciar as atividades sugeridas abaixo.

### 2. METODOLOGIA

#### a) INTRODUÇÃO

A metodologia que será sugerida para o desenvolvimento do assunto **Curva Normal** é chamada **Metodologia de resolução de problema**, seguindo as etapas que estão descritas por (BINOTTO, 2019), e aparecem na Introdução do presente trabalho. O assunto será introduzido a partir do problema matemático gerador 4.1.1, adaptado do livro didático Ser Protagonista (SMOLE, 2020), como ponto de partida para se ensinar os conceitos envolvidos.

**b) DESENVOLVIMENTO****• PREPARAÇÃO DO PROBLEMA**

**Exemplo 4.1.1.** (SMOLE, 2020) Imagine a situação a seguir.

Certo jogo de dados para duas pessoas funciona da seguinte maneira.

- em cada jogada, lançam-se dois dados, um azul e um laranja, e somam-se os pontos de cada um;
- o jogador A ganha se a soma for 6, 7, 8 ou 9;
- o jogador B ganha se a soma for 2, 3, 4, 5, 10, 11 ou 12.

Se você fosse apostar na vitória de um deles, qual escolheria?

**• LEITURA DO PROBLEMA INDIVIDUAL E EM GRUPO** Inicialmente o professor sugerirá que os alunos façam uma leitura individual e depois, em grupos de 3 alunos, será feita a leitura em conjunto.

**• RESOLUÇÃO DO PROBLEMA** Em grupos, sem dúvidas quanto ao enunciado do problema, com uma atitude de cooperação, deverão resolver o problema.

**• PROFESSOR: OBSERVA E INCENTIVA**

O professor, nessa etapa, deixa de ser o transmissor de conteúdo e passa a ter o papel de mediador. Ele deve também incentivar os alunos a trocarem idéias, buscar os conhecimentos prévios, usar as técnicas operatórias conhecidas. Além disso, ele deverá ter um olhar diferenciado para aqueles que possuem mais dificuldades, ajudando-os quando necessário.

**• REGISTRO DAS RESOLUÇÕES NA LOUSA**

Após todos os grupos terem concluídas as resoluções, o professor incentivará um representante de um dos grupos a fazer os registros da sua solução na lousa, estando certa ou errada.

**• PLENÁRIA**

Na próxima etapa dessa metodologia de resolução de problemas, chamada plenária, os alunos irão confrontar a resolução apresentada com as realizadas no seu grupo.

**• BUSCA DE CONSENSO**

O professor e alunos chegam a um consenso sobre o resultado correto (CASTANHÊDE, 2015).

**• FORMALIZAÇÃO DO CONTEÚDO** O professor, então, formalizará a resolução do exercício, mostrando inicialmente a figura 10 com os resultados de todas as somas possíveis.

		DADO LARANJA					
		1	2	3	4	5	6
D A D O  A Z U L	1	2	3	4	5	6	7
	2	3	4	5	6	7	8
	3	4	5	6	7	8	9
	4	5	6	7	8	9	10
	5	6	7	8	9	10	11
	6	7	8	9	10	11	12

Figura 10 – tabela com todos os resultados possíveis da soma dos números que estão nas faces superiores dos dados.

A figura 11 mostra as somas possíveis e com que frequências elas aparecem na figura 10.

Soma dos números das faces superiores de dois dados	frequencia
2	1
3	2
4	3
5	4
6	5
7	6
8	5
9	4
10	3
11	2
12	1

Figura 11 – tabela com todos os resultados possíveis da soma que aparecem na tabela 10, com a frequência com que elas aparecem

Nesse momento, o professor pedirá que os alunos, em grupos e buscando seus conhecimentos prévios calculem as seguintes probabilidades:

- Probabilidade da soma dar 6.
- Probabilidade da soma dar 7.
- Probabilidade da soma dar 8.
- Probabilidade da soma dar 9.
- Qual a probabilidade da soma ser igual a 6 a 7, a 8 ou a 9?
- Probabilidade da soma dar 2.
- Probabilidade da soma dar 3.
- Probabilidade da soma dar 4.
- Probabilidade da soma dar 5.
- Probabilidade da soma dar 10.

- Probabilidade dar soma dar 11
- Probabilidade dar soma dar 12
- Qual a probabilidade da soma ser igual a 2 a 3, a 4, a 5, a 10, a 11 ou a 12?
- Observando os resultados dos itens anteriores, qual dos dois jogadores, A ou B, tem maior probabilidade de vencer o jogo?

### CONCLUSÃO DO PROBLEMA

Assim que os alunos terminarem a resolução, o professor poderá pedir que os alunos socializem a resolução do ítem  $n$  e as conclusões a que chegaram, comparando com o que eles haviam concluído antes da formalização do conteúdo. Ao final da resolução das atividades, eles devem ter chegado ao seguinte resultado: a probabilidade de o jogador A vencer o jogo  $\left(\frac{20}{36}\right)$ , é maior que a probabilidade de o jogador B vencer:  $\left(\frac{16}{36}\right)$ .

## 4.2 Introduzindo a Distribuição Normal

O professor, agora fará a representação gráfica da situação descrita acima, a partir da figura 11, com o auxílio da planilha EXCEL.

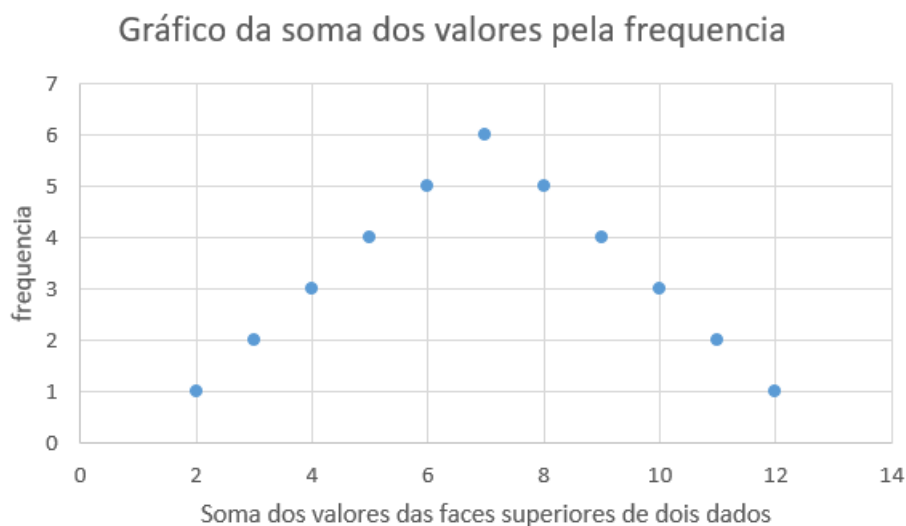


Figura 12 – Gráfico de dispersão relacionando a soma dos valores das faces superiores de dois dados e a frequência com que elas aparecem.

O professor poderá pedir que os alunos que, simulando (em grupos) o lançamento de 3 dados, façam a tabela de frequência e a representação gráfica, usando a ferramenta EXCEL, usando os computadores disponíveis na escola ou os seus respectivos smartphones. Ao propor essa atividade, o professor transfere o foco da atividade para os alunos, deixando, mais uma vez de ser o transmissor do conhecimento. Seu papel, nessa atividade, é o de

mediá-los. O produto final dessa atividade deverá um ser uma tabela e um gráfico parecido com as figuras 13 e 14.

soma dos 3 valores dos dados	frequencia
3	1
4	3
5	6
6	10
7	15
8	21
9	25
10	27
11	27
12	25
13	21
14	15
15	10
16	6
17	3
18	1

Figura 13 – Frequência das somas dos valores obtidos ao lançar três dados



Figura 14 – Gráfico de dispersão relacionando a soma dos valores das faces superiores de três dados e a frequência com que elas aparecem

O professor, então mostrará como ficaria a representação gráfica ao se lançar 4 dados.



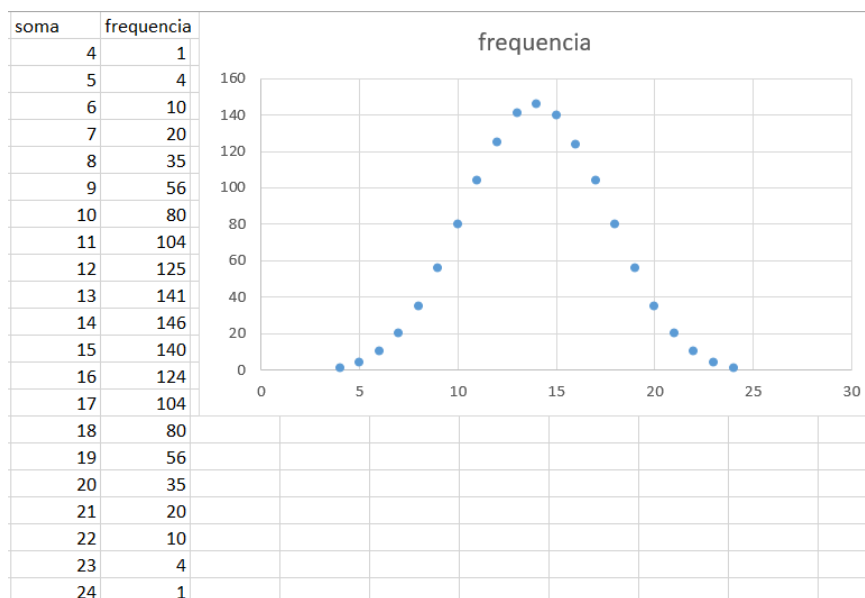


Figura 15 – Tabela e gráfico de linha relacionando a soma dos valores das faces superiores de quatro dados e a frequência com que elas aparecem

A partir da figura 15, o professor irá propor que os alunos calculem a média e o desvio-padrão, usando ou não, os recursos do EXCEL. O resultado esperado é Média  $\approx 14$  e desvio-padrão  $\approx 0,77$ . O professor então, comentará que, se uníssemos os pontos do gráfico 15, teríamos um gráfico parecido com a figura 16, que recebe o nome de **Curva Normal**. Tal curva aparece no estudo de características, como comprimento do pé, tempo de reação a um medicamento, número de filhos, etc (SMOLE, 2020).

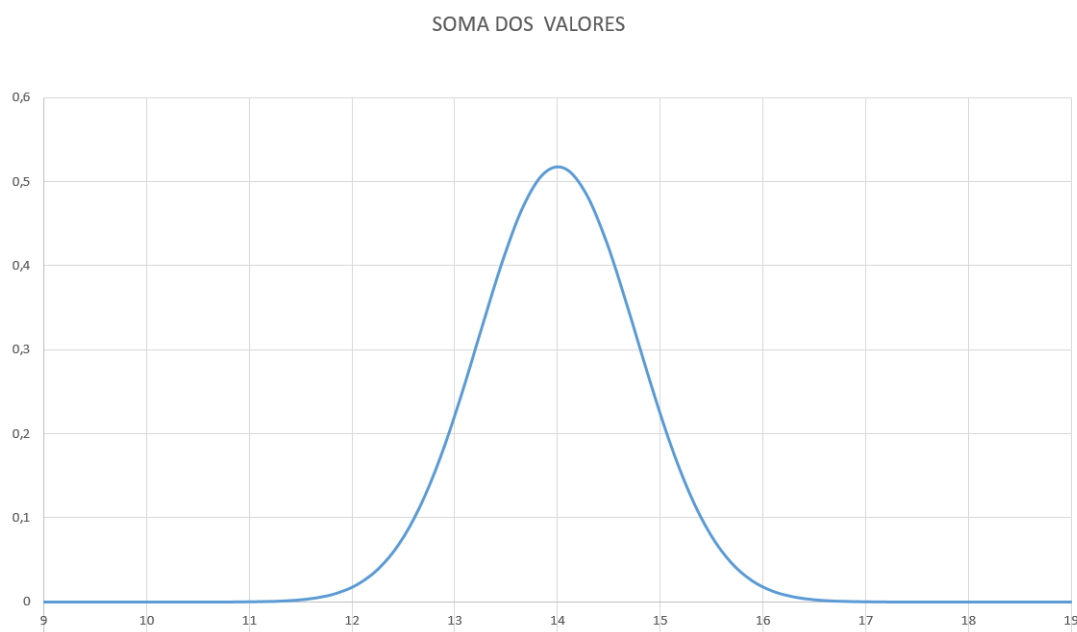


Figura 16 – Representação gráfica de uma distribuição normal com média 14 e desvio-padrão 0,77

Caso o professor tenha interesse em saber como se construir um gráfico da distribuição

normal utilizando os recursos do EXCEL, uma sugestão é assistir ao vídeo (GESTAO, 2021), referenciado no final do trabalho. Tal vídeo ensina como construir o gráfico de uma distribuição normal, dados a média e o desvio-padrão.

### 4.2.1 Características de uma Distribuição Normal

As características da distribuição normal serão discutidas a partir da figura 17.

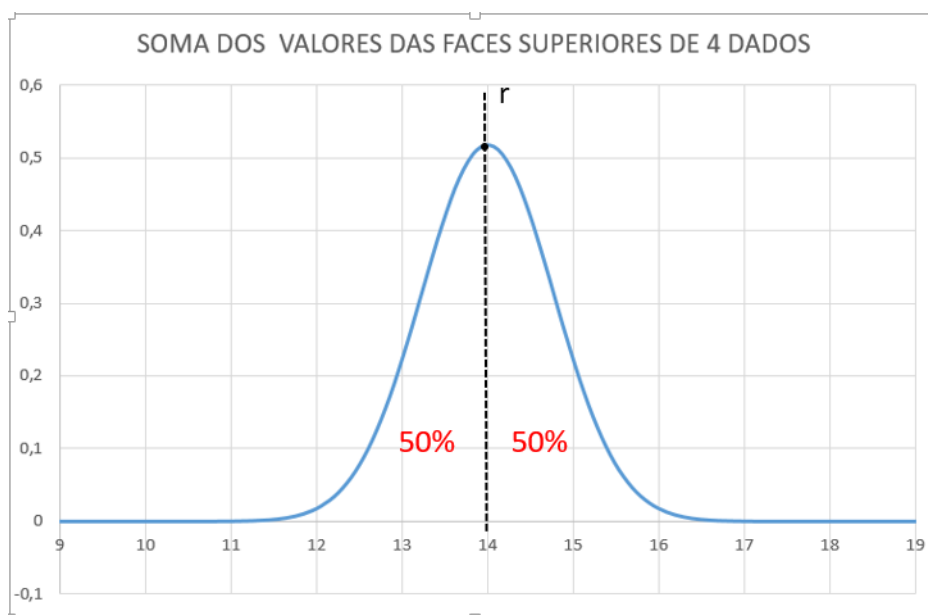


Figura 17 – Representação gráfica de uma distribuição normal para o estudo de suas características

- A curva tem a forma de sino.
- A curva é simétrica em relação à reta **r**.
- A reta **r** intercepta a curva normal no seu ponto máximo, passando pelo eixo **x** no ponto  $x = \mu$ , em que  $\mu$  é o valor médio.

Para discutir a influência da média e do desvio padrão, o professor mostrará 2 figuras. A figura 18 mostra três curvas com mesma média (0), e desvio-padrão com diferentes valores e a figura 19 mostra três curvas com mesmo valor de desvio padrão (1), variando a média. O vídeo citado anteriormente (GESTAO, 2021), dá ao professor uma idéia de como representar graficamente 2 curvas normais com diferentes médias.

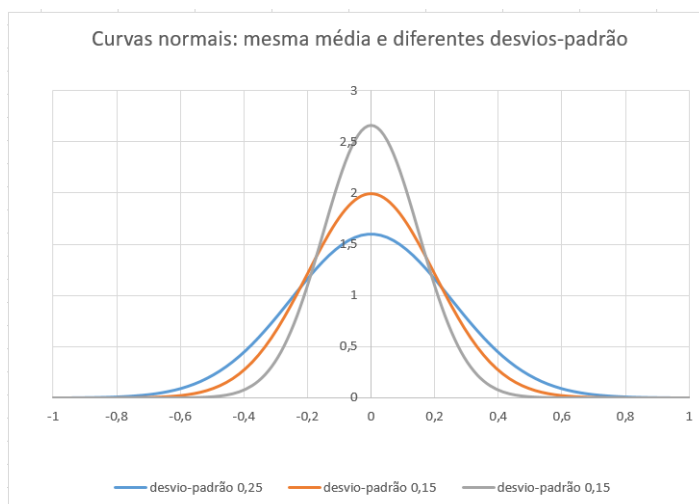


Figura 18 – Representação gráfica de três curvas normais com mesmas médias e desvios-padrão diferentes: 0,25, 0,2 e 0,15

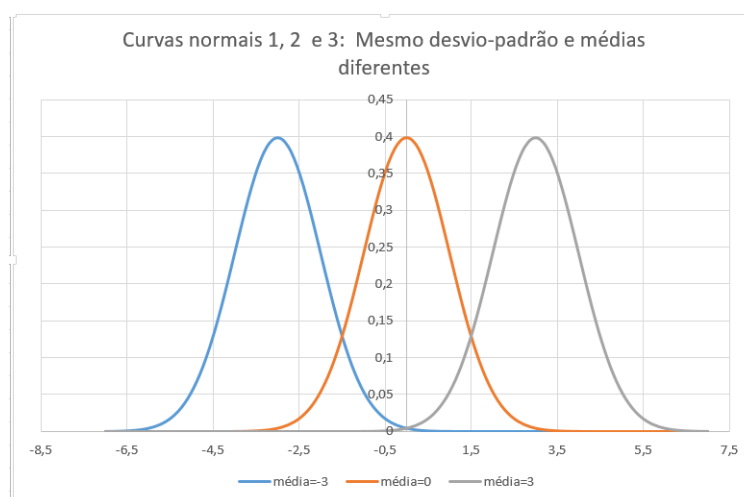


Figura 19 – Representação gráfica de três curvas normais com mesmos desvios-padrão e médias diferentes: -3, 0 e 3

O professor poderá, como atividade avaliativa, pedir que os seus alunos façam um texto respondendo as seguintes questões:

- Qual do comportamento da curva normal quando o desvio-padrão é constante, mas há uma variação na média?
- Modificando-se o desvio-padrão e tornando-se constante a média, como se comporta a curva normal?

### 4.2.2 Calculando Probabilidades

Inicialmente o professor deverá relacionar o cálculo da área sob a curva e o eixo  $x$  com o cálculo da probabilidade, mostrando para os alunos a figura 20.

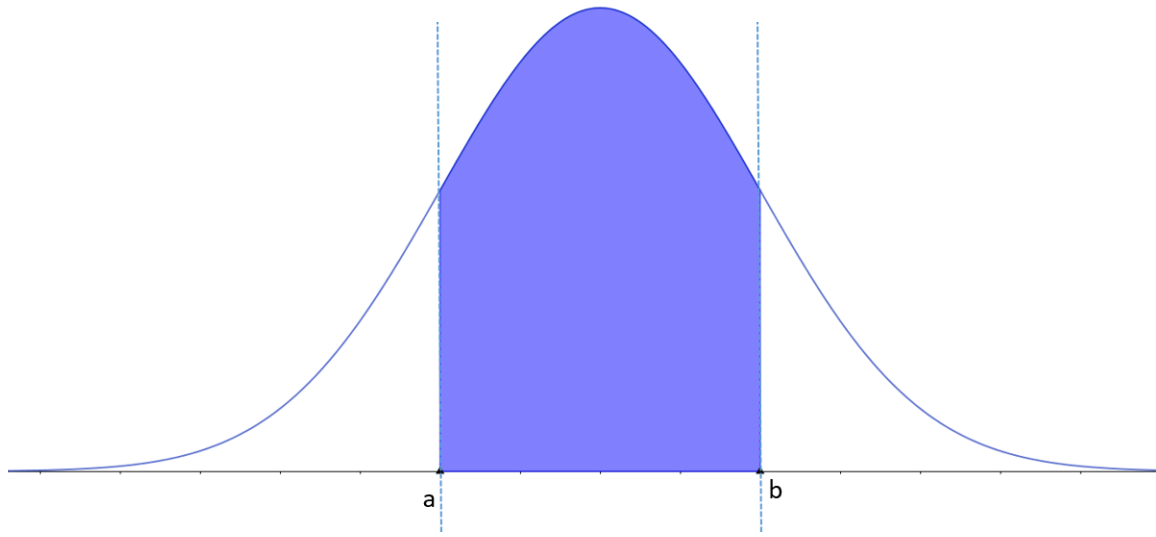


Figura 20 – A área entre a curva e a reta  $x$ , delimitada pelas retas  $x = a$  e  $x = b$ , é igual a  $\mathbb{P}(a \leq X \leq b)$ , em que  $X$  equivale a variável que descreve a característica de interesse.

O modo como se calculam probabilidades em uma população normalmente distribuída é tratado na dissertação de (CARVALHO, 2018) e no livro didático (SMOLE, 2020), por meio das seguintes probabilidades: usaremos, a partir de agora, a notação  $\mu$  é a média e  $\sigma$  é o desvio-padrão.

- 68% da distribuição pertence ao intervalo  $]\mu - \sigma, \mu + \sigma[$ , sendo 34% para o lado esquerdo da média e 34% para o lado direito.
- 96% da distribuição pertence ao intervalo  $]\mu - 2\sigma, \mu + 2\sigma[$ , sendo 48% para o lado esquerdo da média e 48% para o lado direito.
- 4% da distribuição pertence ao intervalo  $]-\infty, \mu - 2\sigma[ \cup ]\mu + 2\sigma, +\infty[$ , sendo 2% para o lado esquerdo de  $]-\infty, \mu - 2\sigma[$  e os outros 2% para o lado direito de  $]\mu + 2\sigma, +\infty[$ .
- 100% da distribuição pertence ao intervalo  $]-\infty, +\infty[$ . Isso implica que a área total sob a curva normal e o eixo  $x$  é igual a 1.

Vamos, por meio do exemplo 4.2.1, adaptado do exercício proposto por (SMOLE, 2020), calcular as probabilidades pedidas.

**Exemplo 4.2.1.** De acordo com um estudo, os "pesos" de jogadores de futebol profissional se distribuem de acordo com uma curva normal, com média de  $\mu = 74$  kg e desvio-padrão de  $\sigma = 4$  kg. Com base nessas informações, vamos calcular as seguintes probabilidades:

- Probabilidade de que ter um jogador de futebol profissional com mais de 78.  
Utilizando as propriedades sobre o cálculo de probabilidades, podemos calcular a probabilidade pedida, por meio da figura 21

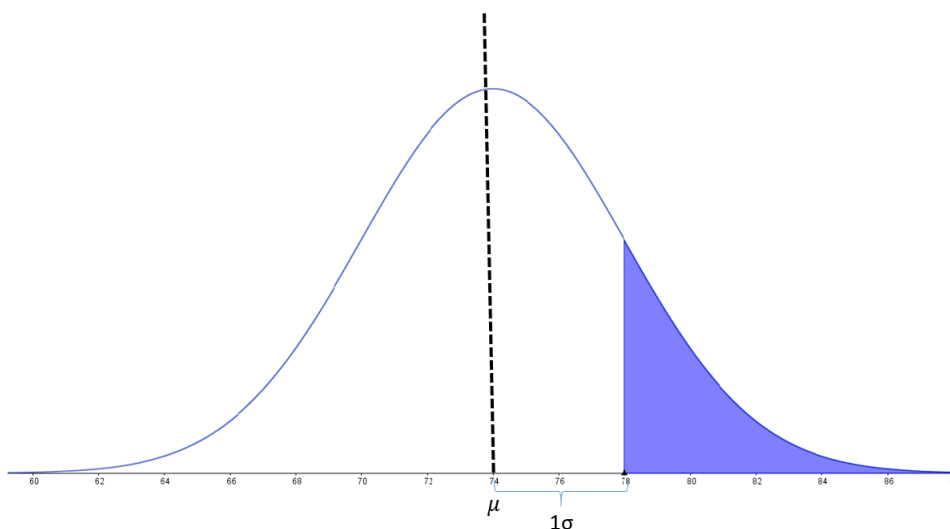


Figura 21 – Representação gráfica para o cálculo da probabilidade pedida no exemplo

#### 4.2.1

A probabilidade de um jogador com mais de 78 kg é de 16%.

- A probabilidade de um jogador ter menos que 80 kg.

Observando a figura 22, podemos verificar que de 74kg até 80 kg temos uma "distância" de  $1,5\sigma$ . Essa probabilidade não pode ser calculada usando as propriedades que já foram mencionadas anteriormente. Nesse caso, é conveniente que o professor ensine aos seus alunos como usar a tabela 7. Assim, eles poderão calcular qualquer probabilidade.

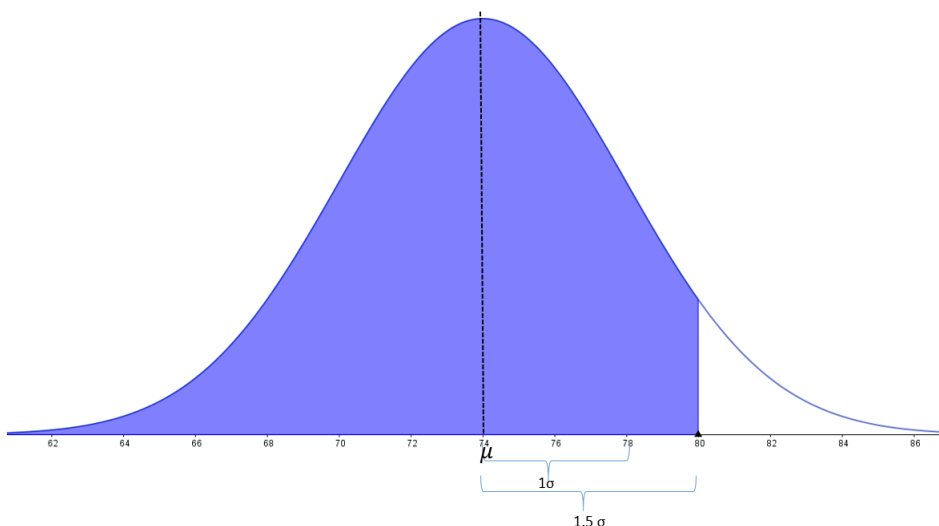


Figura 22 – Representação gráfica da probabilidade com que um jogador de futebol profissional tenha menos que 80 kg

Calcularemos, agora, a probabilidade de um jogador de futebol ter menos que 80 kg, ou seja,  $\mathbb{P}(X < 80)$ , em que  $X$  é o peso do jogador, e corresponde à área da parte

hachurada da figura 22. Para isso, vamos usar a tabela 7, o que só é possível quando a média é 0 e desvio padrão é 1. Para isso usá-la, devemos fazer a transformação:

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

e calcular:

$$\mathbb{P}(X < 80) = \mathbb{P}\left(\frac{X - 74}{4} < \frac{80 - 74}{4}\right) = \mathbb{P}(Z < 1,5)$$

A figura 23 é apenas uma representação gráfica com o objetivo de exemplificar, não condizendo com a realidade.

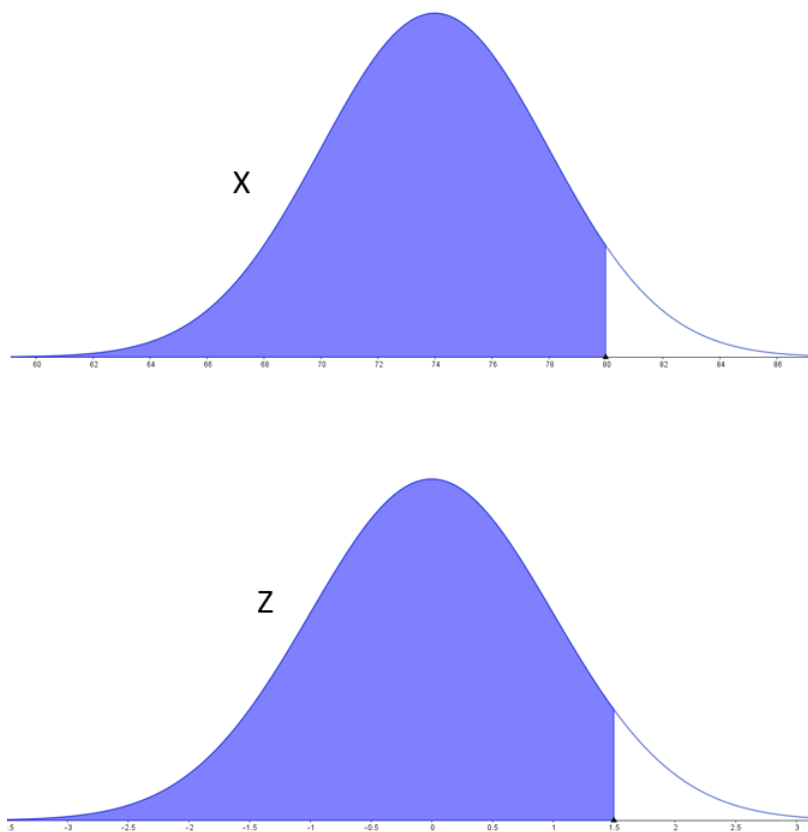


Figura 23 – Representação gráfica das distribuições de  $X$  e  $Z$ , em que  $X$  tem média 74 kg e desvio-padrão 4 kg e  $Z$  tem média 0 e desvio-padrão 1

Vamos agora encontrar, na tabela normal, a  $\mathbb{P}(Z < 1,5)$ . Observando a figura 24, foram traçados dois segmentos: o segmento azul, que passa por 1,5, que corresponde à unidade e primeira casa decimal do número que estamos interessados no cálculo de probabilidades e segmento vermelho, passando pelo 0, que é a segunda casa decimal de 1,5. O ponto de encontro desses dois segmentos nos fornece a probabilidade procurada, que é 0,9332 ou 93,32%.

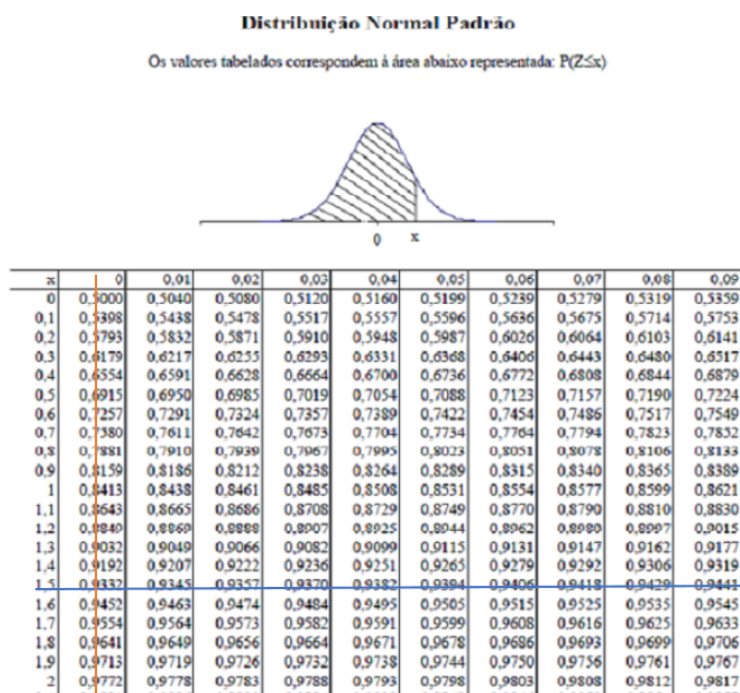


Figura 24 – Uso da tabela normal para o encontro da  $\mathbb{P}(Z < 1,5)$

- Supondo que queiramos saber qual a probabilidade de que, escolhido um jogador, ele tenha menos que 68kg, ou seja, queremos calcular:

$$\mathbb{P}(X < 68) = \mathbb{P}\left(\frac{X - 74}{4}\right) < \left(\frac{68 - 74}{4}\right) \Rightarrow \mathbb{P}(Z < -1,5)$$

Observando a figura 25, podemos concluir que:

$$\mathbb{P}(Z < -1,5) = \mathbb{P}(Z > 1,5) = 1 - \mathbb{P}(Z < 1,5)$$

A partir do ítem anterior, podemos concluir que:

$$\mathbb{P}(Z < -1,5) = 1 - 0,9332 = 0,0668 = 6,68\%$$

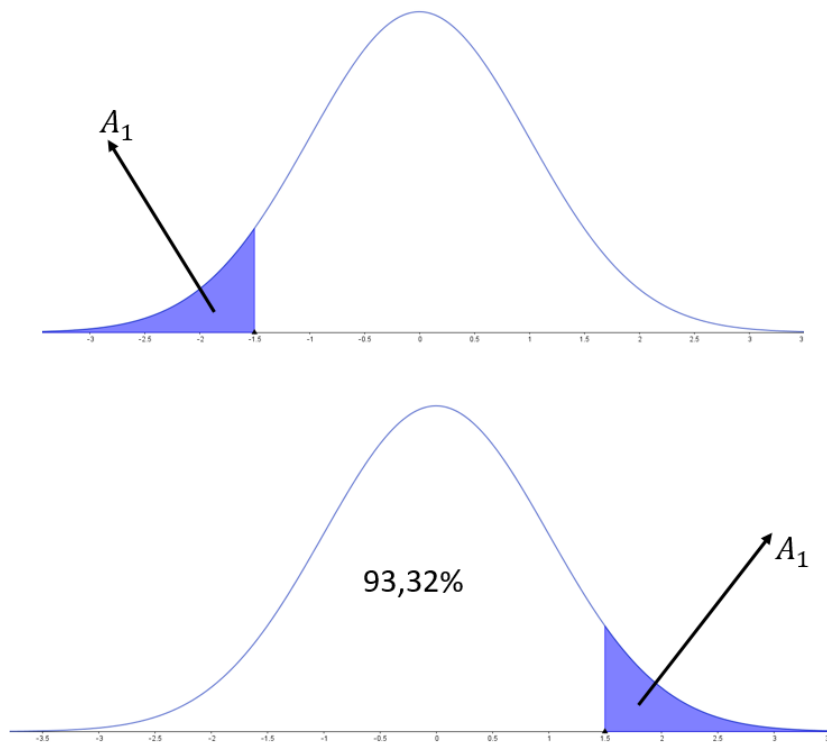


Figura 25 – Uso da tabela normal para o encontro da  $\mathbb{P}(Z < -1.5)$

Uma sugestão é que o professor consulte o livro (SMOLE, 2020), pois são propostas mais atividades envolvendo o cálculo de probabilidades em uma distribuição normal.

### 3. AVALIAÇÃO

A avaliação da aprendizagem deve ocorrer continuamente. Em todas as etapas da Metodologia sugerida nesse trabalho, o professor deve ser aquele que analisa, verifica os avanços dos alunos, auxilia aqueles que apresentam mais dificuldades em realizar as atividades propostas. Como função diagnóstica, a avaliação realizada pelo professor sobre a atitude de cada aluno o norteará para que este planeje uma outra metodologia capaz de sanar as dificuldades deles e atingir o objetivo descrito no início do capítulo: resolver atividades em que os alunos devem saber como resolver probabilidades em uma população normalmente distribuídas.



---

## CONSIDERAÇÕES FINAIS

---

Como professora do Ensino Básico na escola pública estadual, tive uma maior segurança ao elaborar as aulas devido ao aprofundamento dos conteúdos nas disciplinas do Profmat.

Por meio deste trabalho, pude pesquisar caminhos para me auxiliar e auxiliar os demais professores ao desenvolverem suas aulas sobre Probabilidade e Estatística. A busca por ferramentas tecnológicas, que tornam o ensino mais lúdico, foi feita com o objetivo de despertar o interesse dos alunos para esse conteúdo.

Em sala de aula, já tive a experiência de ensinar, com mais segurança, os meus alunos de Ensino Médio usando conteúdo deste trabalho: Distribuição Binomial.

Ao me atualizar sobre as Metodologias Ativas para escrever o capítulo 4, constatei que a Gamificação ([VALERIANI, 2021](#)) faz parte de um outro tipo de Metodologia que pode melhorar a qualidade do ensino básico de Matemática. Sendo assim, pretendo me aprofundar nessa modalidade de ensino, procurando por aplicativos que desenvolvam os conceitos de Estatística no Ensino Básico e por meio de formações continuadas, levar o produto desse estudo para auxiliar os professores a desenvolverem aulas diferentes das expositivas tradicionais.



## REFERÊNCIAS

---

---

ANDRADE, T. M. d. **Matemática interligada: estatística, combinatória e probabilidade**. São Paulo: Scipione, 2020. Citado na página 22.

BINOTTO, C. de O. **Ensino de Estatística Através da Metodologia de Resolução de Problemas - Uma Proposta Aplicada ao Ensino Médio**. Dissertação (Mestrado) — Universidade Tecnológica Federal do Paraná, Pato Branco, 2019. Citado nas páginas 23, 25 e 74.

BITTENCOURT, H. R.; VIALI, L. Contribuições para o ensino da distribuição normal ou curva de gauss em cursos de graduação. **SIPEM (Seminário Internacional de Pesquisa em Educação Matemática)**, 2006. Citado na página 19.

BONJORNO, J. R. **Prisma matemática: estatística, combinatória e probabilidade**. São Paulo: FTD, 2020. Citado na página 22.

BOSCO, G. G. **Introdução à Probabilidade e Inferência Estatística**. Ribeirão Preto, 2010. Citado nas páginas 31, 58 e 59.

BUSSAB, W. d. O. **Estatística Básica**. São Paulo: Saraiva, 2002. Citado nas páginas 41, 53 e 57.

CARVALHO, M. G. de. **Recursos Computacionais Voltados ao Ensino de Probabilidade e Estatística**. Dissertação (Mestrado) — Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 2018. Citado nas páginas 23 e 82.

CASTANHÊDE, R. B. da S. **Ensino-Aprendizagem-Avaliação de Estatística através da Resolução de Problemas: uma experiência com alunos de 3<sup>o</sup> ano do Ensino Médio**. Dissertação (Mestrado) — Universidade Federal do Maranhão, São Luis, 2015. Citado na página 75.

CEVADA, J. **Matemática nos dias de hoje: Probabilidade e Estatística**. São Paulo: Sei, 2020. Citado na página 22.

COPED, C. P. São paulo faz escola. In: EDUCAÇÃO, S. da (Ed.). **Caderno do professor: matemática**. São Paulo: SE, 2020. p. 117. Citado na página 22.

EDUCAÇÃO, R. 2021. Disponível em: <<https://revistaeducacao.com.br/2020/08/04/metodologias-ativas-sponte/>>. Acesso em: 08/07/2021. Citado na página 24.

EFDEPORTES.COM. 2010. Disponível em: <<https://www.efdeportes.com/efd149/crescimento-fisico-de-criancas-de-10-a-14-anos.htm>>. Acesso em: 29/10/2020. Citado nas páginas 13, 68 e 69.

FREIRE, J. L. R. **Introdução à Inferência Estatística**. Dissertação (Mestrado) — Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, 2017. Citado na página 23.

GESTAO, A. 2021. Disponível em: <<https://www.youtube.com/watch?v=01frvYx73xQ>>. Acesso em: 27/06/2021. Citado na página 80.

MARCOMINI, C. C. **A Matemática, a Estatística e o Corte e Costura**. Dissertação (Mestrado) — Universidade de São Paulo, São Carlos, 2020. Citado na página 22.

MEC. Base nacional comum curricular. In: MEC (Ed.). **BNCC - Educação é a Base - Ensino Médio**. Brasília: MEC, 2017. p. 576. Citado nas páginas 20, 21, 73 e 74.

ROSS, S. **PROBABILIDADE: Um curso moderno com aplicações**. Porto Alegre: bookman, 2010. Citado na página 36.

SEE, S. de Educação do Estado de S. P. Currículo paulista. In: EDUCAÇÃO, U. S. U. dos Dirigentes Municipais de (Ed.). **Currículo Paulista: Uma Construção Colaborativa**. São Paulo: SP, 08/2019. p. 526. Citado na página 21.

SMOLE, K. S. **Ser protagonista: matemática e suas tecnologias: estatística e probabilidade: ensino médio**. São Paulo: SM, 2020. Citado nas páginas 22, 74, 75, 79, 82 e 86.

VALERIANI, T. 2021. Disponível em: <<https://querobolsa.com.br/revista/metodologias-ativas-veja-6-exemplos-e-confira-os-seus-beneficios>>. Acesso em: 08/07/2021. Citado na página 87.

WIKIPÉDIA. **Andrei Kolmogorov**. 2019. Disponível em: <[https://pt.wikipedia.org/wiki/Andrei\\_Kolmogorov](https://pt.wikipedia.org/wiki/Andrei_Kolmogorov)>. Acesso em: 15/10/2020. Citado na página 28.

## DEMONSTRAÇÕES

Com a finalidade de tornar a leitura do texto mais fluida, deixaremos para este capítulo as demonstrações das expressões que aparecem no Capítulo Teórico.

1. Demonstração da expressão [A.0.2](#), usando a fórmula da variância ([2.2.4](#)), a propriedade da esperança de  $X$  e [2.2.3](#).

*Demonstração.*

$$\begin{aligned}
 \text{VAR}(X) &= \mathbb{E}[(X - \mathbb{E}(X))^2] \\
 \Rightarrow \text{VAR}(X) &= \sum_x (x - \mathbb{E}(X))^2 \mathbb{P}(X = x) \\
 &= \sum_x (x^2 - 2 \cdot x \cdot \mathbb{E}(X) + (\mathbb{E}(X))^2) \mathbb{P}(X = x) \\
 &= \sum_x x^2 \mathbb{P}(X = x) - 2\mathbb{E}(X) \sum_x x \mathbb{P}(X = x) + (\mathbb{E}(X))^2 \sum_x \mathbb{P}(X = x) \quad (\text{A.0.1})
 \end{aligned}$$

Sabendo que, de [2.2.3](#),  $\sum_x x^2 \mathbb{P}(X = x) = \mathbb{E}(X^2)$ ,  $\sum_x x \mathbb{P}(X = x) = \mathbb{E}(X)$  e  $\sum_x \mathbb{P}(X = x) = 1$ , [A.0.1](#) resulta em:

$$\begin{aligned}
 &\mathbb{E}(X^2) - 2(\mathbb{E}(X))^2 + (\mathbb{E}(X))^2 \\
 \text{VAR}(X) &= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (\text{A.0.2})
 \end{aligned}$$

□

2. Demonstraremos, agora, que  $S^2$ , dado por [3.2.1](#), é um estimador viesado para  $\sigma^2$ .

$$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu + \mu - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu - (\bar{X} - \mu))^2$$

$$= \sum_{i=1}^n (X_i - \mu)^2 - 2 \sum_{i=1}^n (X_i - \mu)(\bar{X} - \mu) + \sum_{i=1}^n (\bar{X} - \mu)^2.$$

Como  $\bar{X} - \mu$  é constante em relação à somatória, já que não são afetados por ela, e  $\sum_{i=1}^n (X_i - \mu) = X_1 - \mu + X_2 - \mu + \dots + X_n - \mu = n(\bar{X} - \mu)$ , então:

$$\begin{aligned} \sum_{i=1}^n (X_i - \bar{X})^2 &= \sum_{i=1}^n (X_i - \mu)^2 - 2(\bar{X} - \mu)n(\bar{X} - \mu) + n(\bar{X} - \mu)^2 \\ &= \sum_{i=1}^n (X_i - \mu)^2 - 2n(\bar{X} - \mu)^2 + n(\bar{X} - \mu)^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2 \end{aligned}$$

Logo,

$$\begin{aligned} E(S^2) &= \frac{1}{n} E\left(\sum_{i=1}^n (X_i - \bar{X})^2\right) = \frac{1}{n} \left(\sum_{i=1}^n E(X_i - \mu)^2 - nE(\bar{X} - \mu)^2\right) \\ &= \frac{1}{n} \left(\sum_{i=1}^n \text{VAR}(X_i) - n\text{VAR}(\bar{X})\right) = \frac{1}{n} (n\sigma^2 - \sigma^2) = \frac{1}{n} \sigma^2 (n-1) = \frac{n-1}{n} \sigma^2 \end{aligned}$$

$$E(S^2) = \frac{n-1}{n} \sigma^2 \tag{A.0.3}$$

Portanto,  $\mathbb{E}(S^2) \neq \sigma^2$  e  $S^2$  é um estimador viesado de  $\sigma^2$

