UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS VII
CENTRO DE CIÊNCIAS EXATAS E SOCIAIS APLICADAS
DEPARTAMENTO DE MATEMÁTICA
CURSO DE LICENCIATURA EM MATEMÁTICA

WALLISSON COSTA FERREIRA

ESTUDO BÁSICO DAS CADEIAS DE MARKOV E APLICAÇÕES DE MATRIZES EM PROBABILIDADE

PATOS
2025

WALLISSON COSTA FERREIRA

ESTUDO BÁSICO DAS CADEIAS DE MARKOV E APLICAÇÕES DE MATRIZES EM PROBABILIDADE

Trabalho de Conclusão de Curso apresentado ao Departamento de Matemática do Centro de Ciências exatas e Sociais Aplicadas da Universidade Estadual da Paraíba como requisito parcial à obtenção do título de Licenciado(a) em Matemática.

Área de concentração: Matemática pura e aplicada

Orientador: Prof. Dr. Ademir Benteus Pampu

PATOS
2025

É expressamente proibida a comercialização deste documento, tanto em versão impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que, na reprodução, figure a identificação do autor, título, instituição e ano do trabalho.

F383e Ferreira, Wallisson Costa.
Estudo basico das cadeias de Markov e aplicações de matrizes em probabilidade [manuscrito] / Wallisson Costa Ferreira. - 2025.
41 f. : il. color.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Matemática) - Universidade Estadual da Paraíba, Centro de Ciências Exatas e Sociais Aplicadas, 2025.

"Orientação : Prof. Dr. Ademir Benteus Pampu, Coordenação do Curso de Matemática - CCEA".

1. Cadeias de Markov. 2. Andrey Andreyevich Markov. 3. Teoria de probabilidades. 4. Distribuição estacionaria. I. Título

21. ed. CDD 519.2

Elaborada por Kaliane Eveny Martins de Oliveira - CRB - 15/986 BSC7

WALLISSON COSTA FERREIRA

ESTUDO BASICO DAS CADEIAS DE MARKOV E APLICAÇÕES DE MATRIZES EM PROBABILIDADE

Trabalho de Conclusão de Curso apresentado à Coordenação do Curso de Matemática da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Licenciado em Matemática

Aprovada em: 04/06/2025.

BANCA EXAMINADORA

Documento assinado eletronicamente por:

- **José Ginaldo de Souza Farias** (***.530.584-**), em **13/06/2025 19:25:25** com chave **4d8907a648a511f0948f1a7cc27eb1f9**.
- **Ademir Benteus Pampu** (***.824.629-**), em **13/06/2025 19:24:41** com chave **333c7e1448a511f0912c1a7cc27eb1f9**.
- **Jean Pereira Soares** (***.449.544-**), em **13/06/2025 23:55:08** com chave **fb0f3ad848ca11f0b8ab1a7cc27eb1f9**.

Documento emitido pelo SUAP. Para comprovar sua autenticidade, faça a leitura do QrCode ao lado ou acesse https://suap.uepb.edu.br/comum/autenticar_documento/ e informe os dados a seguir.
Tipo de Documento: Folha de Aprovação do Projeto Final
Data da Emissão: 15/06/2025
Código de Autenticação: 83f3fd

Dedico este trabalho a José Ferreira Neto e Maria Bernadete de Menezes Costa Ferreira que em nenhum momento mediram esforços para a realização dos meus sonhos, me ensinaram que devemos lutar pelo que queremos apesar das dificuldades.
A eles devo a pessoa que me tornei, sou extremamente feliz e tenho orgulho de chamos-los de pai e mãe.

AGRADECIMENTOS

Em primeiro lugar, agradeço a Deus, por me conceder força, sabedoria e perseverança para concluir esta etapa na minha vida.

Aos meus amados pais, José Neto e Maria Bernadete, sou eternamente grato pelo amor incondicional, pelo constante apoio em meus estudos e por me ensinarem princípios que levarei para toda a vida.

Aos meus irmãos Felicidade, Yasmim, Wellington, Wendel e Wesley por todo amor e dedicação, por terem me apoiado toda minha vida.

Ao meu orientador, Professor Doutor Ademir Benteus Pampu, pela paciência, pelos ensinamentos e incentivo presentes nas discussões teóricas e apoio na construção deste trabalho.

A todos que compõem a Universidade Estadual da Paraíba, Campus VII, pela oportunidade de ingressar nesta instituição como também pela atenção.

A todos os professores que passaram pela minha vida acadêmica e contribuíram para minha formação profissional mostrando o caminho da aprendizagem contínua.

Enfim, a todos que contribuíram de forma direta e indireta para que eu superasse as dificuldades que surgiram ao longo da minha jornada.

RESUMO

As Cadeias de Markov consistem em uma sequência de variáveis aleatórias cuja probabilidade de transição para o estado atual é condicionada exclusivamente pelo estado imediatamente anterior, independentemente de toda a trajetória prévia do processo. Esse modelo probabilístico representa um tópico essencial na teoria das probabilidades, demonstrando ampla aplicabilidade em diversos campos do conhecimento científico. Nesse contexto, o presente trabalho visa realizar o estudo da teoria básica das Cadeias de Markov, com aplicações modeladas por meio de matrizes, apresentadas de uma forma mais didática. O trabalho aborda um panorama histórico da Teoria das Probabilidades, destacando as contribuições fundamentais de seus principais idealizadores, discute também um breve contexto histórico sobre Andrey Andreyevich Markov e apresenta em linhas gerais algumas aplicações modeladas pelos Processos de Markov. Ademais, o trabalho introduz os fundamentos da Teoria das Probabilidades, construindo uma base teórica sólida para a compreensão dos conceitos essenciais ao estudo das Cadeias de Markov, além disso, também apresenta sua teoria básica, introduzindo definições e exemplos ilustrativos que elucidam concretamente definições matemáticas. Por fim, o trabalho propõe inspirar o desenvolvimento futuro na área ao mostrar as Cadeias de Markov através de uma abordagem matricial clara, facilitando a compreensão desses conceitos e suas aplicações.

Palavras-chave: Cadeias de Markov. Andrey Andreyevich Markov. Teoria de Probabilidades. Distribuição estacionaria.

ABSTRACT

Markov chains consist of a sequence of random variables whose transition probability to the current state relies solely on the previous state, irrespective of the entire prior trajectory of the process. This probabilistic model is a crucial subject in probability theory, demonstrating wide applicability across various scientific fields. In this context, the present work aims to explore the basic theory of Markov chains, with applications modeled through matrices, presented in a more accessible manner. The work offers an overview of the history of probability theory, underscoring the fundamental contributions of its key thinkers. It also provides a brief historical context about Andrey Andreyevich Markov and outlines some general applications modeled by Markov processes. Furthermore, the work introduces the principles of probability theory, establishing a solid theoretical foundation for grasping the essential concepts of Markov chains. Additionally, it presents the basic theory, incorporating definitions and illustrative examples that clearly elucidate mathematical definitions. Finally, the work intends to inspire future development in the area by presenting Markov chains through a straightforward, matrix-based approach, enhancing the understanding of these concepts and their applications.

Keywords: Markov Chains. Andrey Andreyevich Markov. Probability Theory. Stationary distribution.

LISTA DE FIGURAS

FIGURA 1 - Problema de Monty Hall 22
FIGURA 2 - Probabilidades de transição 31
FIGURA 3 - Probabilidades de transição entre as regiões 33

SUMÁRIO

Página

1 INTRODUÇÃO 9

2 FUNDAMENTOS DE PROBABILIDADE 14
2.1 EXPERIMENTOS ALEATÓRIOS 14
2.2 ESPAÇO AMOSTRAL E EVENTOS 14
2.2.1 Representação dos eventos 15
2.2.2 Tipos de eventos Probabilisticos 15
2.3 DEFINIÇÃO DE PROBABILIDADE 16
2.4 PROBABILIDADE CONDICIONAL 19
2.5 MATRIZES 25

3 CADEIAS DE MARKOV 28
3.1 CONCEITOS BÁSICOS 28
3.1.1 Sequência de variáveis aleatórias 28
3.1.2 Probabilidade de transição 29
3.1.3 Matriz de transição 30
3.2 DEFINIÇÃO DE UMA CADEIA DE MARKOV 30

4 CONCLUSÃO 39

REFERÊNCIAS 39

1 INTRODUÇÃO

A busca do ser humano em estudar fenômenos ou experimentos da natureza sempre foi impulsionada pela curiosidade e pela necessidade de compreender as leis que regem o mundo ao nosso redor. Essa curiosidade permitiu o desenvolvimento de diversas áreas do conhecimento, entre elas a teoria da probabilidade, que teve início justamente a partir da observação e análise de eventos incertos na natureza e em jogos de azar. No tocante a sua origem, é sabido, que a origem desse conhecimento matemático se iniciou de forma mais profunda a partir do século XVI, onde estudiosos europeus direcionaram suas investigações para estudos envolvendo sorte, jogos e incerteza (Salsa; Moreira, 2014).

Segundo Coutinho (1994), os primeiros registros acerca da teoria da probabilidade apareceram no século XIII, nas cartas de *De Vetula*. Um dos principais nomes que se destacou nessa área foi o de Gerolamo Cardano (1501-1576), que conduziu estudos sobre jogos de azar. Ele foi o primeiro autor a tratar sobre a aleatoriedade, onde foi o autor de O LIVRO DOS JOGOS DE AZAR (1520), apesar de cometer alguns equívocos, deixou seu legado a definição do conceito que hoje conhecemos por espaço amostral.

Gerolamo Cardano (1501-1576) era um renomado e famoso médico de Milão do século XVI, que também se dedicou à matemática e aos jogos de azar e apostas. Ele escreveu um tratado sobre apostas com o título *Liber de Ludo Aleae* "Livro de Jogos de Azar" uma obra que traz as primeiras noções sobre probabilidade. Neste livro, Cardano desenvolve cálculos de expectativas acerca de jogos de dados e também dá conselhos sobre como se proteger de jogadores que queiram trapacear no jogo. Além disso, ele trouxe, primeira vez, uma definição de probabilidade como sendo a razão do número de casos favoráveis e número de casos possíveis (Silva, 2018).

Segundo a história, o início dos registros acerca da Probabilidade como ciência se deu a partir da troca de correspondências entre os estudiosos franceses Blaise Pascal (1623 - 1662) e Pierre de Fermat (1601 - 1655) (Calabria; Cavalari, 2013). Tais cartas se tratavam de discussões a respeito de um problema exposto pelo francês Chevalier de Méré, um jogador da época apaixonado por jogos de azar. O mesmo escreveu ao matemático Blaise Pascal com diversos questionamentos das probabilidades de vitórias em jogos de azar sob diversas condições, sendo publicado então a obra pioneira dos fundamentos da teoria das probabilidades intitulada "raciocínio em jogos de azar" (Morgado et al., 2005).

Uma importante obra do italiano Luca Pacioli (1445 - 1517) conhecida como *Summa de aritmética, geometria, proportioni e proportionalità* ou simplesmente Summa foi publicada em Veneza em 1494, teve grande importância para o desenvolvimento da teoria da probabilidade (Viali, 2008). Desta obra, uma questão discutida por os franceses Pascal e Fermat foi o problema "o problema dos pontos" conhecido como repartição das apostas, em que consistia em qual deveria ser a divisão justa quando o jogo fosse interrompido

antes do final. O cenário era o seguinte: *" Dois jogadores estão participando de um jogo, ambos com a mesma probabilidade de vencer. O jogo é interrompido antes do seu término, e a pontuação de cada jogador nesse momento é conhecida. Nesse caso, como as apostas devem ser dividido o prêmio entre eles?"* (EVES, 2011).

Esse problema em questão envolve determinar as probabilidades de cada jogador vencer e, consequentemente, estabelecer a divisão justa do prêmio de acordo com as chances de vitória de cada jogador. Dessa forma, introduziu-se, de maneira intuitiva, o conceito de esperança matemática, que corresponde ao produto entre o ganho possível e a probabilidade de alcançá-lo (Silva, 2018).

Christian Huygens (1629-1695), foi o primeiro a perceber que não se tratava apenas de jogos de azar, mas que uma grande teoria estava sendo desenvolvida. Em 1655, em sua primeira viajem a Paris, conheceu as correspondências trocadas por Pascal e Fermat. Em 1657 ele publicou um livro intitulado *"De Ratiociniis in Ludo Aleae"*, (O Raciocínio em jogos de azar), considerado um dos primeiros tratados sobre a Teoria de Probabilidade (MORAES, 2014).

Huygens, Pascal e Fermat tiveram grande importância para o desenvolvimento da teoria da probabilidade, seus trabalhos influenciaram outros matemáticos posteriormente (SILVA, 2018). Dentre estes, destaca-se Abraham De Moivre (1667-1754) e Jacob Bernoulli (1654-1705), com o primeiro grande tratado de probabilidade, Ars Conjectand (a arte da conjectura). No entanto, ao falecer, Jacob deixou a obra incompleta, onde seu sobrinho, o também matemático Nicolaus Bernoulli (1687–1759), finalizou o trabalho, sendo publicado postumamente em 1713. Em seu livro, ele provou um teorema de grande relevância para a probabilidade, chamada de LEI DOS GRANDES NÚMEROS (MORAES, 2014).

Já Abraham De Moivre, conhecido principalmente por suas contribuições à teoria das probabilidades, destacou-se com suas obras*De Mensura Sortis* (sobre as Medidas das Possibilidades) em 1711;*Miscellanea analytica* (Miscelânea de análise) de 1721 e *Doctrine of Chances* (Doutrina das Chances) de 1738 (EVES, 2011).

Os livros de Bernoulli e De Moivre foram as contribuições mais importantes no período inicial da teoria da probabilidade antes de Laplace. Além deles, podemos citar outros autores que deram sua parcela de contribuição para o desenvolvimento da probabilidade, como Galileu Galilei (1564–1642), Leonhard Euler (1707-1783), Siméon Denis Poisson (1781–1840), Carl Friedrich Gauss (1777–1855), entre outros.

Em 1763, foi publicado o trabalho *Essay towards solving a problem in the doctrine of chances,* do matemático inglês Thomas Bayes (1702–1761), hoje conhecido como o Teorema de Bayes. Esse teorema é fundamental para o cálculo de probabilidades condicionais, permitindo a atualização de estimativas com base em novas evidências. Bayes entendia a probabilidade como uma medida subjetiva de risco em relação a expectativas, refletindo uma perspectiva axiomática (SILVA; GRANDO, 2024).

Após as publicações de Bernoulli e De Moivre, nenhum livro de maior relevância foi publicado até 1812, quando o renomado matemático francês Pierre Simon de Laplace (1749 - 1827), neste ano, lançou o seu grande tratado *"Théorie Analytique des Probabilitées"* (Teoria Analítica das Probabilidades) com base em trabalhos que desenvolveu entre os anos de 1771 e 1786. Os fundamentos da teoria da probabilidade foram então colocados por Laplace em uma forma conhecida hoje em dia como a "Definição Clássica de Probabilidade", que se manteve praticamente inalterada até o início do século XX (SILVA, 2018).

A partir de Laplace, muitos matemáticos contribuíram para este ramo da matemática. No final do século XIX, o matemático russo Pafnuty L'vovich Tchebyshev (1821-1894) fundou a célebre escola de São Petersburgo que produziu grandes matemáticos russos com contribuições fundamentais à teoria da probabilidade. Dentre eles, podemos citar Kolmorogov, aquele que hoje é considerado o pai da probabilidade moderna (VIALI, 2008).

Em 1933, Andrei Kolmogorov publicou a monografia que em inglês foi batizada de *Foundations of Probability Theory* (Fundamentos da Teoria da Probabilidade), que deu início à etapa moderna da teoria. Ele formulou a primeira apresentação axiomática que tornou a Teoria das Probabilidades uma parte autônoma dentro da Matemática e com isso possibilitou grande avanço científico nesta área, sobretudo sob o aspecto teórico. A utilização de tais modelos como instrumento explicativo voltado ao controle de grande número de sucessos é hoje uma opção cada vez mais utilizada no mundo cientifico, seja nas ciências humanas ou nas políticas. (ROTUNNO, 2007).

Em 1938, Kolmogorov publicou um artigo na área da Teoria das Probabilidades que serviu de base para os posteriores estudos de Andrey Markov.

Andrey Andreyevich Markov (1856-1922) graduou-se na Universidade de São Petersburgo (1878) e começou a atuar como professor na mesma Universidade em 1886 (Silva, 2013). Ele estabeleceu a prova do teorema central do limite em um contexto bastante geral. Markov é particularmente lembrado pelas cadeias que levam seu nome que são sequências de variáveis aleatórias em que o próximo valor é determinado apenas pelo atual, seguindo o princípio de que o estado presente só depende exclusivamente do estado anterior, sendo irrelevante como os estados prévios evoluíram até alcançar o estado atual (Magela, 2015; VIALI, 2008).

As cadeias de Markov tiveram origem em um estudo pioneiro no qual Markov examinou padrões de alternância entre vogais e consoantes no poema 'Eugene Onegin', de Alexander Pushkin (1883). Através de análise empírica, Markov constatou que a probabilidade de uma vogal ser sucedida por uma consoante era de 87%, enquanto uma consoante antecedia uma vogal em 66% dos casos. Esse estudo pioneiro, que originou as Cadeias de Markov, constitui um marco fundamental no desenvolvimento da Teoria das Probabilidades. Dentre os principais matemáticos que contribuíram para o seu avanço dessa teoria

podemos destacar o matemático russo Andrei Kolmorogov (Magela, 2015).

As cadeias de Markov podem ser aplicadas em diversas áreas, atualmente, uma importante aplicação das cadeias de Markov é utilizada pelo Google, o algoritmo Page Rank, usado para apresentar no topo, as páginas da web mais interessantes sobre um determinado assunto pesquisado (Magela, 2015).

As cadeias de Markov também tem aplicabilidade na previsão do tempo como mostra o estudo de Carvalho et al. (2017), no qual era analisada a ocorrência de dias secos e chuvosos na região de Rio Largo–Alagoas. Ao final do estudo concluiu-se que a transição entre dias secos e chuvosos foi baixa ao longo do ano.

Outra aplicação bastante interessante das Cadeias de Markov é no jogo monopoly, mais conhecido no brasil como banco imobiliário. Souza Junior (2016), mostra em seu estudo a modelagem matemática deste jogo e para isso usa as cadeias de Markov, tendo em vista que a cada jogada um jogador salta de uma casa para outra aleatoriamente, de acordo com o resultado da rolagem de 2 dados.

Uma Cadeia de Markov consiste em uma sequência de variáveis aleatórias cuja probabilidade de transição para o estado atual é condicionada exclusivamente pelo estado imediatamente anterior, independentemente de toda a trajetória prévia do processo. Esse modelo probabilístico representa um tópico essencial na teoria das probabilidades, demonstrando ampla aplicabilidade multidisciplinar em diversos campos do conhecimento científico. Nesse contexto, o presente trabalhou tem como objetivo realizar o estudo da teoria básica das Cadeias de Markov com aplicações modeladas por meio de matrizes, apresentando de uma forma mais didática. ilustrando sua aplicação com exemplos práticos.

Na seção 1 traremos uma breve introdução histórica sobre a Teoria da Probabilidade, destacando os principais nomes de vulto na construção científica nesta área. Além disso, falamos um pouco do contexto histórico de Andrey Andreyevich Markov, em seguida, falaremos em linhas gerais a respeito de algumas aplicações que podem ser modeladas por meio das Cadeias de Markov.

A seção 2 introduz os principais conceitos da teoria das probabilidades, estabelecendo as bases teóricas necessárias para os desenvolvimentos posteriores. Inicialmente, serão apresentados os conceitos e exemplos de Experimentos aleatórios, espaço amostral, representação dos eventos, tipos dos eventos. Será desenvolvida a definição axiomática de probabilidade, que generaliza e amplia o conceito clássico, juntamente com as principais propriedades que dela decorrem. Adicionalmente, serão formalizados os conceitos de probabilidade condicional e o Teorema de Bayes, cuja aplicação será ilustrada por meio do paradoxo clássico de Monty Hall. Por fim, introduziremos noções básicas sobre matrizes, incluindo operações de soma e multiplicação por escalar, além do conceito fundamental de convergência de sequências de matrizes, conceito essencial para o estudo do comportamento de longo prazo das Cadeias de Markov.

Na seção 3, apresentaremos os conceitos básicos que fundamentam a teoria das Cadeias de Markov, abordaremos a teoria básica das Cadeias de Markov, , introduzindo definições e exemplos. como sequência de variáveis aleatória, probabilidade de transição, matriz de transição, definição formal de Cadeia de Markov finita, distribuição de probabilidade, convergência da cadeia para o estado estacionário, que permite analisar o comportamento da cadeia no longo prazo. Por fim, traremos na seção 4 as conclusões acerca do que foi visto neste estudo.

2 FUNDAMENTOS DE PROBABILIDADE

2.1 EXPERIMENTOS ALEATÓRIOS

A seção 2.1 foi escrita baseada em Hazzan (2004) e Silva (2020).

Experimentos que, embora repetidos sob as mesmas condições, conduzem a resultados essencialmente distintos são chamados de experimentos aleatórios. Ou seja, é todo experimento cujos resultados só podem ser conhecidos após a sua execução. Por exemplo, o lançamento de um dado usual de "seis faces numeradas de 1 a 6"e equilibrado, ou seja, todas as faces têm a mesma chance de sair. Não há como saber qual face ficará voltada para cima antes do lançamento do dado.

A seguir, veremos exemplos de experimentos aleatórios:

- Lançar uma moeda e observar a face voltada para cima.
- Lançar duas moedas e observar as sequências de caras e coroas obtidas.
- Lançar um dado e observar o número da face de cima.
- Em uma urna contendo 3 bolas pretas e 2 bolas brancas, selecionar uma bola e observar sua cor.
- De um baralho contendo 52 cartas, com os naipes: paus, espadas, ouros e copas, selecionar uma carta e observar seu naipe.

2.2 ESPAÇO AMOSTRAL E EVENTOS

Nesta seção, o autor se baseou-se em Hazzan (2004) e Silva (2020).

Suponhamos que um experimento seja realizado sob certas condições fixas. O conjunto de todos os resultados possíveis de um experimento aleatório é chamado de espaço amostral ou conjunto universo, representado pela letra grega Ω.

Exemplo 2.2.1: O experimento aleatório em que se joga um dado usual para cima e observa-se o número da face superior. É evidente que $\Omega = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$ são os resultados possíveis do experimento. Mas, em alguns casos, o conjunto dos resultados possíveis de um experimento não é fácil de ser definido.

Exemplo 2.2.2 Ao lançar três moedas não viciadas simultaneamente, as faces: cara será representada por (Ca) e coroa (Co). Assim, o espaço amostral Ω é formado por:

$\Omega = \{$(Ca, Ca, Ca), (Ca, Ca, Co), (Ca, Co, Co), (Ca, Co, Ca),
(Co, Co, Co), (Co, Co, Ca), (Co, Ca, Ca), (Co, Ca, Co)$\}$

Seja Ω o espaço amostral de um experimento aleatório. Um subconjunto de Ω é chamado de evento, e denotado por letras do alfabeto (A, B, ..., Y, Z). Dessa forma, para qualquer subconjunto E, se $E \subset \Omega$, então podemos considerar E como um evento de Ω.

2.2.1 Representação dos eventos

Utilizando o exemplo acima representaremos os seguintes eventos:

- **Evento A:** Sair pelo menos uma face "cara".

  $A = \{$(Ca, Ca, Ca), (Ca, Ca, Co), (Ca, Co, Co), (Ca, Co, Ca),
  (Co, Co, Ca), (Co, Ca, Ca), (Co, Ca, Co)$\}$.

- **Evento B:** Sair somente uma face "coroa".

  $B = \{$(Ca, Ca, Co), (Ca, Co, Ca), (Co, Ca, Ca)$\}$.

- **Evento C:** Sair as três faces iguais.

  $C = \{$(Ca, Ca, Ca), (Co, Co, Co)$\}$.

Geralmente, o número de elementos no espaço amostral e em um evento de um experimento aleatório é determinado com base nos conceitos da análise combinatória.

2.2.2 Tipos de eventos Probabilisticos

Considere o lançamento de um dado comum e observe-se a face voltada para cima. Definimos o espaço amostral como o conjunto $\Omega = \{1, 2, 3, 4, 5, 6\}$. A seguir, serão definidos alguns tipos de eventos:

Evento elementar: Seja $A \subset \Omega$, e E um subconjunto unitário, chamamos A de evento elementar, ou seja, é aquele formado por um único elemento de Ω.

Exemplo 2.2.2.1: Definimos o evento A = "sair a face 3 voltada para cima". Então, $A = \{3\}$.

Evento certo: Se $C = \Omega$, então E é um evento certo, ou seja, é aquele que inclui todos os resultados possíveis do espaço amostral em um experimento aleatório.

Exemplo 2.2.2.2: Seja C = "sair um número entre 1 e 6". Então $C = \Omega$, pois E abrange todos os possíveis resultados do experimento.

Evento impossível: Seja $B = \varnothing$. Então, B é chamado de evento impossível, ou seja, é aquele que não pode ocorrer em um experimento aleatório.

Exemplo 2.2.2.3: Seja B = "sair um número maior que 6". Então, $B = \varnothing$, pois nenhum resultado possível atende a essa condição.

Evento união: A reunião de dois ou mais eventos.

Exemplo 2.2.2.4: Sejam A e B dois eventos, com A = "sair um número primo" ($A = \{2,\ 3,\ 5\}$) e B = "sair um número maior que 2" ($B = \{3,\ 4,\ 5,\ 6\}$). O evento $A \cup B$ é dado por $A \cup B = \{2,\ 3,\ 4,\ 5,\ 6\}$.

Evento interseção: Diz-se que dois ou mais eventos são intersecionais quando ocorrem simultaneamente.

Exemplo 2.2.2.5: Sejam L = "sair um número par" ($L = \{2,\ 4,\ 6\}$) e H = "sair um número maior que 3" ($H = \{4,\ 5,\ 6\}$). Então, $L \cap H = \{4,\ 6\}$.

Eventos mutuamente exclusivos: Sejam X e Y dois eventos, dizemos que esses eventos são mutualmente exclusivos quando não ocorrem ao mesmo tempo, isto é, $X \cap Y = \varnothing$.

Exemplo 2.2.2.6: Seja X = "sair um múltiplo de 2" ($X = \{2,\ 4,\ 6\}$) e Y = "sair um múltiplo de 5" ($Y = \{5\}$). Como $X \cap Y = \varnothing$, são eventos mutuamente exclusivos.

Evento complementar: Seja G um evento qualquer. O evento G^c é chamado de complementar de G, se:

$$G^c = \Omega - G = \{x \in \Omega \mid x \notin G\}.$$

Exemplo 2.2.2.7: Seja G = "sair um número par" ($G = \{2,\ 4,\ 6\}$). Seu complementar é $G^c = \{1, 3, 5\}$.

2.3 DEFINIÇÃO DE PROBABILIDADE

Esta seção fundamenta-se nos trabalhos de Amorim e Mozer (2020) e Ross (2010).

Definição 2.3.1 (Definição axiomática de probabilidade): Seja Ω um espaço amostral finito e E o conjunto formado por todos os subconjuntos de Ω. Uma probabilidade é uma função $P: E \to \mathbb{R}$ que associa a cada evento A um número real, chamada probabilidade do evento, satisfazendo as seguintes propriedades:

i) $P(A) \geq 0$ para todo $A \in E$;

ii) $P(\Omega) = 1$;

iii) Se $A, B \subset \Omega$ são eventos mutuamente exclusivos ($A \cap B = \varnothing$), então:

$$P(A \cup B) = P(A) + P(B).$$

Exemplo 2.3.1: Considere o lançamento de um dado honesto de seis faces, onde todas têm igual probabilidade de ocorrência. Neste caso:

$$P(\{1\}) = P(\{2\}) = \cdots = P(\{6\}) = \frac{1}{6}.$$

Aplicando a Propriedade iii), a probabilidade de sair um número ímpar ($A = \{1, 3, 5\}$) será:

$$P(A) = P(\{1\} \cup \{3\} \cup \{5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

Seja Ω um espaço amostral e P uma de probabilidade satisfazendo os axiomas da definição 2.3.1, Valem as seguintes propriedades:

i) $P(\varnothing) = 0$;

ii) Seja $\{A_1, A_2, \ldots, A_n\}$ uma coleção de eventos mutuamente excludentes (isto é, $A_i \cap A_j = \varnothing$ para todo $i \neq j$). Então:

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

iii) Se $A, B \subset \Omega$, então:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B);$$

iv) Para todo evento $A \subset \Omega$, vale:

$$P(A) = 1 - P(A^c);$$

onde A^c é o complementar de A.

v) Se $A \subset B \subset \Omega$, então:

$$P(A) \leq P(B).$$

Demonstração. i) Dado um evento $B \subset \Omega$ qualquer, temos $B \cap \varnothing = \varnothing$ e $B \cup \varnothing = B$. Assim, pelo item iii) da definição 3.2.1, temos:

$$P(A) = P(A \cup \varnothing) = P(A) + P(\varnothing) \Rightarrow P(\varnothing) = 0$$

ii) Note que os eventos A_1 e $B = A_2 \cup \cdots \cup A_n$ são mutuamente excludentes. Então, pelo item iii) da definição 2.3.1, temos:

$$P\left(\bigcup_{i=1}^{n} A_i\right) = P(A_1 \cup B) = P(A_1) + P(B) = P(A_1) + P\left(\bigcup_{i=2}^{n} A_i\right).$$

Aplicando recursivamente o mesmo argumento, obtemos:

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &= P(A_1) + P\left(\bigcup_{i=2}^{n} A_i\right) \\ &= P(A_1) + P(A_2) + P\left(\bigcup_{i=3}^{n} A_i\right) \\ &= \quad \vdots \\ &= P(A_1) + P(A_2) + \cdots + P(A_n) \\ &= \sum_{i=1}^{n} P(A_i). \end{aligned}$$

iii) Sejam $A, B \subset \Omega$. Podemos escrever $A \cup B$ como:

$$A \cup B = (A - B) \cup B$$

e

$$A = (A - B) \cup (A \cap B),$$

como ambas as uniões são disjuntas, segue do item iii) da definição 2.3.1 que:

$$P(A \cup B) = P(A - B) + P(B) = P(A) - P(A \cap B) + P(B).$$

iv) Note que $\Omega = A \cup A^c$, pelos itens ii) e iii) da definição 2.3.1 segue que:

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c) \Rightarrow P(A^c) = 1 - P(A).$$

v) Como $A \subset B$, segue que $B = (B - A) \cup A$, sendo essa união disjunta. Logo, pelo iii) da definição 3.2.1, temos:

$$P(B) = P(B - A) + P(A) \Rightarrow P(A) = P(B) - P(B - A) \leq P(B),$$

já que $P(B - A) \geq 0$.

□

A demonstração das propriedades acima foram baseadas em Amorim e Mozer (2020) e Bezerra (2018).

2.4 PROBABILIDADE CONDICIONAL

As referências bibliográficas que fundamentam esta seção foram extraídas das obras de Salsa e Moreira (2014); Bussab e Morenttin (2010) e Amorim e Mozer (2020).

A probabilidade condicional formaliza a noção de como a ocorrência de um evento pode redefinir a medida de incerteza associada a outro evento. Quando são conhecidas as informações sobre a realização de um evento, a probabilidade original de um segundo evento pode ser aumentada, reduzida ou manter-se inalterada. No caso em que essa informação não altera a probabilidade, diz-se que os eventos são independentes.

Definição 2.4.1: Sejam A e B dois eventos de um espaço amostral Ω com uma probabilidade P, onde $P(B) > 0$. Então, a probabilidade condicional de A dado que B ocorreu, indicada por $P(A \mid B)$, é definida como:

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

Exemplo 2.4.2: Uma urna contém 10 bolas, diferenciadas pelas cores preta e branca, numeradas por 1 ou 2. Sendo:

- **5 bolas pretas:** 3 numeradas por 1 e 2 numeradas por 2.
- **5 bolas brancas:** 4 numeradas por 1 e 1 numerada por 2.

Qual a probabilidade de selecionar uma bola preta, dado que ela tem o número 1?

Solução: Sejam os eventos:

- B = "selecionar uma bola preta",
- $N1$ = "selecionar uma bola com número 1".

Queremos calcular a probabilidade condicional $P(B \mid N1)$. Pela definição:

$$P(B \mid N1) = \frac{P(B \cap N1)}{P(N1)}.$$

- **Quantidade de bolas pretas com número 1:** 3 (ou seja, $B \cap N1$ corresponde a 3 bolas).
- **Total de bolas com número 1:** $3 + 4 = 7$ (3 pretas e 4 brancas).

Substituindo na fórmula:

$$P(B \mid N1) = \frac{\frac{3}{10}}{\frac{7}{10}} = \frac{3}{7}.$$

A probabilidade de selecionar uma bola preta, dado que ela tem o número 1, é 3/7.

Teorema 2.4.4: Sejam A e B dois eventos do espaço amostral Ω, tais que $P(A) \neq 0$ e $P(B) \neq 0$. São equivalentes:

i) Diz-se que A é independente de B se a ocorrência do evento A não afeta a probabilidade do evento B, ou seja:

$$P(A \mid B) = P(A),$$

ii) Diz-se que B é independente de A se a ocorrência do evento B não afeta a probabilidade do evento A, ou seja:

$$P(B \mid A) = P(B).$$

Demonstração. Como A e B são eventos possíveis, com $P(A) > 0$ e $P(B) > 0$. Pela definição de probabilidade condicional, temos:

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

Por hipótese,

$$P(B) = P(B \mid A) = \frac{P(A \cap B)}{P(A)}, \quad \text{ou seja,} \quad P(A \cap B) = P(A) \cdot P(B).$$

Assim,

$$\begin{aligned} P(A \mid B) &= \frac{P(A \cap B)}{P(B)} \\ &= \frac{P(A) \cdot P(B)}{P(B)} \\ &= P(A). \end{aligned}$$

Logo, A é independente de B.

Por outro lado, se B é independente de A, temos:

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

Por hipótese,

$$P(A) = P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad \text{ou seja,} \quad P(A \cap B) = P(A) \cdot P(B).$$

Assim,

$$\begin{aligned} P(B \mid A) &= \frac{P(A \cap B)}{P(A)} \\ &= \frac{P(A) \cdot P(B)}{P(A)} \\ &= P(B). \end{aligned}$$

□

Em resumo, se A é independente de B, então B é independente de A e além disso:

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

Definição 2.4.4: Seja Ω um espaço amostral não vazio. Uma partição de Ω é uma família de eventos $\{C_1, C_2, \ldots, C_n\}$ todos não-vazios, e tais que:

1. $C_i \cap C_j = \varnothing$, sempre que $i \neq j$,
2. $C_1 \cup C_2 \cup \cdots \cup C_n = \Omega$.

Ou seja, os eventos $C_1, C_2, \ldots, C_n$ são disjuntos dois a dois e sua união é igual a Ω.

A seguir será enunciado o Teorema de Bayes, que será fundamental para facilitar o entendimento do exemplo de Monty Hall posteriormente. Contudo, a demonstração desse teorema não será desenvolvida nesse trabalho, podendo ser consultada em Bussab; Morenttin (2010) das referências bibliográficas.

Teorema 2.4.4 (Teorema de Bayes): A probabilidade de ocorrência do evento C_i, supondo-se a ocorrência do evento A, é dada por

$$P(C_i \mid A) = \frac{P(C_i)P(A \mid C_i)}{\sum_{j=1}^{n} P(C_j)P(A \mid C_j)}.$$

Para todo $i = 1, 2, \ldots, n$.

O exemplo a seguir, adaptado de Bussab e Morenttin (2010) ilustra a aplicação do Teorema de Bayes em um contexto de escolha entre urnas com composições distintas.

Exemplo 2.4.5: Temos cinco urnas, cada uma com seis bolas. Duas dessas urnas (tipo C_1) têm 3 bolas brancas, duas outras (tipo C_2) têm 2 bolas brancas, e a última urna (tipo C_3) tem 6 bolas brancas. Escolhemos uma urna ao acaso e dela retiramos uma bola. Qual a probabilidade de a urna escolhida ser do tipo C_3, sabendo-se que a bola sorteada é branca?

Utilizando o teorema de Bayes, calcula-se a probabilidade de escolher a urna do tipo C_3, dado que a bola sorteada foi branca, ou seja, $P(C_3 \mid B)$, onde, definimos B o evento "sortear bola uma branca".
Sabendo que:

- $P(C_1) = \frac{2}{5}$, pois temos 2 urnas do tipo C1 em um total de 5 urnas.
- $P(C_2) = \frac{2}{5}$, pois temos 2 urnas do tipo C2 em um total de 5 urnas.
- $P(C_3) = \frac{1}{5}$, pois temos 1 urna do tipo C3 em um total de 5 urnas.

Calcula-se as probabilidades de sorteio de bola branca em cada tipo de urna, ou seja,

- Para C1: $P(B \mid C_1) = \frac{3}{6} = \frac{1}{2}$,
- Para C2: $P(B \mid C_2) = \frac{2}{6} = \frac{1}{3}$,
- Para C3: $P(B \mid C_3) = \frac{6}{6} = 1$.

Como C_1, C_2 e C_3 formam uma partição de Ω temos que,

$$\begin{aligned} B &= (C_1 \cap B) \cup (C_2 \cap B) \cup (C_3 \cap B) \\ P(B) &= P(C_1 \cap B) + P(C_2 \cap B) + P(C_3 \cap B) \\ &= P(C_1) \cdot P(B \mid C_1) + P(C_2) \cdot P(B \mid C_2) + P(C_3) \cdot P(B \mid C_3) \\ &= \frac{2}{5} \cdot \frac{1}{2} + \frac{2}{5} \cdot \frac{1}{3} + \frac{1}{5} \cdot 1 \\ &= \frac{8}{15}. \end{aligned}$$

Logo, pelo teorema de Bayes, temos

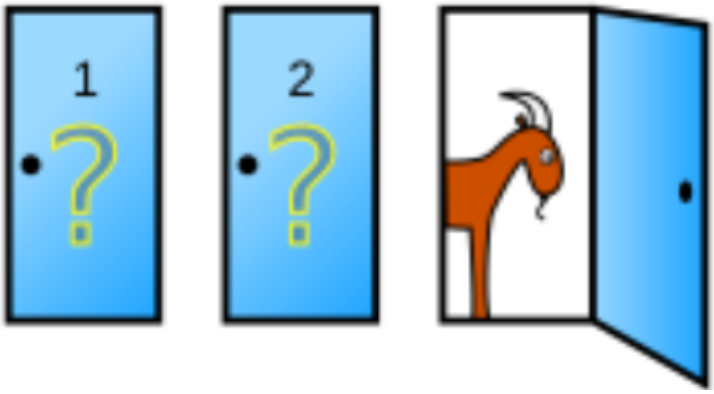
$$\begin{aligned} P(C_3 \mid B) &= \frac{P(C_3)P(C_3 \mid B)}{P(B)} \\ &= \frac{\frac{1}{5} \cdot 1}{\frac{8}{15}} \\ &= \frac{3}{8}. \end{aligned}$$

Assim, a probabilidade de a urna escolhida ser do tipo C_3, dado que a bola sorteada é branca, é 3/8.

O próximo exemplo apresemta uma aplicação do exemplo de monty Hall e esse exemplo é adaptado de Amorim; Mozer (2020).

Exemplo 2.4.5 (Problema de Monty Hall): O Problema de Monty Hall é um célebre enigma de probabilidade que se popularizou na década de 1990, após a colunista Marilyn vos Savant discutir o tema em sua seção de perguntas e respostas, gerando intensa controvérsia. No jogo, inspirado no programa de televisão apresentado por Monty Hall, um participante enfrenta três portas fechadas: atrás de uma há um carro (o prêmio desejado), e atrás das outras duas, bodes.

Figura 1 – Problema de Monty Hall



Fonte: (AMORIM; MOZER, 2020).

O processo segue três etapas:

1ª etapa: O participante escolhia uma das três portas, que permanecia fechada.
2ª etapa: O apresentador, que sabia onde estava o prêmio, abria uma das outras duas portas, sempre revelando um bode.
3ª etapa: Após o apresentador abrir uma das portas e mostrar um bode, ele perguntava ao participante se ele queria manter a escolha inicial ou trocar de porta e ficar com a outra porta fechada.

A grande questão é: qual estratégia maximiza a chance de vitória? Manter a porta inicial, trocar de porta, ou ambas são equivalentes?

Em 1990, Marilyn vos Savant argumentou que trocar de porta dobra as chances de ganhar o prêmio, elevando a probabilidade de sucesso de $\frac{1}{3}$ para $\frac{2}{3}$. Essa conclusão, embora correta, foi inicialmente rejeitada por muitos matemáticos e leitores, que a consideraram contraintuitiva. Acreditava-se que, após a revelação de um bode, as duas portas restantes teriam probabilidades iguais (50% cada), mas essa intuição ignorava a informação adicional introduzida pela ação estratégica do apresentador.

A solução emerge da probabilidade condicional e do Teorema de Bayes. Suponha que o participante escolha a porta 1 (sem perda de generalidade). Sejam os eventos C_1 = "o carro está da portas 1", C_2 = "o carro está da porta 2" e C_3 = "o carro está da porta 3". Em que cada evento tem probabilidade inicial $\frac{1}{3}$. Definimos o evento A = "O apresentador abre a

Porta 3". Note que, as probabilidades condicionais ajustam-se conforme o conhecimento prévio do apresentador:

- Se o carro estiver na porta 1 (C_1), o apresentador escolhe aleatoriamente entre as portas 2 ou 3, logo $P(A \mid C_1) = \frac{1}{2}$.
- Se o carro estiver na porta 2 (C_2), o apresentador deve abrir a porta 3, pois não pode revelar o prêmio, então $P(A \mid C_2) = 1$.
- Se o carro estiver na porta 3 (C_3), o apresentador nunca a abrirá, portanto $P(A \mid C_3) = 0$.

Logo, pelo Teorema de Bayes temos que,

$$P(C_2 \mid A) = \frac{P(C_2) \cdot P(A \mid C_2)}{P(A)}$$

$$P(C_1 \mid A) = \frac{P(C_1) \cdot P(A \mid C_1)}{P(A)}$$

Antes de calcularmos $P(A)$, note que C_1, C_2 e C_3 são eventos mutuamente exclusivos, e reunidos formam o espaço amostral completo, então, podemos decompor o evento A na reunião de três outros, também mutuamente exclusivos, como segue:

$$\begin{aligned} A &= (C_1 \cap A) \cup (C_2 \cap A) \cup (C_3 \cap A) \\ P(A) &= P(C_1 \cap A) + P(C_2 \cap A) + P(C_3 \cap A) \\ &= P(C_1) \cdot P(A \mid C_1) + P(C_2) \cdot P(A \mid C_2) + P(C_3) \cdot P(A \mid C_3). \end{aligned}$$

Dessa forma,

$$\begin{aligned} P(A) &= P(C_1) \cdot P(A \mid C_1) + P(C_2) \cdot P(A \mid C_2) + P(C_3) \cdot P(A \mid C_3) \\ &= \frac{1}{3} \cdot \frac{1}{2} + \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 0 \\ &= \frac{1}{2}. \end{aligned}$$

Logo, pelo Teorema de Bayes, a probabilidade de o carro estar na porta 2 após o apresentador abrir a porta 3 é:

$$
\begin{aligned}
P(C_2 \mid A) &= \frac{P(C_2) \cdot P(A \mid C_2)}{P(A)} \\
&= \frac{\frac{1}{3} \cdot 1}{\frac{1}{2}} \\
&= \frac{2}{3}.
\end{aligned}
$$

A probabilidade de o carro estar na porta (C_1, escolhida inicialmente pelo participante permanece 1/3, pois

$$
\begin{aligned}
P(C_1 \mid A) &= \frac{P(C_1) \cdot P(A \mid C_1)}{P(A)} \\
&= \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{2}} \\
&= \frac{1}{3}.
\end{aligned}
$$

Essa análise mostra que a estratégia de trocar de porta aproveita a informação implícita na ação do apresentador, redistribuindo a probabilidade inicialmente uniforme para concentrar 2/3 na porta não escolhida e não revelada. A generalização para qualquer porta inicial (2 ou 3) mantém o mesmo raciocínio, confirmando que a troca sempre oferece vantagem probabilística, resolvendo o paradoxo aparente e mostrando que a resposta de Marilyn vos Savant estava certa.

2.5 MATRIZES

Essa seção foi escrita de acordo com Bussab; Morettin (2010) e Boldrini et al. (1980).

Nesta seção, serão introduzidos alguns conceitos de matrizes, sendo abordados os conceitos de operações básicas como: a soma de matrizes, definida apenas para matrizes de mesma ordem e a multiplicação por escalar, que multiplica todos os elementos da matriz por um mesmo valor escalar, além disso, discutiremos noções de convergência de sequências de matrizes, em que, uma sucessão de matrizes tende a uma matriz limite, sendo este, um conceito fundamental para análise de sistemas que evoluem ao longo do tempo.

Definição 2.5.1 (Soma de matrizes): Sejam $A = [a_{ij}]_{m\times n}$ e $B = [b_{ij}]_{m\times n}$ matrizes de mesma ordem. A soma $A+B$ é definida como a matriz cujas entradas são obtidas ao somar cada elemento a_{ij} da matriz A ao elemento correspondente b_{ij} da matriz B. Matrizes de ordens distintas não podem ser somadas.

Exemplo 2.5.1: Considere as matrizes A e B de ordem 4×4:

$$A = \begin{bmatrix} 7 & -9 & 0 & -2 \\ 8 & 3 & 4 & 10 \\ -2 & 1 & 6 & -1 \\ 1 & 2 & 0 & 11 \end{bmatrix}, \quad B = \begin{bmatrix} -1 & 4 & 2 & -2 \\ 8 & 7 & 1 & -6 \\ -1 & 6 & 4 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$$

A soma $A + B$ resulta em:

$$A + B = \begin{bmatrix} 6 & -5 & 2 & -4 \\ 16 & 10 & 5 & 4 \\ -3 & 7 & 10 & -1 \\ 2 & 2 & 0 & 13 \end{bmatrix}$$

Definição 2.5.2 (Multiplicação por escalar): Seja $B = [b_{ij}]_{m \times n}$ uma matriz e c um escalar. A multiplicação por escalar $c \cdot B$ é definida como a matriz cujas entradas são obtidas multiplicando cada elemento b_{ij} de B por c. Ou seja:

$$c \cdot B = [c \cdot b_{ij}]_{m \times n}$$

Exemplo 2.5.2: Dadas as matrizes A e B:

$$A = \begin{bmatrix} 3 & -2 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 1 \\ -4 & -6 & 1 \end{bmatrix}$$

Temos:

$$2A = 2 \cdot \begin{bmatrix} 3 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & -4 \\ 0 & 2 \end{bmatrix}$$

$$(-5)B = (-5) \cdot \begin{bmatrix} 1 & 3 & 0 \\ 2 & -1 & 1 \\ -4 & -6 & 1 \end{bmatrix} = \begin{bmatrix} -5 & -15 & 0 \\ -10 & 5 & -5 \\ 20 & 30 & -5 \end{bmatrix}$$

Definição 2.5.3 (Convergência de sequência de matrizes): Seja $\{A_k\}_{k \in \mathbb{N}}$ uma sequência de matrizes em $\mathbb{R}^{m \times n}$. Dizemos que $\{A_k\}$ converge para a matriz $A = [a_{ij}]_{m \times n}$ se, para todo $i = 1, 2, \ldots, m$ e $j = 1, 2, \ldots, n$, o elemento (i, j) da matriz A_k se aproxima do elemento (i, j) correspondente da matriz A, isto é,

$$\lim_{k \to \infty} a_{ij}^{(k)} = a_{ij}, \quad \forall\, i \in \{1, 2, \ldots, m\},\ \forall\, j \in \{1, 2, \ldots, n\},$$

onde $a_{ij}^{(k)}$ denota o elemento na linha i, coluna j da matriz A_k.

Exemplo 2.5.3: Considere a sequência $\{A_k\}_{k\in\mathbb{N}}$, onde:

$$A_k = \begin{bmatrix} \dfrac{k^2-1}{k^2+2} & \dfrac{1}{k} \end{bmatrix},$$

Para entender o comportamento da sequência, vamos calcular alguns termos específicos e observar o padrão:

- Para $k = 1$:
$$A_1 = \begin{bmatrix} 0 & 1 \end{bmatrix},$$
- Para $k = 2$:
$$A_2 = \begin{bmatrix} \dfrac{2^2-1}{2^2+2} & \dfrac{1}{2} \end{bmatrix} = \begin{bmatrix} \dfrac{3}{6} & \dfrac{1}{2} \end{bmatrix} = \begin{bmatrix} 0{,}5 & 0{,}5 \end{bmatrix},$$
- Para $k = 100$:
$$A_{100} = \begin{bmatrix} \dfrac{100^2-1}{100^2+2} & \dfrac{1}{100} \end{bmatrix} \approx \begin{bmatrix} 0{,}9997 & 0{,}01 \end{bmatrix},$$
- Para $k = 1000$:
$$A_{1000} = \begin{bmatrix} \dfrac{1000^2-1}{1000^2+2} & \dfrac{1}{1000} \end{bmatrix} \approx \begin{bmatrix} 0{,}9999 & 0{,}001 \end{bmatrix}.$$

Embora a matriz inicial $A_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$ possa sugerir um comportamento diferente, o limite da sequência não é $\begin{bmatrix} 0 & 1 \end{bmatrix}$. À medida que calculamos A_k para valores grandes de k, as componentes a_{11} e a_{12} da matriz A_k se aproximam dos valores 1 e 0, respectivamente, ou seja:

$$\lim_{k\to\infty} \frac{k^2-1}{k^2+2} = 1 \quad \text{e} \quad \lim_{k\to\infty} \frac{1}{k} = 0.$$

Assim,

$$\lim_{k\to\infty} A_k = \lim_{k\to\infty} \begin{bmatrix} \dfrac{k^2-1}{k^2+2} & \dfrac{1}{k} \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

3 CADEIAS DE MARKOV

As seções 3.1 e 3.2 foram escritas baseadas de acordo com os trabalhos de Magela (2015); Amorim; Mozer (2020); Ross (2010) e Costa (2017).

O estudo das cadeias de Markov é um tópico relevante e moderno na teoria das probabilidades, que envolve sequências de variáveis aleatórias. Nesta seção propõe-se apresentar conceitos e definições relacionados as cadeias de Markov, ilustrando as ideias por meio da exposição de exemplos. Inicialmente, é necessário apresentar o conceito de sequências de variáveis aleatórias. Em seguida, será apresentada definição formal de uma cadeia de Markov finita. A partir daí, definiremos uma matriz, denominada matriz de transição de uma cadeia de Markov, cuja utilidade é fundamental na teoria.

Com base na matriz de transição, podemos analisá-la e estudar sua distribuição de probabilidade estacionária utilizando operações básicas com matrizes. Esta é uma característica importante na aplicação da teoria.

3.1 CONCEITOS BÁSICOS

3.1.1 Sequência de variáveis aleatórias

Definição 3.1.1.1: Uma sequência de variáveis aleatórias $(X_n)_{n\geq 0}$ indexadas por n, em um conjunto A não vazio, é uma aplicação

$$X : \mathbb{N} \to A,$$

onde tem-se apenas a probabilidade que X_n, $n > 0$, pode assumir em A.

De modo geral, elas são usadas em estudos de fenômenos observados ao longo de um período.

Exemplo 3.1.1.1: Considere um experimento no qual um dado honesto de seis faces (numeradas de 1 a 6) é lançado repetidamente. Cada lançamento gera um resultado que pode ser representado por uma sequência variável aleatória $(X_n)_{n\geq 0}$, onde n indica a posição na sequência de lançamentos. Por exemplo, uma possível sequência de resultados após múltiplos lançamentos é:

$$(2, 4, 5, 3, 6, 1, 1, 3, 4, 6, 5, 1, 2, 5, 5, 6, 1, 2, 5, 3, 4, 6, 1, \dots).$$

Essa sequência é uma sequência de variáveis aleatórias, pois cada termo de $(X_n)_{n\geq 0}$ é imprevisível e obtido aleatoriamente. Além disso, observa-se que essas sequências se diferenciam de uma sequência determinística. Como exemplo, considere a sequência $(X_n)_{n\geq 0}$ definida por $X_n = 2n$, que gera os termos:

$$(2, 4, 6, 8, \dots),$$

neste caso, para cada valor de n, os termos são previsíveis e determinados pela regra $X_n = 2n$, contrastando com a aleatoriedade e imprevisibilidade das sequências de variáveis aleatórias.

No exemplo do dado, a sequência $(X_n)_{n\geq 0}$ possui duas propriedades fundamentais: são independentes e identicamente distribuídas (i.i.d.). A independência significa que o resultado de um lançamento não influencia os resultados subsequentes, e a distribuição idêntica garante que cada $(X_n)_{n\geq 0}$ mantém a mesma probabilidade para cada face, independentemente de n, ou seja, a variável aleatória X_n no n-ésimo lançamento do dado é igual a um valor especifíco k, isto é,

$$P(X_n = k) = \frac{1}{6}, \quad \text{para todo } n \geq 0 \text{ e } k \in \{1, 2, 3, 4, 5, 6\}.$$

No entanto, as cadeias de Markov representam uma classe mais específica de sequências de variáveis aleatórias. Nelas, a dependência entre os termos não é totalmente ausente como em processos (i.i.d.), mas condicionada ao estado atual. Formalmente, uma cadeia de Markov satisfaz a propriedade de Markov:

$$P(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \dots, X_n = i) = P(X_{n+1} = j \mid X_n = i),$$

ou seja, o estado futuro X_{n+1} depende apenas do estado presente X_n e não de estados passados $X_0, X_1, X_2, \dots, X_{n-1}$. Essa memória limitada diferencia cadeias de Markov de sequências (i.i.d.), onde há independência total entre todos os termos.

As Cadeias de Markov são modelos probabilísticos amplamente utilizados para representar sistemas que evoluem ao longo do tempo, onde o próximo estado depende apenas do estado atual, sem influência de estados anteriores, tornando assim essenciais em diversas áreas, como teoria das probabilidades, física, economia e ciência da computação, devido à sua simplicidade e capacidade de modelar dinâmicas complexas.

3.1.2 Probabilidade de transição

Quando o sistema encontra-se no estado i no tempo n, ele pode permanecer no estado i ou transitar para o estado j no instante $n + 1$. Essas mudanças entre os estados são denominadas probabilidade de transição, representadas por p_{ij}, que corresponde a probabilidade de transição do estado i para o estado j. ou seja,

$$p_{ij}(n) = P(X_{n+1} = j \mid X_n = i),$$

onde $p_{ij}(n)$ representa a probabilidade condicional de o sistema transitar do estado i no instante n para o estado j no instante $n+1$. Quando as probabilidades de transições $P_{ij}(n)$ não dependem do instante n, ou seja, são constantes ao longo do tempo, denotamos por:

$$P_{ij}(n) = P_{ij}.$$

3.1.3 Matriz de transição

Uma Cadeia de Markov com n estados possíveis possui sua matriz de transição definida como uma matriz quadrada de ordem $n \times n$, denotada por $Q = (p_{ij})$, em que p_{ij} representa a probabilidade de transição do estado i para o estado j, com $1 \leq i, j \leq n$. Essa matriz pode ser representada da seguinte forma:

$$Q = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix}$$

onde Q satisfaz as seguintes propriedades:

1) $p_{ij} \geq 0$

2) $\sum_{j=1}^{n} p_{ij} = 1 \quad$ para $i = 1, 2, \ldots, n$.

3.2 DEFINIÇÃO DE UMA CADEIA DE MARKOV

Definição 3.2.1: Uma cadeia de Markov é uma sequência de variáveis aleatórias $(X_n)_{n\geq 0}$ em que cada X_n assume valores no conjunto A, satisfazendo a propriedade de Markov:

$$P(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \ldots, X_n = i) = P(X_{n+1} = j \mid X_n = i) = q_{ij}$$

para todo $n \geq 0$ e para todos $i, j \in A$. Além disso, a cadeia de Markov começa em um estado inicial $X_0 = a \in A$, com probabilidade $P(X_0 = a)$ especificada.

Observação 3.2.1: Considere uma cadeia de Markov definida sobre o conjunto finito de estados A com matriz de transição $Q = (q_{ij})$, onde q_{ij} representa a probabilidade de transição do estado i para o estado j.

A seguir, enuncia-se um teorema fundamental para o estudo de Cadeias de Markov, que estabelece propriedades essenciais sobre sua dinâmica e convergência:

Teorema 3.2.1: Considere uma cadeia de Markov com matriz de transição Q e seja D_n a matriz linha de distribuição de probabilidades que descreve descreve como as probabilidades estão distribuídas entre os possíveis estados da variável X_n em n passos. Então:

1) $D_n = D_{n-1}Q$ e $D_n = D_0 \cdot Q^n$ para todo $n \geq 1$.

2) Se Q^k tem todas as entradas positivas para algum $k \geq 1$, então:

$$\lim_{n\to\infty} D_n = E \quad \text{e} \quad \lim_{n\to\infty} Q^n = S,$$

onde E é a única matriz linha (de probabilidades) que satisfaz a equação:

$$EQ = E,$$

denominada *matriz de estado estacionário* e S é a matriz de mesma ordem de Q na qual todas as linhas são iguais a E.

Entretanto, a demonstração deste teorema não será apresentada neste trabalho, ficando como um ponto de interesse para o leitor; a demonstração pode ser encontrada em Levin; Peres e Wilmer (2020).

Os exemplos 3.2.1 e 3.2.2 a seguir, são adaptados dos trabalhos de Costa (2017) e Amorim e Mozer (2020). Esses exemplos que ilustram a construção e análise de Cadeias de Markov, explorando conceitos como probabilidade de transição, distribuição de probabilidades, matriz de transição e comportamento a longo prazo.

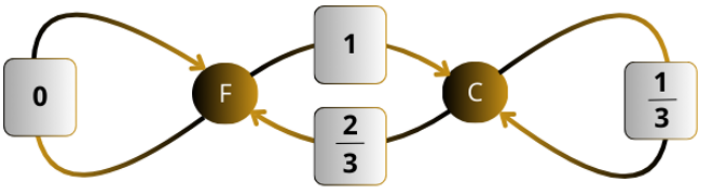
Exemplo 3.2.1: Toda quarta-feira à noite, Mikael realiza uma atividade de lazer: assiste a uma partida de futebol ou vai ao cinema, de modo que ele nunca assiste a partidas de futebol duas quartas seguidas. No entanto, se ele decidir ir ao cinema em uma quarta-feira, na próxima quarta a probabilidade de repetir a ida ao cinema é de $\frac{1}{3}$.

A escolha que Mikael faz cada quarta-feira, baseia-se exclusivamente na sua decisão tomada na quarta-feira anterior. Logo, essa mudança de hábito que Mikael faz cada quarta-feira é caracterizada como uma Cadeia de Markov onde os estados da Cadeia são representados por (F - Mikael assistir jogo de futebol) e (C – Mikael ir ao cinema). Desse modo, vamos representar as probabilidades transição por meio de grafos.

A probabilidade de Mikael assistir um jogo de futebol nessa quarta-feira, e na próxima quarta ele assistir futebol novamente é zero, pois, ele nunca assiste a um jogo de futebol duas semanas seguidas, ou seja, $p_{11} = 0$.

A probabilidade de Mikael assistir um jogo de futebol nessa quarta-feira e na próxima ele ir ao cinema é 1, pois ele nunca assiste patidas de futebol duas semanas seguidas, ou seja, $p_{12} = 1$.

Figura 2 – Probabilidades de Transição.



Fonte: Adaptado de Costa (2017).

Partindo do princípio estabelecido pelo problema que a probabilidade de Mikael repetir a ida ao cinema em quartas consecutivas é de $\frac{1}{3}$, a probabilidade de Mikael ir ao cinema nessa quarta feira e na próxima quarta ele assistir ao jogo de futebol, tem probabilidade de $\frac{2}{3}$. Esse valor é obtido através da propriedade 2) da matriz de transição, que exige que a soma das probabilidades de transição a partir de qualquer estado seja igual a 1. Como $p_{11} = \frac{1}{3}$ por definição, segue que $p_{21} = 1 - p_{11} = 1 - \frac{1}{3} = \frac{2}{3}$.

Nota-se que a probabilidade de Mikael ir ao cinema nessa semana e na próxima semana ele ir ao cinema novamente é $\frac{1}{3}$, pois, de acordo com o enunciado do problema, a entrada $p_{22} = \frac{1}{3}$ da matriz Q.

Assim, obtemos a matriz de transição Q:

$$Q = \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

A probabilidade de Mikael de assistir ao jogo de futebol ou ir ao cinema é a mesma e as duas somadas são iguais a 100%, ou seja, a distribuição de probabilidade do sistema no instante inicial é 50% para cada estado:

$$D_0 = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Note que,

$$Q^2 = \begin{bmatrix} 0.66667 & 0.33333 \\ 0.22222 & 0.77778 \end{bmatrix}$$

Após três quarta-feira a probabilidade de Mikael está assistindo futebol é calculada utilizando utilizando o item 1) do teorema 3.2.1 temos,

$$
\begin{aligned}
D_n &= D_0 \cdot Q^n \\
D_3 &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}^3 \\
D_3 &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \cdot \begin{bmatrix} 0.2222 & 0.7778 \\ 0.5185 & 0.4815 \end{bmatrix} \\
D_3 &= \begin{bmatrix} 0.3704 & 0.6296 \end{bmatrix}.
\end{aligned}
$$

Desse modo, após três quartas-feiras a probabilidade de Mikael de ir assistir ao futebol é de 37.04% e 62.96% dele ir ao cinema.

No estudo de Cadeias de Markov, analisamos as distribuições de probabilidade para as quais o sistema converge a longo prazo. Neste exemplo, verificamos a distribuição de probabilidade após três quartas-feiras para as escolhas de Mikael entre ir ao cinema ou assistir ao jogo de futebol. Porém, ao analisarmos o comportamento de Mikael após n quartas-feiras, qual será o resultado?

Inicialmente, devemos avaliar a existência ou não da matriz de estado estacionário para a qual D_n converge. De acordo com o item 2) do Teorema 3.2.1, devemos verificar se existe uma potência da matriz Q com entradas todas positivas. De fato, calculando Q elevado a n para diversos valores de n, verificamos que Q^2 tem todas as entradas positivas. Logo, D_n converge para uma matriz de estado estacionário E. Assim, uma aproximação da matriz de estado estacionário com precisão de cinco casas decimais é dada por:

$$
D_{21} = D_0 \cdot Q^{21} \cong \begin{bmatrix} 0,39998 & 0,60002 \end{bmatrix}.
$$

Dessa forma, verifica-se que as entradas da matriz D_n permanecem iguais com o arredondamento de cinco casas decimais, a partir de $n = 21$.

Ainda de acordo com item 2) do teorema 3.2.1 podemos obter de forma precisa a matriz E resolvendo o sistema linear $EQ = E$. Assim, Seja $\mathbf{E} = \begin{bmatrix} F & C \end{bmatrix}$, temos

$$
\begin{bmatrix} F & C \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix} = \begin{bmatrix} F & C \end{bmatrix}
$$

O que implica em:

$$
\begin{cases} F = \dfrac{2C}{3} \\ F + \dfrac{C}{3} = C \end{cases}
$$

Ao resolver o sistema indeterminado, obtêm-se infinitas soluções. Mas, apenas uma

delas satisfaz a condição necessária para a matriz E, que a soma de todas entradas deve ser igual a 1, pois se trata de uma matriz de probabilidade. Logo,

$$\Rightarrow \begin{cases} F = \dfrac{2C}{3} \\ F + \dfrac{C}{3} = C \\ F + C = 1 \end{cases} \Rightarrow F = \frac{2}{5},\ C = \frac{3}{5}.$$

Assim, a matriz de estado estacionário é:

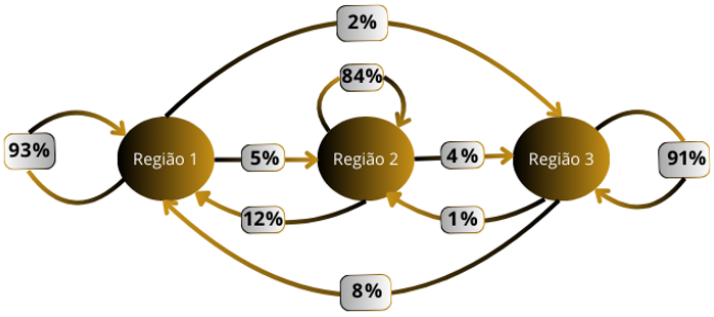
$$E = \begin{bmatrix} \frac{2}{5} & \frac{3}{5} \end{bmatrix}.$$

Para todo $n \geq 21$, a matriz D_n coincide com E arredondada com cinco casas decimais. O valor exato de E pode ser comparado com aproximações numéricas calculadas para diferentes valores de n. Nesse caso, após 21 iterações (para $n = 21$), obtém-se uma aproximação com uma precisão de cinco casas decimais.

Além disso, o resultado obtido para **E** indica que, no equilíbrio, as probabilidades de Mikael optar por assistir uma partida de futebol e ir ao cinema são de 40% e 60%, respectivamente.

Exemplo 3.2.2: Um país está dividido em três regiões demográficas, Constatou-se que todos os anos os moradores mudam de uma região para outra conforme a figura abaixo:

Figura 3 – Probabilidades de transição entre as regiões.



Fonte: O autor.

De acordo com a figura, nota-se que:

- Entre os moradores da região 1, 5% mudam-se para região 2 e 2% para a região 3.
- Entre os moradores da região 2, 12% mudam-se para a região 1 e 4% para a região 3.
- Entre os moradores da região 3, 8% vão para a região 1 e 1% para a região 2.

Para definir a matriz de transição $Q = (p_{ij})_{3\times3}$, considere que p_{ii} representa a probabilidade de um morador permanecer na região i, enquanto p_{ij} com $i \neq j$ corresponde à probabilidade de transição do estado i para o estado j. Nesse caso, a probabilidade do morador sair da região 1 e ir para região 2 é p_{12}, ou seja, $p_{12} = 5\%$ ou $0,05$. já a probabilidade do morador sair da região 1 e ir para região 3 é $p_{13} = 2\%$ ou $0,02$. E a probabilidade p_{11} do morador permanecer na região 1 é obtida por meio da propriedade 2) da matriz de transição, em que a soma das entradas da matriz Q é igual a 1, ou seja,

$$\sum_{j=1}^{3} p_{1j} = 1 \iff p_{11} + p_{12} + p_{13} = 1$$

$$\begin{aligned} p_{11} &= 1 - p_{12} - p_{13} \\ &= 0{,}93. \end{aligned}$$

De maneira análoga, pela a propriedade 2) da matriz de transição encontramos as probabilidades $p_{22} = 84\%$ ou $0,84$ e $p_{33} = 91\%$ ou $0,91$. Assim, obtemos a matriz de transição Q:

$$Q = \begin{bmatrix} 93\% & 5\% & 2\% \\ 12\% & 84\% & 4\% \\ 8\% & 1\% & 91\% \end{bmatrix} \quad \text{ou} \quad Q = \begin{bmatrix} 0,93 & 0,05 & 0,02 \\ 0,12 & 0,84 & 0,04 \\ 0,08 & 0,01 & 0,91 \end{bmatrix}$$

A sequência de variáveis aleatórias $(X_n)_{n\geq0}$, que representa a região ocupada pelos moradores no ano n, satisfaz a propriedade de Markov, ou seja, a probabilidade de transição para o próximo estado depende apenas do estado atual, e não dos estados anteriores. Por exemplo, a probabilidade de um morador da Região 2 migrar para a Região 1 no quinto ano, não depende dos anos anteriores ao quarto ano, ou seja:

$$P(X_5 = 1 \mid X_4 = 2) = 0,12.$$

Para analisar a distribuição populacional ao longo do tempo, consideramos um estado inicial D_0. Se a cadeia inicia na Região 1, temos 100% de chance que o processo comece no primeiro estado, sem possibilidade de começar em qualquer outro, o que é representado por

$$D_0 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}.$$

Após um ano, a distribuição evolui para D_1, onde há 93% de probabilidade de permanecer na Região 1, 5% de migrar para a Região 2 e 2% para a Região 3, dada por

$$D_1 = \begin{bmatrix} 0{,}93 & 0{,}05 & 0{,}02 \end{bmatrix}.$$

Após dois anos, a distribuição D_2 é calculada multiplicando D_1 pela matriz Q, com aproximação de cinco casas decimais:

$$D_2 = \begin{bmatrix} 0{,}93 & 0{,}05 & 0{,}02 \end{bmatrix} \cdot \begin{bmatrix} 0{,}93 & 0{,}05 & 0{,}02 \\ 0{,}12 & 0{,}84 & 0{,}04 \\ 0{,}08 & 0{,}01 & 0{,}91 \end{bmatrix} = \begin{bmatrix} 0{,}8725 & 0{,}0887 & 0{,}0388 \end{bmatrix}.$$

Isso implica $P(X_2 = 1) = 0{,}8725$, $P(X_2 = 2) = 0{,}0887$ e $P(X_2 = 3) = 0{,}0388$.

Após três anos, a distribuição D_3 é calculada multiplicando D_2 pela matriz Q, com aproximação de cinco casas decimais:

$$D_3 = \begin{bmatrix} 0{,}8725 & 0{,}0887 & 0{,}0388 \end{bmatrix} \cdot \begin{bmatrix} 0{,}93 & 0{,}05 & 0{,}02 \\ 0{,}12 & 0{,}84 & 0{,}04 \\ 0{,}08 & 0{,}01 & 0{,}91 \end{bmatrix} = \begin{bmatrix} 0{,}82517 & 0{,}11852 & 0{,}05631 \end{bmatrix}.$$

De forma geral, usando argumentos semelhantes podemos deduzir que, para todo $n \geq 1$ temos:

$$D_n = D_{n-1}Q \implies \begin{cases} D_1 = D_0 \cdot Q \\ D_2 = D_1 \cdot Q = D_0 \cdot Q^2 \\ \quad \vdots \\ D_n = D_{n-1} \cdot Q = D_0 \cdot Q^n \end{cases}$$

Assim, por meio da relação $D_n = D_0 \cdot Q^n$, dependemos apenas da distribuição inicial da população D_0 para determinar a proporção de moradores em cada região após n anos.

Se assumirmos uma distribuição inicial em que cada um dos três estados tem a mesma probabilidade de ocorrer, $D_0 = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$, as distribuições após n anos são obtidas iterativamente. Por exemplo, para $n = 1$, $n = 2$, $n = 3$, e $n = 4$, temos:

$$D_1 = \begin{bmatrix} 0,38 & 0,30 & 0,32 \end{bmatrix}, \quad D_2 = \begin{bmatrix} 0,41 & 0,27 & 0,31 \end{bmatrix},$$

$$D_3 = \begin{bmatrix} 0,44 & 0,25 & 0,30 \end{bmatrix}, \quad D_4 = \begin{bmatrix} 0,47 & 0,24 & 0,30 \end{bmatrix}.$$

Após quatro anos, observa-se que aproximadamente 30% da população estará na Região 3.

No estudo das cadeias de Markov, essa análise é particularmente relevante, pois permite identificar padrões e tendências estáveis à medida que o sistema evolui. Nesse caso, esse comportamento pode ser avaliado através da existência (ou não) da matriz de estado estacionário para qual D_n converge. Essa matriz representa uma distribuição de probabilidades que se mantém inalterada sob a aplicação repetida da matriz de transição associada à cadeia. De acordo com o item 1) do Teorema 3.2.1, é necessário analisar se existe uma potência da matriz Q com todas as entradas positivas.

De fato, pode-se verificar calculando as potências de Q para diversos valores de n, verificando que Q^5 tem todas as entradas positivas. Observe a seguir:

$$Q^2 = \begin{bmatrix} 0,87 & 0,09 & 0,04 \\ 0,22 & 0,71 & 0,07 \\ 0,15 & 0,02 & 0,83 \end{bmatrix}, \quad Q^3 = \begin{bmatrix} 0,83 & 0,12 & 0,06 \\ 0,29 & 0,61 & 0,10 \\ 0,21 & 0,03 & 0,76 \end{bmatrix},$$

$$Q^4 = \begin{bmatrix} 0,79 & 0,14 & 0,07 \\ 0,35 & 0,53 & 0,12 \\ 0,26 & 0,05 & 0,70 \end{bmatrix}, \quad Q^5 = \begin{bmatrix} 0,75 & 0,16 & 0,09 \\ 0,40 & 0,46 & 0,14 \\ 0,30 & 0,06 & 0,64 \end{bmatrix}.$$

Logo, D_n converge para uma matriz de estado estacionário E. Assim, utilizando aproximações de cinco casas decimais, a matriz de estado estacionário é dada por:

$$D_{91} = D_0 \cdot Q^{91} \cong \begin{bmatrix} 0,58577 & 0,19665 & 0,21757 \end{bmatrix}.$$

Ademais, de acordo com o item (2) do teorema 3.2.1, pode-se determinar com precisão a matriz E de estado estacionário resolvendo o sistema linear $EQ = E$. Seja $E = \begin{bmatrix} x & y & z \end{bmatrix}$ a matriz de estado estacionário. Segue que:

$$\begin{bmatrix} x & y & z \end{bmatrix} \cdot \begin{bmatrix} 0,93 & 0,05 & 0,02 \\ 0,12 & 0,84 & 0,04 \\ 0,08 & 0,01 & 0,91 \end{bmatrix} = \begin{bmatrix} x & y & z \end{bmatrix}.$$

Esta equação recai no sistema:

$$\begin{cases} 0,93x + 0,12y + 0,08z = x, \\ 0,05x + 0,84y + 0,01z = y, \\ 0,02x + 0,04y + 0,91z = z, \end{cases}$$

que simplifica para:

$$x = \frac{35}{13}z, \quad y = \frac{47}{52}z.$$

Ao resolver o sistema indeterminado, obtêm-se infinitas soluções em função de z. No entanto, somente uma delas cumpre a condição $x + y + z = 1$, necessária para que E seja uma matriz de probabilidades. Logo:

$$\begin{cases} 0{,}93x + 0{,}12y + 0{,}08z = x, \\ 0{,}05x + 0{,}84y + 0{,}01z = y, \\ 0{,}02x + 0{,}04y + 0{,}91z = z, \\ x + y + z = 1, \end{cases} \iff \begin{cases} x = \dfrac{140}{239}, \\ y = \dfrac{47}{239}, \\ z = \dfrac{52}{239}. \end{cases}$$

Assim, a matriz de estado estacionário é:

$$E = \begin{bmatrix} \frac{140}{239} & \frac{47}{239} & \frac{52}{239} \end{bmatrix}.$$

A distribuição coincide com D_n para todo $n \geq 91$, calculada após 91 iterações, confirmando a convergência para o estado estacionário.

Este exemplo ilustra como as Cadeias de Markov permitem prever tendências populacionais e identificar como o sistema evolui ao longo do tempo. A distribuição estacionária indica que, a longo prazo, cerca de $58,6\%$ da população residirá na Região 1, $19,58\%$ na Região 2 e $21,75\%$ na Região 3, demonstrando a utilidade do modelo na análise de mobilidade geográfica e planejamento demográfico.

4 CONCLUSÃO

O presente estudo demonstrou a evolução da Teoria da Probabilidade evidenciando os diversos estudiosos que contribuíram para o seu desenvolvimento, desde os jogos de azar no Renascimento, até a formalização rigorosa no século XX.

A teoria das Cadeias de Markov descreve sistemas cuja evolução depende apenas do estado atual e não do histórico de estados anteriores, sendo amplamente aplicadas em diversas áreas do conhecimento. Assim, o presente estudo trouxe um panorama dessa teoria de maneira clara e aplicável com a utilização de matrizes permitindo uma abordagem mais didática e compreensível de conceitos que, à primeira vista, podem parecer complexos.

Por fim, este estudo serviu não apenas para consolidar o conhecimento a respeito dessa temática, mas também para incentivar investigações futuras. Acredita-se que a abordagem didática adotada nesse trabalho possa auxiliar outros pesquisadores e estudantes a se familiarizarem com as Cadeias de Markov.

REFERÊNCIAS

[1] AMORIM, V.; MOZER, G. **Probabilidade além da combinatória: tópicos e problemas reais com foco no raciocínio probabilístico**. 1. ed. Rio de Janeiro: ANPMat, 2020.

[2] BEZERRA, N. **Análise Combinatória e Probabilidade**. 1. ed. Belém: Editaedi, 2018.

[3] BOLDRINI, J. L. et al. **Álgebra Linear**. 2. ed. São Paulo: Harbra, 1980.

[4] BUSSAB, W. O.; MORETTIN, P. A. **Estatística Básica.** 6. ed. São Paulo : Saraiva, 2010.

[5] CALABRIA, A. R.; CAVALARI, M. F. **Um passeio histórico pelo início da teoria das probabilidades**. Campinas: SBHMAT, 2013.

[6] CARVALHO, A. L. de et al. Aplicação da cadeia de Markov para dias secos e chuvosos. **Revista Brasileira de Meteorologia**, v. 32, p. 207-214, 2017.

[7] COUTINHO, C. Q. S. **Introdução ao conceito de probabilidade por uma visão frequentista: estudo epistemológico e didático**. 1994. 151 f. Dissertação (Mestrado em Matemática) - Pontifícia Universidade Católica de São Paulo, São Paulo, 1994.

[8] COSTA, F. S. **Cadeias de Markov regulares: Uma abordagem para alunos e professores do ensino médio.** 100 f. Dissertação (Mestrado profissional em Matemática em Rede Nacional) - Universidade Federal do Amazonas, Manaus, 2017.

[9] DA SILVA, G. S.; GRANDO, R. C. O que a história nos conta sobre a Teoria da Probabilidade? **Anais do ENAPHEM - Encontro Nacional de Pesquisa em História da Educação Matemática**, n. 7, 2024.

[10] DA SILVA, C. E. V. **Aplicações da álgebra linear nas Cadeias de Markov**. 39 f. Dissertação (Mestrado) – Universidade Federal de Goiás, Instituto de Matemática e Estatística, 2013.

[11] EVES, H. **Introdução à história da matemática**. Tradução de Hygino H. Domingues. 5. ed. Campinas, SP: Editora da Unicamp, 2011.

[12] HAZZAN, S. **Fundamentos de matemática elementar, 5: combinatória e probabilidade**. 7. ed. São Paulo: Atual, 2004.

[13] LEVIN, D. A.; PERES, Y. WILMER, E.L. **Markov Chains and Mixing Times** Providence, R.I.: American Mathematical Society, 2009.

[14] MAGELA, M. **Teoria básica das cadeias de Markov**. Vitória: Universidade Federal do Esprito Santo, 2015.

[15] MORGADO, A. C. et al. **Análise combinatória e probabilidade**. 7. ed. Rio de Janeiro: SBM, 2005.

[16] MORAES, L. C. L. et al. **Ensino de probabilidade: historicidade e interdisciplinaridade**. 2014.

[17] ROSS, S. **Probabilidade: um curso moderno com aplicações [recurso eletrônico]**. 8. ed. Tradução de Alberto Resende De Conti. Porto Alegre: Bookman, 2010.

[18] ROTUNNO, S. A. M. **Estatística e probabilidade: um estudo sobre a inserção desses conteúdos no ensino fundamental**. Dissertação (mestrado profissional)-Universidade Federal do Paraná, Curitiba. 117 p., 2007.

[19] SALSA, I. S.; MOREIRA, J. A. **Probabilidade e estatística**. 2. ed. Natal: EDUFRN, 2014.

[20] SILVA, V. D. M. **A abordagem da história da matemática no ensino da probabilidade nos livros didáticos do ensino médio** [manuscrito]. 2018. 62 p.: il., color.

[21] SILVA, F. C. A. **Cadeias de Markov - Uma Sequência Didática para o Ensino Médio**. Dissertação (mestrado profissional)-Universidade Federal de Lavras (UFLA), Lavras. 51 p., 2020.

[22] SOUZA JUNIOR, F. L. **Cadeias de Markov e o jogo Monopoly**. Dissertação (mestrado profissional)-Universidade Federal do ABC, Santo André. 110 p., 2020.

[23] VIALI, L. Algumas considerações sobre a origem da teoria da probabilidade. **Revista Brasileira de História da Matemática**, v. 8, n. 16, p. 143-153, 2008.