Marilda Medeiros de Almeida dos Santos

Probabilidade Condicional e suas aplicações

Santo André, 2019

UNIVERSIDADE FEDERAL DO ABC

CENTRO DE MATEMÁTICA, COMPUTAÇÃO E COGNIÇÃO

MARILDA MEDEIROS DE ALMEIDA DOS SANTOS

PROBABILIDADE CONDICIONAL E SUAS APLICAÇÕES

Orientador: Prof. Dr. Sinuê Dayan Barbero Lodovici

Dissertação de mestrado apresentada ao Centro de Matemática, Computação e Cognição para obtenção do título de Mestre

ESTE EXEMPLAR CORRESPONDE A VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELA ALUNA MARILDA MEDEIROS DE ALMEIDA DOS SANTOS, E ORIENTADA PELO PROF. DR. SINUÊ DAYAN BARBERO LODOVICI.

SANTO ANDRÉ, 2019

Sistema de Bibliotecas da Universidade Federal do ABC
Elaborada pelo Sistema de Geração de Ficha Catalográfica da UFABC
com os dados fornecidos pelo(a) autor(a).

Medeiros de Almeida dos Santos, Marilda
PROBABILIDADE CONDICIONAL E SUAS APLICAÇÕES / Marilda Medeiros de Almeida dos Santos. — 2019.

82 fls. : il.

Orientador: Sinuê Dayan Barbero Lodovici

Dissertação (Mestrado) — Universidade Federal do ABC, Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, Santo André, 2019.

1. Probabilidade Condicional. I. Dayan Barbero Lodovici, Sinuê. II. Mestrado Profissional em Matemática em Rede Nacional - PROFMAT, 2019. III. Título.

Este exemplar foi revisado e alterado em relação à versão original, de acordo com as observações levantadas pela banca no dia da defesa, sob responsabilidade única do(a) autor(a) e com a anuência do(a) orientador(a).

Santo André/ SP **24 de Junho de 2019**

Assinatura do(a) autor(a):

Assinatura do(a) orientador(a):

MINISTÉRIO DA EDUCAÇÃO
Fundação Universidade Federal do ABC
Programa de Pós-Graduação em Mestrado Profissional em Matemática em Rede Nacional
Avenida dos Estados, 5001 – Bairro Santa Terezinha – Santo André – SP
CEP 09210-580 · Fone: (11) 4996-0017
profmat@ufabc.edu.br

FOLHA DE ASSINATURAS

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado da candidata Marilda Medeiros de Almeida dos Santos, realizada em 11 de abril de 2019:

____

Prof.(a) Dr.(a) **Sinuê Dayan Barbero Lodovici** (Universidade Federal do ABC) – Presidente

____

Prof.(a) Dr.(a) **Rafael de Mattos Grisi** (Universidade Federal do ABC) – Membro Titular

Gleiciane da Silva Aragão
____

Prof.(a) Dr.(a) **Gleiciane da Silva Aragão** (Universidade Federal de São Paulo) – Membro Titular

____

Prof.(a) Dr.(a) **Márcio Fabiano da Silva** (Universidade Federal do ABC) – Membro Suplente

____

Prof.(a) Dr.(a) **Alexandre Lymberopoulos** (Universidade de São Paulo) – Membro Suplente

Dedico este trabalho à minha família querida, principalmente ao meu marido David que, com seu amor e paciência, foi meu maior incentivador; aos meus filhos, David e Fernanda, que me tornam uma pessoa melhor a cada dia. E, à minha mãe Jairinete, que me ensinou a nunca desistir.

AGRADECIMENTOS

Agradeço a Deus em primeiro lugar, que me presenteou com esse mestrado; aos meus colegas do PROFMAT, especialmente Raphael Escorse, Juliana Resende, Eduardo Colle, que me ajudaram tanto nos momentos cruciais. Também aos meus colegas do colégio e curso Objetivo, pelo incentivo e companheirismo.

A todos os meus queridos e maravilhosos professores do PROFMAT, cujo aprendizado que me permitiram levarei para toda a vida, especialmente Ana Carolina Boero, Rafael Grisi e Maurício Lima.

E, um agradecimento especial ao meu orientador, Sinuê, que se dedicou à minha orientação com atenção, paciência e carinho.

Um querido professor disse-nos, no início do mestrado, que: “só ficariam ao nosso lado, ao final desse mestrado, as pessoas que realmente nos amassem”. (Rafael Grisi)

A essas pessoas minha sincera gratidão e amor.

Por fim, gostaria de salientar que o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)-Código de Financiamento 001.

"A melhor mesa é aquela que você não é o mais inteligente. A mesa de pessoas que te puxam para cima. Essa é a mesa que eu quero..."

(Mateus Costa Ribeiro, primeiro aluno brasileiro a conseguir entrar no Mestrado em Direito em Harvard)

RESUMO

Nesse trabalho desenvolvemos alguns conteúdos de Teoria de Conjuntos e Análise Combinatória. Apresentamos a definição de Probabilidade segundo Kolmogorov e, então, nos dedicamos ao estudo de Probabilidades Condicionais. Introduzimos o Teorema de Bayes junto com uma larga gama de exemplos, tratando condicionais simples até problemas mais complexos de condicionais múltiplas. Descrevemos então os Paradoxos de Bertrand e de Simpson. E, por fim, usamos tais paradoxos e o conhecido problema de Monty Hall como ferramentas educacionais para a busca de uma decisão ou resultado em um experimento.

A abordagem simples tem como público alvo professores e alunos do ensino médio, com a finalidade de mostrar uma matemática prática e aplicável em situações da vida.

Palavras Chaves: combinatória, probabilidade condicional, Teorema de Bayes, paradoxo de Simpson, paradoxo de Bertrand, problema de Monty Hall

ABSTRACT

On this work we developed some contents from Set Theory and Combinatorial Analysis. We presented the definition of probability as done by Kolmogorov and, then, we put our efforts on the study of Conditional Probability. We introduced the Bayes Theorem with a large array of examples, from simple conditionals problems to more complex problems involving multiple conditionals. We describe Bertrand's and Simpson's Paradoxes. And, finally, we use such paradoxes and the well known Monty Hall problem as educational tools in the search for a decision or result in an experiment.

The simple approach has as audience teachers and high school students, with the purpose of showing a practical and applicable mathematics in real life situations.

Keywords: combinatorics, conditional probability, Bayes' Theorem, Simpson's paradox, Bertrand's paradox, Monty Hall's problem

CONTEÚDO

INTRODUÇÃO 1

1 UMA NOÇÃO DE CONTAGEM 3
1.1 Princípio fundamental da contagem 3
1.1.1 Princípio Multiplicativo 3
1.1.2 Princípio Aditivo 6
1.2 Arranjos, Permutações e Combinações 7

2 CONCEITOS INICIAIS DE PROBABILIDADE 13
2.1 Função 13
2.2 Probabilidade como Função 14
2.2.1 Operações entre Conjuntos 15
2.3 Experimento Aleatório, Espaço Amostral e Evento 19
2.4 Equiprobabilidade 21
2.5 Teoria de Conjuntos Aplicada a Eventos 24
2.5.1 Probabilidade da União de Eventos 25
2.6 Axiomas de Probabilidade ou Axiomas de Kolmogorov 26
2.7 Primeira Lei de Mendel e a Função Probabilidade 26
2.8 Função Probabilidade e Paradoxo de Bertrand’s 29

3 PROBABILIDADE CONDICIONAL E O TEOREMA DE BAYES 35
3.1 Probabilidade Condicional 35
3.2 Dependência e Independência de Eventos 38
3.3 Probabilidade Condicional como Probabilidade 39
3.4 Thomas Bayes (1701-1761) 40
3.5 O Teorema de Bayes 42
3.5.1 Múltiplas Condicionais 43
3.6 Aplicações do Teorema de Bayes 43
3.6.1 O Teorema de Bayes Aplicado às Ervilhas de Mendel 48

4 PARADOXOS DA PROBABILIDADE CONDICIONAL 53

4.1 Paradoxo de Bertrand 53
4.2 Paradoxo de Simpson 55
4.2.1 Caso Geral 58

5 SUGESTÕES DE ATIVIDADES PARA PRÁTICA DO ENSINO DA PROBABILIDADE CONDICIONAL 63
5.1 Probabilidade Condicional- Teorema de Bayes 63
5.1.1 Exemplo Prático 64
5.2 Paradoxo de Simpson 66
5.3 Paradoxo de Monty Hall 68

6 CONCLUSÃO 73

Bibliografia 75

INTRODUÇÃO

No século XVI, Girolamo Cardano, um dos médicos mais conhecidos da Europa, em sua autobiografia “De Propria Vita”, confessa ter jogado diariamente xadrez por mais de quarenta anos e dados por mais de vinte e cinco anos. No século XVI, o jogo era passatempo dominante e se jogava a dinheiro. Assim atribui-se a Cardano, um pequeno manual do jogador, intitulado “Liber de Ludo Alea” (o livro dos jogos de azar), que ele sequer considerava digno de publicação, mas que pode ter sido sua contribuição maior à matemática. Na parte técnica do livro, Cardano discute equiprobabilidade, esperança (o montante correto da aposta a ser feita por um jogador que tem probabilidade p de ganhar a importância s), estabeleceu a lei $p_n = p^n$, que dá a probabilidade de que um evento de probabilidade p ocorra independentemente n sucessivas vezes. Cardano escreveu o livro “Liber de Ludo Aleae” (em torno de 1550), considerado o marco inicial da Teoria das Probabilidades (ver [6]).

A proposta desse trabalho é abordar a área da probabilidade, seus pré-requisitos e especificamente a probabilidade condicional, com o objetivo de auxiliar a didática desse conteúdo para o “Ensino Médio”, tanto para os colegas professores como para os alunos. A intenção é estimular os alunos a desenvolverem o pensamento estatístico e probabilístico, conforme o objetivo proposto pela área de Matemática, segundo o PCN (Programa Curricular Brasileiro), BNCC (Base Nacional Comum Curricular) e ENEM (Exame Nacional do Ensino Médio). Usaremos para isso diversos exemplos.

Ainda nesse trabalho vamos explorar a probabilidade condicional de uma maneira mais abrangente. Tendo em vista que essa probabilidade é um forte instrumento de decisão, que permite um conhecimento mais concreto dos riscos e possibilidades em um experimento, se faz necessário o estudo de várias condicionais encontradas na situação. Observemos um exemplo simples:

Exemplo 0.1. Maria deseja fazer uma cirurgia plástica com o doutor Pedro mas, antes de tomar a decisão de se submeter a essa intervenção, gostaria de saber a chance de ser bem sucedida.

No caso da Maria, por exemplo, analisarmos em quantas das n cirurgias feitas pelo doutor Pedro houve sucesso seria apenas uma das condicionais a influenciar a decisão. Seria essa condicional suficiente?

Em uma situação como essa, observamos como o uso da Probabilidade,especialmente a Probabilidade Condicional, é uma ferramenta que pode auxiliar bastante nas escolhas assertivas. .

Esse tipo de probabilidade, a probabilidade condicional de ocorrência de um determinado evento, nos permite usar o chamado "enfoque Bayesiano" que calcula a probabilidade de sucesso de um determinado evento, com base na análise de mais do que uma informação à priori, isto é, informações anteriores à esse evento.

Para melhor desenvolver os exemplos de probabilidade e probabilidade condicional, se faz necessário introduzir o conceito de probabilidade com a definição dos conceitos de Espaço Amostral e de Eventos em um experimento e, para tal vamos rever algumas definições em Análise Combinatória, principalmente Princípio Fundamental da Contagem, que é a base para contagem de elementos dos conjuntos acima citados.

Por isso, nos capítulos iniciais, foram abordados: uma noção básica de Análise Combinatória, uma revisão rápida de Operações entre Conjuntos, e a definição de Probabilidade como função, pretendendo assim dar um suporte básico para a evolução gradual do raciocínio probabilístico, até o estudo do Teorema de Bayes e os Paradoxos da Probabilidade. Houve também uma preocupação em estabelecer uma interdisciplinaridade no uso do Cálculo de Probabilidades, representada pelo uso de conhecimentos da área da Genética e na Saúde, com foco na Primeira Lei de Mendel e no diagnóstico de doenças como Câncer. Tentamos ao longo de todo o trabalho, criar e adaptar problemas e exemplos, tratando-os de forma exaustiva, para que auxiliassem o leitor na compreensão e aplicação da teoria apresentada.

1

UMA NOÇÃO DE CONTAGEM

1.1 PRINCÍPIO FUNDAMENTAL DA CONTAGEM

Nesse capítulo será trabalhada a Análise Combinatória com ênfase em Princípio Fundamental da Contagem e com exemplos adaptados de Vestibulares e ENEMs.

Análise Combinatória é o ramo da matemática que desenvolve métodos para que possamos contar o número de elementos ou agrupamentos de um determinado conjunto finito, sob determinadas condições.

O Princípio Fundamental da Contagem **PFC** é uma forma de conseguir contar o número de agrupamentos possíveis de serem definidos dentro de determinadas situações do dia a dia, como por exemplo senhas bancárias, placas de carros, etc... O **PFC** permite uma certa praticidade, em termos de contagem, caso o número de elementos do grupo a ser formado seja pequeno e, permite também deduzir fórmulas e notações, que possam ser usadas como ferramentas para condensação, e representação de operações, em exercícios de contagem mais complexos.

De uma forma simples o Princípio Fundamental da Contagem divide-se em duas situações:

1.1.1 *Princípio Multiplicativo*

Ao longo do texto, usaremos a notação $n(A)$ representando o número de elementos de um conjunto A.

O princípio multiplicativo permite que, considerando n conjuntos:

$$N_1 = \{a_1^1, a_2^1, a_3^1, \ldots, a_{p1}^1\},$$

com $n(N_1) = p_1$, isto é, número de elementos de N_1 igual de a p_1,

$$N_2 = \{a_1^2, a_2^2, a_3^2, \ldots, a_{p2}^2\},$$

com $n(N_2) = p_2$,

$$N_3 = \{a_1^3, a_2^3, a_3^3, \ldots, a_{p3}^3\},$$

com $n(N_3) = p_3$,

$$\vdots$$

$$N_n = \{a_1^n, a_2^n, a_3^n, \ldots, a_{pn}^n\},$$

com $n(N_n) = p_n$,

temos que o número de n-uplas ordenadas (sequências de n elementos) do tipo:

$$(a_i^1, a_j^2, \ldots, a_p^n),$$

em que $a_i^1 \in N_1, a_j^2 \in N_2, \ldots, a_p^n \in N_n$, é:

$$p_1 \cdot p_2 \cdot p_3 \cdot p_n.$$

Podemos demonstrar o princípio multiplicativo pelo princípio da indução finita, onde à partir de $n = 2$, é fácil notar que para cada um dos p_1 elementos do conjunto $N_1 = \{a_1^1, a_2^1, a_3^1, \ldots, a_{p1}^1\}$, podemos dispor p_2 elementos do conjunto $N_2 = \{a_1^2, a_2^2, a_3^2, \ldots, a_{p2}^2\}$, concluindo assim um total de $p_1 \cdot p_2$ pares ordenados.

Vamos supor válido o princípio para n conjuntos e demonstrá-lo válido para $n + 1$ conjuntos. Para n conjuntos tomemos as sequências de n elementos do tipo

$$(a_i^1, a_j^2, \ldots, a_p^n),$$

em que o número de elementos será

$$p_1 \cdot p_2 \cdot p_3 \cdot p_n.$$

Por Hipótese de Indução, existem $p_1 \cdot p_2 \cdot p_3 \cdots p_n$ sequências e p_{n+1} elementos pertencentes ao conjunto N_{n+1}.

Portanto o número de sequências do tipo

$(p_1 \cdot p_2 \cdot p_3 \cdots p_n) \cdot p_{n+1}$ é igual a

$$p_1 \cdot p_2 \cdot p_3 \cdot p_n \cdot p_{n+1},$$

pois para cada n-upla de $(p_1, p_2, p_3, \ldots, p_n)$ temos p_{n+1} opções, provando assim, por indução, o Princípio Multiplicativo.

Exemplo 1.1. Uma pessoa vai a uma lanchonete para comer um só tipo de sanduíche e tomar um só tipo de bebida. A lanchonete dispõe de cinco tipos diferentes de sanduíches e quatro tipos diferentes de refrigerante. Quantas opções essa pessoa tem para se servir nas condições mencionadas?

Cada opção pode ser representada por um par ordenado (x, y), em que x representa o tipo de lanche e y representa o tipo de refrigerante, concluímos assim que para x existem apenas 5 opções e para y, apenas 4 opções. Portanto para cada uma das cinco opções x, existem 4 opções y, totalizando $5 \cdot 4$ opções da pessoa escolher sua refeição.

Exemplo 1.2. De quantas maneiras podemos responder à 10 questões de uma prova, cujas perguntas devem ser respondidas apenas com V (Verdadeiro) ou F (Falso)?

Observe que cada pergunta dessa prova tem apenas duas opções de resposta:

$$P_1 = \{F, V\}, \text{ duas opções}$$
$$P_2 = \{F, V\}, \text{ duas opções}$$
$$P_3 = \{F, V\}, \text{ duas opções}$$
$$\vdots$$
$$P_{10} = \{F, V\}, \text{ duas opções:}$$

logo, pelo princípio multiplicativo o número de opções para se responder às 10 questões da prova é:

$$2 \cdot 2 \cdot 2 \cdot \ldots \cdot 2 = 2^{10} = 1024$$

1.1.2 *Princípio Aditivo*

À partir dos exemplos abaixo, verificamos que o caso nem sempre é de multiplicar possibilidades, mas sim, de somá-las. Observe, por exemplo, que se em um grupo de pessoas com h homens e m mulheres precisarmos escolher um casal, um homem **e** uma mulher, pelo princípio multiplicativo, temos a possibilidade de formar $h \cdot m$ casais. Porém se quisermos escolher apenas uma pessoa, teremos a opção de escolher entre os h homens **ou** as m mulheres, assim o número de opções para a escolha é :

$$h + m,$$

e assim usamos o chamado Princípio Aditivo.

Exemplo 1.3. Um caixa eletrônico dispõe apenas de notas de cinco reais e notas de dez reais. De quantas maneiras diferentes uma pessoa pode sacar cem reais?

Observe que nessa situação de contagem precisamos enumerar as possibilidades de se fazer um saque de cem reais apenas em notas de cinco e de dez reais:(ver tabela 1)

Tabela 1: Tabela de Opções para Saque

nota de cinco	nota de dez
0	10
2	9
4	8
6	7
8	6
10	5
12	4
14	3
16	2
18	1
20	0

Percebemos então que não é o caso de multiplicar opções, mais sim de somá-las, totalizando assim 11 opções, já que podemos escolher a primeira $(0, 10)$, **ou** a segunda $(2, 9)$ **ou** a terceira $(4, 8)$ **ou** ... até a décima primeira opção, caracterizando assim o Princípio Aditivo.

Exemplo 1.4. Para ir de uma cidade A até uma cidade D, pode-se escolher passar pela cidade B ou pela cidade C. De A até B, temos duas opções de caminho, de B até D são três opções de caminho e de A até C pode-se escolher entre três opções de caminhos, sendo que de C a D podemos optar por dois caminhos diferentes. De quantas maneiras diferentes podemos ir de A até D?

Podemos representar as opções dadas, com a figura 1 :

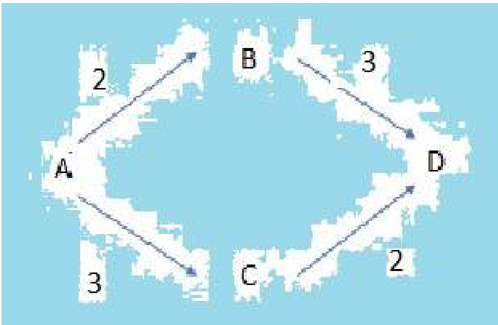

Figura 1: Opções de Caminhos

Pelo Princípio Multiplicativo, resulta em:

$$(2 \cdot 3) + (3 \cdot 2) = 6 + 6 = 12$$

Assim serão 12 opções de caminho para ir da cidade A até a cidade D.

1.2 ARRANJOS, PERMUTAÇÕES E COMBINAÇÕES

Vejamos agora exemplos de problemas que tomaremos como base para definir e classificar Arranjos, Combinações ou Permutações:

Exemplo 1.5. Quantas senhas bancárias podemos formar usando quatro algarismos distintos do conjunto $\{1, 2, 3, 4, 5\}$?

Observemos que pelo princípio fundamental da contagem, conseguimos:

- cinco possibilidades para o algarismo das unidades de milhar,
- como não podemos repetir algarismos, teremos quatro possibilidades para o algarismo das centenas,
- três para o algarismo das dezenas e,
- duas para algarismo das unidades.

Assim conseguimos formar $5 \cdot 4 \cdot 3 \cdot 2 = 120$ possibilidades. Algumas delas são: $2345, 1234, 1345, 4325 \ldots 5432$.

À partir desse exemplo podemos definir **Arranjos Simples** como sendo agrupamentos de objetos, números, pessoas, ... que se diferenciam pela ordem ou natureza dos elementos, ou seja, a mudança de posição ou ordem dos elementos, muda o agrupamento.

Exemplo 1.6. Ao usarmos todos os elementos disponíveis no conjunto para a formação das senhas, percebemos que as mesmas só se alteram com a mudança de ordem ou posição dos elementos. Podemos formar então: senha 12345, senha 34512, senha 15432, enfim, as senhas são obtidas apenas com a PERMUTAÇÃO dos elementos do conjunto $\{1, 2, 3, 4, 5\}$, gerando assim as Permutações Simples.

Portanto **Permutações Simples** são agrupamentos que diferem entre si apenas pela ordem dos elementos.

Note que, dessa forma, o número de Permutações Simples de n elementos é dado por:

$$n \cdot (n-1) \cdot (n-2) \ldots 3 \cdot 2 \cdot 1.$$

Exemplo 1.7. Usando o mesmo conjunto do Exemplo 1.5, se precisarmos formar senhas, com 4 algarismos distintos e em ordem estritamente crescente, conseguiremos formar apenas as senhas: $1234, 2345, 1345, 1245, 1235$, pois não poderíamos inverter a posição dos algarismos, já que precisam estar em ordem crescente. Nesse caso, poderíamos contar o número de senhas, ainda usando o princípio fundamental da contagem, porém dividindo o total de opções obtidas pelo número de permutações entre os quatro elementos, que não podemos fazer, pois devemos mantê-los em ordem crescente, obtendo:

$$\frac{5 \cdot 4 \cdot 3 \cdot 2}{4 \cdot 3 \cdot 2 \cdot 1} = \frac{120}{24} = 5,$$

que são: $1234, 2345, 1345, 1245, 1235$, conforme já havíamos concluído.

Comparando os Exemplos 1.5 e 1.7, verificamos que, pelo fato de não podermos permutar os algarismos que formam a senha, obteremos senhas que se distinguem apenas pela natureza de seus elementos, obtendo-se assim um novo tipo de caso, a que chamamos de Combinações, tão necessário na área da Probabilidade.

Portanto **Combinações Simples** são agrupamentos de objetos, números ou pessoas, que diferem entre si apenas pela natureza de seus elementos, isto é, a mudança de ordem entre os elementos não altera o agrupamento.

Calculamos o número de combinações de n elementos em conjuntos de k em k elementos por:

$$C_{n,k} = \binom{n}{k} = \frac{n!}{(n-k)! \cdot k!},$$

para $k \leq n$.

E dizemos que $\binom{n}{k}$ representa o número de combinações de n elementos k a k.

Por convenção, define-se $0! = 1$. Com isso, $\binom{n}{0} = \binom{n}{n} = 1$. Além disso assume-se que $\binom{n}{i} = 0$ quando $0 < n < i$.

Assim, $\binom{n}{k}$ representa o número de **subconjuntos** diferentes com k elementos que podemos formar através de um conjunto de n elementos.

Observemos algumas aplicações nos exemplos abaixo.

Exemplo 1.8. (Puccamp-2018-adaptado) Admita que certa cidade brasileira tenha 8 canais de TV aberta, todos com transmissões diárias. Se uma pessoa pretende assistir três dos oito canais em um mesmo dia, ela pode fazer isso de x maneiras diferentes sem levar em consideração a ordem em que assiste os canais, e pode fazer de y maneiras diferentes levando em consideração a ordem em que assiste os canais. Sendo assim, calcule o valor de $y - x$.

De acordo com o enunciado, como nas x maneiras de assistir, a ordem não altera a programação, concluímos que se trata de combinação de 8 canais agrupados de 3 em 3, assim há

$$x = C_{8,3} = \binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = 56$$

maneiras de assistir 3 canais dos 8 disponíveis independente da ordem em que são escolhidos.

Em contra partida, as y maneiras de se escolher os três canais de forma que a mudança de ordem entre eles altere a programação, transforma essas escolhas em Arranjos, de tal forma que obtemos:

$$y = 8 \cdot 7 \cdot 6 = 336$$

opções de escolha dos três canais. Portanto, $y - x = 336 - 56 = 280$.

Exemplo 1.9. Um torneio de xadrez tem cinco competidores, dos quais dois são russos, dois são dos Estados Unidos, e um é do Brasil. Se o resultado do torneio listar apenas a nacionalidade dos jogadores em sua ordem de colocação, quantos resultados serão possíveis?

Usando o mesmo raciocínio do Exemplo 1.7 para resolver, observemos que se pensarmos em todos os resultados independentes da nacionalidade de cada competidor, teríamos $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$ possibilidades de resultados, porém como o resultado do torneio deve listar apenas a nacionalidade dos jogadores, permutar entre as nacionalidades não mudaria o resultado, por exemplo: $R_1R_2A_1A_2B = R_2R_1A_2A_1B$; onde R_i, A_i, B representam as nacionalidades russa, norte americana e brasileira respectivamente.

Portanto, para não repetir resultados, dividimos o total de trocas ou permutações pela permuta entre as nacionalidades iguais, dessa maneira temos:

$$\frac{5!}{2! \cdot 2! \cdot 1!} = \frac{120}{4} = 30$$

resultados possíveis e distintos, listando apenas a nacionalidade de cada jogador, ou seja: $RRAAB \neq RARAB \neq BRRAA \neq RBARA$, onde R representa a nacionalidade russa, A representa a nacionalidade norte americana e B a nacionalidade brasileira.

Observemos que podemos calcular a Permutação com Repetição da seguinte maneira:

$$P_n^{\alpha,\beta,\ldots,\omega} = \frac{n!}{\alpha! \cdot \beta! \ldots \cdot \omega!},$$

onde:

n = total de elementos a permutar e

$\alpha, \beta, \ldots, \omega$ representam o número de repetições de cada elemento.

E assim com esses conceitos gerais de Análise Combinatória pode-se dar início ao conceito de Probabilidade.

2

CONCEITOS INICIAIS DE PROBABILIDADE

Nesse capítulo estudaremos o conceito de *probabilidade* segundo Kolmogorov. Para tal, recordemos inicialmente o conceito de função.

2.1 FUNÇÃO

Uma função $f : A \to B$ é uma relação entre dois conjuntos A e B, denominados domínio e contra-domínio de f, em que, para cada elemento x de A, temos, obrigatoriamente, um único elemento y que se relaciona com ele em B. Nesse caso escrevemos, $f(x) = y$ e dizemos que y é a imagem de x pela função f. O conjunto $f(A) = \{f(x) \in B; x \in A\}$ é conhecido como imagem da função f.

Vejamos o exemplo prático da função C que associa à distância x percorrida pelo taxi o custo $C(x)$ de tal corrida na cidade de São Paulo.

O valor cobrado é calculado à partir de uma taxa fixa, mais um valor por quilômetro rodado. O valor pago por um usuário dentro da cidade de São Paulo, na Bandeira 1, é $R\$4,10$ de taxa fixa, mais $R\$2,75$ por quilômetro rodado, em média (www.tarifadetaxi.com).

Podemos calcular o custo de um passeio de táxi em São Paulo ($C(x)$), segundo dados acima, pela "lei de formação":

$$C(x) = 4,10 + 2,75 \cdot x,$$

onde x é o número de quilômetros rodados.

Veja na Tabela abaixo alguns exemplos da aplicação da função C em casos concretos:

Percurso	*Km*	*custo*
Vila Pires/S.André $\rightarrow$ UFABC/S.André	$3,8Km$	$R\$14,55$
Vila Pires/S.André $\rightarrow$ UFABC/S.B.C	$4,8Km$	$R\$17,30$
Vila Pires/S.André $\rightarrow$ Rua Senador Fláquer/S.André	$2,6km$	$R\$11,25$

Tabela 2: Tabela Função $C(x)$

2.2 PROBABILIDADE COMO FUNÇÃO

Sinteticamente podemos dizer que *probabilidade* é uma medida da chance de, num experimento aleatório, um determinado evento ocorrer.
A ideia de probabilidade pode ser vista como uma função ao relacionar a um determinado evento (em um conjunto de resultados possíveis de um experimento), sua chance de ocorrer. Tal chance é representada usualmente por um número real entre 0 e 1, onde 0 representa a total impossibilidade de ocorrência do evento e 1 representa um evento certo, ou seja que, independentemente do acaso, ocorrerá com total certeza.
Ao conjunto S dos resultados possíveis para um dado experimento, damos o nome de **Espaço Amostral**. Um **Evento** pode ser então entendido como um subconjunto de S.
Assim, o conceito de probabilidade pode ser entendido como uma função P que associa cada evento E contido no espaço amostral S, a um único número real $P(E)$ entre 0 e 1.
Assim, define-se:

$$P : \mathbb{P}(S) \longrightarrow [0,1] \subset \mathbb{R},$$

sendo $\mathbb{P}(S)$ o conjunto das partes de S, quer dizer, o conjunto de todos os subconjuntos de S.
Para que uma tal função P represente aquilo que intuitivamente entendemos por "probabilidade" precisamos que esta função satisfaça alguns pré-requisitos:

- $P(S) = 1$;
- $P(A \cup B) = P(A) + P(B)$ se A e B são dois eventos disjuntos, ou seja, que não contém um mesmo resultado (elemento de S).

Antes, porém, de continuarmos o estudo de probabilidade, lembremos alguns fundamentos da teoria de conjuntos.

2.2.1 *Operações entre Conjuntos*

Quando se fala em relações entre dois ou mais eventos, que é o que acontece frequentemente no estudo de probabilidade, é de fundamental importância o exato entendimento das operações entre conjuntos: união, intersecção, diferença e complementar de um conjunto.

União entre conjuntos A e B

A união entre os conjuntos A e B, $A \cup B$, resulta em um conjunto com os elementos que pertencem ao conjunto A **ou** ao conjunto B(ver figuras 2 e 3). Assim:

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

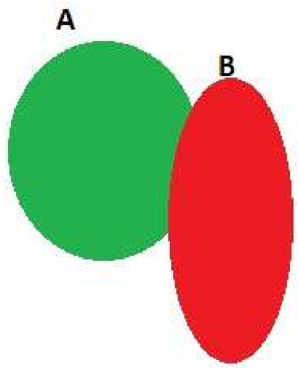

Figura 2: União dos conjuntos A e B não disjuntos.

. Observa-se que será definido na próxima seção o significado de conjuntos disjuntos.

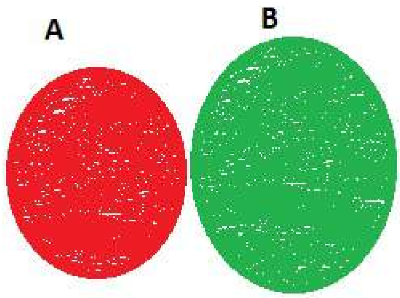

Figura 3: União dos conjuntos A e B disjuntos.

Intersecção entre os conjuntos A e B

A intersecção entre os elementos dos conjuntos A e B, $A \cap B$, resulta em um conjunto com apenas os elementos comuns aos conjuntos envolvidos, isto é, elementos pertencentes ao conjunto A **e** ao conjunto B, ao mesmo tempo (Veja Figura 4).
Assim:

$$A \cap B = \{x : x \in A \text{ e } x \in B\}.$$

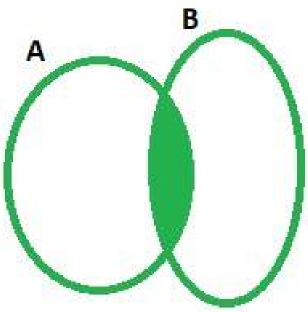

Figura 4: Intersecção entre os conjuntos A e B.

Conjuntos Disjuntos

Nesse momento, vamos abrir um parênteses para definir o que são conjuntos disjuntos, mencionados em Figura 3. Conjuntos disjuntos são conjuntos que não possuem elementos em comum, portanto, conforme definição de intersecção de conjuntos, a intersecção entre os conjuntos A e B disjuntos é vazia, isto é, $A \cap B = \emptyset$.

Diferença ou subtração entre dois conjuntos A e B

A diferença entre os conjuntos A e B, $A - B$, resulta em um conjunto com os elementos que pertencem ao primeiro conjunto (A), mas não pertencem ao segundo conjunto (B), quer dizer que excluem-se os elementos do conjunto B, sobrando apenas os elementos que estão em A, mas não estão em B (ver Figura 5).
Assim: $A - B = \{x : x \in A \text{ e } x \notin B\}$

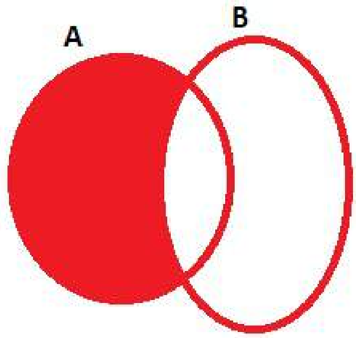

Figura 5: Diferença $A - B$.

Complementar de um conjunto A em relação à S

Dados os conjuntos A e S, tal que $A \subset S$, representa-se o conjunto complementar de A em relação à S por $\overline{A}$, cuja operação é $S - A$, quer dizer, diferença $S - A$ sob a condição de $A \subset S$, sendo o conjunto resultado formado por elementos que estão em S excluídos os elementos de A (ver Figura 6).
$\overline{A} = S - A = \{x : x \in S \text{ e } x \notin A \text{ e } A \subset S\}$

Cardinalidade de um conjunto finito

Seja $I_k = \{1, 2, 3, \ldots, k\}$ o conjunto formado pelos números naturais de 1 à k. Dado um conjunto A, se houver uma bijeção, ou seja função um-a-um, entre A e I_k dizemos que A é um conjunto finito e tem k elementos. Nesse caso escrevemos $n(A) = k$ e dizemos que k é a cardinalidade do conjunto A.

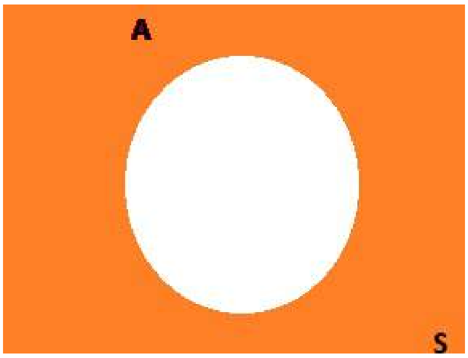

Figura 6: $\overline{A} = S - A$.

Exemplo 2.1. Observemos os resultados ocorridos no lançamento de um dado. Chamemos de:
A=conjunto dos números pares, portanto $A = \{2, 4, 6\}$;
B=conjunto dos números primos, portanto $B = \{2, 3, 5\}$;
S=conjunto de todos os resultados possíveis, portanto $S = \{1, 2, 3, 4, 5, 6\}$.
Operando os conjuntos A, B e S, obtemos:
$A \cup B = \{2, 3, 4, 5, 6\}$ (ver Figura 7)

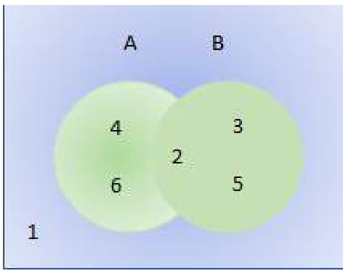

Figura 7: $A \cup B$

$A \cap B = \{2\}$ (ver Figura 8)
$\overline{A} = S - A = \{1, 3, 5\}$ (ver Figura 9)
Denotemos por $n(A)$, o número de elementos de um conjunto A, assim no nosso exemplo:
$n(A) = 3$;
$n(A \cup B) = 5$;

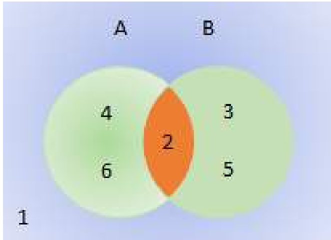

Figura 8: $A \cap B$

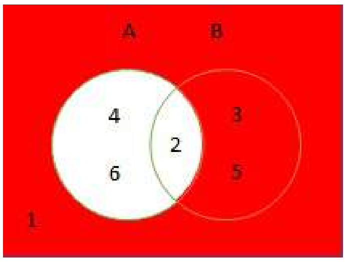

Figura 9: $\overline{A} = S - A$

$n(A \cap B) = 1$;
$n(\overline{A}) = 3$.

2.3 EXPERIMENTO ALEATÓRIO, ESPAÇO AMOSTRAL E EVENTO

O conceito de probabilidade caracteriza um modelo matemático que estuda os riscos ou chances da ocorrência de determinados resultados em um experimento aleatório.
Nomeia-se **Experimento Aleatório** qualquer experimento que repetido em idênticas condições, produzem resultados que não podem ser previstos com certeza, mas embora não saibamos qual resultado vai ocorrer, conseguimos descrever o conjunto de todos os resultados possíveis que podem ocorrer.
Segundo Samuel Hazzan, no livro Fundamentos da Matemática Elementar:

> "Experimento significa (de experiência) observar ou fazer alguma coisa sob determinada "condição", o que resultará em um resultado ou estado final de acontecimentos que não são previsíveis." [6]

O lançamento de uma moeda é um experimento aleatório, pois todos os possíveis resultados são conhecidos $\{cara, coroa\}$, porém não se pode prever o resultado do experimento (pode acontecer qualquer um desses dois resultados).
Embora o resultado do experimento não seja previamente conhecido, o conjunto de todos os resultados possíveis é conhecido como o **Espaço Amostral** do experimento e vamos representá-lo, nesse trabalho, com a letra maiúscula S.
Notamos que, de modo geral, S pode ser finito ou infinito (enumerável ou não). Os resultados no experimento aleatório "lançamento de um dado", por exemplo, são finitos, nesse caso $S = \{1, 2, 3, 4, 5, 6\}$; no lançamento de um "ponto material" no interior de um círculo temos que S é um infinito (não-enumerável) formado por todos os pontos no interior do círculo.
Observemos os exemplos:

Exemplo 2.2. Se o resultado de um experimento consiste na determinação do sexo de um bebê recém nascido, então:
$S = \{m, f\}$, onde o resultado m significa que o bebê é do sexo masculino e f que o bebê é do sexo feminino.

Exemplo 2.3. Se o experimento consiste em retirar uma bola de uma urna que possui quatro bolas numeradas de 1 a 4, temos como espaço amostral

$$S = \{1, 2, 3, 4\}.$$

Exemplo 2.4. Se o experimento consiste em retirar duas bolas, simultaneamente, de uma urna que possui quatro bolas numeradas de 1 a 4, temos como espaço amostral o conjunto $S = \{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$.
Percebe-se então que o número de elementos desse espaço amostral S pode ser obtido por Análise Combinatória, conforme visto na Seção (1.2), mais especificamente com a "Combinação de 4 elementos agrupados de 2 em 2":

$$n(S) = C_{4,2} = \frac{4 \cdot 3}{2 \cdot 1} = 6,$$

sendo $n(S)$ = número de elementos do espaço amostral S.

Exemplo 2.5. Se o resultado de um experimento é a ordem em que podemos enfileirar 5 pessoas, representadas pelas letras A,B,C,D,E, então S é o conjunto com todas as 5! permutações de A,B,C,D e E, conforme visto no Exemplo 1.6.
Representamos um resultado denotando a ordem dos elementos na fileira (B, C, D, E, A), por exemplo, significa que colocamos a pessoa B em primeiro lugar, a pessoa C em segundo, a pessoa D em terceiro, e assim por diante.

Quando os resultados de um experimento aleatório puderem ser univocamente associados a um subconjunto de um espaço numérico, denominamos o valor associado ao resultado de **Variável Aleatória** do experimento.
Intitula-se **Evento** (E) de um Espaço Amostral a qualquer subconjunto de (S).
Assim no Exemplo 2.5, temos que se E é o conjunto de todos os resultados de S começados com a pessoa B, teremos:

$$E = \{(B, C, D, E, A), (B, C, D, A, E), (B, C, A, D, E), \ldots, (B, A, C, D, E)\}.$$

Nesse caso, $n(E) = 4 \cdot 3 \cdot 2 \cdot 1 = P_4 = 24$, onde $n(E)$ significa número de elementos de E e pode ser representado pela permutação das pessoas A,C,D,E, já que B deve ficar em primeiro lugar.

2.4 EQUIPROBABILIDADE

Um experimento aleatório com resultados equiprováveis são experimentos com espaço amostral S finito onde cada subconjunto unitário de S tem a mesma probabilidade de ocorrer.
Definimos nesse caso a função probabilidade $p : \mathbb{P}(S) \to [0, 1] \subset \mathbb{R}$ por:

$$p(E) = \frac{n(E)}{n(S)}, \tag{2.1}$$

onde $E \subset S$.
Nesse caso, se a $n(S) = k$, temos que a probabilidade da ocorrência qualquer resultado (evento unitário) é igual a $\frac{1}{k}$.

Exemplo 2.6. Tomemos como exemplo o lançamento de um dado não viciado, cujas probabilidades de acontecer cada uma de suas faces é sempre $\frac{1}{6}$, temos então um espaço amostral equiprovável.

Notamos para experimentos com resultados equiprováveis o seguinte resultado:

Proposição 2.1. *Se $p : \mathbb{P}(S) \to [0,1] \subset \mathbb{R}$ é definida pela igualdade* (2.1)*, então valem:*

(i) $P(S) = 1$

(ii) $0 \leq P(E) \leq 1$*, qualquer que seja* $E \subset S$*.*

(iii) $P(A \cup B) = P(A) + P(B)$*, se* $A, B \subset S$ *são dois eventos disjuntos.*

Demonstração. (i) Se $E = S$, então $n(E) = n(S)$ e portanto:

$$P(E) = \frac{n(S)}{n(S)} = 1.$$

(ii) Qualquer que seja $E \subset S$ temos que $0 \leq n(E) \leq n(S)$ (ver [8]), daí:

$$0 \leq \frac{n(E)}{n(S)} \leq \frac{n(S)}{n(S)} = 1,$$

ou seja, $0 \leq p(E) \leq 1$.

(iii) Sejam $A, B \subset S$ são dois conjuntos disjuntos, temos:

$$p(A \cup B) = \frac{n(A \cup B)}{n(S)} = \frac{n(A) + n(B)}{n(S)} = \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} = p(A) + p(B).$$

□

Observação 2.4.1. *Se E é um evento que representa um conjunto vazio dentro do espaço amostral, teremos* $P(E) = 0$*. Denominamos tal evento de* ***Evento Impossível****.*

Exemplo 2.7. O evento que representa resultados maiores do que seis no lançamento de um dado é um evento impossível de acontecer, e portanto com probabilidade igual a 0 de ocorrer.

Notamos que para espaços amostrais finitos nem sempre temos resultados equiprováveis, observe:

Exemplo 2.8. Se, no entanto, lançarmos um dado alterado, ou seja viciado, com faces numeradas de 1 a 4, onde a probabilidade de ocorrer cada face seja proporcional ao número da face, na razão de $\frac{1}{10}$ teremos:
P(ocorrer face 1)$= 1 \cdot \frac{1}{10} = \frac{1}{10}$;
P(ocorrer face 2)$= 2 \cdot \frac{1}{10} = \frac{2}{10}$;
P(ocorrer face 3)$= 3 \cdot \frac{1}{10} = \frac{3}{10}$;
P(ocorrer face 4)$= 4 \cdot \frac{1}{10} = \frac{4}{10}$.
E assim cada face tem probabilidade de ocorrer diferente uma das outras, o que exemplifica uma situação de espaço amostral não equiprovável.
Note que aqui ainda são válidas as três condições necessárias no conceito de probabilidade:

(i) $P(S) = 1$.

(ii) $0 \leq P(E) \leq 1$, qualquer que seja $E \subset S$.

(iii) $P(A \cup B) = P(A) + P(B)$, se $A, B \subset S$ são dois eventos disjuntos.

Observe o exemplo citado na dissertação [13]:

Exemplo 2.9. Quais seriam as possíveis respostas à pergunta : "Vai chover amanhã?"
Pensamos em apenas duas opões de respostas: "Sim ou Não" nos levando à ideia de que a probabilidade de chover amanhã é de 50%.
Porém um universo de possíveis eventos futuros representado pelo espaço amostral {chover; não chover} não é equiprovável. Apesar de o conjunto possuir apenas dois elementos, o acontecimento de cada um depende de uma série de fatores fora do nosso alcance.

Observação 2.4.2. *Vale a pena notar que a função probabilidade definida nessa seção é um caso particular da chamada* ***Distribuição Uniforme de Probabilidade****, onde temos um espaço amostral, uma medida e uma distribuição de probabilidade tais que se A e B são eventos de mesma medida, vale que $p(A) = p(B)$ (ver seção "A Variável Aleatória Uniforme" de [14]).*
Na Seção 2.8 apresentaremos alguns exemplos onde o espaço amostral é infinito e fazemos a escolha de um ponto no interior de um círculo utilizando uma distribuição uniforme de probabilidade.

2.5 TEORIA DE CONJUNTOS APLICADA A EVENTOS

Sabendo que eventos são subconjuntos de um conjunto chamado Espaço Amostral, fica fácil definir a relação entre eventos e suas respectivas probabilidades, com as operações entre conjuntos (vistas na subseção 2.2.1).
Dois ou mais eventos são classificados como **mutuamente exclusivos** ou **excludentes** ou **disjuntos**, quando a intersecção entre eles é vazia, isto é, não há possibilidade de ocorrerem todos ao mesmo tempo.

Exemplo 2.10. Ao usarmos como exemplo o lançamento de um dado honesto, cujo espaço amostral é $S = \{1,2,3,4,5,6\}$, os eventos $A = \{\text{ocorrer número par}\}$ e $B = \{\text{ocorrer número 3}\}$ são mutuamente exclusivos ou disjuntos, pois $A \cap B = \emptyset$.

Exemplo 2.11. Ainda no lançamento de um dado calculando a probabilidade de ocorrer evento $A = \{\text{ocorrer número par}\}$ ou evento $B = \{\text{ocorrer número primo}\}$, temos que $A = \{2,4,6\}$ e $B = \{2,3,5\}$ e que o resultado 2 se encontra tanto em A como em B.
Dessa forma existe um elemento de ambos que ocorre simultaneamente em A e em B, o que faz com esses eventos não sejam mutuamente exclusivos e portanto, “não"disjuntos, ou seja, a intersecção entre eles não é vazia, $A \cap B \neq \emptyset$.

Como vemos no exemplo acima, fica claro que dois eventos podem ter elementos comuns, e o conjunto que os representa é o conjunto $A \cap B$, cujos elementos são apenas os elementos comuns a A e a B.
Ficam, assim definidas então as **uniões e intersecções de mais de dois eventos**.
Se $E_1, E_2 ...$ são eventos, então a união desses eventos, representada por

$$\bigcup_{n=1}^{\infty} E_n,$$

é definida como o evento formado por todos os resultados que aparecem em E_n para pelo menos um valor de $n = 1, 2, ...$
Similarmente, a intersecção dos eventos E_n, representada por

$$\bigcap_{n=1}^{\infty} E_n,$$

é definida como o evento formado pelos resultados que aparecem em todos os eventos E_n, $n = 1, 2, \ldots.$

Finalmente, para qualquer evento E, definimos um novo evento $\overline{E}$, chamado de **evento complemento ou complementar** de E em relação ao espaço amostral S.

$\overline{E}$ representa o evento formado por todos os resultados do espaço amostral que não estão contidos em E, isto é, que completam E em relação a S. Portanto, $\overline{E}$ ocorrerá se, e somente se E não ocorrer, podemos associar à seguinte notação tirada da Teoria dos Conjuntos:

$\overline{E} = C_S^E = S - E = \{x : x \in S \text{ e } x \notin E\}$, e sob a condição de $E \subset S$.

Note que se $\overline{E}$ complementa E em relação a S, então:

$$P(\overline{E}) + P(E) = P(S) = 1.$$

2.5.1 *Probabilidade da União de Eventos*

A operação de união entre dois conjuntos em Teoria dos Conjuntos se aplica a relação de probabilidade da união entre dois ou mais eventos (subseção 2.2.1), com o cuidado de observar se ambos são mutuamente exclusivos ou não.

Nos eventos A e B escolhidos no Exemplo 2.10, observemos que a união de A com B resulta em um conjunto com os elementos de um *ou* de outro, obtendo a junção de todos os elementos, portanto:

$A \cup B = \{2, 3, 4, 6\}$, assim obtemos:

$$n(A \cup B) = 4,$$

donde podemos concluir que:

$$P(A \cup B) = P(A) + P(B) = \frac{3}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3} \approx 0,67 = 67\%.$$

Porém, no Exemplo 2.11, como os eventos A e B não são disjuntos, devemos tomar o cuidado de não repetir duas vezes o elemento comum aos dois, o número 2, que é ao mesmo tempo, par e primo. Assim, sendo $A = \{2, 4, 6\}$ e $B = \{2, 3, 5\}$ temos $A \cup B = \{2, 3, 4, 5, 6\}$ e $A \cap B = 2$, consequentemente,

$$n(A \cup B) = n(A) + n(B) - n(A \cap B) = 3 + 3 - 1 = 5.$$

Assim:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{3}{6} + \frac{3}{6} - \frac{1}{6} = \frac{5}{6}.$$

2.6 AXIOMAS DE PROBABILIDADE OU AXIOMAS DE KOLMOGOROV

Por volta de 1930, o russo A.N.Kolmogorov apresentou um conjunto de axiomas matemáticos para definir probabilidades, proporcionando à Teoria da Probabilidade uma base matemática firme.
Segundo Kolomogorov para que uma dada função $P : \mathbb{P}(S) \longrightarrow [0,1] \subset \mathbb{R}$ seja uma **função probabilidade** é necessário que esta obedeça às seguintes condições ou Axiomas:

Axioma 1. Se S denota o espaço amostral temos:

$$P(S) = 1.$$

Axioma 2. Para todo evento $E \subset S$ deve valer:

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

Axioma 3. Dada uma sequência de eventos mutuamente exclusivos $E_1, E_2, \dots$ temos:

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

Como num experimento de resultados equiprováveis o espaço amostral é necessariamente finito, temos que toda sequência de eventos mutuamente exclusivos tem $E_i \neq \emptyset$ apenas para um conjunto finito de índices. Assim sendo, é fácil ver que a Proposição 2.1 mostra que a função probabilidade definida na Seção 2.4 é, de fato, uma função probabilidade segundo Kolmogorov.
É fácil ver também que a função apresentada no Exemplo 2.8, apesar de não se tratar de um experimento de resultados equiprováveis, é também uma função probabilidade segundo Kolmogorov.

2.7 PRIMEIRA LEI DE MENDEL E A FUNÇÃO PROBABILIDADE

Segundo [2], Gregor Johan Mendel foi um monge agostiniano nascido no ano de 1822 que se interessou em explicar como as características dos pais são passadas aos seus descendentes. Mendel realizou todas as suas pesquisas com ervilhas de cheiro (*Pisum Sativa*) por suas diversas vantagens, como: fácil cultivo, produção de grande quantidade de sementes, ciclo de vida curto, além de características contrastantes e de fácil identificação. Essa escolha foi consequência de um longo processo de acompanhamento cuidadoso da germinação das sementes, por aproximadamente dois anos.

Em seus experimentos Mendel utilizou apenas plantas de linhagem pura, isto é, plantas de sementes verdes que só originassem sementes verdes e plantas de sementes amarelas que só originassem sementes amarelas.
Fazendo o cruzamento dessas linhagens puras, Mendel percebeu que ao cruzar plantas de semente amarela com plantas de semente verde, na primeira geração formavam-se apenas indivíduos com sementes amarelas. Mendel notou, no entanto, que fazendo o cruzamento entre os descendentes diretos das linhagens puras, não mais tínhamos todos os descendentes com sementes amarelas. Na verdade Mendel percebeu que para cada 3 descendentes com sementes amarelas nascia 1 com sementes verdes. A partir dessa observação Mendel tentou explicar a transmissão de características entre os seres vivos.
Sua hipótese então formulada, mais tarde chamada de Primeira Lei de Mendel, dizia que todas as características de um indivíduo eram determinadas por pares de genes que se segregam, isto é, se separam, durante a formação dos gametas, assim pai e mãe, para uma dada característica, transmitiam apenas um gene para os seus descendentes.
Para explicar o resultado de seus experimentos, Mendel usou uma simbologia com dois "fatores", como ele chamou. Usou as letras VV para representar a cor amarela da semente e vv para representar a cor verde nos indivíduos de linhagem pura.
Hoje sabemos que os "fatores" determinantes da hereditariedade a que Mendel se referia são os genes, segmentos de "DNA" codificadores de proteínas correspondentes às características observadas.
Assumindo a sua hipótese sobre a transmissão dos "fatores", ele obteve sementes amarelas, com características Vv, o que já mostrava que o indivíduo, quer dizer a ervilha gerada, herdava um gene da semente amarela e um gene da semente verde.
Continuando com sua hipótese, se realizássemos uma auto fecundação nessa nova geração de sementes amarelas, Mendel obteria três sementes amarelas com características VV, Vv, vV, e uma semente verde, "fatores" vv, o que explicava a proporção por ele encontrada no experimento.(Ver figura 10)
Mendel notou que a semente verde não apareceu na primeira geração, mas reapareceu na segunda. Por esse motivo chamou as sementes verdes de **recessivas** e a sementes amarelas de **dominantes**. Diante desses resultados, Mendel concluiu que:

1. cada ser vivo é único e possui um par de fatores, que hoje entendemos representar os genes, para cada característica,

2. as características hereditárias são herdadas metade do pai e metade da mãe os descendentes herdarão apenas um gene de cada característica de seus pais,

3. genes são transmitidos através de genes;
4. os descendentes herdarão apenas um gene de cada característica de seus pais, ou seja, para uma determinada característica, haverá apenas um gene do par, tanto da mãe quanto do pai.

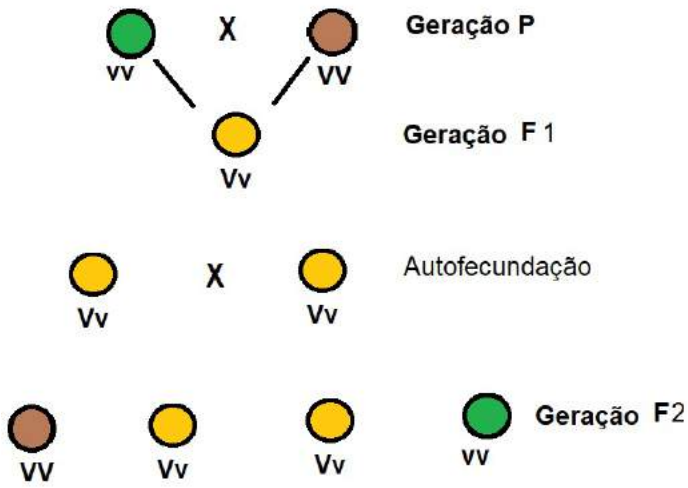

Figura 10: Ervilhas de Mendel.

E foi assim que Mendel pode enunciar a sua primeira Lei.
Para entendermos o cruzamento das plantas só a óptica da probabilidade trabalhamos com o espaço amostral $S = \{VV, vv, Vv, vV\}$, que tem como elementos os genótipos correspondentes às plantas amarela pura, verde pura, híbrida amarela com gene primeiro gene dominante, e híbrida amarela com segundo gene dominante, respectivamente.
O experimento correspondente ao cruzamento de pais híbridos é então descrita pela função $P : \mathbb{P}(S) \longrightarrow [0, 1] \subset \mathbb{R}$ que descreve um experimento de resultados equiprováveis.
O quadro de Punnet é uma ferramenta valiosa para o estudo genético dos cruzamentos, melhor, de eventos de fertilização.
Vamos usar e adaptar ao conceito de função probabilidade, a experiência de Mendel observando o **Quadro de Punnet** abaixo (ver Tabela 3) que representará os resultados da autofecundação, de ervilhas amarelas com gene Vv:

Tabela 3: Quadro de Punnet

Variável	V	v
V	VV	Vv
v	vV	vv

Vamos tomar X_i como o evento que representa a ocorrência do "fator" v.
Nossa função probabilidade é a função $P(X_i)$ que relaciona a probabilidade de ocorrer o evento X_i, onde $i = \{0, 1, 2\}$, representará o número de genes recessivos v, no espaço amostral S. Com o quadro de Punnet o resultado da probabilidade pedida é obtido de forma imediata e visual:
Para $X_0 = \{VV\}$ temos $P(X_0) = \frac{1}{4}$ que corresponde à probabilidade de ocorrer o gene v zero vezes, isto é, nenhuma vez.
Para $X_1 = \{Vv, vV\}$ temos $P(X_1) = \frac{2}{4} = \frac{1}{2}$ é a probabilidade de ocorrer o gene v apenas uma vez.
E, para $X_2 = \{vv\}$ temos $P(X_2) = \frac{1}{4}$ corresponde à probabilidade de ocorrer gene vv.
Vemos aqui um uso do Princípio Fundamental da Contagem **PFC** visto no Capítulo 1, mais especificamente Princípio Multiplicativo visto na subseção1.1.1, nos assessorando no número de elementos do Espaço Amostral $S = \{VV, Vv, vV, vv\}$, observe árvore de possibilidades abaixo,(ver Figura 11). Nela representamos na primeira bifurcação as opções para o primeiro gene do indivíduo e, nas bifurcações seguintes, as opções para o segundo gene.
Verificamos que $n(S)$ equivale ao produto do número de possibilidades V pelo número de possibilidades v das ervilhas, isto é, $2 \cdot 2 = 4$ possibilidades.

2.8 FUNÇÃO PROBABILIDADE E PARADOXO DE BERTRAND'S

O paradoxo de Bertrand's é um paradoxo da teoria elementar de probabilidade apresentada por Joseph Bertrand em seu trabalho Calcul des Probabilités. Apresentamos abaixo esse problema adaptado de [12] e [10].
O Paradoxo de Bertrand, consta no livro Calcul des Probabilités de onde obtemos o problema a seguir, passível de diferentes interpretações, de acordo com o **Espaço Amostral** considerado e por isso aceito como um paradoxo.

Exemplo 2.12. O problema é formulado da seguinte maneira:

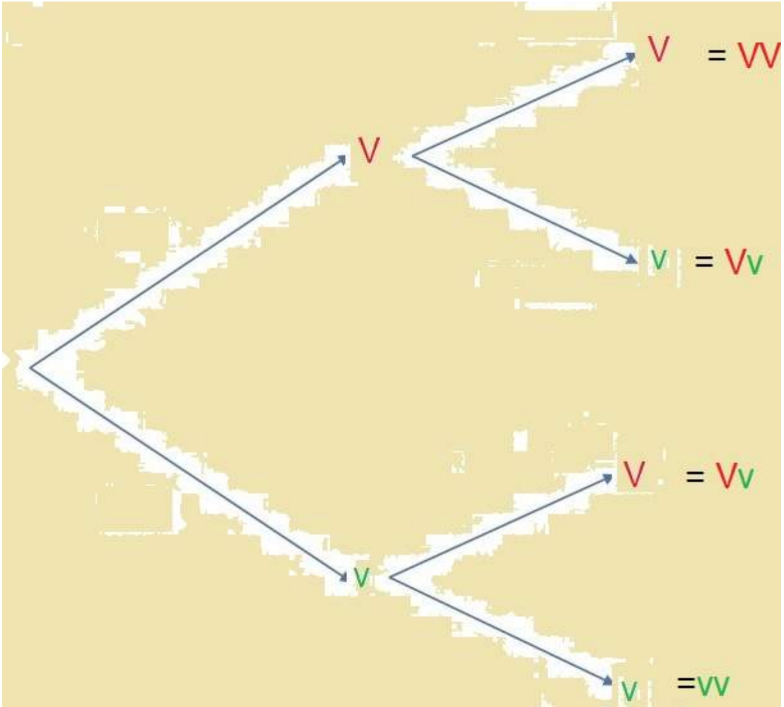

Figura 11: Árvore de possibilidades para **S**.

Em um círculo Ω de raio unitário está inscrito um triângulo equilátero ABC de lado l. Qual é a probabilidade de uma corda desse círculo escolhido ao acaso, ter comprimento maior do que o lado l do triângulo ABC?

Interpretação 1:

A corda $\overline{DE}$ é escolhida ao acaso da seguinte forma, escolhemos aleatoriamente um ponto P dentro do círculo Ω o ligamos ao centro O através de um segmento de reta $\overline{OP}$. A corda é traçada passando pelo ponto P de forma a ser perpendicular ao segmento $\overline{OP}$. Observe que a interpretação geométrica nos permite concluir que a circunferência δ de raio $\overline{OP}$ é uma circunferência concêntrica a Ω e inscrita no triângulo equilátero.(Ver Figura 12)

Ainda, em uma menção geométrica, o centro O de ambas as circunferências é o chamado Baricentro do Triângulo ABC, inscrito em Ω e circunscrito a δ, e divide a altura do triângulo na razão de $2:1$, consequentemente o raio r de δ é metade da medida do raio $R = 1$, assim $r = \frac{1}{2}$.

Interpretando $P(A)$ como a probabilidade da corda $\overline{CD}$ ter medida maior que lado l do triângulo ABC, observamos que $P(A)$ é então a razão entre a área do círculo δ e a área do círculo Ω, pois as cordas de medidas maiores que os lados do triângulo ABC são cordas internas ou tangentes ao círculo δ. Assim:

$$P(A) = \frac{\pi r^2}{\pi R^2} = \frac{r^2}{R^2} = \frac{(\frac{1}{2})^2}{1^2} = \frac{1}{4}.$$

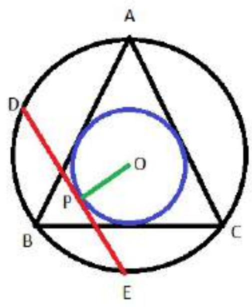

Figura 12: Interpretação 1.

Interpretação 2:
Escolhemos agora, um ponto ao acaso sobre a circunferência e em seguida escolhemos um segundo ponto. Sabemos que os vértices de um triângulo equilátero dividem a circunferência em três arcos congruentes, isto é, de mesma medida. Se escolhermos o primeiro ponto sobre um dos vértices do triângulo e o segundo ponto estiver sobre um desses arcos, oposto ao vértice escolhido, o comprimento da corda será maior do que o do lado do triângulo, porém, se estiver sobre os outros dois arcos, adjacentes aos arcos escolhidos, será menor. Logo, a probabilidade da corda escolhida ser maior que o lado do triângulo é igual a $1/3$, uma opção de arco conveniente num total de três arcos.(Ver Figura 13)

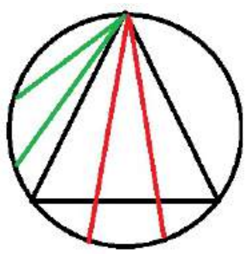

Figura 13: Interpretação 2.

Interpretação 3:

Podemos ainda escolher um ponto e em seguida observarmos o raio por ele determinado na circunferência circunscrita ao triângulo equilátero. Consideremos as cordas perpendiculares a esse raio. Podemos considerar duas regiões definidas por esse raio, tomando como referência uma região entre o centro da circunferência e o lado do triângulo e, uma região abaixo do lado do triângulo, tal que, as cordas que interceptam esse raio contidas na região entre o centro e o lado do triângulo, têm seus comprimentos maiores do que o lado do triângulo; e as cordas que cortam o raio na outra região de igual tamanho (garantido geometricamente pelo Baricentro mencionado na Interpretação 1 do exemplo 2.12, têm seu comprimento menor do que o lado do triângulo. Dessa forma concluímos que a probabilidade se escolher uma corda da circunferência com medida maior que o lado do triângulo equilátero inscrito é $\frac{1}{2}$.(Ver Figura 14).

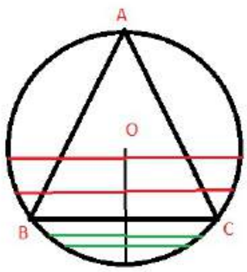

Figura 14: Interpretação 3.

Bertrand, diante desse paradoxo, postulou que, das soluções acima, nenhuma estava correta e nenhuma estava errada. Na verdade a situação paradoxal se dá em condicional ao espaço amostral observado.
Observe na Interpretação 1, o espaço amostral é a área do círculo Ω, na Interpretação 2, o espaço amostral é representado pelos três arcos congruentes obtidos pela intersecção da circunferência com os três vértices do triângulo equilátero. E, finalmente, na Interpretação 3, o espaço amostral é composto por duas regiões, uma região entre o centro da circunferência e o lado do triângulo e, uma região abaixo do lado do triângulo, limitada pela circunferência Ω.

Nota-se a importância do reconhecimento do espaço amostral S em cada experimento. Conforme a Observação 2.4.2, podemos aplicar o conceito de probabilidade usando comprimento de intervalos, medidas de áreas ou similares, caindo por exemplo numa “Probabilidade Geométrica” como no Exemplo 2.8, onde na Interpretação 1 da resolução, observamos que $P(A)$ é então a razão entre a área do círculo δ e a área do círculo Ω.

3

PROBABILIDADE CONDICIONAL E O TEOREMA DE BAYES

3.1 PROBABILIDADE CONDICIONAL

Ao falarmos de probabilidade condicional, nos referimos à probabilidade de ocorrer determinado evento sob a condição da ocorrência de algum outro. Por exemplo, qual seria a probabilidade de um meteorologista acertar a previsão do tempo de amanhã, com base no tempo que fez hoje?
Note que nesse exemplo foram usados dois eventos:

- Evento A, previsão do tempo de amanhã,
- Evento B, tempo de hoje,

os quais, teoricamente um pode não interferir no outro, ou seja, o tempo ocorrido hoje, pode não influenciar o tempo previsto para amanhã.
A notação usual para o que chamamos de Probabilidade Condicional entre os eventos A e B é:

$$P(A/B) = \frac{P(A \cap B)}{P(B)},$$

cuja interpretação é:
Probabilidade de ocorrer o evento A, sabendo que ocorreu o evento B é igual à razão entre a probabilidade de ocorrer A e B e a probabilidade de ocorrer B.
Observe que ao condicionarmos a ocorrência de A ao evento B, restringimos o espaço amostral de S para B.
Notamos que, em experimentos com resultados equiprováveis, obtemos:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{n(A \cap B)}{n(S)}}{\frac{n(B)}{n(S)}} = \frac{n(A \cap B)}{n(B)},$$

ou seja, a probabilidade condicional de ocorrer A dado que B ocorreu é a razão entre o número de resultados comuns a A e B, e o número de resultados em B.
Observemos os exemplos que tratam de experimentos com resultados equiprováveis:

Exemplo 3.1. Em uma sala de aula há 50 alunos, dos quais, 28 são meninas e 20 alunos usam óculos. Sabemos que, entre os meninos, apenas 8 usam óculos. Na turma em questão existe apenas um representante de sala.

(a) Qual a probabilidade de o representante de classe ser um aluno que usa óculos?

(b) Sabendo que o representante de sala é uma menina, qual a probabilidade de o representante de classe ser um aluno que usa óculos?

Considere os seguintes eventos:

- O: Usar óculos;
- F: Ser do sexo feminino.

Para responder a primeira pergunta observamos que, do total de 50 alunos, apenas 20 usam óculos. Logo:

$$P(O) = \frac{n(O)}{n(S)} = \frac{20}{50} = \frac{2}{5},$$

onde, como de costume, denotamos por S o espaço amostral composto pelos 50 alunos. Porém, se tivermos a informação adicional de que o representante de turma é uma menina, como no item (b), estamos condicionando, e portanto, restringindo nossa escolha ao grupo de meninas e não à sala toda, assim o espaço amostral passa a ser o grupo de meninas e não a sala toda e por isso obteremos:

$$P(O/F) = \frac{n(O \cap F)}{n(F)} = \frac{12}{28} = \frac{3}{7},$$

ou seja, probabilidade do representante usar óculos sabendo que é menino é igual ao número de pessoas com óculos e que são meninas entre o número total de meninas.

Note que o grande contraste da probabilidade condicional é restringir o total de resultados possíveis, isto é, restringir o espaço amostral.

Exemplo 3.2. Nas ervilhas, a textura pode ser lisa ou rugosa. O cruzamento de plantas que geram ervilhas de textura lisa com plantas cujas ervilhas possuem textura rugosa resultou uma geração que Mendel chamou de F_1, constituídas todas de sementes lisas. Isso caracteriza que o gene do tipo **lisa** é o que chamamos de **gene dominante** e o gene do tipo **rugosa**, se intitula **gene recessivo**.
Vamos usar a seguinte nomenclatura para calcular as probabilidades à seguir: representaremos com **R** para o gene dominante **lisa** e **r** para o gene recessivo **rugosa**.
Em um cruzamento entre plantas de ervilhas lisas **heterozigóticas**, isto é, que possui os dois genes (R e r), ou mesmo na autofecundação de plantas de ervilhas lisas heterozigóticas, a probabilidade esperada de os descendentes terem sementes lisas é de $\frac{3}{4}$.
Observe o resultado obtido, no quadro abaixo:(Ver tabela 4)

Tabela 4: Quadro de Punnet-Ervilhas Lisas

Variável	R	r
R	RR	Rr
r	rR	rr

No quadro verificamos, um espaço amostral $S = \{RR, Rr, rR, rr\}$ e o evento L que representa a ocorrência de ervilhas lisas é $L = \{RR, Rr, rR\}$.
Portanto:

$$P(L) = \frac{3}{4}.$$

Tomando agora, ao acaso um descendente desse cruzamento cujo genótipo é liso, qual a probabilidade de ele ser homozigótico?
Observamos no quadro que esse probabilidade é de $\frac{1}{3}$, pois a característica lisa é dominante. O espaço amostral passou a ser: $\{RR, Rr, rR\}$ e o evento lisa e homozigótica é $\{RR\}$.
Se M é o evento "ser homozigótico", podemos representar essa condicional da seguinte maneira:

$$P(M/L) = \frac{P(L \cap M)}{P(L)} = \frac{1/4}{3/4} = \frac{1}{3}.$$

Note que, como deveria:

$$P(M/L) = \frac{n(L \cap M)}{n(L)} = \frac{1}{3}.$$

3.2 DEPENDÊNCIA E INDEPENDÊNCIA DE EVENTOS

Alterar ou não a ocorrência de determinada situação em função de outra, permite que, à partir da relação entre dois eventos que podem ou não influenciar um ao outro, definamos a chamada Relação de Independência ou Dependência entre Eventos.
Sendo assim, é a Probabilidade Condicional que define se dois eventos independem ou não.
Segundo Abraham de Moivre:

> "Dois eventos são independentes quando eles não têm influência um sobre o outro, a probabilidade de realização de um não altera a realização do outro." [15]

Na prática, consideramos que dois eventos independem quando "um não interfere" na ocorrência do outro, dessa forma, podemos dizer que A e B são dois eventos independentes se:

$$P(A/B) = P(A),$$

pois ocorrer B ou não ocorrer B, não interfere de fato na ocorrência de A.
A independência de A e B pode ser reinterpretada pela seguinte proposição:

Proposição 3.1. *São equivalentes:*

(i) $P(A/B) = P(A)$;

(ii) $P(B/A) = P(B)$;

(iii) $P(A \cap B) = P(A) \cdot P(B)$.

Demonstração. Mostremos que (i) ocorre se, e somente se, ocorre (iii).
Se $P(A/B) = P(A)$, temos, pela definição de probabilidade condicional, que:

$$\frac{P(A \cap B)}{P(B)} = P(A).$$

Logo $P(A \cap B) = P(A) \cdot P(B)$.
Reciprocamente, se $P(A \cap B) = P(A) \cdot P(B)$, temos:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \cdot P(B)}{P(B)} = P(A).$$

Analogamente, prova-se que (ii) e (iii) são equivalentes. □

Observação 3.2.1. *A relação:*

$$\boxed{P(A \cap B) = P(A) \cdot P(B)}$$

*é usualmente conhecida como **Lei do Produto entre dois eventos independentes**.*

É importante notar que nem sempre dois eventos são independentes. Observe:

Exemplo 3.3. No lançamento de um dado sejam $A = \{2, 4, 6\}$ e $B = \{2, 3, 5\}$, ou seja "ser par" e "ser primo", respectivamente.
Temos $A \cap B = \{2\}$ donde $P(A \cap B) = 1/6$. No entanto, $P(A) = 1/2$ e $P(B) = 1/2$ donde $P(A) \cdot P(B) = 1/4$.
Logo, nesse exemplo, A e B não são independentes.

3.3 PROBABILIDADE CONDICIONAL COMO PROBABILIDADE

Pretendemos, em breve, tratar de problemas com múltiplas condicionais. Para tal é de fundamental importância o seguinte resultado:

Proposição 3.2. *Se $P : \mathbb{P}(S) \longrightarrow \mathbb{R}$ é função probabilidade segundo Kolmogorov e B é um evento no espaço amostral S então $P(\cdot/B) : \mathbb{P}(B) \longrightarrow \mathbb{R}$ é função probabilidade segundo Kolmogorov no espaço amostral B.*

Demonstração. Precisamos mostrar que $P(\cdot/B)$ satisfaz os três axiomas descritos na Seção 2.6.

1. Observe que
$$P(B/B) = \frac{P(B \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$
   Logo $P(\cdot/B)$ satisfaz o primeiro axioma.

2. Como para qualquer $A \subset B$ temos $P(B) = P(B \cap A) + P(B \cap \overline{A})$, pelo Axioma 3 satisfeito por P, e $0 \leq P(E)$ para todo evento E em S (Axioma 2), temos que $P(A \cap B) \leq P(B)$. Logo, dividindo $0 \leq P(A \cap B) \leq P(B)$ por $P(B)$, temos:
$$0 \leq \frac{P(A \cap B)}{P(B)} \leq 1,$$
   ou seja, $0 \leq P(A/B) \leq 1$ para qualquer evento A em B.

   Logo vale o Axioma 2 para $P(\cdot/B)$.

3. Dada uma sequência de eventos mutuamente exclusivos $E_1, E_2, \ldots$ em B temos:

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

onde a segunda igualdade segue pelo Axioma 3 para P. Dividindo a igualdade por $P(B)$ segue:

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

Portanto o Axioma 3 é satisfeito por $P(\cdot/B)$.

□

Observação 3.3.1. *Não é, em geral, verdade que se para uma dada probabilidade $P(\cdot)$ os eventos B e C são independentes, então para B e C são também independentes para a probabilidade $P(\cdot/A)$, qualquer que seja $A \subset S$. Ou seja, se $P(B \cap C) = P(B) \cdot P(C)$ então pode ocorrer $P(B \cap C/A) \neq P(B/A) \cdot P(C/A)$.*

Exemplo 3.4. Considere o lançamento de duas moedas. Denotemos cara por c e coroa por k. O espaço amostral pode ser representado então por $S = \{(c,c),(c,k),(k,c),(k,k)\}$. Considere os eventos $B = \{(c,c),(c,k)\}$, isto é, tirar cara na primeira moeda; $C = \{(c,c),(k,c)\}$, isto é, tirar cara na segunda moeda, e $A = \{(c,k),(k,c)\}$, ou seja, tirar cara apenas em uma moeda.

Observamos que:

$$P(B \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B)P(C),$$

donde temos B e C eventos independentes.

Porém:

$$P(B \cap C/A) = 0 \neq \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B/A)P(C/A),$$

ou seja, B e C não são independentes para $P(\cdot/A)$.

3.4 THOMAS BAYES (1701-1761)

O TEOREMA DE BAYES (alternativamente Lei de Bayes ou a Regra de Bayes), descreve a probabilidade de um evento, baseado em um conhecimento a priori de tudo o que pode estar relacionado ao evento, isto é, no caso de Maria, no exemplo dado na introdução desse trabalho, o exemplo 0.1, a probabilidade de sucesso em uma cirurgia seria calculada com base em mais do que uma condicional, por exemplo, hospital escolhido para se fazer a cirurgia, o número de dias da internação. Enfim, Maria pode ter uma visão mais segura de suas chances na cirurgia se a probabilidade de sucesso for calculada analisando-se duas ou mais condicionais e é isso que o Teorema de Bayes defende.

O nome do teorema se deve ao inglês Thomas Bayes, pastor protestante, provavelmente nascido em 1701. Sua obra mais famosa foi Essay Toward Solving a Problem in the Doctrine of Chances (Tentativa de elucidar um problema na doutrina ao acaso), publicada somente depois de sua morte.

Seu tema tratava de como formular convicções probabilísticas a respeito do mundo quando diante de novos dados, porém, quem desenvolveu o Teorema de Bayes foi o brilhante matemático e astrônomo francês Pierre-Simon Laplace. Ele argumentava ser possível prever o universo de forma perfeita, mas para isto, deveríamos saber a posição de cada partícula em seu interior, dando assim o passo inicial para a noção de Probabilidade Condicional.

O foco do Teorema de Bayes é a probabilidade condicionada, ou seja, probabilidade de uma teoria ou hipótese ser verdadeira condicionada a um ou mais acontecimentos ou informações.

Define-se assim um tipo particular de probabilidade, a Probabilidade Condicional para a qual podemos usar a seguinte notação:

$$P(A/B) = \frac{P(A \cap B)}{P(B)},$$

que nos dá a probabilidade de ocorrer um evento A sabendo que já ocorreu o evento B, ou tal que já tenha ocorrido o evento B, onde $P(A \cap B)$ representa a probabilidade de ocorrer elementos comuns aos conjuntos A e B, simultaneamente, e $P(B)$ representa probabilidade de ocorrer B.

O Teorema de Bayes ou Teorema da Inversão comumente usado para realizar inferências, isto é, conclusões, é um teorema com base em um certo número de condicionais que podem interferir em um resultado ou efeito.

Segundo Michel Henry em [7], a noção de **causa** ordena a equiprobabilidade a priori dos valores possíveis para uma probabilidade desconhecida, parte para reajustá-la a posteriori, introduzindo assim duas noções de probabilidade: a primeira a que procuramos estimar, é objetiva, a segunda, apurando o valor possível da precedente, colocada a priori, é subjetiva.

3.5 O TEOREMA DE BAYES

Teorema 3.3 (Bayes)**.** *Sejam P uma função probabilidade segundo Kolmogorov de espaço amostral S e A, B dois eventos em S. Vale então que:*

$$\boxed{P(A/B) = \frac{P(B/A)\cdot P(A)}{P(B/A)\cdot P(A) + P(B/\overline{A})\cdot P(\overline{A})}}.$$

Demonstração. Diretamente das definições de $P(A/B)$ e $P(B/A)$ temos:

$$P(A \cap B) = P(A/B)\cdot P(B),$$
$$P(B \cap A) = P(B/A)\cdot P(A).$$

Como $P(A \cap B) = P(B \cap A)$, temos que: $P(A/B)\cdot P(B) = P(B/A)\cdot P(A)$.
Portanto:

$$P(A/B) = \frac{P(B/A)\cdot P(A)}{P(B)}.$$

Por outro lado temos:

$$P(B) = P(B \cap (A \cup \overline{A})) = P((B \cap A) \cup (B \cap \overline{A})) = P(B/A)\cdot P(A) + P(B/\overline{A})\cdot P(\overline{A}),$$

conhecida como probabilidade total de ocorrer o evento B.
Dessa forma segue o **Teorema de Bayes**:

$$\boxed{P(A/B) = \frac{P(B/A)\cdot P(A)}{P(B/A)\cdot P(A) + P(B/\overline{A})\cdot P(\overline{A})}}.$$

□

Observação 3.5.1. *$P(A/B)$ é conhecida a **Probabilidade a Posteriori** (de A) e $P(A)$ é a **Probabilidade a Priori** de A.*
O Teorema de Bayes é uma das principais ferramentas que temos para "refinar" uma previsão probabilística (Probabilidade a Priori) quando novos eventos, em geral não independentes, acontecem. A Probabilidade a Posteriori é justamente o resultado desse "refinamento" frente às novas informações.

O estudo da probabilidade $P(A/B)$, pelo Teo. de Bayes, passa a levar em conta as intituladas probabilidades:

3.5.1 *Múltiplas Condicionais*

Observação 3.5.2. *Não é, em geral, verdade que se para uma dada probabilidade $P(\cdot)$ os eventos B e C são independentes, então para B e C são também independentes para a probabilidade $P(\cdot/A)$, qualquer que seja $A \subset S$. Ou seja, se $P(B \cap C) = P(B) \cdot P(C)$ então pode ocorrer $P(B \cap C/A) \neq P(B/A) \cdot P(C/A)$.*

Exemplo 3.5. Considere o lançamento de duas moedas. Denotemos cara por c e coroa por k. O espaço amostral pode ser representado então por $S = \{(c,c),(c,k),(k,c),(k,k)\}$. Considere os eventos $B = \{(c,c),(c,k)\}$, isto é, tirar cara na primeira moeda; $C = \{(c,c),(k,c)\}$, isto é, tirar cara na segunda moeda, e $A = \{(c,k),(k,c)\}$, ou seja, tirar cara apenas em uma moeda.
Observamos que:

$$P(B \cap C) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B)P(C),$$

donde temos B e C eventos independentes.
Porém:

$$P((B \cap C)/A) = 0 \neq \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P(B/A)P(C/A),$$

ou seja, B e C não são independentes para $P(\cdot/A)$.

3.6 APLICAÇÕES DO TEOREMA DE BAYES

O uso da probabilidade condicional de forma mais aprofundada na programação do ensino médio e na cobrança dos vestibulares tem ocorrido de forma mais frequente. Os números apresentados nos exemplos a seguir buscam evidenciar a importância de uma correta aplicação da teoria de Bayes, no entanto, são valores que não necessariamente correspondem a valores reais e experimentalmente testados.
Observemos a seguir um exemplo adaptado do material do curso pré-vestibular Objetivo:

Exemplo 3.6. Sabe-se que 80% dos pênaltis marcados a favor do Brasil são cobrados por jogadores que jogam no exterior. A probabilidade de um pênalti ser convertido é de 40% se o jogador joga no exterior e é de 70% em caso contrário. Um pênalti a favor do Brasil acabou de ser marcado:

(a) Qual é a probabilidade do pênalti ser cobrado por um jogador que joga no exterior ser convertido?

(b) Qual a probabilidade do pênalti ser convertido?

(c) Um pênalti foi marcado a favor do Brasil e acabou de ser desperdiçado. Qual a probabilidade de que o cobrador jogue no exterior?

Para auxiliar o item (a), veja Figura 15 nomeamos por A o evento "jogador jogar no exterior", e por B o evento "pênalti ser convertido".
Temos que:

$$P(A \cap B) = 0,80 \cdot 0,40 = 0,32 = 32\%.$$

Sendo que $P(A \cap B)$ é a probabilidade de ocorrer A e B, isto é, do jogador jogar no exterior e o pênalti ser convertido.
Respondendo então a pergunta (b), mantemos a mesma notação acima e observamos que:

$$P(B) = P(A \cap B) + P(\overline{A} \cap B) = 0,80 \cdot 0,40 + 0,20 \cdot 0,70 = 0,32 + 0,14 = 0,46 = 46\%.$$

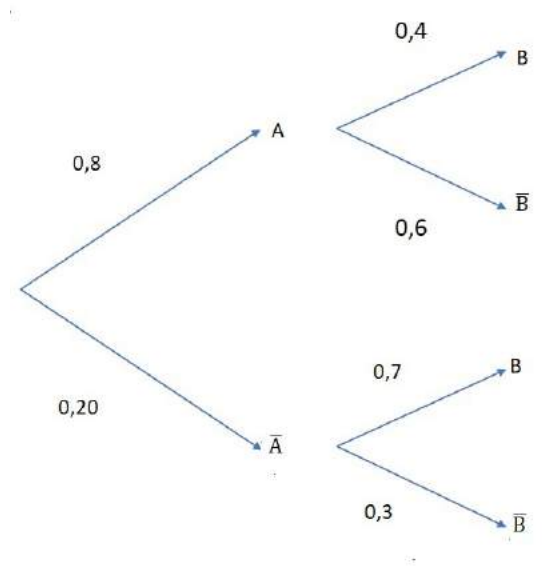

Figura 15: Diagrama das Probabilidades

Por fim, observamos que a questão (c) é uma clássica aplicação do Teorema de Bayes, apresentado na seção anterior.
Temos:

$$P(A/\overline{B}) = \frac{P(\overline{B}/A)\cdot P(A)}{P(\overline{B}/A)\cdot P(A) + P(\overline{B}/\overline{A})\cdot P(A)}$$

$$= \frac{0,6\cdot 0,8}{0,6\cdot 0,8 + 0,3\cdot 0,2} = \frac{0,48}{0,48+0,06} = \frac{0,48}{0,54} = \frac{8}{9} \approx 88,9\%.$$

Exemplo 3.7. Um bom exemplo para analisar o raciocínio bayesiano é o exemplo médico, relacionado com a chance de uma mulher ter câncer de mama, usando dados de um artigo do norte americano Eliezer Yudkowsky, pesquisador da inteligência artificial. Recomenda-se que, a partir dos 40 anos, as mulheres façam mamografias anuais. Nessa idade 1% das mulheres são portadoras de um tumor assintomático de mama.
Sabe-se que a mamografia apresenta resultado positivo em 80% das mulheres com câncer de mama, mas esse mesmo resultado ocorre também com $9,6\%$ das mulheres sem o câncer (falso positivo).
Imagine agora uma pessoa desesperada porque fez uma mamografia de rotina e o resultado foi positivo. Qual a probabilidade de ser um câncer de mama?
Pelo Teorema de Bayes:

$$P(A/M_+) = \frac{P(M_+/A)\cdot P(A)}{P(M_+/A)\cdot P(A) + P(M_+/\overline{A})\cdot P(\overline{A})},$$

onde o evento A é "ter câncer" e M_+ é "ter mamografia positiva".
Segundo dados observados, temos: $P(A) = 1\% = 0,01$ e, consequentemente, $P(\overline{A}) = 99\% = 0,99$, já que $P(A) + P(\overline{A}) = 1$. Assim, $P(A) = 0,01$ é a probabilidade À PRIORI da mulher ter câncer.
O resultado de uma mamografia dá resultado positivo em 80% das mulheres que realmente estão com câncer, isto é $P(M_+/A) = 0,8$, e dá um resultado falso positivo em $9,6\%$ das mulheres que não estão doentes, ou seja $P(M_+/\overline{A}) = 0,096$. Desses dados obtemos as probabilidades conjuntas:
$P(M_+ \cap A) = P(M_+/A)\cdot P(A) = 0,8\cdot 0,01 = 0,008$ é a probabilidade de ocorrer mamografia positiva e a pessoa estar com câncer e,
$P(M_+ \cap \overline{A}) = P(M_+/\overline{A})\cdot P(\overline{A}) = 0,096\cdot 0,99 = 0,09504$ é a probabilidade de ocorrer mamografia positiva e a pessoa não ter câncer.

Temos assim como probabilidade total $P(M_+)$, a soma das duas probabilidades conjuntas calculadas:
$0,008 + 0,09504 = 0,10304$.
Portanto, pelo Teorema de Bayes:
$P(A/M_+) = \frac{P(M_+ \cap A)}{P(M_+)} = \frac{0{,}008}{0{,}1034} \approx 0,0774 = 7,74\%$;
e $P(\overline{A}/M_+) = \frac{P(M_+ \cap \overline{A})}{P(M_+)} = \frac{0{,}09504}{0{,}1034} \approx 0,92 = 92\%$.
Observemos que o fato da mamografia apresentar nódulo, quer dizer, resultado positivo, não podemos concluir que a pessoa está doente, aliás as probabilidades acima mostram que é muito pequena a chance desse nódulo ser um câncer $(7,74\%)$.

Se faz necessário incluir mais condicionais, isto é, mais informações que vão desde hereditariedade à exames mais invasivos para um diagnóstico mais preciso.
Podemos por exemplo anexar aos resultados da mamografia, os resultados de uma biopsia.

Exemplo 3.8. À Priori, vamos supor os seguintes dados:
$P(B_+/A) = 0,99$ significa que a biópsia acusa resultado positivo com segurança de 99% no caso da pessoa estar com câncer.
$P(B_+/\overline{A}) = 0,005$ significa que, no caso de um falso positivo, a biópsia dá uma chance de $0,5\%$ do nódulo ser câncer.
Portanto, calculando a probabilidade total $P(B_+)$ dessa biópsia dar positiva para um câncer, temos:

$$P(B_+) = P(B_+/A) \cdot P(A) + P(B_+/\overline{A}) \cdot P(\overline{A})$$

$$\Rightarrow P(B_+) = 0,99 \cdot 0,01 + 0,005 \cdot 0,99 = 0,01485$$

Sendo $P(B_+/A) \cdot P(A) = 0,0099$ a probabilidade conjunta de ocorrer resultado positivo entre as pessoas que têm câncer e,
$P(B_+/\overline{A}) \cdot P(\overline{A}) = 0,00495$, a probabilidade conjunta de ocorrer resultado positivo entre as pessoas que não têm câncer.
Observe que estamos relatando resultados apenas da biópsia como um único exame, independente da mamografia.
Pelo Teorema de Bayes concluímos:

$$P(A/B_+) = \frac{P(B_+/A) \cdot P(A)}{P(B_+)} = \frac{0,0099}{0,01485} = 0,667 = 66,7\%.$$

Note que chegamos a um resultado mais significativo do que o resultado da mamografia.

Agora, conjugando os resultados dos dois exames, fazemos a seguinte pergunta: qual é a probabilidade da pessoa estar com câncer, dado que mamografia e biópsia deram positivas?
Para responder essa pergunta é necessário alguma informação sobre como os dois exames se relacionam. Assim propomos a seguir duas variações do mesmo problema.

Exemplo 3.9. Suponha que, como nos exemplos anteriores, 1% das mulheres são portadoras de câncer, que a mamografia apresenta resultado positivo em 80% das mulheres com câncer, mas esse mesmo resultado ocorre também com $9,6\%$ das mulheres sem o câncer.
Responda qual a probabilidade da pessoa estar com câncer, dado que mamografia e biópsia deram positivas nos seguintes casos

(a) Mamografia e biópsia positivas são eventos independentes para a probabilidade $P(\cdot/A)$ (probabilidade quando a pessoa tem câncer), que a biópsia tem resultado positivo em 99% no caso da pessoa estar com câncer e a probabilidade de ambos os resultados darem positivos e a pessoa não estar com câncer é de $0,05\%$.

(b) A probabilidade da biópsia dar positiva no grupo das pessoas com câncer e mamografia positivas é de $99,5\%$, e biópsia dar positiva no grupo das pessoas sem câncer, mas com mamografia positiva é de $0,1\%$.

Para responder o item (a), observamos que, pelo Teorema de Bayes, vale:

$$P(A/(M_+ \cap B_+)) = \frac{P((M_+ \cap B_+)/A) \cdot P(A)}{P((M_+ \cap B_+)/A) \cdot P(A) + P((M_+ \cap B_+)/\overline{A}) \cdot P(\overline{A})}.$$

Como M_+ e B_+ são eventos independentes para a probabilidade $P(\cdot/A)$, a Lei do Produto garante que:

$$P((M_+ \cap B_+)/A) = P(M_+/A) \cdot P(B_+/A).$$

Do enunciado temos que $P((M_+ \cap B_+)/\overline{A}) = 0,0005$.
Portanto:

$$\begin{aligned}
P(A/(M_+ \cap B_+)) &= \frac{P(M_+/A) \cdot P(B_+/A) \cdot P(A)}{P(M_+/A) \cdot P(B_+/A) \cdot P(A) + P((M_+ \cap B_+)/\overline{A}) \cdot P(\overline{A})} \\
&= \frac{0,8 \cdot 0,99 \cdot 0,01}{0,8 \cdot 0,99 \cdot 0,01 + 0,0005 \cdot 0,99} = 94,12\%.
\end{aligned}$$

Logo, mamografia e biópsia juntas, nos passam um diagnóstico mais preciso, confirmando com o Teorema de Bayes que é de $99,94\%$ a probabilidade dessa pessoa estar realmente doente, caso ambas apresentem resultados positivos na análise do nódulo encontrado.
Para resolver (b), observamos que podemos aplicar Bayes utilizando a função probabilidade $Q(\cdot) = P(\cdot/M_+)$, pois $Q(A/B_+) = P(A/(M_+ \cap B_+))$:

$$Q(A/B_+) = \frac{Q(B_+/A) \cdot Q(A)}{Q(B_+/A) \cdot Q(A) + Q(B_+/\overline{A}) \cdot Q(\overline{A})}.$$

Observamos então que $Q(A) = P(A/M_+) \approx 0,0774$ (Exemplo 3.7).
Pelo enunciado temos que probabilidade da biópsia dar positiva no grupo das pessoas com câncer e mamografia positivas é de $99,5\%$, isto é, $Q(B_+/A) = P(B_+/A \cap M_+) = 0,995$ e biópsia dar positiva no grupo das pessoas sem câncer, mas com mamografia positiva é de $0,1\%$, ou seja, $Q(B_+/\overline{A}) = P(B_+/\overline{A} \cap M_+) = 0,001$.
Assim, temos:

$$\begin{aligned} P(A/(M_+ \cap B_+)) = Q(A/B_+)) &= \frac{Q(B_+/A) \cdot Q(A)}{Q(B_+/A) \cdot Q(A) + Q(B_+/\overline{A}) \cdot Q(\overline{A})} \\ &= \frac{P(B_+/(A \cap M_+)) \cdot P(A/M_+)}{P(B_+/(A \cap M_+)) \cdot P(A/M_+) + P(B_+/(\overline{A} \cap M_+)) \cdot P(\overline{A}/M_+)} \\ &= \frac{0,995 \cdot 0,0774}{0,995 \cdot 0,0774 + 0,001 \cdot (1 - 0,0774)} \approx 98,82\%. \end{aligned}$$

3.6.1 *O Teorema de Bayes Aplicado às Ervilhas de Mendel*

Comecemos nosso exemplo supondo uma população de 4000 ervilhas, sendo assim distribuídos seus genótipos:

- 1000 ervilhas RR;
- 2000 ervilhas Rr;
- 1000 ervilhas rr;

onde R significa gene dominante liso e r, gene recessivo rugoso.
Adotando como representação do evento a própria descrição do genótipo da ervilha temos:

- $P(RR) = \frac{1}{4}$,

- $P(Rr) = \frac{2}{4} = \frac{1}{2}$,
- $P(rr) = \frac{1}{4}$.

Montamos abaixo a Tabela 5 onde representamos nas linhas o genótipo do primeiro genitor e nas colunas o genótipo do segundo genitor, evidenciando então a probabilidade de se formar um "casal" com os genótipos indicados:

mãe/pai	RR	Rr	rr
RR	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$
Rr	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
rr	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

Tabela 5: Genitores

Assim, a probabilidades de obter o evento **genitores lisos**, denotado aqui por G, é dada por:

$$\begin{aligned}P(G) &= P(RR \times RR \cup Rr \times RR \cup Rr \times Rr \cup RR \times Rr)\\ &= P(RR \times RR) + P(Rr \times RR) + P(Rr \times Rr) + P(RR \times Rr)\\ &= \frac{1}{16} + \frac{1}{8} + \frac{1}{4} + \frac{1}{8} = \frac{9}{16}.\end{aligned}$$

Denotemos então por f o evento "ocorrer filhos rugosos", e portanto rr, já que r é o gene recessivo. Vamos calcular a probabilidade de ocorrer genitores lisos (G) sabendo que ocorreu ervilha rugosa (f), isto é, $P(G/f)$.

Pelo Teorema de Bayes resolvemos essa questão levando em conta as seguintes condicionais:

$$P(G/f) = \frac{P(f/G) \cdot P(G)}{P(f)},$$

sendo $P(f/G)$ a probabilidade de ocorrer filhos rugosos tal que os genitores sejam lisos e $P(f)$ a probabilidade total de se obter ervilhas rugosas. Sabemos que:

$$P(f/G) = \frac{P(f \cap G)}{P(G)}.$$

Quando falamos em $P(f \cap G)$, o evento G se subdivide nos seguintes casos de ocorrer pais lisos (vistos anteriormente na subseção 5): $G = G_1 \cup G_2 \cup G_3 \cup G_4$, onde $G_1 = (RR \times RR)$, $G_2 = (Rr \times RR)$, $G_3 = (RR \times Rr)$ e $G_4 = (Rr \times Rr)$.

Assim

$$\begin{aligned} P(f \cap G) &= P(f \cap (G_1 \cup G_2 \cup G_3 \cup G_4)) \\ &= P(f \cap G_1) + P(f \cap G_2) + P(f \cap G_3) + P(f \cap G_4) \\ &= P(f/G_1) \cdot P(G_1) + P(f/G_2) \cdot P(G_2) + P(f/G_3) \cdot P(G_3) + P(f/G_4) \cdot P(G_4). \end{aligned}$$

Observemos as tabelas de cruzamento abaixo para encontrar $P(f/G_1)$, $P(f/G_2)$, $P(f/G_3)$ e $P(f/G_4)$.

Do cruzamento $G_1 = (RR \times RR)$ temos:

	R	R
R	RR	RR
R	RR	RR

Portanto $P(f/G_1 = 0)$, já que não ocorre rr.

Dos cruzamentos $G_2 = (Rr \times RR)$ e $G_3 = (RR \times Rr)$ temos:

	R	r
R	RR	Rr
R	RR	Rr

Portanto $P(F/G_2) = P(F/G_2) = 0$, já que aqui também não ocorre rr.

Por fim, do cruzamento $G_4 = (Rr \times Rr)$ temos:

	R	r
R	RR	Rr
r	Rr	rr

Portanto $P(f/G_4) = \frac{1}{4})$.

Assim:

$$\begin{aligned} P(f \cap G) &= P(f/G_1) \cdot P(G_1) + P(f/G_2) \cdot P(G_2) + P(f/G_3) \cdot P(G_3) + P(f/G_4) \cdot P(G_4) \\ &= 0 \cdot \frac{1}{16} + 0 \cdot \frac{1}{8} + \frac{1}{4} \cdot \frac{1}{4} + 0 \cdot \frac{1}{8} = \frac{1}{16}. \end{aligned}$$

Donde temos:

$$P(f/G) = \frac{P(f \cap G)}{P(G)} = \frac{1/16}{9/16} = \frac{1}{9}.$$

Como $P(f)$ é a probabilidade total de obter filhos rugosos, denotando por G_5 o evento pais $(Rr \times rr)$, G_6 pais $(rr \times Rr)$, G_7 genitores $(RR \times rr)$, G_8 genitores $(rr \times RR)$ e G_9 genitores $(rr \times rr)$ temos:

$$\begin{aligned} P(f) &= P(f/G_1) \cdot P(G_1) + P(f/G_2) \cdot P(G_2) + P(f/G_3) \cdot P(G_3) \\ &+ P(f/G_4) \cdot P(G_4) + P(f/G_5) \cdot P(G_5) + P(f/G_6) \cdot P(G_6) \\ &+ P(f/G_7) \cdot P(G_7) + P(f/G_8) \cdot P(G_8) + P(f/G_9) \cdot P(G_9) \\ &= 0 + 0 + 0 + \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{8} + \frac{1}{2} \cdot \frac{1}{8} + 0 + 0 + + \frac{1}{16} \cdot 1 = \frac{1}{4}. \end{aligned}$$

E assim:

$$\begin{aligned} P(G/f) &= \frac{P(f/G) \cdot P(G)}{P(f)} \\ &= \frac{\frac{1}{9} \cdot \frac{9}{16}}{\frac{1}{4}} = \frac{1}{4}. \end{aligned}$$

Podemos confirmar esses resultados calculando $P(\overline{X}/A)$, ou seja, probabilidade dos pais não serem lisos, sabendo que os filhos são rugosos.
Sendo $\overline{G}$ o evento ocorrer pais não lisos, e como já mencionado, evento f, representando a ocorrência de ocorrer filhos rugosos, temos que $\overline{G} = G_5 \cup G_6 \cup G_7 \cup G_8 \cup G_9$, donde:

$$\begin{aligned} P(\overline{G}) &= P(G_5 \cup G_6 \cup G_7 \cup G_8 \cup G_9) \\ &= P(G_5) + P(G_6) + P(G_7) + P(G_8) + P(G_9) = \frac{1}{16} + \frac{1}{8} + \frac{1}{16} + \frac{1}{8} + \frac{1}{16} = \frac{7}{16}, \end{aligned}$$

probabilidades essas representadas na tabela 5.
Assim:

$$\begin{aligned} P(f \cap \overline{G}) &= P(f \cap G_5) + P(f \cap G_6) + P(f \cap G_7) + P(f \cap G_8) + P(f \cap G_9) \\ &= P(f/G_5) \cdot P(G_5) + P(f/G_6) \cdot P(G_6) + P(f/G_7) \cdot P(G_7) \\ &+ P(f/G_8) \cdot P(G_8) + P(f/G_9) \cdot P(G_9) \\ &= \frac{1}{16} \cdot 0 + \frac{1}{8} \cdot \frac{1}{2} + \frac{1}{16} \cdot 0 + \frac{1}{8} \cdot \frac{1}{2} + \frac{1}{16} \cdot 1 = \frac{3}{16}. \end{aligned}$$

Observe que, de fato, temos:

$$P(f) = P(f \cap G) + P(f \cap \overline{G}).$$

Além disso:

$$P(f/\overline{G}) = \frac{P(f/\overline{G})}{P(\overline{G})} = \frac{\frac{3}{16}}{\frac{7}{16}} = \frac{3}{7}.$$

Então, por Bayes:

$$P(\overline{G}/f) = P(f/\overline{G}) \cdot P(\overline{G}) = \frac{\frac{3}{7} \cdot \frac{7}{16}}{\frac{1}{4}} = \frac{3}{4}.$$

Verificando assim o resultado de $P(X/A) = \frac{1}{4}$:

$$P(G/f) + P(\overline{G}/f) = \frac{1}{4} + \frac{3}{4} = 1.$$

4

PARADOXOS DA PROBABILIDADE CONDICIONAL

4.1 PARADOXO DE BERTRAND

O exemplo a seguir representa um Paradoxo de Bertrand, tendo sido Bertrand citado no Capítulo 2, mais especificamente na Seção 2.8.
Dadas três caixas:

- A primeira caixa, contém duas moedas de ouro;
- A segunda caixa, contém duas moedas de prata;
- A terceira caixa, contém uma moeda de ouro e uma moeda de prata.

Observa-se que, ao se escolher uma caixa ao acaso, e se retirar uma moeda aleatoriamente, caso aconteça uma moeda de ouro, temos a tendência de concluir que em um próximo sorteio será de $\frac{1}{2}$ a probabilidade de ocorrer moeda de ouro novamente. Essa tendência decorre do fato de existirem apenas duas caixas com moedas de ouro, uma dessas caixas tem uma segunda moeda de ouro e a outra uma moeda de prata. Uma análise superficial do problema nos leva a crer que a ocorrência das caixas 1 e 3 é equiprovável, o que, de fato, ocorre quando não há evento conhecido. No entanto a ocorrência de uma moeda de ouro "quebra" a equiprobabilidade e, na verdade, a probabilidade de ocorrer moeda de ouro novamente torna-se $\frac{2}{3}$.
Representando-se G como evento "ocorrer no sorteio de uma moeda uma moeda de ouro" e por G_2 como evento "ocorrer no sorteio de duas moedas, duas de ouro".
Vamos combinar os seguintes eventos para prosseguir com o raciocínio:
C_1= evento ocorrer caixa 1,
C_2= evento ocorrer caixa 2,
C_3= evento ocorrer caixa 3.

Ao analisarmos a probabilidade $P(G_2/G)$, melhor representada como $P(C_1/G)$ podemos trabalhar com o Teorema de Bayes, para comprovarmos que o resultado, é $P(C_1/G) = \frac{2}{3}$.
Observemos o raciocínio segundo Teorema de Bayes:

$$P(C_1/G) = \frac{P(G/C_1) \cdot P(C_1)}{P(G/C_1) \cdot P(C_1) + P(G/C_2) \cdot P(C_2) + P(G/C_3) \cdot P(C_3)} =$$

$$\frac{1 \cdot \frac{1}{3}}{1 \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{1}{3}} =$$

$$\frac{1 \cdot \frac{1}{3}}{\frac{1}{3} \cdot (1 + 0 + \frac{1}{2})} = \frac{1}{\frac{3}{2}} = \frac{2}{3}.$$

De fato, as probabilidades aqui envolvidas inicialmente são equiprováveis, por exemplo, a probabilidade de ocorrer cada caixa é a mesma, e igual a $\frac{1}{3}$, e a probabilidade de acontecer cada moeda em cada caixa é $\frac{1}{6}$, que pode ser representada com a seguinte notação de eventos:

- Na primeira caixa temos duas moedas a (ouro) e b (ouro).
- Na segunda caixa temos duas moedas c (prata) e d (prata).
- Na terceira caixa temos duas moedas d (ouro) e f (prata).

Assim temos um total de seis moedas e: $P(a) = P(b) = P(c) = P(d) = P(e) = P(f) = \frac{1}{6}$.
O evento "ocorrer no sorteio de uma moeda uma moeda de ouro" é então $G = \{a, b, d\}$.
Ocorrer a caixa 1 é $C_1 = \{a, b\}$, ocorrer a caixa 2 é $C_2 = \{c, d\}$, Ocorrer a caixa 3 é $C_3 = \{d, e\}$.
Daí, como os eventos são equiprováveis temos:

$$P(C_1/G) = \frac{P(C_1 \cap G)}{P(G)} = \frac{n(C_1 \cap G)}{n(G)} = \frac{2}{3}.$$

A "quebra" da equiprobabilidade está no fato de que, ocorrendo uma moeda de ouro, temos probabilidade maior de termos tirado essa moeda da caixa 1, pois a primeira moeda poderia ser a, b ou c (com igual probabilidade), porém tanto a quanto b estão na caixa 1.

4.2 PARADOXO DE SIMPSON

O Paradoxo de Simpson representa uma importante ferramenta para comparar resultados que apresentam uma certa tendência em diferentes grupos de dados, contudo, quando considerados os mesmos grupos simultaneamente, as tendências não se confirmam. Observe o seguinte exemplo cujos números foram tirados do site [3]:

Exemplo 4.1. Imagine que uma pessoa precise realizar uma cirurgia e que para tal ela tenha que optar entre dois Hospitais. No Hospital A o paciente recebe a informação de que o Hospital A tem maior taxa de sucesso que o Hospital B em todos os tipos de cirurgia. Chegando no Hospital B, por sua vez, o atendente informa ao paciente que a taxa de sucesso em cirurgias é maior no Hospital B. Isso é possível? Alguém tem que estar necessariamente mentindo?
Observe a Tabela 6:

	Hospital A		Hospital B	
Tipo	número	sucesso	número	sucesso
Cirurgia Geral	359	292	88	70
Cirurgia Plástica	1836	1449	514	391
Cirurgia Maxilo-facial	299	178	222	113
Otorrinolaringologia	2086	434	86	12
Ortopedia	149	13	45	2

Tabela 6: Dados numéricos, por especialidades, dos Hospitais A e B

Resumidamente apresentamos os dados na Tabela 7:

	Hospital A		Hospital B	
	número	sucesso	número	sucesso
Total de cirurgias	4729	2366	955	588

Tabela 7: Tabela Resumo de Sucesso em Cirurgias dos Hospitais A e B

Denotemos o evento "cirurgia no Hospital A" por A e "cirurgia no Hospital B" por B. Seja S o evento "ter cirurgia bem sucedida". Sejam G, Pl, M, Ol e Op os eventos "a cirurgia é do tipo: Geral, Plástica, Maxilo-facial, Otorrinolaringologia e Ortopedia", respectivamente.

Da Tabela 7 observamos:

$$P(S/A) = \frac{2366}{4729} \approx 50,03\% < 61,57\% \approx \frac{588}{955} = P(S/B).$$

Logo a probabilidade de sucesso de cirurgia no Hospital B é, de fato, maior que se a cirurgia ocorresse no Hospital A.

Por outro lado temos:

$$P(S/(A \cap G)) = \frac{292}{359} \approx 81,34\% > 79,55\% \approx \frac{70}{88} = P(S/(B \cap G));$$

$$P(S/(A \cap Pl)) = \frac{1449}{1836} \approx 78,92\% > 76,07\% \approx \frac{391}{514} = P(S/(B \cap Pl));$$

$$P(S/(A \cap M)) = \frac{178}{299} \approx 59,53\% > 50,90\% \approx \frac{113}{222} = P(S/(B \cap M));$$

$$P(S/(A \cap Ol)) = \frac{434}{2086} \approx 20,81\% > 13,95\% \approx \frac{12}{86} = P(S/(B \cap Ol));$$

$$P(S/(A \cap Op)) = \frac{13}{149} \approx 8,72\% > 4,44\% \approx \frac{2}{45} = P(S/(B \cap Op)).$$

Portanto concluímos que o Hospital A tem maior taxa de sucesso que o Hospital B em todos os tipos de cirurgia. Ou seja, ambos os atendentes passaram informações corretas.

Qual é, então, a relação entre essas probabilidades?

Observe:

$$
\begin{aligned}
P(S/A) &= \frac{P(A/S)P(S)}{P(A)} \\
&= \frac{P(A \cap (G \cup Pl \cup M \cup Ol \cup Op)/S)P(S)}{P(A)} \\
&= \frac{P\left((A \cap G) \cup (A \cap Pl) \cup (A \cap M) \cup (A \cap Ol) \cup (A \cap Op)/S\right) P(S)}{P(A)} \\
&= (P(A \cap G/S) + P(A \cap Pl/S) + P(A \cap M/S) \\
&\quad + P(A \cap Ol/S) + P(A \cap Op/S))\frac{P(S)}{P(A)} \\
&= \left(\frac{P(S/A \cap G)P(A \cap G)}{P(S)} + \frac{P(S/A \cap Pl)P(A \cap Pl)}{P(S)} + \frac{P(S/A \cap M)P(A \cap M)}{P(S)}\right. \\
&\quad \left. + \frac{P(S/A \cap Ol)P(A \cap Ol)}{P(S)} + \frac{P(S/A \cap Op)P(A \cap Op)}{P(S)}\right)\frac{P(S)}{P(A)} \\
&= \Big(P(S/A \cap G)P(A \cap G) + P(S/A \cap Pl)P(A \cap Pl) + P(S/A \cap M)P(A \cap M) \\
&\quad + P(S/A \cap Ol)P(A \cap Ol) + P(S/A \cap Op)P(A \cap Op)\Big)/P(A) \\
&= P(S/A \cap G)P(G/A) + P(S/A \cap Pl)P(Pl/A) + P(S/A \cap M)P(M/A) \\
&\quad + P(S/A \cap Ol)P(Ol/A) + P(S/A \cap Op)P(Op/A).
\end{aligned}
$$

Analogamente:

$$
\begin{aligned}
P(S/B) = P(S/B \cap G)P(G/B) + P(S/B \cap Pl)P(Pl/B) + P(S/B \cap M)P(M/B) \\
+ P(S/B \cap Ol)P(Ol/B) + P(S/B \cap Op)P(Op/B).
\end{aligned}
$$

Note por fim que:

- $$P(G/A) = \frac{359}{4729} \approx 7,59\% \qquad P(G/B) = \frac{88}{955} \approx 9,21\%;$$
- $$P(Pl/A) = \frac{1836}{4729} \approx 38,82\% \qquad P(Pl/B) = \frac{514}{955} \approx 53,82\%;$$
- $$P(M/A) = \frac{299}{4729} \approx 6,32\% \qquad P(M/B) = \frac{222}{955} \approx 23,24\%;$$
- $$P(Ol/A) = \frac{2086}{4729} \approx 44,11\% \qquad P(Ol/B) = \frac{86}{955} \approx 9,01\%;$$

- $$P(Ol/A) = \frac{149}{4729} \approx 3,15\% \qquad P(Ol/B) = \frac{45}{955} \approx 4,71\%.$$

Daí, para o Hospital A temos a Tabela 8:

	$P(S/(A \cap \cdot))$	$P(\cdot/A)$	$P(S/(A \cap \cdot)) \cdot P(\cdot/A)$
G	$81,34\%$	$7,59\%$	$0,06173706$
Pl	$78,92\%$	$38,82\%$	$0,30636744$
M	$59,53\%$	$6,32\%$	$0,03762296$
Ol	$20,81\%$	$44,11\%$	$0,09179291$
Op	$8,72\%$	$3,15\%$	$0,0027468$
		Soma:	$50,03\%$

Tabela 8: Hospital A

E, para o Hospital B temos a Tabela 9:

	$P(S/(A \cap \cdot))$	$P(\cdot/A)$	$P(S/(A \cap \cdot)) \cdot P(\cdot/A)$
G	$79,55\%$	$9,21\%$	$0,07326555$
Pl	$76,07\%$	$53,82\%$	$0,40940874$
M	$50,90\%$	$23,24\%$	$0,1182916$
Ol	$13,95\%$	$9,01\%$	$0,01256895$
Op	$4,44\%$	$4,71\%$	$0,00209124$
		Soma:	$61,56\%$

Tabela 9: Hospital B

Esse é um exemplo do Paradoxo de Simpson, que ocorre quando dados informados escondem variáveis condicionais importantes.
Observa-se que os dados se comportam de determinada forma quando divididos em grupos e de forma diferente quando combinados os grupos em função de determinadas condicionais.
Para evitar o Paradoxo de Simpson, os dados devem ser divididos cuidadosamente em variáveis as mais detalhadas possíveis.

4.2.1 *Caso Geral*

Consideremos eventos E, F e G, dentro de um Espaço Amostral S, tal que a probabilidade total de ocorrer o evento E é:

$$P(E) = P(E/F) \cdot P(F) + P(E/\overline{F}) \cdot P(E/\overline{F}).$$

Consideramos agora, a probabilidade condicional:

$$P(E/G) = \frac{P(E \cap G)}{P(G)} = \frac{P(E \cap F \cap G) + P(E \cap \overline{F} \cap G)}{P(G)}.$$

Como:

$$P(E \cap F \cap G) = P(E/(F \cap G)) \cdot P(F \cap G)$$

e:

$$P(E \cap \overline{F} \cap G) = P(E/(\overline{F} \cap G)) \cdot P(\overline{F} \cap G),$$

sabendo também que:

$$\frac{P(F \cap G)}{P(G)} = P(F/G)$$

$$\frac{P(\overline{F} \cap G)}{P(G)} = P(\overline{F}/G).$$

Obtemos:

$$\begin{aligned} P(E/G) &= \frac{P(E/(F \cap G)) \cdot P(F \cap G)) + P(E/(\overline{F} \cap G)) \cdot P(\overline{F} \cap G)}{P(G)} \\ &= \left(\frac{P(E/(F \cap G)) \cdot P(F \cap G)}{P(G)}\right) + \left(\frac{P(E/(\overline{F} \cap G)) \cdot P(\overline{F} \cap G)}{P(G)}\right). \end{aligned}$$

Podemos concluir então:

$$P(E/G) = P(E/(F \cap G) \cdot P(F/G) + P(E/(\overline{F} \cap G) \cdot P(\overline{F}/G).$$

Analogamente:

$$P(E/\overline{G}) = P(E/(F \cap \overline{G}) \cdot P(F/\overline{G}) + P(E/(\overline{F} \cap \overline{G}) \cdot P(\overline{F}/\overline{G}).$$

Qualquer grupo de dados estatísticos pode estar escondendo alguma informação que pode contrariar resultados já concluídos.

Dadas as definições acima pode-se concluir o Paradoxo de Simpson à partir das seguintes comparações:

Temos $P(E/(F \cap G)) > P(E/(\overline{F} \cap G))$ e $P(E/(F \cap \overline{G})) > P(E/(\overline{F} \cap \overline{G}))$ mas,

$$P(E/G) < P(E/\overline{G})$$

e:
$P(E/(F \cap G)) < P(E/(\overline{F} \cap G))$ e $P(E/(F \cap \overline{G})) < P(E/(\overline{F} \cap \overline{G}))$ mas,

$$P(E/G) > P(E/\overline{G}).$$

Analisemos o seguinte exemplo à partir das comparações definidas acima para o Paradoxo de Simpson:

Exemplo 4.2. A fim de observar o aproveitamento dos estudantes de uma classe com 80 estudantes, cada estudante é convocado a responder a 40 questões de múltipla escolha e 40 questões dissertativas. Verificam-se os seguintes resultados com base nos eventos abaixo destacados:
Evento M para representar a ocorrência de estudantes mulheres.
Evento H para representar a ocorrência de estudantes homens.
Evento E para representar a ocorrência de estudantes aprovados.
Evento $\overline{E}$ para representar a ocorrência de estudantes não aprovados.
Observemos tabela 10:

Eventos	E	$\overline{E}$	T
M	20	20	40
H	24	16	40
T	44	36	80

Tabela 10: Distribuição de Estudantes

$$P(E/M) = \frac{P(E \cap M)}{P(T)} = \frac{20}{40} = \frac{1}{2};$$

$$P(E/H) = \frac{P(E \cap H)}{P(T)} = \frac{24}{40} = \frac{3}{5}.$$

Assim:

$$P(E/M) < P(E/H),$$

logo conclui-se que a taxa de aprovação entre mulheres é menor do que a taxa de aprovação entre homens.
Porém, os resultados considerados separadamente contradizem essa conclusão.
Observemos primeiro os resultados nas questões de múltiplas escolhas, num total de 40 questões, ver Tabela 11:

Eventos	E	$\overline{E}$	T
M	8	2	10
H	21	9	30
T	29	11	40

Tabela 11: Resultados Questões Múltipla Escolha

$$P(E/M) = \frac{P(E \cap M)}{P(M)} = \frac{8}{10} = 80\%;$$

$$P(E/H) = \frac{P(E \cap H)}{P(H)} = \frac{21}{30} = \frac{7}{10} = 70\%.$$

Assim:

$$P(E/M) > P(E/H),$$

portanto a probabilidade dos aprovados serem do sexo feminino é maior do que de serem do sexo masculino, contrariando o resultado anterior.
Analogamente, observemos os resultados nas questões dissertativas, também 40 questões, ver Tabela 12:

Eventos	E	$\overline{E}$	T
M	12	18	30
H	3	7	10
T	15	25	40

Tabela 12: Resultados Questões Dissertativas

$$P(E/M) = \frac{P(E \cap M)}{P(M)} = \frac{12}{30} = 40\%;$$

$$P(E/H) = \frac{P(E \cap H)}{P(H)} = \frac{3}{10} = 30\%;$$

e novamente obtivemos um resultado contraditório ao inicial:

$$P(E/M) > P(E/H).$$

Concluímos assim, que a análise em separado do desempenho dos alunos nas questões dissertativas e nas questões de múltipla escolha, que o desempenho dos estudantes de

sexo feminino foi melhor do que o desempenho dos resultados dos estudantes do sexo masculino. Podemos resolver o Paradoxo se acrescentarmos mais alguma condicional na análise das tabelas em separado, por exemplo, a probabilidade condicional sobre um evento F que poderia ser o estudante optar pelas questões de múltipla escolha, consequentemente, $\overline{F}$ a opção pelas questões dissertativas.
$P(E/F \cap M) = \frac{8}{10}$,
$P(E/\overline{F} \cap M) = \frac{4}{10}$,
$P(E/F \cap H) = \frac{21}{30} = \frac{7}{10}$,
$P(E/\overline{F} \cap H) = \frac{3}{10}$.
O que nos leva à confirmação de que o desempenho das meninas nas questões dissertativas é melhor do que o desempenho dos meninos e em separado, o desempenho das meninas nas questões de múltipla escolha também é melhor do que o desempenho dos meninos, confirmando os resultados analisados no Exemplo12.

Logo, à partir do Paradoxo exemplificado em Exemplo12, o Paradoxo de Simpson, percebemos a necessidade de ter cautela na análise de dados agrupados, pois a conclusão geral inicial, de uma taxa de aprovação masculina mais elevada, resulta das escolhas desiguais de mulheres e homens para os testes, e portanto:

$$P(E/M) = \frac{1}{4} < P(E/H) = \frac{3}{4}.$$

O Paradoxo de Simpson para ser resolvido sugere incluir mais condicionais à análise dos fatos.
O que podemos comprovar também com o Teorema de Bayes:

$$P(E/M) = \frac{P(E/F \cap M) + P(E/\overline{F} \cap M)}{P(E/F \cap M) + P(E/\overline{F} \cap M) + P(E/F \cap H) + P(E/\overline{F} \cap H)}$$

$$P(E/M) = \frac{\frac{8}{10} + \frac{4}{10}}{\frac{8}{10} + \frac{4}{10} + \frac{7}{10} + \frac{3}{10}} = \frac{\frac{12}{10}}{\frac{22}{10}} = \frac{12}{22} \approx 55\%.$$

5

SUGESTÕES DE ATIVIDADES PARA PRÁTICA DO ENSINO DA PROBABILIDADE CONDICIONAL

Levando-se em conta as Probabilidades Condicionais mencionadas nesse Trabalho, seguem as seguintes sugestões de Atividades para o ensino médio.
Público Alvo: 2º Ciclo do Ensino Médio

5.1 PROBABILIDADE CONDICIONAL- TEOREMA DE BAYES

Esse é um Projeto de Sala de Aula com a proposta de exemplificar Probabilidade Condicional
Composição: Professor mediador e sala de aula com no máximo 40 alunos participando ativamente.
Objetivos:

- Reconhecer da restrição do Espaço Amostral como fator condicional;
- Verificar o uso da Análise Combinatória no cálculo de probabilidade;
- Comprovar resultados com aplicação das Fórmulas de Probabilidades Condicionais, inclusive o Teorema de Bayes.

Procedimentos:

- O professor mediador pede para a sala sugestões de uma palavra com poucas letras, sendo elas distintas,
- É solicitado aos alunos que com a ajuda de um dicionário em celular ou tablet, eles procurem palavras curtas que permitam várias permutações com significado,

- Com base na Análise Combinatória eles devem calcular o total de permutações possíveis, mesmo que sem significado, e registrar cada anagrama em cartões feitos por eles,

- Inicialmente, sorteando um cartão, trabalhar com a probabilidade clássica de ocorrer, por exemplo palavras com significado, ou palavras que se iniciem com vogal, ..., contando os casos nos cartões e depois associando ao cálculo com envolvimento de análise combinatória,

- À seguir, começar a brincar com a probabilidade condicional simples (apenas uma condicional), por exemplo, calculando a probabilidade de ocorrer palavra que comece com vogal sabendo que termina em consoante, observando os resultados com os cartões e confirmando com a fórmula.

$$P(A/B) = \frac{n(A \cap B)}{n(B)};$$

- Introduzir a aplicação do Teorema de Bayes aumentando o número de condicionais para duas, por exemplo, probabilidade de ocorrer, no sorteio, uma palavra começada por consoante, sabendo que devem ter significado e terminar em vogal. Fazer essa probabilidade empiricamente com os cartões e comprovar o resultado com o Teorema de Bayes.

5.1.1 *Exemplo Prático*

Exemplo 5.1.

- Os alunos escolhem a palavra AMOR

- Colocam todos os anagramas em cartões recortados por eles:

AMOR	AMRO	AOMR	AORM	AROM	ARMO
MAOR	MARO	MORA	MOAR	MRAO	MROA
RAMO	RAOM	RMOA	RMAO	ROMA	ROAM
OARM	OAMR	OMAR	OMRA	ORMA	ORAM

Contando percebemos que são 24 anagramas, que podem ser obtidos com:

$P_4 = 4! = 24$

Observamos que, com significado temos:

$\boxed{\text{AMOR}}$ $\boxed{\text{RAMO}}$ $\boxed{\text{ROMA}}$ $\boxed{\text{MORA}}$

Portanto, a probabilidade de, ao sortearmos um dos 24 cartões, obtermos o evento $E = \{\text{anagramas com significado}\}$ é:

$$P(E) = \frac{4}{24} \approx 0,17 = 17\%.$$

- Dados os eventos:

  $A = \{\text{um cartão que comece com consoante}\}$,

  $B = \{\text{ocorrer anagrama terminado em vogal}\}$, ou seja:

  Calculemos a probabilidade condicional de, ao retirarmos um cartão, ele conter um anagrama começado por consoante, sabendo que é um anagrama terminado em vogal.

  Qual o valor de $P(A \cap B)$?

  Os elementos do evento A são:

  $\boxed{\text{MAOR}}$ $\boxed{\text{MARO}}$ $\boxed{\text{MORA}}$ $\boxed{\text{MOAR}}$ $\boxed{\text{MRAO}}$ $\boxed{\text{MROA}}$ $\boxed{\text{RAMO}}$ $\boxed{\text{RAOM}}$
  $\boxed{\text{RMOA}}$ $\boxed{\text{RMAO}}$ $\boxed{\text{ROMA}}$ $\boxed{\text{ROAM}}$, portanto: $n(A) = 12$.

  Os elementos do evento B são:

  $\boxed{\text{AMRO}}$ $\boxed{\text{ARMO}}$ $\boxed{\text{MARO}}$ $\boxed{\text{MORA}}$ $\boxed{\text{MRAO}}$ $\boxed{\text{MROA}}$ $\boxed{\text{RAMO}}$ $\boxed{\text{RMOA}}$
  $\boxed{\text{RMAO}}$ $\boxed{\text{ROMA}}$ $\boxed{\text{OMRA}}$ $\boxed{\text{ORMA}}$, portanto: $n(B) = 12$.

  Os elementos de $(A \cap B)$ são:

  $\boxed{\text{MARO}}$ $\boxed{\text{MORA}}$ $\boxed{\text{MRAO}}$ $\boxed{\text{MROA}}$ $\boxed{\text{RAMO}}$ $\boxed{\text{RMOA}}$ $\boxed{\text{RMAO}}$ $\boxed{\text{ROMA}}$,
  portanto $n(A \cap B) = 8$.

  Deve-se aproveitar para mostrar aos alunos que o número de elementos de cada conjunto trabalhado pode ser obtido por análise combinatória, ou seja,

  $n(A \cap B) = 8$, por análise combinatória temos:

$$(\underbrace{\text{consoantes}}_{2}) \cdot (\underbrace{\text{vogais}}_{2}) \cdot (\underbrace{\text{permutação entre as letras restantes}}_{2!}).$$

$$n(A \cap B) = 2 \cdot 2 \cdot 2! = 8.$$

  Portanto:

$$P(A/B) = \frac{n(A \cap B)}{n(B)} = \frac{8}{12} \approx 0,67 = 67\%.$$

- Acrescentando mais uma condicional:

  Sendo evento $E = \{$ocorrer anagrama com significado$\}$, analisemos agora a probabilidade de, no sorteio de um cartão, ocorrer anagrama começados com consoante, sabendo que ocorreu um resultado terminado em vogal e com significado.

  $$P(A/B \cap E) = 100\%.$$

  Comprovação com Teorema de Bayes

  $$P(A/B \cap E) = \frac{P((B \cap E)/A) \cdot P(A)}{P((B \cap E)/A) \cdot P(A) + P((B \cap E)/\overline{A}) \cdot P(\overline{A})}$$

  $$= \frac{\frac{8}{12} \cdot \frac{3}{12} \cdot \frac{12}{24}}{\frac{8}{12} \cdot \frac{3}{12} \cdot \frac{12}{24} + \frac{4}{12} \cdot 0 \cdot \frac{12}{24}} = 1 = 100\%.$$

5.2 PARADOXO DE SIMPSON

Proposta: Induzir o aluno, à partir da resolução do exercício a reconhecer o que significa um paradoxo, especificamente o Paradoxo de Simpson:
Onde:
$P(E/(F \cap G)) > P(E/(\overline{F} \cap G))$ e $P(E/(F \cap \overline{G})) > P(E/(\overline{F} \cap \overline{G}))$ mas,

$$P(E/G) < P(E/\overline{G}),$$

ou $P(E/(F \cap G)) < P(E/(\overline{F} \cap G))$ e $P(E/(F \cap \overline{G})) < P(E/(\overline{F} \cap \overline{G}))$ mas,

$$P(E/G) > P(E/\overline{G}).$$

A tabela abaixo (Ver Tabela 13) nos dá dados hipotéticos de uma pesquisa para comparar a eficiência de dois remédios.

pílulas vermelhas //////// pílulas amarelas				
pacientes	sobreviveram	morreram	sobreviveram	morreram
Homens	80 (80%)	20(20%)	78 (78%)	22 (22%)
Mulheres	20(50%)	20(50%)	2 (40%)	3 (60%)
Todos	100 (71,4%)	40(28,6%)	80(76,2%)	25 (23,8%)

Tabela 13: Dados da Pesquisa com as Pílulas Vermelhas e Amarelas

A Tabela 13 foi citado em [5]

Proposta:

O aluno deve tentar tirar conclusões sobre a eficiência das pílulas vermelhas e amarelas analisando os resultados na tabela, de ambas as pílulas, para o sexo feminino e para o sexo masculino, em separado. Em seguida, analisar o resultado ignorando os sexos dos pacientes.

Pode-se trabalhar com os seguintes eventos:

Evento E= ocorrência de sobreviventes;

Evento H= pacientes do sexo masculino;

Evento M= pacientes do sexo feminino;

Evento V= pílulas vermelhas;

Evento A= pílulas amarelas.

Observamos então que:

$$\underbrace{P(E/(H \cap V))}_{80\%} > \underbrace{P(E/(H \cap A))}_{78\%}.$$

Da mesma maneira:

$$\underbrace{P(E/(M \cap V))}_{50\%} > \underbrace{P(E/(M \cap A))}_{40\%},$$

o que nos levaria a acreditar que a pílula vermelha tem uma eficiência maior do que a pílula amarela.

No entanto, ao ignorarmos o sexo dos pacientes, obtivemos o seguinte resultado:

$$\underbrace{P(E/V)}_{71,4\%} < \underbrace{P(E/A)}_{76,2\%}.$$

Contrariando o resultado anterior, a pílula amarela demonstra ser mais eficiente do que a pílula vermelha.

À partir desse exemplo pode-se observar o sentido do Paradoxo de Simpson, que muda resultados em função de cada variável escolhida na análise e que, para evitar esse paradoxo, as variáveis a se incluir no estudo devem ser criteriosamente selecionadas.
No exemplo citado os motivos da aparente contradição, podem ser justificados, por exemplo, pelo fato da distribuição de pílulas amarelas ter sido muito desigual entre homens e mulheres, ou ainda por ambos os medicamentos terem sido mais eficientes em homens.

5.3 PARADOXO DE MONTY HALL

O paradoxo de Monty Hall é um problema matemático que surgiu à partir de um programa de televisão nos Estados Unidos chamado "Let's Make a Deal ", exibido na década de 70. O jogo acontecia da seguinte maneira:
O apresentador Monty Hall, apresentava três portas aos concorrentes. Atrás de uma delas estava um prêmio (um carro) e as outras duas dois bodes.
Na primeira etapa o concorrente escolhe uma das três portas, sem ser revelado a ele o que tem atrás dela.
Na segunda etapa Monty pergunta ao concorrente se quer permanecer com a escolha ou se quer trocar por uma das outras duas portas ainda fechadas.
Monty abre uma das portas, sabendo que nessa não está o prêmio e oferece a troca novamente ao concorrente, pela porta fechada.(Ver Figura 16)

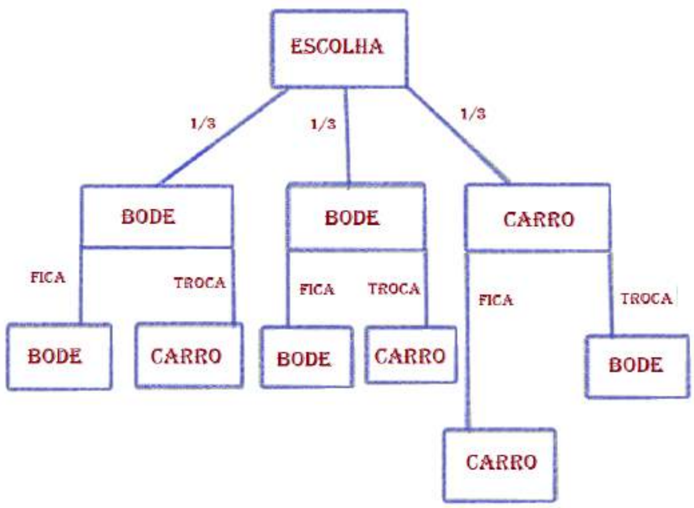

Figura 16: Monty Hall

Exemplo 5.2. A questão é a seguinte, é mais vantajoso para o concorrente manter a porta escolhida ou trocá-la pela porta oferecida?
Essa resposta pode ser obtida através do Teorema de Bayes.
Vamos estabelecer os seguintes eventos:
A_1= escolher porta número 1 e carro estar na porta número 1;
A_2= escolher porta número 1 e o carro estar na porta número 2;
A_3= escolher porta número 1 e o carro estar na porta 3;
B= evento o apresentador abrir a porta 2.
$P(A_1) = \frac{1}{3}$, pois a probabilidade do carro estar em cada é das três portas é 1/3.
Assim:
$P(B/A_1) = \frac{1}{2}$, pois a probabilidade do apresentador abrir a porta 2 dado que o carro está na porta 1 é uma entre as duas portas, a 2 e a 3.
$P(B/A_2) = 0$, pois o apresentador sabe que o carro se encontra na porta 2, logo não vai abrir essa porta.
$P(B/A_3) = 1$, o carro está na porta 3 e a porta 1 já foi escolhida, portanto o apresentador só vai escolher a porta 2.
Portanto:

$$P(A_3/B) = \frac{P(B/A_3) \cdot P(A_3)}{P(B/A_1) \cdot P(A_1) + P(B/A_2) \cdot P(A_2) + P(B/A_3) \cdot P(A_3)}$$

$$= \frac{1 \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{2}{3}.$$

Conclusão, trocar para a porta 3 é a melhor opção.

Atividade de Sala

Objetivo: Comprovar com a atividade que, trocar é sempre a melhor opção nesse tipo de jogo e, apresentar essa conclusão com o cálculo da probabilidade de ganhar condicionada à troca, com o Teorema de Bayes.
Material: três cartões da mesma cor, dos quais dois contém a frase "não foi dessa vez "e um deles um prêmio de $R\$10,00$.
Composição: A sala será dividida em 2 grupos.
O professor será o apresentador no jogo, e o jogo se realizará entre o apresentador e cada um dos dois grupos.
Cada grupo receberá os três cartões.

Desenvolvimento do jogo

- O apresentador embaralha os três cartões e os coloca com os dizeres voltados para baixo.
- é feito um sorteio para a escolha de um dos dois grupos para o início do jogo. O primeiro grupo sorteado vai trabalhar com a opção da troca da carta escolhida.
- O grupo sorteado escolhe carta sem ver o resultado
- O apresentador, que sabe qual é o cartão premiado, abre uma das outras cartas e pergunta se o grupo quer trocar ou manter o cartão escolhido pela carta que sobrou.
- Esse primeiro grupo optará sempre por trocar a carta. Após a decisão do grupo, o apresentador revela o cartão premiado.

  E assim se repete o jogo com esse grupo,uma certa quantidade de vezes, podem ser, por exemplo 5 rodadas. Enquanto, isso o outro grupo vai anotando em quantas situações a troca foi favorável.

- O apresentador então muda de grupo e começa o jogo novamente,em mais 5 rodadas, sendo que esse grupo deverá optar por manter a escolha em todas as rodadas.

  Após jogar a mesma quantidade de vezes que o primeiro grupo, que estará anotando os resultados, compara-se em quantas vezes manter foi ou não vantajoso. Anotar os resultados de todas as rodadas em uma tabela, por exemplo, em 10 rodadas observou-se:(Ver Tabela 14)

Decisão/Resultado	Acertou	Errou
Trocou	7	3
Não Trocou	4	6

Tabela 14: Resultados nas Rodadas

- Finalizar a brincadeira comparando os resultados e verificando qual foi o grupo que obteve mais sucesso em suas opções, constatando assim se trocar é um bom negócio discutindo o cálculo da probabilidade de ganhar.

  Para os cálculos das probabilidades de ganhar ou de perder o prêmio, vamos supor que o prêmio esteja na carta 1.

  Podemos esquematizar as situações da seguinte forma:

Tabela 1:SEMPRE TROCAR

Escolha inicial	P(Escolher carta não premiada)	Probabilidades para troca	Probabilidade
$P(C_1) = \frac{1}{3}$	$P(C_2) = \frac{1}{2}$	$P(C_3) = 1$	$\frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{6}$
$P(C_1) = \frac{1}{3}$	$P(C_3) = \frac{1}{2}$	$P(C_2) = 1$	$\frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{6}$
$P(C_2) = \frac{1}{3}$	$P(C_3) = 1$	$P(C_1) = 1$	$\frac{1}{3} \cdot 1 \cdot 1 = \frac{1}{3}$
$P(C_3) = \frac{1}{3}$	$P(C_2) = 1$	$P(C_1) = 1$	$\frac{1}{3} \cdot 1 \cdot 1 = \frac{1}{3}$

Na tabela, em **P(Escolher carta não premiada)**, devemos entender que o apresentador já sabe qual é a carta não premiada, por isso, por exemplo, se na escolha inicial o grupo optou pela carta C_2, a carta escolhida pelo apresentador só poderá ser a C_3, já que ele sabe que a C_1 é premiada e assim $P(C_3) = 1$.

Os resultados obtidos pela Tabela 1 devem ser assim interpretados: A soma das probabilidades das duas primeiras linhas, $\frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$, representa a probabilidade de perder ao optar por sempre trocar.

A soma das probabilidades das linhas 3 e 4, $\frac{1}{3} + \frac{1}{3} = \frac{2}{3}$, representa a probabilidade de ganhar ao optar pela troca.

Observamos que na opção de sempre trocar, a probabilidade de ganhar já é maior do que a probabilidade de perder.

Ainda supondo que o prêmio está na Carta 1, se o jogador opta por nunca trocar de carta, a situação pode ser representada pela Tabela 2:

Tabela 2:NUNCA TROCAR

Escolha inicial	P(Escolher carta não premiada)	Escolha final sem troca	Probabilidade
$P(C_1) = \frac{1}{3}$	$P(C_2) = \frac{1}{2}$	$P(C_1) = 1$	$\frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{6}$
$P(C_1) = \frac{1}{3}$	$P(C_3) = \frac{1}{2}$	$P(C_1) = 1$	$\frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{6}$
$P(C_2) = \frac{1}{3}$	$P(C_3) = 1$	$P(C_2) = 1$	$\frac{1}{3} \cdot 1 \cdot 1 = \frac{1}{3}$
$P(C_3) = \frac{1}{3}$	$P(C_2) = 1$	$P(C_3) = 1$	$\frac{1}{3} \cdot 1 \cdot 1 = \frac{1}{3}$

Observamos nas duas primeiras linhas da Tabela 2 que a soma das probabilidades $\frac{1}{6} + \frac{1}{6} = \frac{1}{3}$, representa a probabilidade de ganhar ao manter a escolha e nas linhas 3 e 4, a soma das probabilidades $\frac{1}{3} + \frac{1}{3} = \frac{2}{3}$, representa a probabilidade de perder ao manter a escolha. Assim a tabela 2 nos mostra que ao optar em não trocar a escolha inicial, a probabilidade de vencer ($\frac{1}{3}$) é menor do que a probabilidade de perder ($\frac{2}{3}$).

Pode-se estender a brincadeira trocando a carta premiada e repetindo todo o procedimento, assegurando assim os resultados.

6

CONCLUSÃO

Atualmente, o Ensino Médio tem como metas, dentre outras, formar cidadãos éticos e autônomos, capazes de compreender os processos produtivos e prepará-los para o mercado de trabalho, segundo a Base Nacional Comum Curricular (BNCC).
Assim, considero, indispensável tomar os PCNEMs (Parâmetros Curriculares Nacionais para o Ensino Médio) e o ENEM (Exame Nacional do Ensino Médio), além da BNCC, homologada em dezembro de 2018 e ainda por ser sancionada, como documentos referenciais para o desenvolvimento desse trabalho.
Os PCNs nos coloca diante de um problema antigo na área do currículo: as disciplinas tradicionais não dão conta de um conjunto de questões postas pela realidade vivida pelos alunos, vindo desse problema a eterna pergunta feita por eles: "Para o que serve isso"?
Esse documento, publicado em 1997, justifica a inclusão de conteúdos em função da demanda social. Segundo o mesmo, um olhar atento para a sociedade mostra a necessidade de o cidadão saber interpretar informações apresentadas por meio de dados estatísticos, tabelas e gráficos e a raciocinar, utilizando ideias relativas à probabilidade e à combinatória (BRASIL, 1998) (ver [9]).
Seus elaboradores ainda acrescentam que tais temas propiciam aos alunos estabelecer relações entre a Matemática e os conteúdos de outras áreas, à medida que os percebem como ferramentas essenciais para a formação de uma atitude crítica diante de questões sociais, políticas, culturais e científicas da atualidade (BRASIL, 1998).
Assim, a probabilidade e a estatística se apresentam como ferramentas na análise e valorização de informações provenientes de diferentes fontes, a fim de permitir que o indivíduo forme uma opinião própria que permita expressar-se criticamente sobre problemas inclusive, de outras áreas do conhecimento, além de facilitar tomada de decisões.

Os Parâmetros Curriculares Nacionais (PCN) e as diretrizes do Exame Nacional do Ensino Médio (ENEM) apontam para a contextualização do ensino nos universos do trabalho, da cidadania, da cultura, da tecnologia e da ciência, e demonstram, principalmente, um incentivo para a interdisciplinaridade (ver [1]).

Qualquer representação de dados, como um gráfico, uma tabela, uma reta de regressão ou um resumo, é um modelo representativo da realidade e permite observar o comportamento da variável em estudo, o que possibilita buscar a aproximação deste com um modelo estatístico e assim podem representar um ambiente de aprendizagem.

A BNCC do Ensino Médio se organiza em continuidade ao proposto para a Educação Infantil e o Ensino Fundamental, centrada no desenvolvimento de competências e orientada pelo princípio da educação integral.

A área de Matemática, no Ensino Fundamental, centra-se no desenvolvimento da compreensão de conceitos e procedimentos em seus diferentes campos, visando à resolução de situações-problema. No Ensino Médio, na área de Matemática e suas Tecnologias, os estudantes devem utilizar conceitos, procedimentos e estratégias não apenas para resolver problemas, mas também para formulá-los, descrever dados, selecionar modelos matemáticos e desenvolver o pensamento computacional, por meio da utilização de diferentes recursos da área. No tocante à Probabilidade, os estudantes do Ensino Fundamental têm a possibilidade, desde os anos iniciais, de construir o espaço amostral de eventos equiprováveis, utilizando a árvore de possibilidades, o princípio multiplicativo ou simulações, para estimar a probabilidade de sucesso de um dos eventos.

Então, dentro das habilidades propostas a serem desenvolvidas pelo aluno, acredito que podemos incluir a Probabilidade, seus pré-requisitos e as probabilidades condicionais, tema desse Trabalho, como uma extensão, dos conteúdos trabalhados no ensino fundamental, para o Ensino Médio.

BIBLIOGRAFIA

[1] Barbosa, J. C.: *A "Contextualização" e a Modelagem na educação matemática do ensino médio*. Encontro Nacional de Educação Matemática, 8:1–8, 2004.

[2] Carvalho, E. de: *Apostila do Objetivo*. Editora CERED, 2018.

[3] Cruz, A.: *Paradoxo De Simpson*. `http://value-from-data.com/blog/2017/03/simpson-paradox/`, 2017. [Online; acesso 20-03-2019].

[4] Cunha, S. B. da e S. Carvajal: *Estatistica Basica-a Arte de Trabalhar com Dados*. Elsevier Brasil, 2009.

[5] Gama, L. B. *et al.*: *Paradoxo de Banach-Tarski*. 2017.

[6] Hazzan, S.: *Fundamentos de matemática elementar: combinatória, probabilidade*. Atual, 1993.

[7] Henry, Y., P. Vain e J. De Buyser: *Genetic analysis of in vitro plant tissue culture responses and regeneration capacities*, vol. 79. Springer, 1994.

[8] Lima, E. L. e I. de Matematica Pura e Aplicada (Brasil): *Análise real*. Nº v. 1 em *Análise real*. IMPA, 1989, ISBN 9788524400483. `https://books.google.com.br/books?id=U6NAPwAACAAJ`.

[9] Macedo, E. F. de: *Parâmetros Curriculares Nacionais: a falácia de seus temas transversais*. Currículo: políticas e práticas, pp. 43–citation_lastpage, 1999.

[10] Magalhães, M. N.: *Probabilidade e variáveis aleatórias*. Edusp, 2006.

[11] Oliveira Mendonça, L. de e C. Espasandin Lopes: *Modelagem Matemática: um ambiente de aprendizagem para a implementação da Educação Estatística no Ensino Médio*. Boletim de Educação Matemática, 24(40), 2011.

[12] Possani, C.: *Probabilidade geométrica: história, paradoxos e rigor*. ComCiência, (143):0–0, 2012.

[13] Raquel, R. F.: *Princípios de probabilidade*. Tese de Mestrado, Universidade Federal do Rio Grande do Norte, 2014.

[14] Ross, S. M.: *Probabilidade: Um Curso Moderno com Aplicações*. Bookman, Porto Alegre, 8ª ed., 2010, ISBN 9788577806881.

[15] Totohasina, A.: *Méthode implicative en analyse de données et application à l'analyse de conceptions d'étudiants sur la notion de probabilité conditionnelle*. Tese de Doutoramento, Rennes I, 1992.