UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO, DE CIÊNCIAS EXATAS E EDUCAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO MESTRADO PROFISSIONAL EM MATEMÁTICA EM REDE NACIONAL – PROFMAT

Camila Werner

Variáveis aleatórias e Cadeias de Markov:
O problema de completar um álbum de figurinhas e uma proposta de aplicação no Ensino Médio

Blumenau
2020

Camila Werner

Variáveis aleatórias e Cadeias de Markov:
O problema de completar um álbum de figurinhas e uma proposta de aplicação no Ensino Médio

Dissertação submetida ao Mestrado Profissional em Matemática em Rede Nacional – PROFMAT – Campus Blumenau da Universidade Federal de Santa Catarina para a obtenção do Título de Mestre em Matemática.
Orientador: Prof. Dr. Eleomar Cardoso Júnior.

Blumenau
2020

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Werner , Camila
Variáveis aleatórias e Cadeias de Markov : O problema de completar um álbum de figurinhas e uma proposta de aplicação no Ensino Médio / Camila Werner ; orientador, Eleomar Cardoso Júnior, 2020.
97 p.

Dissertação (mestrado profissional) - Universidade Federal de Santa Catarina, Campus Blumenau, Programa de Pós Graduação em Matemática, Blumenau, 2020.

Inclui referências.

1. Matemática. 2. Probabilidade condicional. 3. Variáveis aleatórias. 4. Cadeias de Markov. I. Cardoso Júnior, Eleomar. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Matemática. III. Título.

Camila Werner

Variáveis aleatórias e Cadeias de Markov:

O problema de completar um álbum de figurinhas e uma proposta de aplicação no Ensino Médio

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Dr. Hugo José Lara Urdaneta
Universidade Federal de Santa Catarina – Campus Blumenau

Profª. Drª. Louise Reips.
Universidade Federal de Santa Catarina – Campus Blumenau

Prof. Dr. Leandro Batista Morgado
Universidade Federal de Santa Catarina – Campus Florianópolis

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Mestrado Profissional em Matemática em Rede Nacional – PROFMAT.

Documento assinado digitalmente
Marcio de Jesus Soares
Data: 05/11/2020 22:48:13-0300
CPF: 270.670.858-10

Coordenação do Programa

Documento assinado digitalmente
Eleomar Cardoso Junior
Data: 05/11/2020 14:03:48-0300
CPF: 056.666.459-35

Prof. Dr. Eleomar Cardoso Júnior
Orientador

Blumenau, 2020.

Este trabalho é dedicado aos meus queridos pais ao meu irmão e colegas de classe.

AGRADECIMENTOS

Agradeço aos meus pais que me ensinaram desde pequena a importância dos estudos, e por sempre me apoiarem nas minhas escolhas e compreender essa minha paixão pela matemática.

Ao meu orientador Prof. Dr. Eleomar Cardoso Júnior que me ajudou imensamente na realização desse trabalho.

A todos os professores do curso. Aos professores Hugo, Louise e Leandro membros da banca examinadora.

Aos meus colegas de curso, que me ajudaram e me apoiaram nos momentos bons e ruins que passamos nesses 2 anos.

Por fim, agradeço a Deus que me proporcionou muita fé e força e tornou este sonho de fazer um mestrado realidade.

RESUMO

Esta dissertação tem seus estudos baseados nas variáveis aleatórias e Cadeias de Markov. Inicialmente, apresentam-se as definições de elementos chave para compreensão do texto, como probabilidade condicional, variáveis aleatórias e cadeias de Markov. Além disso, mostram-se as definições de matriz de transição e vetor-estado para efetuar os produtos matriciais que levam a probabilidades de experimentos aleatórios sucessivos. Dá-se ênfase nas classificações de cadeias de Markov em absorventes ou regulares. Na sequência, abordam-se três situações-problema que são interpretadas via utilização de Cadeias de Markov. O primeiro problema descreve o lançamento de dados associados a números primos e o segundo está associado a uma hipotética migração em regiões da cidade de Blumenau. Por sua vez, o último traz o problema de completar uma coleção de figurinhas, que consiste em estimar a quantidade de compras necessárias para obter uma chance razoável de completar uma coleção de trinta elementos. Por fim, detalha-se uma proposta de abordagem de Cadeias de Markov no Ensino Médio vinculada aos dois primeiros problemas acima citados.

Palavras-chave: Probabilidade condicional, variáveis aleatórias, cadeias de Markov.

ABSTRACT

This dissertation concerns about random variables and Markov chains. Initially, we talk about preliminary definitions associated with conditional probability, random variables and Markov chains. Next, we study stochastic matrix and other tools used to find properties associated with successive random experiments. We classify regular and absorbing Markov chains. We also study three problems by using Markov chains' theory. The first one concerns prime numbers in a dice and the second concerns with a hypothetical migration among the regions of Blumenau. The last one addresses the sticker album collector problem. In our study, we estimate the number of cards that should be bought to guarantee reasonable chances to complete a collection with thirty elements. Finally, based on the first two problems, we detail a proposal of treatment of Markov chains in Brazilian secondary school.

Keywords: Conditional probability, random variables, Markov chains.

LISTA DE FIGURAS

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

O estudo de probabilidades tem suas raízes numa teoria matemática elementar dos jogos de azar. Os matemáticos Cardano (1501 – 1576) e Galileu Galilei (1564 – 1642) foram os responsáveis pelos primeiros estudos matemáticos relacionados à probabilidade e tratavam dos jogos de dados. Outros matemáticos se comunicavam por correspondências, Pascal (1612 – 1668) e Fermat (1601 – 1665), quando solucionaram um famoso problema da divisão das apostas num jogo de azar.

Por volta do ano de 1700, tem início um período importante de desenvolvimento para a teoria das probabilidades. Com os matemáticos Bernoulli (1667 – 1748) e De Moivre (1667 – 1754), a teoria dos jogos de azar foi desenvolvida baseada na definição clássica de probabilidade e vários métodos combinatórios foram utilizados como ferramenta.

Laplace (1749 – 1827) discutiu as aplicações de probabilidades em erros de observações, na determinação da massa de Júpiter, Saturno e Urano. Até Laplace, a probabilidade estava essencialmente voltada aos jogos de azar, assim, ele ampliou o campo de aplicações da teoria para outras áreas, como para a teoria dos erros. Foi a partir do trabalho de Laplace que os estudos na área de probabilidade se ampliaram e chamaram a atenção de matemáticos como Gauss (1777 – 1855) e Andrei Markov (1856 – 1922). Este último teve sua importância no estudo de probabilidades condicionais e ficou conhecido por criar o processo das Cadeias de Markov.

O presente trabalho tem seus estudos baseados nas variáveis aleatórias e Cadeias de Markov. O estudo de variáveis aleatórias nos permite trabalhar com experimentos aleatórios de forma numérica. Estes números, chamados de estados de experimento, tornam mais simples a descrição de situações que envolvem eventos que estão associados a experimentos aleatórios sucessivos, tais como: lançamentos de um dado ou de uma moeda e retiradas sucessivas, com reposição, de bolas de uma urna.

As cadeias de Markov permitem calcular certas probabilidades de experimentos aleatórios sucessivos, fazendo utilização de produtos matriciais. Para estas multiplicações são inicialmente requisitadas duas matrizes: uma matriz quadrada, a qual chamamos de matriz de transição, em que todos os seus elementos são probabilidades condicionais relacionadas aos estados do problema estudado; e uma matriz linha que denominamos de vetor estado inicial, onde seus elementos são as probabilidades iniciais de cada estado do experimento aleatório.

A depender das propriedades da matriz de transição, é possível prever o comportamento do modelo estudado. Se a matriz de transição apresentar um estado absorvente, por exemplo, as probabilidades tendem a se estabilizar se tal estado for, de alguma forma, atingido.

Este trabalho aborda assuntos discutidos no âmbito do Ensino Médio. O estudo de probabilidades é um componente curricular escolar muito relevante. Esta dissertação traz avanços no sentido que permite relacionar componentes como probabilidade e matrizes para compreender situações mais sofisticadas associadas a experimentos aleatórios. Vale ressaltar, contudo, que matrizes e suas propriedades é um tópico assumido como um conhecimento prévio do leitor.

O trabalho está desenvolvido da seguinte forma.

No Capítulo 2, apresentamos os elementos indispensáveis para a compreensão do restante do texto. Como pré-requisitos são detalhados a definição de experimentos aleatórios, espaço amostral, eventos, probabilidade e probabilidade condicional. Além disso, nesse capítulo, falamos de variáveis aleatórias e cadeias de Markov e do método que permite calcular os produtos matriciais que levam a probabilidades de experimentos aleatórios sucessivos através da matriz de transição e do vetor estado. Finalmente, são tratadas as classificações das Cadeias de Markov em Absorventes e Regulares, o que permite prever o comportamento das probabilidades vinculadas ao modelo estudado.

No Capítulo 3, abordamos duas situações-problema que são interpretadas via utilização de Cadeias de Markov. O primeiro exemplo descreve o lançamento de dados associado a números primos. Aqui, analisamos a probabilidade de aparecer 0, 1, 2 ou 3 números primos distintos, após vários lançamentos de um dado. Nosso segundo exemplo está associado à migração entre regiões da cidade de Blumenau. É importante destacar que apesar do segundo exemplo tratar de uma situação fictícia, a ideia pode ser facilmente adaptada para o estudo de outras situações, adequando-o à realidade dos estudantes do Ensino Médio.

O Capítulo 4 está destinado ao problema de completar uma coleção de figurinhas. Esse capítulo é inteiramente baseado no artigo **Álbuns de figurinhas: uma abordagem via Cadeias de Markov** (ver ref. [15]). O problema detalhado consiste em identificar a quantidade de compras necessárias para obter uma chance razoável de completar uma coleção de 30 figurinhas, adquiridas na compra dos chocolates Surpresa. Além de uma abordagem que envolve cadeias de Markov, é estabelecida uma estimativa baseada em propriedades de funções exponenciais e logarítmicas.

Por fim, no Capítulo 5, detalhamos uma proposta de abordagem de Cadeias de Markov no Ensino Médio. Os dois exemplos detalhados no Capítulo 3 são transmitidos para os alunos com conhecimentos em nível escolar. Além disso, será acrescido um terceiro exemplo relativo à retirada de bolas de uma urna. A aplicação é fundamentada na BNCC – Base Nacional Comum Curricular e nos PCNs – Parâmetros Curriculares Nacionais e incita que os alunos participem ativamente do processo de construção do conhecimento e que desenvolvam um conhecimento mais aprimorado sobre os temas probabilidades e matrizes.

Este trabalho, portanto, busca oferecer ao leitor um estudo inicial sobre Cadeias de Markov e algumas aplicações. Aos professores de Ensino Médio, apresenta uma proposta de aplicação – mostrando a possibilidade de tratar um tema teoricamente mais avançado no âmbito escolar.

2 VARIÁVEIS ALEATÓRIAS E CADEIAS DE MARKOV

O objetivo deste capítulo é introduzir Variáveis Aleatórias e Cadeias de Markov. Para tal, faz-se necessário a apresentação de alguns conceitos elementares como Espaço Amostral e Eventos, Probabilidade e Probabilidade Condicional, permitindo um maior entendimento da caracterização do objeto de estudo principal aqui abordado. As definições e propriedades apresentadas ao longo deste capítulo são baseadas nas referências [2], [7], [11] e [12].

2.1 ESPAÇO AMOSTRAL E EVENTOS

Inicialmente, apresentamos as definições de experimentos aleatórios e determinísticos, espaço amostral e eventos. Para todas as situações, é exposto um exemplo para sua melhor compreensão.

2.1.1 Experimento Aleatório

Um experimento é dito aleatório se, repetido nas mesmas condições, apresenta resultados imprevisíveis, que dependem do acaso.

Apesar de não se conhecer previamente o resultado de um experimento aleatório é, geralmente, possível estabelecer um conjunto com todos os resultados possíveis de serem observados.

Exemplo 2.1. No lançamento de um dado não viciado, não sabemos exatamente qual face aparecerá voltada para cima. Mas, devido à baixa complexidade do problema, podemos estabelecer todos os resultados possíveis de serem observados, que neste caso são os números $1, 2, 3, 4, 5$ ou 6.

2.1.2 Experimento Determinístico

Um experimento é dito determinístico se, repetido nas mesmas condições, obtém-se sempre os mesmos resultados.

Exemplo 2.2. Ao somarmos os números 2 e 3, o resultado é sempre 5. O ato de somar essas duas parcelas é um experimento determinístico.

2.1.3 Espaço Amostral

O conjunto Ω formado por todas as possibilidades de resultados de um experimento aleatório é denominado espaço amostral.

Exemplo 2.3. No lançamento de um dado não viciado, qual face aparecerá voltada para cima? Neste caso, temos como Espaço Amostral o conjunto $\Omega = \{1, 2, 3, 4, 5, 6\}$, que mostra todas as possibilidades de resultados desse experimento aleatório.

2.1.4 Evento

Seja Ω o espaço amostral associado a um experimento aleatório. Qualquer subconjunto de Ω define um evento.

O evento cujo conjunto de elementos (resultados) corresponde ao conjunto constituído pelo espaço amostral é definido como o evento certo. Como o conjunto vazio é também um subconjunto do espaço amostral Ω, o evento cujo conjunto de resultados é vazio é chamado de evento impossível.

Exemplo 2.4 Após o lançamento de um dado comum, observa-se o número voltado para cima. O espaço amostral é o conjunto $\Omega = \{1, 2, 3, 4, 5, 6\}$, no qual estão listadas todas as possibilidades de resultados para o lançamento de um dado. Considere os eventos A, B e C:

Evento A → ocorrência de um número ímpar: $A = \{1, 3, 5\}$.

Evento B → ocorrência de um número par e primo: $B = \{2\}$.

Evento C → ocorrência de um número múltiplo de 7: $C = \emptyset$.

Sejam A e B dois eventos. Então $A \cup B$ é também um evento que ocorre se, e somente se, A ou B (ou ambos ocorrem). Dizemos que $A \cup B$ é a união entre o evento A e o evento B. Seja $\{A_0, A_1, \ldots, A_n\}$ uma coleção finita de eventos. Então,

$$\bigcup_{i=0}^{n} A_i = A_0 \cup A_1 \cup \ldots \cup A_n$$

é também um evento que ocorre se, e somente se, ao menos um dos eventos A_j ocorre.

Sejam A e B dois eventos. Segue que $A \cap B$ é também um evento que ocorre se, e somente se, A e B ocorrem simultaneamente.

Em particular, se $A \cap B = \emptyset$, A e B são eventos mutuamente exclusivos dois a dois.

2.2 PROBABILIDADE

Nesta seção, abordaremos algumas propriedades e exemplos de probabilidade e probabilidade condicional.

Definição 2.1. Consideremos um experimento aleatório que tem Ω como seu espaço amostral. Uma probabilidade é uma função P que associa cada evento $A \subset \Omega$ a um número real $P(A)$ que caracteriza a probabilidade de ocorrência do evento A.

A função probabilidade satisfaz às seguintes condições:

- $0 \leq P(A) \leq 1$, para todo $A \subset \Omega$;
- $P(\emptyset) = 0$ e $P(\Omega) = 1$.
- Consideremos que $\{A_i\}_{i \in I}$ e $I \subset \mathbb{N}$ seja uma família finita ou infinita de eventos, mutuamente exclusivos dois a dois, isto é, $A_i \cap A_j = \emptyset$ se $i \neq j$. Então,

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i).$$

Definição 2.2. O espaço amostral Ω é dito equiprovável quando todos os resultados são igualmente prováveis. Quando o espaço amostral é finito e equiprovável, a probabilidade de ocorrência de um evento A (representado por $P(A)$) é o quociente entre o número de elementos do evento A (representado por $n(A)$) pelo número de elementos do espaço amostral Ω (representado por $n(\Omega)$). Em resumo,

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

Definição 2.3. Seja $A \subset \Omega$ um evento. Então, A^C é também um evento que ocorre se, e somente se, A não ocorre. Ou ainda dois eventos são ditos complementares quando são mutuamente exclusivos e sua união é igual ao espaço amostral Ω.

Sejam A e B dois eventos no espaço amostral Ω. P é uma função probabilidade.

Proposição 2.1. São válidas as seguintes propriedades:

i. $P(A^C) = 1 - P(A)$, em que A^C é o complementar de A em Ω.
ii. Se $A \subset B$, então, $P(A) = P(B) - P(B - A)$.
iii. Se $A \subset B$, então, $P(A) \leq P(B)$.
iv. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Demonstração.

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

ii. Como $B = A \cup (B - A)$, segue que:

$$P(B) = P\big(A \cup (B - A)\big) = P(A) + P(B - A)$$

$$P(A) = P(B) - P(B - A).$$

iii. Como $P(A) = P(B) - P(B - A)$ e $P(B - A) \geq 0$, pois, P é uma probabilidade, segue que $P(A) \leq P(B)$.

iv. $A \cup B = A \cup (B \cap A^C)$. Como os conjuntos A e $(B \cap A^C)$ são mutuamente exclusivos, então, por (iii), $P(A \cup B) = P(A) + P(B \cap A^C)$. Além disso, $B = (A \cap B) \cup (B \cap A^C)$ é união de dois conjuntos mutuamente exclusivos e, $P(B) = P(A \cap B) + P(B \cap A^C)$. Logo, $P(B \cap A^C) = P(B) - P(A \cap B)$ e, finalmente,

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

■

2.2.1 Probabilidade Condicional

Definição 2.4. Seja Ω um espaço amostral, em que A e B sejam seus eventos. O símbolo $P(B\,|\,A)$ caracteriza a probabilidade de ocorrência do evento B, dada a condição de que o evento A já tenha ocorrido. Nessa situação, o espaço amostral é um subconjunto de Ω, constituído pelos elementos do conjunto A.

Assumindo que Ω seja finito e equiprovável e que $n(A) > 0$, estabelece-se que

$$P(B\,|\,A) = \frac{n(A \cap B)}{n(A)} = \frac{\frac{n(A \cap B)}{n(\Omega)}}{\frac{n(A)}{n(\Omega)}} = \frac{P(A \cap B)}{P(A)}.$$

Logo,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}. \tag{2.1}$$

Da relação (2.1) obtemos a chamada regra do produto de probabilidades,

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

Aqui, temos duas formas de calcular probabilidades condicionais $P(B \mid A)$:

- Reduzindo o espaço amostral para os resultados possíveis considerados do evento A e calculando $P(B \mid A)$ através das ocorrências de B dentro de A.
- Empregando a equação (2.1) com as probabilidades $P(A)$ e $P(A \cap B)$ consideradas em relação ao espaço amostral Ω original.

Segue um exemplo de probabilidade condicional utilizando os dois métodos de resolução.

Exemplo 2.5. No lançamento de dois dados comuns, não viciados, determine a probabilidade de que a soma das duas faces voltadas para cima seja maior do que 7, sabendo que ambas as faces voltadas para cima são números primos.

Resolução: O conjunto *Ω* caracteriza o espaço amostral de todas as possíveis faces dos dois dados após o lançamento:

$$\Omega = \begin{Bmatrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{Bmatrix}.$$

Veja que o número de elementos do espaço amostral Ω é 36. O evento A caracteriza os casos em que as duas faces voltadas para cima são números primos:

$$A = \{(2,2),(2,3),(2,5),(3,2),(3,3),(3,5),(5,2),(5,3),(5,5)\}.$$

Note que o número de elementos do evento A é 9, ou seja, $n(A) = 9$.

O evento B mostra os casos em que a soma das faces é maior que 7:

$$B = \begin{Bmatrix} (2,6),(3,5),(3,6),(4,4),(4,5),(4,6),(5,3),(5,4), \\ (5,5),(5,6)(6,2),(6,3),(6,4),(6,5),(6,6) \end{Bmatrix}.$$

Agora, liste os casos em que a soma é maior que sete dentre os elementos do espaço amostral A. Vê-se $A \cap B = \{(3,5), (5,3), (5,5)\}$. Aqui,

$$n(A \cap B) = 3 \text{ e } P\big(B \,\big|\, A\big) = \frac{n(A \cap B)}{n(A)} = \frac{3}{9} = \frac{1}{3}.$$

Então, a probabilidade de ocorrência do evento B, dada a condição de que o evento A já tenha ocorrido é $\frac{1}{3}$. Em outras palavras, a probabilidade de que a soma das duas faces voltadas para cima seja maior do que 7, dada a condição de que ambas as faces voltadas para cima sejam números primos é de $\frac{1}{3}$.

Segue outra estratégia de resolução:

Inicialmente, lembramos que $n(\Omega) = 36$. Já sabemos que o evento A caracteriza os casos em que as duas faces voltadas para cima são números primos:

$$A = \{(2,2), (2,3), (2,5), (3,2), (3,3), (3,5), (5,2), (5,3), (5,5)\}.$$

Perceba que $n(A) = 9$ e $P(A) = \frac{n(A)}{n(\Omega)} = \frac{9}{36} = \frac{1}{4}$.

Agora o evento B mostra os casos em que a soma das faces é maior que 7:

$$B = \left\{ \begin{matrix} (2,6), (3,5), (3,6), (4,4), (4,5), (4,6), (5,3), (5,4), \\ (5,5), (5,6)(6,2), (6,3), (6,4), (6,5), (6,6) \end{matrix} \right\}.$$

Com efeito de informação, $n(B) = 15$ e, portanto, $P(B) = \frac{n(B)}{n(\Omega)} = \frac{15}{36} = \frac{5}{12}$.

Buscando os elementos que os eventos A e B têm em comum, estabelecemos o conjunto $A \cap B = \{(3,5), (5,3), (5,5)\}$ e identificamos que:

$$P(A \cap B) = \frac{n(A \cap B)}{n(\Omega)} = \frac{3}{36} = \frac{1}{12}.$$

Portanto,

$$P\big(B \,\big|\, A\big) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{1}{12}}{\frac{1}{4}} = \frac{1}{3}.$$

Definição 2.5. Dois eventos A e B de um espaço amostral Ω são ditos independentes entre si se $P(A \cap B) = P(A) \cdot P(B)$. Caso contrário, os eventos são ditos dependentes entre si.

Sejam A e B dois eventos independentes entre si (em nossa abordagem, Ω é finito e equiprovável). Se $P(A) \neq 0$, tem-se que

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

Consequentemente,

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

Analogamente, se $P(B) \neq 0$,

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

Em resumo, dois eventos são independentes entre si quando a ocorrência ou a não ocorrência de um não altera a ocorrência do outro.

Exemplo 2.6. Uma urna contém duas bolas brancas (B) e três vermelhas (V). Suponha que sejam sorteadas duas bolas ao acaso, com reposição. A primeira bola retirada é reposta na urna antes da extração da segunda. Nessas condições, qual a probabilidade de ocorrência de sair uma bola branca na primeira retirada e uma bola vermelha na segunda?

Resolução: Há cinco bolas na urna: B_1, B_2, V_1, V_2, V_3. O espaço amostral das retiradas das duas bolas, com reposição, é constituído por 25 elementos:

$$\Omega = \left\{ \begin{matrix} (B_1, B_1), (B_2, B_1), (V_1, B_1), (V_2, B_1), (V_3, B_1), \\ (B_1, B_2), (B_2, B_2), (V_1, B_2), (V_2, B_2), (V_3, B_2), \\ (B_1, V_1), (B_2, V_1), (V_1, V_1), (V_2, V_1), (V_3, V_1), \\ (B_1, V_2), (B_2, V_2), (V_1, V_2), (V_2, V_2), (V_3, V_2), \\ (B_1, V_3), (B_2, V_3), (V_1, V_3), (V_2, V_3), (V_3, V_3) \end{matrix} \right\}.$$

Considere E o evento de ter uma bola branca na primeira retirada. Observa-se que

$$E = \left\{ \begin{matrix} (B_1, B_1), (B_2, B_1), (B_1, B_2), (B_2, B_2), (B_1, V_1), \\ (B_2, V_1), (B_1, V_2), (B_2, V_2), (B_1, V_3), (B_2, V_3) \end{matrix} \right\}.$$

Aqui, $n(E) = 10$ e $P(E) = \frac{n(E)}{n(\Omega)} = \frac{10}{25} = \frac{2}{5}$.

Na sequência, considere F o evento de ter uma bola vermelha na segunda retirada. Vê-se que

$$F = \left\{ \begin{matrix} (B_1, V_1), (B_2, V_1), (V_1, V_1), (V_2, V_1), (V_3, V_1), \\ (B_1, V_2), (B_2, V_2), (V_1, V_2), (V_2, V_2), (V_3, V_2), \\ (B_1, V_3), (B_2, V_3), (V_1, V_3), (V_2, V_3), (V_3, V_3) \end{matrix} \right\}.$$

Neste caso, $n(F) = 15$ e $P(F) = \frac{n(F)}{n(\Omega)} = \frac{15}{25} = \frac{3}{5}$.

Finalmente, considere $E \cap F$ o evento de ter uma bola branca na primeira retirada e uma bola vermelha na segunda. Identifica-se que

$$E \cap F\ = \{(B_1, V_1), (B_2, V_1), (B_1, V_2), (B_2, V_2), (B_1, V_3), (B_2, V_3)\}.$$

Consequentemente, $n(E \cap F\ ) = 6$ e $P(E \cap F) = \frac{n(E \cap F)}{n(\Omega)} = \frac{6}{25}$ é a resposta que buscamos.

Perceba que $P(E \cap F) = \frac{6}{25} = \frac{2}{5} \cdot \frac{3}{5} = P(E) \cdot P(F)$. E e F são, portanto, exemplos de eventos independentes. A saída de uma bola vermelha na segunda retirada independe de ter saído ou não uma bola branca na primeira. Em particular,

$$P(F \,|\, E) = \frac{6}{10} = \frac{3}{5} = P(F).$$

Observação 2.1. Se, nas mesmas condições do Exemplo 2.6 trabalhássemos com bolas sorteadas sem reposição, perderíamos a independência entre os eventos. Com efeito, há cinco bolas na urna: B_1, B_2, V_1, V_2, V_3. O espaço amostral das retiradas das duas bolas, sem reposição, é constituído por 20 elementos:

$$\Omega = \begin{Bmatrix} (B_2, B_1), (V_1, B_1), (V_2, B_1), (V_3, B_1), \\ (B_1, B_2), (V_1, B_2), (V_2, B_2), (V_3, B_2), \\ (B_1, V_1), (B_2, V_1), (V_2, V_1), (V_3, V_1), \\ (B_1, V_2), (B_2, V_2), (V_1, V_2), (V_3, V_2), \\ (B_1, V_3), (B_2, V_3), (V_1, V_3), (V_2, V_3) \end{Bmatrix}.$$

Considere E o evento de ter uma bola branca na primeira retirada. Observa-se que

$$E = \begin{Bmatrix} (B_2, B_1), (B_1, B_2), (B_1, V_1), (B_2, V_1), \\ (B_1, V_2), (B_2, V_2), (B_1, V_3), (B_2, V_3) \end{Bmatrix}.$$

Nesta conjuntura, $n(E) = 8$ e $P(E) = \frac{n(E)}{n(\Omega)} = \frac{8}{20} = \frac{2}{5}$.

Seja F o evento de ter uma bola vermelha na segunda retirada. Vê-se que

$$F = \begin{Bmatrix} (B_1, V_1), (B_2, V_1), (V_2, V_1), (V_3, V_1), \\ (B_1, V_2), (B_2, V_2), (V_1, V_2), (V_3, V_2), \\ (B_1, V_3), (B_2, V_3), (V_1, V_3), (V_2, V_3) \end{Bmatrix}.$$

Neste caso, $n(F) = 12$ e $P(F) = \frac{n(F)}{n(\Omega)} = \frac{12}{20} = \frac{3}{5}$.

Finalmente, considere $E \cap F$ o evento de ter uma bola branca na primeira retirada e uma bola vermelha na segunda. Identifica-se que

$$E \cap F\ = \{(B_1, V_1), (B_2, V_1), (B_1, V_2), (B_2, V_2), (B_1, V_3), (B_2, V_3)\}.$$

Consequentemente, $n(E \cap F\ ) = 6$ e $P(E \cap F) = \frac{n(E \cap F)}{n(\Omega)} = \frac{6}{20} = \frac{3}{10}$.

Perceba que $P(E \cap F) = \frac{3}{10}$ e $P(E \cap F) \neq P(E) \cdot P(F)$. Isso é suficiente para entender que os eventos E e F não são independentes. Ou seja, sem reposição, os eventos E e F são dependentes.

2.3 VARIÁVEL ALEATÓRIA

Suponhamos um experimento aleatório de lançar uma moeda sobre uma mesa e observar a face que fica voltada para cima. Imaginemos que teremos sucesso se observarmos a face cara e fracasso se observarmos a coroa. Podemos simplificar o procedimento via uma representação numérica quando atribuímos, por exemplo, 1 ao sucesso e 0 ao fracasso.

Em síntese, num experimento aleatório, o resultado de uma operação individual pode ser escrito de uma forma numérica. Esta representação numérica facilita o entendimento e manipulação. O tratamento aqui proposto de representar experimento aleatório a uma descrição numérica está intimamente relacionado ao conceito de variável aleatória.

Nesta seção, apresentamos as definições e exemplos de variáveis aleatórias, além dos conceitos de função de probabilidade e esperança.

Definição 2.6. Sejam ε um experimento aleatório e Ω o espaço amostral associado. A função $X\colon \Omega \to \mathbb{R}$ que associa cada elemento $\omega \in \Omega$ a um número real $X(\omega)$ é denominada uma variável aleatória.

Exemplo 2.7. No lançamento de um dado, observa-se a face voltada para cima. O espaço amostral é o conjunto $\Omega = \{1, 2, 3, 4, 5, 6\}$. Seja X a variável aleatória que associa cada elemento do espaço amostral Ω aos números 0 ou 1, de tal forma que é 0 se a face for um número par e 1 se for um número ímpar. A tabela a seguir associa o domínio à imagem da função.

Tabela 1 – Lançamento de um dado

ω	1	2	3	4	5	6
$X(\omega)$	1	0	1	0	1	0

Fonte: Elaborada pela autora.

Exemplo 2.8. No lançamento de duas moedas, o espaço amostral determina o conjunto constituído por quatro elementos $\Omega = \{(Ca, Ca), (Ca, Co), (Co, Ca), (Co, Co)\}$, de forma que o símbolo Ca representa o aparecimento de cara e Co o aparecimento de coroa.

A variável aleatória X será definida por $X(\omega)$ = número de coroas em $\omega \in \Omega$. Sendo assim, a tabela abaixo associa $\omega \in \Omega$ à sua imagem $X(\omega)$.

Tabela 2 – Lançamento de moedas

ω	(Ca, Ca)	(Ca, Co)	(Co, Ca)	(Co, Co)
$X(\omega)$	0	1	1	2

Fonte: Elaborada pela autora.

Exemplo 2.9. Uma urna contém uma bola azul e uma bola branca. Para efeito de notação, usa-se A para representar a bola azul e B para representar a bola branca. Com reposição, são feitas três retiradas de bola da urna. O espaço amostral das possíveis sequências de cores na retirada das três bolas é descrito pelo conjunto:

$$\Omega = \{(A, A, A), (A, A, B), (A, B, A), (B, A, A), (A, B, B), (B, A, B), (B, B, A), (B, B, B)\}.$$

Se for considerada a variável aleatória X que associa cada elemento $\omega \in \Omega$ à quantidade de cores distintas vistas nas três retiradas, é possível construir a seguinte tabela que relaciona o domínio à imagem:

Tabela 3 – Retiradas de bolas da urna

ω	(A, A, A)	(A, A, B)	(A, B, A)	(B, A, A)	(A, B, B)	(B, A, B)	(B, B, A)	(B, B, B)
$X(\omega)$	1	2	2	2	2	2	2	1

Fonte: Elaborada pela autora.

Observação 2.2. Neste texto, os exemplos usados para descrever variáveis aleatórias caracterizam exemplos discretos, em que há uma quantidade contável de elementos no espaço amostral. Cabe informar que também existem variáveis aleatórias contínuas, mas, tal abordagem não é de interesse à nossa proposta.

2.3.1 Função de Probabilidade

Sejam ε um experimento aleatório qualquer e Ω o espaço amostral associado. Consideremos que $X: \Omega \to \mathbb{R}$ seja uma variável aleatória e que $R_X = X(\Omega)$ caracterize o conjunto imagem de Ω através de X. A notação x_i para $i \in \mathbb{N}^*$ será utilizada para representar os elementos do conjunto R_X, ou seja, $R_X = \{x_1,\ x_2,\ x_3, \dots, x_n, \dots\}$. Em verdade, assumimos o conjunto R_X como contável. Segue abaixo a definição de função de probabilidade:

Definição 2.7. A função $p: R_X \to [0, 1]$ que associa todo resultado elementar x_i à sua probabilidade $p(x_i) = p[X = x_i]$ é chamada de função de probabilidade da variável aleatória X. A coleção de pares $\{(x_i, p(x_i))\}$, por sua vez, é denominada distribuição de probabilidade de X.

É possível estabelecer que:

- $p(x_i) \geq 0$, para todo $i \in \mathbb{N}^*$;
- $\sum_i p(x_i) = 1$.

Convenciona-se que se x não for um elemento de R_X, então, $p(x) = 0$.

Vamos fazer algumas reflexões acerca dos Exemplos 2.7, 2.8 e 2.9.

No Exemplo 2.7 sobre o lançamento de um dado vê-se que $R_X = \{0,1\}$. O $zero$ está associado a três elementos do espaço amostral. E o 1 está associado, por sua vez, a outros três elementos de Ω. É possível construir a seguinte tabela que associa x_i a $p(x_i)$.

Tabela 4 – Lançamento de um dado: associa x_i a $p(x_i)$

x_i	0	1
$p(x_i)$	$\frac{1}{2}$	$\frac{1}{2}$

Fonte: Elaborada pela autora.

No Exemplo 2.8 sobre o lançamento de duas moedas vê-se que $R_X = \{0, 1, 2\}$. Neste caso, $X^{-1}(0) = \{(Ca, Ca)\}$, $X^{-1}(1) = \{(Ca, Co), (Co, Ca)\}$ e $X^{-1}(2) = \{(Co, Co)\}$. Obviamente, a imagem que tem maior probabilidade de aparecer é a de número 1. A tabela a seguir associa x_i a $p(x_i)$.

Tabela 5 – Lançamento de moedas: associa x_i a $p(x_i)$

x_i	0	1	2
$p(x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Fonte: Elaborada pela autora.

Finalmente, no Exemplo 2.9 sobre as bolas retiradas da urna tem-se que $R_X = \{1,2\}$. Baseada na construção, a tabela a seguir associa x_i a $p(x_i)$. Claramente após três retiradas de bolas é mais provável que sejam vistas duas cores distintas que apenas uma.

Tabela 6 – Retiradas de bolas da urna: associa x_i a $p(x_i)$

x_i	1	2
$p(x_i)$	$\frac{1}{4}$	$\frac{3}{4}$

Fonte: Elaborada pela autora.

2.3.1.1 Esperança

Sejam ε um experimento aleatório qualquer e Ω o espaço amostral associado. Consideremos que $X: \Omega \rightarrow \mathbb{R}$ seja uma variável aleatória. A esperança identifica um valor médio dos elementos do conjunto R_X através da função de probabilidade da variável aleatória. Segue abaixo a definição de esperança.

Definição 2.8. A esperança de uma variável aleatória X é determinada por

$$E[X] = \sum_{k=1}^{\infty} x_k \cdot p[X = x_k].$$

No Exemplo 2.7, percebemos que $E[X] = 0{,}5$, pois,

$$E[X] = 0 \cdot p[X = 0] + 1 \cdot p[X = 1] = 0 \cdot 0{,}5 + 1 \cdot 0{,}5 = 0{,}5.$$

Como os valores 0 e 1 aparecem na mesma quantidade de vezes como imagens do espaço amostral, a esperança acaba se tornando uma média aritmética desses dois números.

No Exemplo 2.8, temos que $E[X] = 1$, pois,

$$E[X] = 0 \cdot p[X = 0] + 1 \cdot p[X = 1] + 2 \cdot p[X = 2] = 0 \cdot 0{,}25 + 1 \cdot 0{,}5 + 2 \cdot 0{,}25 = 1.$$

Percebe-se que o valor 1 aparece o dobro de vezes que os valores 0 e 2, que aparecem na mesma quantidade. Logo, a esperança é o próprio 1 – a imagem mais provável de ser obtida.

E, no Exemplo 2.9, percebemos que $E[X] = 1{,}75$, pois,

$$E[X] = 1 \cdot p[X = 1] + 2 \cdot p[X = 2] = 1 \cdot 0{,}25 + 2 \cdot 0{,}75 = 1{,}75.$$

Como o valor 2 aparece mais vezes do que o 1, isso acaba refletindo na esperança, que gera um valor mais próximo do 2 que do 1.

2.4 PROCESSOS ESTOCÁSTICOS

Neste momento, é muito importante abordar a definição de processos estocásticos, para posteriormente compreender a definição de Cadeias de Markov. Portanto, segue abaixo a definição de processos estocásticos, com a inclusão de dois exemplos para ajudar no entendimento desse assunto.

Definição 2.9. Um processo estocástico consiste de um espaço amostral Ω, um conjunto não-vazio T de índices e uma família de aplicações $X_t : \Omega \to \mathbb{R}$, $t \in T$. Definimos formalmente um processo estocástico $\tilde{X} = \{X_t : t \in T\}$ como uma família de variáveis aleatórias indexadas por T, que, em geral, são usadas no estudo de fenômenos que são observados ao longo do tempo.

Os processos estocásticos podem ser classificados de acordo com a cardinalidade de seus conjuntos indexados, usualmente interpretados como o tempo e o espaço de estados – que está definido abaixo.

Definição 2.10. O conjunto $\Gamma = \bigcup_{t \in T} Im(X_t)$ é denominado o espaço de estados do processo estocástico proposto na Definição 2.9.

Observação 2.3. Em nossa abordagem, o conjunto T proposto na Definição 2.9 deve ser considerado enumerável, já que estamos interessados em uma abordagem discreta. Vamos trabalhar somente com parâmetro de tempo discreto, portanto,

$$T = \{0, 1, 2, 3, \ldots\}.$$

Exemplo 2.10. Seja um dado usual, com 6 faces distintas: $1, 2, 3, 4, 5$ e 6. Suponhamos que este dado seja lançado n vezes. A função $X_n: \Omega_n \to \mathbb{R}$ associa os elementos do espaço amostral aos números 1 ou 2. Se o produto das "faces para cima" após n lançamentos for ímpar, a imagem será 1. Se o produto das "faces para cima" após n lançamentos for par, a imagem será 2. Neste exemplo, portanto, espera-se que $\Gamma = \{1,2\}$. Notemos que $\{X_n; n \in \mathbb{N}\}$ pode constituir uma sequência aleatória a depender da escolha de Ω (contemplando uma "espécie" de união de Ω_n com n variando sobre o conjunto dos números naturais).

Entendemos que Ω_n descreve a sequência dos espaços amostrais que listam as faces que ficam expostas "para cima'', lançamento a lançamento.

Em caso de apenas 1 lançamento, o espaço amostral é constituído por 6 elementos:

$$\Omega_1 = \{1, 2, 3, 4, 5, 6\}.$$

Em caso de 2 lançamentos, o espaço amostral é constituído por $6^2 = 36$ elementos:

$$\Omega_2 = \left\{\begin{matrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{matrix}\right\}.$$

Nessa segunda situação, entendemos que $X_2\big((3,2)\big) = 2$, pois, o produto das faces 2 e 3 resulta um número par. Por sua vez, $X_2\big((3,5)\big) = 1$, pois, o produto das faces 3 e 5 resulta um número ímpar.

Em caso de 3 lançamentos, o espaço amostral é constituído por $6^3 = 216$ elementos. Aqui, vemos, por exemplo, $X_3((1,1,1)) = 1$, $X_3((1,2,1)) = 2$ e $X_3((1,2,5)) = 2$. Naturalmente, notamos que o espaço amostral Ω_n é constituído por 6^n elementos.

Exemplo 2.11. Uma urna contém três bolas nas cores azul, branca e preta. Sempre, com reposição, são retiradas da urna n bolas. A variável aleatória $X_n: \Omega_n \to \mathbb{R}$ é a função que associa o espaço amostral ao número de cores distintas de bolas observadas após as n retiradas. Espera-se $\Gamma = \{1,2,3\}$ e $\{X_n; n \in \mathbb{N}\}$ pode constituir uma sequência aleatória a depender da construção de Ω.

Neste caso, entendemos que Ω_n descreve os espaços amostrais que listam as retiradas de bolas da urna, após n retiradas. Sejam A = azul, B = branca e P = preta as três bolas. Em caso de uma única retirada, o espaço amostral é constituído por 3 elementos:

$$\Omega_1 = \{A, B, P\}.$$

Aqui, entendemos que

$$X_1(A) = X_1(B) = X_1(P) = 1.$$

Em caso de 2 retiradas, o espaço amostral é constituído por $3^2 = 9$ elementos:

$$\Omega_2 = \{(A,B),(A,P),(A,A),(B,A),(B,P),(B,B),(P,A),(P,B),(P,P)\}.$$

Vemos que $X_2\big((A,A)\big) = X_2\big((B,B)\big) = X_2\big((P,P)\big) = 1$. Por sua vez, $X_2\big((A,B)\big) = 2$.

Em caso de 3 retiradas, o espaço amostral é constituído por $3^3 = 27$ elementos:

$$\Omega_3 = \left\{\begin{matrix} (A,B,A),(A,B,B),(A,B,P),(A,P,A),(A,P,B),(A,P,P),(A,A,A), \\ (A,A,B),(A,A,P),(B,A,A),(B,A,B),(B,A,P),(B,P,A),(B,P,B), \\ (B,P,P),(B,B,A),(B,B,B),(B,B,P),(P,A,A),(P,A,B),(P,A,P), \\ (P,B,A),(P,B,B),(P,B,P),(P,P,A),(P,P,B),(P,P,P) \end{matrix}\right\}$$

Claramente, $X_3((P,P,P)) = 1$, $X_3((P,B,B)) = 2$ e $X_3((P,B,A)) = 3$.

Percebemos que o espaço amostral Ω_n é constituído por 3^n elementos. Seguindo o raciocínio adotado, identificamos que $X_4((P,P,P,P)) = 1$, $X_5((P,B,B,B,P)) = 2$ e $X_8((P,B,A,P,P,P,B,P)) = 3$.

2.5 CADEIAS DE MARKOV

Inicialmente, propomos um breve histórico sobre a vida de Andrei Markov.

Figura 1 – Andrei Andreyevich Markov (1856-1922)

Fonte: Andrei Markov em [24].

Andrei Andreyevich Markov foi um matemático russo, que nasceu em 14 de junho de 1856 na cidade de Ryazan e faleceu em 20 de julho de 1922 em São Petersburgo.

Markov estudou na Universidade de São Petersburgo, onde veio a atuar como professor. Os seus primeiros trabalhos foram nas áreas de teoria dos números, análise, frações contínuas, limites de integrais, teoria da aproximação e convergência de séries.

Este matemático ficou conhecido por seus estudos sobre as Cadeias de Markov. Os estudos iniciais envolvendo cadeias surgiram em um trabalho onde Markov estudava a probabilidade de ocorrer uma consoante em uma determinada posição de uma palavra qualquer. Para ele, tal probabilidade dependeria apenas da letra anterior ser uma vogal ou outra consoante.

Cadeia de Markov foi o nome dado a um determinado tipo de processo aleatório cuja principal propriedade é a de não ter nenhuma dependência com fatos ocorridos anteriormente, isto é, apenas o estado atual do processo pode influenciar no que ocorrerá posteriormente. Os primeiros resultados sobre o tema foram publicados em 1906.

Atualmente, as Cadeias de Markov são amplamente utilizadas em diversas áreas do conhecimento como física atômica, teoria quântica, biologia, genética, comportamento social, economia e finanças. Neste trabalho, nos capítulos 3 e 4 apresentamos algumas aplicações envolvendo Cadeias de Markov.

Apresentamos, abaixo, a definição formal de Cadeias de Markov.

Definição 2.11. Considere um espaço de estados com um número finito (ou enumerável) de elementos Γ. Um processo estocástico discreto $(X_n)_{n\in\mathbb{N}}$ é uma cadeia (ou processo) de Markov se a probabilidade condicional satisfizer

$$P\left[X_{n+1} = j \,\middle|\, X_0 = x_0, X_1 = x_1, \dots, X_n = i\right] = P\left[X_{n+1} = j \,\middle|\, X_n = i\right],$$

para todo $n \in \mathbb{N}$ e para todos $i, j \in \Gamma$. Lembremos que Γ é o espaço de estados.

Em uma linguagem mais simples, a cadeia de Markov é um processo estocástico caracterizado por seu estado futuro depender apenas do seu estado no presente, não importando aquilo que ocorreu no passado. Dizemos então que um processo de Markov "não tem memória alguma".

2.5.1 Probabilidade de Transição

Na sequência, apresentamos a definição de probabilidade de transição.

Definição 2.12. A probabilidade de transição em um passo n, representada por $p_{ij}(n)$, é definida com a seguinte probabilidade condicional:

$$p_{ij}(n) = P\left[X_{n+1} = j \,\middle|\, X_n = i\right].$$

Em outras palavras, $p_{ij}(n)$ é a probabilidade condicional de estar no estado j no tempo $n + 1$, dado que estava no estado i no tempo anterior n, $\forall\, i, j \in \Gamma$.

Esta probabilidade condicional é denominada uma probabilidade de transição de um passo do processo de Markov. Neste trabalho, estamos interessados nas probabilidades de transição estacionárias.

2.5.1.1 Probabilidade de transição estacionária

Definição 2.13. As probabilidades de transição são ditas estacionárias, se para cada $i, j \in \Gamma$ e para todo natural n, temos

$$p_{ij} = P\left[X_{n+1} = j \,\middle|\, X_n = i\right] = P\left[X_1 = j \,\middle|\, X_0 = i\right].$$

A probabilidade de o sistema estar no estado j em qualquer observação se na observação imediatamente precedente estava no estado i, é representada por p_{ij}. Perceba que as probabilidades de transição estacionárias são probabilidades condicionais que não mudam em relação ao momento em que a observação é feita.

Uma cadeia de Markov que possui apenas probabilidades de transição estacionárias, ou seja, aquelas que não mudam ao longo da variação de n, será chamada de cadeia de Markov estacionária. Nesta dissertação, abordaremos apenas o estudo das cadeias de Markov estacionárias.

Os Exemplos 2.10 e 2.11 configuram cadeias de Markov estacionárias. Suponhamos, que no Exemplo 2.10, até o $(m$-1)-ésimo lançamento do dado, o produto das "faces voltadas para cima" seja um número ímpar. A probabilidade de que no m-ésimo lançamento o produto se torne par independe do número natural m considerado. Uma observação similar pode ser feita no Exemplo 2.11. Consideremos que até a $(m$-1)-ésima retirada de bola da urna tenha sido registrado o aparecimento de duas cores distintas de bolas. A probabilidade de aparecer a terceira cor de bola na m-ésima retirada independe de quão grande seja esse número m.

2.5.2 Matriz de Transição

Considerando todas as probabilidades p_{ij} já citadas anteriormente, podemos determinar uma matriz. Tal matriz será chamada de matriz de transição.

Definição 2.14. Sejam $\{X_n; n \in \mathbb{N}\}$ uma cadeia de Markov estacionária e $\Gamma = \{1, 2,3, \dots, k\}$. Define-se $P = (p_{ij})$ como a matriz de transição de um passo da cadeia.

Para um processo com k estados, a matriz das probabilidades de transição, ou simplesmente, matriz de transição, é estabelecida como:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1k} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2k} \\ \vdots & \vdots & & \ddots & \\ p_{k1} & p_{k2} & p_{k3} & \dots & p_{kk} \end{pmatrix}.$$

A matriz de transição será sempre quadrada. Cada elemento p_{ij} é um número real. Lembrando que se trata de uma probabilidade, $p_{ij} \in [0, 1]$, se $i, j \in \{1, 2, 3, \dots, k\}$.

Proposição 2.2. Seja a matriz de transição $P = (p_{ij})$ de uma Cadeia de Markov estacionária com k estados. Considere o conjunto $\Gamma = \{1, 2, 3, \dots, k\}$ e $i \in \Gamma$. Então,

$$p_{i1} + p_{i2} + p_{i3} + \cdots + p_{ik} = 1.$$

Em outras palavras, a soma dos elementos de uma linha qualquer de uma matriz de transição é sempre igual a 1.

Demonstração.

Sejam $i, j \in \Gamma$. Então, para m no conjunto de índices, temos que

$$p_{ij} = P\big[X_{m+1} = j \,\big|\, X_m = i\big].$$

Desta forma, para cada $i \in \Gamma$,

$$\sum_{j=1}^{k} p_{ij} = \sum_{j=1}^{k} P\big[X_{m+1} = j \,\big|\, X_m = i\big].$$

Para cada $j \in \Gamma$, defina $B_j = [X_{m+1} = j]$. Note que $B_l \cap B_j = \emptyset, \quad \forall\, l, j \in \Gamma, l \neq j$. Além disso, $\Omega = \bigcup_{j=1}^{k} B_j$. Definindo $B = [X_m = i]$, temos que

$$\sum_{j=1}^{k} p_{ij} = \sum_{j=1}^{k} P\big[X_{m+1} = j \,\big|\, X_m = i\big] = \sum_{j=1}^{k} \frac{P\big[B_j \cap B\big]}{P[B]}.$$

Isto é,

$$\sum_{j=1}^{k} p_{ij} = \frac{1}{P[B]} P\left[\bigcup_{j=1}^{k} B_j \cap B\right] = \frac{1}{P[B]} P\left[B \bigcap \left(\bigcup_{j=1}^{k} B_j\right)\right] = \frac{1}{P[B]} \cdot P[B] = 1.$$

■

A seguir, vejamos dois exemplos de construção da matriz de transição.

Exemplo 2.12. Vejamos agora a matriz de transição de um passo da cadeia referente à situação proposta no Exemplo 2.10. Obtemos

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}.$$

De fato, observamos que

$$p_{11} = P\big[X_{n+1} = 1 \,\big|\, X_n = 1\big] = \frac{1}{2}.$$

Isto é, a probabilidade do sistema estar no estado 1 em qualquer observação, se na observação imediatamente anterior já estava no estado 1, é igual a $\frac{1}{2}$. A probabilidade condicional de em qualquer lançamento de um dado o produto de todas as faces voltadas para cima ser ímpar, dado que no lançamento imediatamente anterior o produto das faces voltadas para cima já era ímpar, é igual a $\frac{1}{2}$. Suponhamos que temos um número ímpar. Sendo assim, só podemos escolher números ímpares para gerar, via produto, um novo número ímpar. Das 6 faces existentes no dado, se quisermos um produto que resulte ímpar só podemos escolher 3 delas, as de números 1, 3 ou 5.

Por sua vez,

$$p_{12} = P\big[X_{n+1} = 2 \,\big|\, X_n = 1\big] = \tfrac{1}{2}.$$

Ou seja, a probabilidade de o sistema estar no estado 2 em qualquer observação, se na observação imediatamente anterior estava no estado 1, é igual a $\frac{1}{2}$. A probabilidade condicional de em qualquer lançamento de um dado o produto de todas as faces voltadas para cima ser par, dado que no lançamento imediatamente anterior o produto das faces voltadas para cima era ímpar, é igual a $\frac{1}{2}$. Perceba que se temos um número ímpar, só podemos escolher números pares para gerar, via produto, um número par. Das 6 faces existentes no dado comum, portanto, apenas 3 interessam, aquelas com os números 2, 4 ou 6.

Segue que

$$p_{21} = P\left[X_{n+1} = 1 \,\middle|\, X_n = 2\right] = 0.$$

Agora, a probabilidade de o sistema estar no estado 1 em qualquer observação, se na observação imediatamente precedente estava no estado 2, é igual a 0. A probabilidade condicional de em qualquer lançamento de um dado o produto de todas as faces voltadas para cima ser ímpar, dado que no lançamento imediatamente anterior o produto das faces voltadas para cima era par, é igual a 0. Perceba que não existe um número natural que multiplicado por um número par, resulte em um número ímpar.

Finalmente,

$$p_{22} = P\left[X_{n+1} = 2 \,\middle|\, X_n = 2\right] = 1.$$

A probabilidade de o sistema estar no estado 2 em qualquer observação, se na observação imediatamente precedente estava no estado 2, é 1. Em outras palavras, a probabilidade condicional de em qualquer lançamento de um dado o produto de todas as faces voltadas para cima ser par, dado que no lançamento imediatamente anterior o produto das faces voltadas para cima já era par, é igual 1, ou seja, 100%. Perceba que qualquer número natural multiplicado por um número par gera um número par. Logo, se já foi identificado um produto que resulte em número par, no próximo lançamento pode aparecer qualquer uma das 6 faces que teremos novamente o produto de todas as faces voltadas para cima par.

Exemplo 2.13. No Exemplo 2.11, é possível determinar a seguinte configuração:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 1 \end{pmatrix}.$$

Analisando algumas probabilidades de transição, temos:

$$p_{11} = P\left[X_{n+1} = 1 \,\middle|\, X_n = 1\right] = \frac{1}{3}.$$

Aqui, tratamos a probabilidade de o sistema estar no estado 1 em qualquer observação se na observação imediatamente anterior já estava no estado 1. Neste caso, a probabilidade de transição de 1 cor distinta para continuar com 1 cor distinta é de $\frac{1}{3}$. Perceba que se temos 3 bolas com cores diferentes, a chance de retirar a mesma cor de bola, é 1 de 3, ou seja, a probabilidade é $\frac{1}{3}$.

Por sua vez,

$$p_{12} = P[X_{n+1} = 2 \,|\, X_n = 1] = \frac{2}{3}.$$

Agora, tratamos a probabilidade de o sistema estar no estado 2 em qualquer observação se na observação imediatamente anterior estava no estado 1. Neste caso, a probabilidade de transição de 1 cor distinta para 2 cores distintas é de $\frac{2}{3}$. Note que como temos 3 bolas com cores diferentes, sendo que só foi retirada uma cor de bola, a chance de retirar uma bola com uma cor distinta é 2 em 3, portanto, a probabilidade é $\frac{2}{3}$.

Finalmente,

$$p_{13} = P[X_{n+1} = 3 \,|\, X_n = 1] = 0.$$

Para p_{13}, tratamos a probabilidade de o sistema estar no estado 3 em qualquer observação se na observação imediatamente anterior estava no estado 1. Neste caso, a probabilidade de transição de 1 cor distinta para 3 cores distintas é 0. Perceba que na retirada de uma bola não é possível passar de 1 cor para 3 cores distintas – podemos ter 1 ou 2 cores apenas. Com raciocínio análogo, encontramos as outras probabilidades de transição dessa matriz.

2.5.3 Vetor Estado

Definição 2.15. O vetor estado de uma observação de uma cadeia de Markov estacionária, com k estados, é um vetor linha x cujo i-ésimo componente x_i é a probabilidade do sistema estar, naquela observação, no i-ésimo estado, dado que $\Gamma = \{1, 2, 3, \ldots, k\}$ e $i \in \Gamma$. Segue abaixo o vetor linha x:

$$x = (x_1 \quad x_2 \quad x_3 \quad \cdots \quad x_k).$$

Observação 2.4. Note que o vetor linha x é uma matriz linha com k entradas e $\sum_{i=1}^{k} x_i = 1$.

Sabendo o vetor estado $x^{(0)}$ de uma cadeia de Markov, é possível determinar os vetores estado $x^{(1)}$, $x^{(2)}$, $x^{(3)}, \dots, x^{(n)}, \dots$ nas observações seguintes, aplicando o teorema abaixo.

Teorema 2.1. Se P for a matriz de transição de uma cadeia de Markov estacionária e $x^{(n)}$ for o vetor estado da n-ésima observação, então,

$$x^{(n+1)} = x^{(n)} \cdot P.$$

Demonstração: Seja P uma matriz de transição de uma cadeia de Markov de k estados, onde cada p_{ij} é a probabilidade de, estando em i, passar para o estado j após uma unidade de tempo (um passo),

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1k} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2k} \\ \vdots & \vdots & & \ddots & \\ p_{k1} & p_{k2} & p_{k3} & \dots & p_{kk} \end{pmatrix}.$$

E seja $x^{(n)}$ o vetor estado, da mesma cadeia de Markov, em uma observação qualquer $t = n$,

$$x^{(n)} = \begin{pmatrix} x_1{}^{(n)} & x_2{}^{(n)} & x_3{}^{(n)} & \cdots & x_k{}^{(n)} \end{pmatrix}.$$

Aqui,

$x_1{}^{(n)}$ é a probabilidade de o sistema estar no estado 1 no instante $t = n$;

$x_2{}^{(n)}$ é a probabilidade de o sistema estar no estado 2 no instante $t = n$;

$x_3{}^{(n)}$ é a probabilidade de o sistema estar no estado 3 no instante $t = n$;

$\vdots$

$x_k{}^{(n)}$ é a probabilidade de o sistema estar no estado k no instante $t = n$.

Sendo que o vetor $x^{(n+1)}$ o vetor estado no instante $t = n + 1$, temos

$$x^{(n+1)} = \begin{pmatrix} x_1{}^{(n+1)} & x_2{}^{(n+1)} & x_3{}^{(n+1)} & \cdots & x_k{}^{(n+1)} \end{pmatrix}.$$

Levando em consideração as probabilidades condicionais, identificamos que

$$x_1{}^{(n+1)} = P\left[X_{n+1} = 1 \,\middle|\, X_n = 1\right] \cdot P[X_n = 1] + P\left[X_{n+1} = 1 \,\middle|\, X_n = 2\right] \cdot P[X_n = 2] + \cdots +$$
$$+ P\left[X_{n+1} = 1 \,\middle|\, X_n = k\right] \cdot P[X_n = k];$$

$$x_2{}^{(n+1)} = P\left[X_{n+1} = 2 \,\middle|\, X_n = 1\right] \cdot P[X_n = 1] + P\left[X_{n+1} = 2 \,\middle|\, X_n = 2\right] \cdot P[X_n = 2] + \cdots +$$
$$+ P\left[X_{n+1} = 2 \,\middle|\, X_n = k\right] \cdot P[X_n = k];$$

$\vdots$

$$x_k{}^{(n+1)} = P\big[X_{n+1} = k \,\big|\, X_n = 1\big] \cdot P[X_n = 1] + P\big[X_{n+1} = k \,\big|\, X_n = 2\big] \cdot P[X_n = 2] + \cdots +$$
$$+ P\big[X_{n+1} = k \,\big|\, X_n = k\big] \cdot P[X_n = k].$$

Em outros termos,

$$x_1{}^{(n+1)} = p_{11} \cdot x_1{}^{(n)} + p_{21} \cdot x_2{}^{(n)} + p_{31} \cdot x_3{}^{(n)} + \cdots + p_{k1} \cdot x_k{}^{(n)}$$
$$x_2{}^{(n+1)} = p_{12} \cdot x_1{}^{(n)} + p_{22} \cdot x_2{}^{(n)} + p_{32} \cdot x_3{}^{(n)} + \cdots + p_{k2} \cdot x_k{}^{(n)}$$
$$\vdots$$
$$x_k{}^{(n+1)} = p_{1k} \cdot x_1{}^{(n)} + p_{2k} \cdot x_2{}^{(n)} + p_{3k} \cdot x_3{}^{(n)} + \cdots + p_{kk} \cdot x_k{}^{(n)}.$$

Agora partindo de $x^{(n+1)}$ e desenvolvendo a expressão, temos:

$$x^{(n+1)} = \begin{pmatrix} x_1{}^{(n+1)} & x_2{}^{(n+1)} & x_3{}^{(n+1)} & \cdots & x_k{}^{(n+1)} \end{pmatrix}$$

$$= \begin{pmatrix} p_{11} \cdot x_1{}^{(n)} + \cdots + p_{k1} \cdot x_k{}^{(n)} & p_{12} \cdot x_1{}^{(n)} + \cdots + p_{k2} \cdot x_k{}^{(n)} & \cdots & p_{1k} \cdot x_1{}^{(n)} + \cdots + p_{kk} \cdot x_k{}^{(n)} \end{pmatrix}$$

$$= \begin{pmatrix} x_1{}^{(n)} & x_2{}^{(n)} & x_3{}^{(n)} & \cdots & x_k{}^{(n)} \end{pmatrix} \cdot \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1k} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2k} \\ \vdots & \vdots & & \ddots & \\ p_{k1} & p_{k2} & p_{k3} & \cdots & p_{kk} \end{pmatrix}$$

$$= x^{(n)} \cdot P.$$

Logo, $x^{(n+1)} = x^{(n)} \cdot P$, como queríamos demonstrar. ■

Devido ao Teorema 2.1, temos:

$$x^{(1)} = x^{(0)} \cdot P \tag{2.2}$$
$$x^{(2)} = x^{(1)} \cdot P \tag{2.3}$$
$$x^{(3)} = x^{(2)} \cdot P. \tag{2.4}$$

Substituindo (2.2) em (2.3), temos que:

$$x^{(2)} = x^{(1)} \cdot P = x^{(0)} \cdot P \cdot P = x^{(0)} \cdot P^2$$
$$x^{(2)} = x^{(0)} \cdot P^2. \tag{2.5}$$

Substituindo (2.5) em (2.4), temos:

$$x^{(3)} = x^{(2)} \cdot P = x^{(0)} \cdot P^2 \cdot P = x^{(0)} \cdot P^3$$

$x^{(3)} = x^{(0)} \cdot P^3$.

Continuando o mesmo raciocínio obtemos, por indução, que para cada $n \in \mathbb{N}$,

$$x^{(n)} = x^{(0)} \cdot P^n. \tag{2.6}$$

Analogamente, se $n \in \mathbb{N}$,

$$x^{(n+1)} = x^{(0)} \cdot P^{n+1}.$$

Portanto, a expressão $x^{(n+1)} = x^{(0)} \cdot P^{n+1}$ é equivalente a $x^{(n+1)} = x^{(n)} \cdot P$ apresentada no Teorema 2.1. A potência P^n é importante, por exemplo, para que possamos classificar uma cadeia de Markov em Regular.

Na sequência, apresentamos duas classificações de Cadeias de Markov estacionárias. Tais classificações são utilizadas para prever o comportamento assintótico de $x^{(n)}$.

2.5.4 Cadeias de Markov Regulares

Definição 2.16. Uma matriz de transição P é dita regular quando para alguma n −ésima potência, P^n apresentar todas as probabilidades de transição (elementos da matriz) estritamente positivas. Uma cadeia de Markov estacionária que é regida por uma matriz de transição regular é chamada de Cadeia de Markov Regular.

Um estudo mais avançado envolvendo álgebra linear permite enunciar o seguinte resultado:

Teorema 2.2. Se P for uma matriz de transição de uma cadeia de Markov regular, então:

I. Existe um único vetor estado q tal que $q \cdot P = q$;
II. Para qualquer vetor estado inicial $x^{(0)}$, a sequência de vetores estado $x^{(0)}$, $x^{(0)} \cdot P$, $x^{(0)} \cdot P^2$, $x^{(0)} \cdot P^3, \dots, x^{(0)} \cdot P^n$ tende a q como um limite, ou seja, $x^{(0)} \cdot P^n \rightarrow q$ quando $n \rightarrow \infty$. O vetor q é chamado de vetor de estado estacionário.

Demonstração.

Ver teoremas da **Seção 10.5** da Ref. [2].

■

Qualquer Cadeia de Markov Regular possui um vetor estado fixo q tal que, para qualquer escolha de $x^{(0)}$, o vetor $P^n x^{(0)}$ converge para q quando n tende ao infinito. Este sistema sempre acaba convergindo, portanto, para um vetor estado q fixo. No decorrer do Capítulo 3, apresentamos um exemplo de cadeia de Markov regular.

2.5.5 Cadeias de Markov Absorventes

Definição 2.17. Um estado i é chamado de estado absorvente se $p_{ii} = 1$. Um estado que não for absorvente é chamado de estado de transição. Dizemos que uma cadeia de Markov é absorvente se nela existir ao menos um estado absorvente, além disso, deve ser possível, partindo de qualquer estado, atingir um estado absorvente, independentemente do número de passos necessários para isso.

Uma das propriedades mais relevantes das cadeias de Markov absorventes é que, uma vez atingido um estado absorvente, não é mais possível sair dele. A referência [22] apresenta uma abordagem detalhada sobre o tema cadeias de Markov Absorventes.

O resultado abaixo garante, por sua vez, que cadeias de Markov absorventes não são cadeias de Markov regulares.

Teorema 2.3. Se uma matriz de transição P, de ordem n ≥ 2, tiver 1 na diagonal principal, então, P não é regular. Podemos dizer, em outras palavras que, se uma cadeia de Markov possuir pelo menos um estado absorvente, então, a cadeia não é regular.

Demonstração.

Ver Ref. [8], Teorema 5.7.1.

■

Os Exemplos 2.10 e 2.11 caracterizam Cadeias de Markov Absorventes.

Veja no Exemplo 2.10 (e no Exemplo 2.12) que $p_{22} = 1$. Neste caso, o estado 2 é chamado de estado absorvente. Quando o estado 2 for atingido é impossível sair dele. Note que, nos lançamentos de um dado, no momento em que aparecer uma face par voltada para cima, o produto de todas as faces será par, e após tal observação, não importa mais o número que vai aparecer nos lançamentos do dado: a paridade do produto estará, portanto, determinada.

No Exemplo 2.11 (e Exemplo 2.13), temos que $p_{33} = 1$, então, o estado 3 é absorvente. Aliás, 3 é o único estado absorvente da cadeia. Perceba que na retirada de bolas de dentro da urna, temos 3 bolas com cores distintas. Atingir o estado 3 é retirar da urna as 3 bolas de cores diferentes. Quando atingir as 3 cores distintas não será mais possível sair deste estado. Por mais retiradas que sejam feitas, as 3 cores já foram obtidas e não será permitido retroceder para uma quantidade inferior de cores observadas em retiradas.

Mais exemplos de cadeias de Markov absorventes são trabalhados e detalhados nos Capítulos 3 e 4.

3 APLICAÇÕES DE CADEIAS DE MARKOV

Neste capítulo, vamos abordar duas situações-problema que podem ser interpretadas via utilização de Cadeias de Markov. Esses exemplos embasarão uma proposta de aplicação dos assuntos Variáveis Aleatórias e Cadeias de Markov no Ensino Médio, a ser retomada no Capítulo 5. Aqui, trabalha-se detalhadamente inúmeros aspectos, de modo a se adquirir o domínio esperado por quem eventualmente se disponha a transmitir tais conhecimentos em nível escolar.

As definições e propriedades vistas anteriormente são de fundamental importância para a compreensão dos dois exemplos. Inicialmente, vamos analisar o exemplo do lançamento de dados associado a números primos. Depois, um exemplo de Cadeia de Markov regular associado à migração em regiões da cidade de Blumenau.

3.1 EXEMPLO I – LANÇAMENTO DE DADOS E NÚMEROS PRIMOS

Suponhamos que um dado hexaédrico usual, com 6 faces distintas $1, 2, 3, 4, 5$ e 6, seja lançado n vezes sobre uma mesa e que se observe a face que fica exposta para cima em cada lançamento.

Consideremos que Ω_n descreva a sequência dos espaços amostrais que listam as faces que ficam expostas "para cima'', lançamento a lançamento.

Em caso de apenas 1 lançamento, o espaço amostral é constituído por 6 elementos:

$$\Omega_1 = \{1, 2, 3, 4, 5, 6\}.$$

Em caso de 2 lançamentos, o espaço amostral é constituído por $6^2 = 36$ elementos:

$$\Omega_2 = \left\{\begin{matrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{matrix}\right\} = \Omega_1 \times \Omega_1.$$

Em caso de 3 lançamentos, o espaço amostral é constituído por $6^3 = 216$ elementos:

$$\Omega_3 = \left\{ \begin{array}{l}
(1,1,1),(1,1,2),(1,1,3),(1,1,4),(1,1,5),(1,1,6),(1,2,1),(1,2,2),(1,2,3),\\
(1,2,4),(1,2,5),(1,2,6),(1,3,1),(1,3,2),(1,3,3),(1,3,4),(1,3,5),(1,3,6),\\
(1,4,1),(1,4,2),(1,4,3),(1,4,4),(1,4,5),(1,4,6),(1,5,1),(1,5,2),(1,5,3),\\
(1,5,4),(1,5,5),(1,5,6),(1,6,1),(1,6,2),(1,6,3),(1,6,4),(1,6,5),(1,6,6),\\
(2,1,1),(2,1,2),(2,1,3),(2,1,4),(2,1,5),(2,1,6),(2,2,1),(2,2,2),(2,2,3),\\
(2,2,4),(2,2,5),(2,2,6),(2,3,1),(2,3,2),(2,3,3),(2,3,4),(2,3,5),(2,3,6),\\
(2,4,1),(2,4,2),(2,4,3),(2,4,4),(2,4,5),(2,4,6),(2,5,1),(2,5,2),(2,5,3),\\
(2,5,4),(2,5,5),(2,5,6),(2,6,1),(2,6,2),(2,6,3),(2,6,4),(2,6,5),(2,6,6),\\
(3,1,1),(3,1,2),(3,1,3),(3,1,4),(3,1,5),(3,1,6),(3,2,1),(3,2,2),(3,2,3),\\
(3,2,4),(3,2,5),(3,2,6),(3,3,1),(3,3,2),(3,3,3),(3,3,4),(3,3,5),(3,3,6),\\
(3,4,1),(3,4,2),(3,4,3),(3,4,4),(3,4,5),(3,4,6),(3,5,1),(3,5,2),(3,5,3),\\
(3,5,4),(3,5,5),(3,5,6),(3,6,1),(3,6,2),(3,6,3),(3,6,4),(3,6,5),(3,6,6),\\
(4,1,1),(4,1,2),(4,1,3),(4,1,4),(4,1,5),(4,1,6),(4,2,1),(4,2,2),(4,2,3),\\
(4,2,4),(4,2,5),(4,2,6),(4,3,1),(4,3,2),(4,3,3),(4,3,4),(4,3,5),(4,3,6),\\
(4,4,1),(4,4,2),(4,4,3),(4,4,4),(4,4,5),(4,4,6),(4,5,1),(4,5,2),(4,5,3),\\
(4,5,4),(4,5,5),(4,5,6),(4,6,1),(4,6,2),(4,6,3),(4,6,4),(4,6,5),(4,6,6),\\
(5,1,1),(5,1,2),(5,1,3),(5,1,4),(5,1,5),(5,1,6),(5,2,1),(5,2,2),(5,2,3),\\
(5,2,4),(5,2,5),(5,2,6),(5,3,1),(5,3,2),(5,3,3),(5,3,4),(5,3,5),(5,3,6),\\
(5,4,1),(5,4,2),(5,4,3),(5,4,4),(5,4,5),(5,4,6),(5,5,1),(5,5,2),(5,5,3),\\
(5,5,4),(5,5,5),(5,5,6),(5,6,1),(5,6,2),(5,6,3),(5,6,4),(5,6,5),(5,6,6),\\
(6,1,1),(6,1,2),(6,1,3),(6,1,4),(6,1,5),(6,1,6),(6,2,1),(6,2,2),(6,2,3),\\
(6,2,4),(6,2,5),(6,2,6),(6,3,1),(6,3,2),(6,3,3),(6,3,4),(6,3,5),(6,3,6),\\
(6,4,1),(6,4,2),(6,4,3),(6,4,4),(6,4,5),(6,4,6),(6,5,1),(6,5,2),(6,5,3),\\
(6,5,4),(6,5,5),(6,5,6),(6,6,1),(6,6,2),(6,6,3),(6,6,4),(6,6,5),(6,6,6)
\end{array} \right\}$$

Nota-se que o espaço amostral Ω_n é constituído por 6^n elementos.

Define-se a variável aleatória $X_n\colon \Omega_n \to \mathbb{R}$, a função que associa o espaço amostral à quantidade de números primos distintos nas "faces para cima" após n lançamentos.

Abaixo, analisamos resultados de probabilidade, caso a caso.

Seja $n = 1$, ao lançar o dado uma única vez, o número de elementos do espaço amostral é $n(\Omega_1) = 6$. Neste lançamento pode aparecer 0 ou 1 número primo. A variável aleatória $X_1\colon \Omega_1 \to \mathbb{R}$ é, portanto, uma função tal que $Im(X_1) = \{0,1\}$.

Neste caso, identifica-se que $X_1^{-1}(0) = \{1,4,6\}$. Em outras palavras, são três os números não primos possíveis de aparecer. Por sua vez, $X_1^{-1}(1) = \{2,3,5\}$, o que mostra que há três jogadas que podem resultar em número primo.

Com isso, as probabilidades de aparecer um número não primo e de aparecer um número primo, após um lançamento, são dadas, respectivamente, por

$$P[X_1 = 0] = \frac{n\left(X_1^{-1}(0)\right)}{n(\Omega_1)} = \frac{3}{6} = \frac{1}{2} \text{ e } P[X_1 = 1] = \frac{n\left(X_1^{-1}(1)\right)}{n(\Omega_1)} = \frac{3}{6} = \frac{1}{2}.$$

Essas probabilidades mostram que ao lançar o dado uma única vez, a chance de aparecer um número primo é 50% e de não aparecer um número primo também é 50%.

Ao lançar o dado por duas vezes, o número de elementos do espaço amostral, como já registrado, é $n(\Omega_2) = 6^2 = 36$. Nesta situação, é possível que apareçam 0, 1 ou 2 números primos distintos nas faces voltadas para cima. Aqui, a variável aleatória $X_2 \colon \Omega_2 \to \mathbb{R}$ satisfaz $Im(X_2) = \{0, 1, 2\}$. Sendo $n = 2$, identifica-se que

$$X_2^{-1}(0) = \{(1,1),(1,4),(1,6),(4,1),(4,4),(4,6),(6,6),(6,1),(6,4)\},$$

$$X_2^{-1}(1) = \begin{Bmatrix} (1,2),(1,3),(1,5),(2,1),(2,2),(2,4),(2,6), \\ (3,1),(3,3),(3,4),(3,6),(4,2),(4,3),(4,5), \\ (5,1),(5,4),(5,5),(5,6),(6,2),(6,3),(6,5) \end{Bmatrix}$$

e

$$X_2^{-1}(2) = \{(2,3),(2,5),(3,2),(3,5),(5,2),(5,3)\}.$$

Notamos que há 9, 21 e 6 elementos em cada imagem inversa, respectivamente. Então, após os dois lançamentos, as probabilidades de aparecer 0, 1 ou 2 números primos distintos são:

$$P[X_2 = 0] = \frac{n\left(X_2^{-1}(0)\right)}{n(\Omega_2)} = \frac{9}{36} = \frac{1}{4},\ P[X_2 = 1] = \frac{n\left(X_2^{-1}(1)\right)}{n(\Omega_2)} = \frac{21}{36} = \frac{7}{12} \text{ e}$$

$$P[X_2 = 2] = \frac{n\left(X_2^{-1}(2)\right)}{n(\Omega_2)} = \frac{6}{36} = \frac{1}{6}.$$

Em outras palavras, após dois lançamentos de um dado, a probabilidade de não aparecer número primo é $\frac{1}{4}$. A chance de aparecer apenas 1 número primo distinto é $\frac{7}{12}$. A probabilidade das duas faces serem números primos distintos é $\frac{1}{6}$.

Seguindo a mesma proposta, vamos tratar a situação em que são efetuados três lançamentos do dado. De forma análoga, tem-se que $X_3 \colon \Omega_3 \to \mathbb{R}$ satisfaz à igualdade $Im(X_3) = \{0, 1, 2, 3\}$, pois, para 3 lançamentos de um dado é possível que apareçam 0, 1, 2 ou 3 números primos distintos nas faces voltadas para cima. Estabelece-se que:

$$X_3^{-1}(0) = \begin{Bmatrix} (1,1,1),(1,1,4),(1,1,6),(1,4,1),(1,4,4),(1,4,6),(1,6,1), \\ (1,6,4),(1,6,6),(4,1,1),(4,1,4),(4,1,6),(4,4,1),(4,4,4), \\ (4,4,6),(4,6,1),(4,6,4),(4,6,6),(6,1,1),(6,1,4),(6,1,6), \\ (6,4,1),(6,4,4),(6,4,6),(6,6,1),(6,6,4),(6,6,6) \end{Bmatrix},$$

$$X_3^{-1}(1) = \left\{\begin{array}{c} (1,1,2\ ),(1,1,3),(1,1,5),(1,2,1),(1,2,2),(1,2,4),(1,2,6),(1,3,1), \\ (1,3,3),(1,3,4),(1,3,6),(1,4,2),(1,4,3),(1,4,5),(1,5,1),(1,5,4), \\ (1,5,5),(1,5,6),(1,6,2),(1,6,3),(1,6,5),(2,1,1),(2,1,2),(2,1,4), \\ (2,1,6),(2,2,1),(2,2,2),(2,2,4),(2,2,6),(2,4,1),(2,4,2),(2,4,4), \\ (2,4,6),(2,6,1),(2,6,2),(2,6,4),(2,6,6),(3,1,1),(3,1,3),(3,1,4), \\ (3,1,6),(3,3,1),(3,3,3),(3,3,4),(3,3,6),(3,4,1),(3,4,3),(3,4,4), \\ (3,4,6),(3,6,1),(3,6,3),(3,6,4),(3,6,6),(4,1,2),(4,1,3),(4,1,5), \\ (4,2,1),(4,2,2),(4,2,4),(4,2,6),(4,3,1),(4,3,3),(4,3,4),(4,3,6), \\ (4,4,2),(4,4,3),(4,4,5),(4,5,1),(4,5,4),(4,5,5),(4,5,6),(4,6,2), \\ (4,6,3),(4,6,5),(5,1,1),(5,1,4),(5,1,5),(5,1,6),(5,4,1),(5,4,4), \\ (5,4,5),(5,4,6),(5,5,1),(5,5,4),(5,5,5),(5,5,6),(5,6,1),(5,6,4), \\ (5,6,5),(5,6,6),(6,1,2),(6,1,3),(6,1,5),(6,2,1),(6,2,2),(6,2,4), \\ (6,2,6),(6,3,1),(6,3,3),(6,3,4),(6,3,6),(6,4,2),(6,4,3),(6,4,5), \\ (6,5,1),(6,5,4),(6,5,5),(6,5,6),(6,6,2),(6,6,3),(6,6,5) \end{array}\right\},$$

$$X_3^{-1}(2) = \left\{\begin{array}{c} (1,2,3),(1,2,5),(1,3,2),(1,3,5),(1,5,2),(1,5,3),(2,1,3),(2,1,5), \\ (2,2,3),(2,2,5),(2,3,1),(2,3,2),(2,3,3),(2,3,4),(2,3,6),(2,4,3), \\ (2,4,5),(2,5,1),(2,5,2),(2,5,4),(2,5,5),(2,5,6),(2,6,3),(2,6,5), \\ (3,1,2),(3,1,5),(3,2,1),(3,2,2),(3,2,3),(3,2,4),(3,2,6),(3,3,2), \\ (3,3,5),(3,4,2),(3,4,5),(3,5,1),(3,5,3),(3,5,4),(3,5,5),(3,5,6), \\ (3,6,2),(3,6,5),(4,2,3),(4,2,5),(4,3,2),(4,3,5),(4,5,2),(4,5,3), \\ (5,1,2),(5,1,3),(5,2,1),(5,2,2),(5,2,4),(5,2,5),(5,2,6),(5,3,1), \\ (5,3,3),(5,3,4),(5,3,5),(5,3,6),(5,4,2),(5,4,3),(5,5,2),(5,5,3), \\ (5,6,2),(5,6,3),(6,2,3),(6,2,5),(6,3,2),(6,3,5),(6,5,2),(6,5,3) \end{array}\right\}$$

e

$$X_3^{-1}(3) = \{(2,3,5),(2,5,3),(3,2,5),(3,5,2),(5,2,3),(5,3,2)\}.$$

Note que o número de elementos desses conjuntos é 27, 111, 72 e 6, respectivamente. Já registramos que, para o caso em que $n = 3$, o número de elementos do espaço amostral é $n(\Omega_3) = 6^3 = 216$. As probabilidades de aparecerem 0, 1, 2 ou 3 números primos distintos, após três lançamentos do dado, estão listadas abaixo:

$$P[X_3 = 0] = \frac{n\left(X_3^{-1}(0)\right)}{n(\Omega_3)} = \frac{27}{216} = \frac{1}{8}, P[X_3 = 1] = \frac{n\left(X_3^{-1}(1)\right)}{n(\Omega_3)} = \frac{111}{216} = \frac{37}{72},$$

$$P[X_3 = 2] = \frac{n\left(X_3^{-1}(2)\right)}{n(\Omega_3)} \frac{72}{216} = \frac{1}{3} \text{ e } P[X_3 = 3] = \frac{n\left(X_3^{-1}(3)\right)}{n(\Omega_3)} = \frac{6}{216} = \frac{1}{36}.$$

Ao lançar 3 vezes um dado, a probabilidade de não aparecer número primo nos 3 lançamentos é $\frac{1}{8}$. A probabilidade de aparecer apenas 1 número primo distinto é $\frac{37}{72}$, de aparecerem 2 números primos distintos é $\frac{1}{3}$ e das 3 faces serem números primos distintos é $\frac{1}{36}$.

Uma pergunta natural a ser feita é: qual a probabilidade de aparecer $0, 1, 2$ ou 3 números primos distintos após n lançamentos de um dado, para $n > 3$? Perceba que conforme aumentamos o valor de n, fica mais trabalhoso e inviável de se encontrar o espaço amostral, os eventos e calcular as probabilidades. Pode-se dizer que manualmente é um processo que se torna exaustivo já que aparecem 6^n elementos só no que diz respeito ao espaço amostral. Apesar de possivelmente existirem outras estratégias associadas à análise combinatória, vamos recorrer às Cadeias de Markov para buscar respostas à pergunta levantada.

Daqui em diante, considere $X_n: \Omega \to \mathbb{R}$, onde Ω é o espaço amostral (comum), para todo $n \in \mathbb{N}$. Omitimos a representação de Ω – que segue a estratégia dos Exemplos 2.10 e 2.11. $\{X_n\}_{n\in\mathbb{N}} = \{X_n: n \in \mathbb{N}\}$ é uma sequência aleatória, ou seja, uma família de variáveis aleatórias tal que

$$\Gamma = \bigcup_{n\in\mathbb{N}} Im(X_n) = \{0, 1, 2, 3\}.$$

Note que os elementos do conjunto Γ (espaço de estados) estão relacionados à quantidade de números primos distintos que podem aparecer após n lançamentos de um dado.

Esta situação configura uma Cadeia de Markov estacionária, pois considere que os índices $i, j \in \Gamma = \{0, 1, 2, 3\}$ e que até o m-ésimo lançamento do dado, tenhamos i números primos distintos observados. A probabilidade condicional de que no $(m + 1)$-ésimo lançamento vejamos j números primos distintos independe do número natural m considerado como passo.

Sejam $i, j \in \Gamma$ fixados. Vamos determinar as probabilidades de o sistema estar no estado j em qualquer observação dado que na observação imediatamente precedente estava no estado i. Representamos essas probabilidades de transição por p_{ij}, em que

$$p_{ij} = P\big[X_{m+1} = j \,\big|\, X_m = i\big].$$

Devemos calcular, ao todo, dezesseis probabilidades de transição para que possamos tratar a matriz de transição. Apresentamos, abaixo, a reflexão sobre a dedução de algumas dessas probabilidades de transição.

Inicialmente, veja que:

$$p_{00} = P[X_{m+1} = 0|X_m = 0] = \tfrac{3}{6} = \tfrac{1}{2}.$$

Ou seja, a probabilidade condicional de termos 0 números primos distintos nas faces para cima após o $(m+1)$-ésimo lançamento de um dado, sendo que no m-ésimo já existia 0 números primos, é de $\frac{1}{2}$. De forma mais clara, assumamos que até o m-ésimo lançamento tenha aparecido apenas os números que não são primos. A chance de continuar a não aparecer número primo no $(m+1)$-ésimo lançamento é de 3 em 6, pois, das seis faces, $1, 4$ e 6 são aquelas que realmente interessam nesta caracterização.

Para p_{12} temos, por inspeção,

$$p_{12} = P[X_{m+1} = 2|X_m = 1] = \frac{2}{6} = \frac{1}{3}.$$

A probabilidade condicional de termos 2 números primos distintos nas faces para cima após o $(m+1)$-ésimo lançamento de um dado, sendo que no m-ésimo existia 1 único número primo, é de $\frac{1}{3}$. Em outras palavras, digamos que até o m-ésimo lançamento tenha aparecido apenas o número primo 5. A chance de aparecer um primo distinto no $(m+1)$-ésimo lançamento é de 2 em 6, pois, das seis faces, 2 e 3 são aquelas que nos interessam.

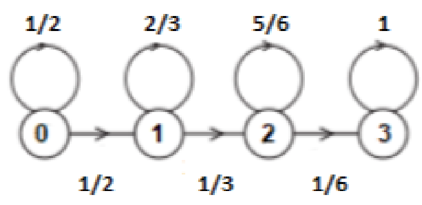
Para finalizar, no caso em que tratamos p_{13}, segue que

$$p_{13} = P[X_{m+1} = 3|X_m = 1] = 0.$$

A probabilidade condicional de termos 3 números primos distintos nas faces para cima após o $(m+1)$-ésimo lançamento de uma dado, sendo que no m-ésimo existia 1 único número primo, é 0. Isto, pois, não é possível em um único lançamento passar de 1 número primo distinto para 3 primos distintos. Poderia apenas manter 1 ou passar para 2.

Na verdade, todas as probabilidades de transição, inclusive as detalhadas no texto, estão sintetizadas na figura abaixo.

Figura 2 – Transição da Cadeia de Markov associada ao problema do lançamento de dados e números primos



Fonte: Elaborada pela autora.

A estrutura detalhada na Figura 2 mostra as probabilidades de transição em um determinado lançamento do dado, podendo aparecer um número primo inédito ou não. As transições omitidas têm, obviamente, probabilidade nula de ocorrência.

Na sequência, construímos a matriz de transição $P = (p_{ij})_{\substack{0 \le i \le 3, \\ 0 \le j \le 3}}$

$$P = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 \\ 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Usando a matriz de transição P e conhecendo o vetor estado $x^{(0)}$, pode-se determinar os vetores estado $x^{(1)}$, $x^{(2)}$, $x^{(3)}, \dots, x^{(n)}, \dots$ de acordo com o Teorema 2.1. Se o dado ainda não foi lançado significa que nenhum número primo foi observado. Sendo assim, o vetor estado inicial é dado por

$$x^{(0)} = (1 \quad 0 \quad 0 \quad 0).$$

O primeiro elemento desse vetor linha representa a probabilidade de ter aparecido zero números primos distintos, que é de 100% nesta altura. A segunda entrada representa a probabilidade de ter aparecido 1 número primo distinto, a terceira entrada mostra a probabilidade de ter aparecido 2 números primos distintos e a última entrada apresenta a probabilidade de ter aparecido 3 números primos distintos.

Pelo Teorema 2.1, se $n \in \mathbb{N}$,

$$x^{(n+1)} = x^{(0)} \cdot [P]^{n+1} = x^{(n)} \cdot [P].$$

Sendo assim, determina-se

$$x^{(1)} = x^{(0)} \cdot [P] = (1 \quad 0 \quad 0 \quad 0) \cdot [P] = \left(\tfrac{1}{2} \quad \tfrac{1}{2} \quad 0 \quad 0\right),$$

$$x^{(2)} = x^{(1)} \cdot [P] = \left(\frac{1}{2} \quad \frac{1}{2} \quad 0 \quad 0\right) \cdot [P] = \left(\frac{1}{4} \quad \frac{7}{12} \quad \frac{1}{6} \quad 0\right)$$

e

$$x^{(3)} = x^{(2)} \cdot [P] = \left(\tfrac{1}{4} \quad \tfrac{7}{12} \quad \tfrac{1}{6} \quad 0\right) \cdot [P] = \left(\tfrac{1}{8} \quad \tfrac{37}{72} \quad \tfrac{1}{3} \quad \tfrac{1}{36}\right).$$

Em 3 lançamentos, temos $\frac{1}{8}$ de chance de não ter aparecido número primo, $\frac{37}{72}$ de chance de ter aparecido exatamente um número primo (podendo ser 2, 3 ou 5), $\frac{1}{3}$ de chance de ter aparecido exatamente 2 números primos distintos e $\frac{1}{36}$ de chance de ter aparecido os 3 números primos distintos. Isto confere com o que observamos manualmente no início desta seção.

Retomando a pergunta, qual a probabilidade de aparecer $i \in \Gamma$ números primos distintos após n lançamentos de um dado, para $n > 3$?

Já comentamos que encontrar o espaço amostral para $n > 3$ e calcular as probabilidades de forma manual é exaustivo e inviável. Aplicando o Teorema 2.1, encontramos os vetores estado para $n > 3$. A tabela a seguir mostra as entradas dos vetores estado com n variando entre 0 e 20. Para facilitar os cálculos, informamos que todos os produtos matriciais foram feitos com auxílio do programa Excel.

Tabela 7 – Entradas dos vetores estado com n variando entre 0 e 20. Situação-problema: Lançamentos de dados e números primos.

	$P[X_n = 0]$	$P[X_n = 1]$	$P[X_n = 2]$	$P[X_n = 3]$
$n = 0$	1	0	0	0
$n = 1$	0,500	0,500	0	0
$n = 2$	0,250	0,583	0,167	0
$n = 3$	0,125	0,514	0,333	0,028
$n = 4$	0,063	0,405	0,449	0,083
$n = 5$	0,032	0,301	0,509	0,158
$n = 6$	0,016	0,216	0,525	0,243
$n = 7$	0,008	0,152	0,510	0,330
$n = 8$	0,004	0,105	0,475	0,416
$n = 9$	0,002	0,072	0,431	0,495
$n = 10$	0,001	0,049	0,383	0,567
$n = 11$	~ 0	0,033	0,336	0,631
$n = 12$	~ 0	0,022	0,291	0,687
$n = 13$	~ 0	0,015	0,250	0,735
$n = 14$	~ 0	0,010	0,213	0,777
$n = 15$	~ 0	0,007	0,181	0,812
$n = 20$	~ 0	~ 0	0,077	0,923

Fonte: Elaborada pela autora.

Todas as probabilidades da tabela são valores aproximados com 3 casas decimais. Note que para 4 lançamentos de um dado, a probabilidade de aparecer 0 números primos é $0{,}063$, 1 número primo distinto é $0{,}405$, 2 números primos distintos é $0{,}449$ e 3 números primos distintos é $0{,}083$.

Para $n = 5$, temos que as chances de aparecer 0, 1, 2 ou 3 números primos distintos são, respectivamente, $0{,}032$; $0{,}301$; $0{,}509$ e $0{,}158$.

Perceba que para $n > 10$, ou seja, acima de 10 lançamentos, a probabilidade de aparecer 0 números primos é muito pequena, se tornando muito próxima de zero. Para $n > 15$, a probabilidade de aparecer apenas 1 número primo distinto também se torna muito pequena, ou seja, próxima a zero.

Conforme os valores de n aumentam, a probabilidade de aparecer 3 números primos distintos cresce. Para $n = 20$, veja que a chance de aparecer 3 números distintos é de $0{,}923$. Percebe-se que a tendência é que com mais lançamentos, a probabilidade de termos os 3 números primos distintos "voltados para parte de cima do dado" fica maior e tão próximo de 1 (100%) o quanto desejarmos.

Na Subseção 2.5.5, vimos que um estado j é absorvente se retornar para ele mesmo com certeza em uma transição, isto é, se $p_{jj} = 1$. O exemplo tratado nessa seção configura uma Cadeia de Markov Estacionária Absorvente, já que no caso $j = 3$, $p_{33} = 1$. Ressalta-se que 3 é o único estado absorvente da cadeia. Quando o estado 3 for atingido não será mais possível sair dele. Notamos que nos lançamentos do dado, quando aparecerem os números 2, 3 e 5 – que são os três números primos distintos – não será mais possível sair desse estado de observação. Por mais lançamentos que sejam efetuados, os 3 números primos distintos já foram obtidos e não será permitido retroceder para uma quantidade inferior de números primos.

Lembrando que

$$x^{(0)} = (1 \quad 0 \quad 0 \quad 0),$$

percebe-se que $x^{(n)} = x^{(0)} \cdot [P]^n \to (\, 0 \quad 0 \quad 0 \quad 1\,)$, quando $n \to \infty$. Isto indica que, quando n tende ao infinito, a probabilidade de aparecer os 3 números primos distintos tende a 1. Essa situação relatada aqui pode ser igualmente interpretada quando levamos em consideração a noção de esperança de uma variável aleatória.

Suponhamos que $E[X_n]$ seja a esperança da variável aleatória X_n, em que $n \in \mathbb{N}$. Na sequência, calculamos a esperança relativa a alguns lançamentos do dado.

Para um lançamento:

$$E[X_1] = 0 \cdot P[x = 0] + 1 \cdot P[x = 1] = 0 \cdot 0{,}5 + 1 \cdot 0{,}5 = 0{,}5.$$

Para 2 lançamentos:

$$E[X_2] = 0 \cdot P[x = 0] + 1 \cdot P[x = 1] + 2 \cdot P[x = 2] = 0 \cdot 0{,}25 + 1 \cdot 0{,}583 + 2 \cdot 0{,}167$$
$$= 0{,}583 + 0{,}334 = 0{,}917.$$

Para 3 lançamentos:

$$E[X_3] = 0 \cdot P[x = 0] + 1 \cdot P[x = 1] + 2 \cdot P[x = 2] + 3 \cdot P[x = 3]$$
$$= 0 \cdot 0{,}125 + 1 \cdot 0{,}514 + 2 \cdot 0{,}333 + 3 \cdot 0{,}028 = 0{,}514 + 0{,}666 + 0{,}084$$
$$= 1{,}264.$$

Para 4 lançamentos:

$$E[X_4] = 0 \cdot P[x = 0] + 1 \cdot P[x = 1] + 2 \cdot P[x = 2] + 3 \cdot P[x = 3]$$
$$= 0 \cdot 0{,}063 + 1 \cdot 0{,}405 + 2 \cdot 0{,}449 + 3 \cdot 0{,}083 = 0{,}405 + 0{,}898 + 0{,}249$$
$$= 1{,}552.$$

Para 10 lançamentos:

$$E[X_{10}] = 0 \cdot P[x = 0] + 1 \cdot P[x = 1] + 2 \cdot P[x = 2] + 3 \cdot P[x = 3]$$
$$= 0 \cdot 0 + 1 \cdot 0{,}049 + 2 \cdot 0{,}384 + 3 \cdot 0{,}567 = 0{,}049 + 0{,}768 + 1{,}701 = 2{,}518.$$

Para 20 lançamentos:

$$E[X_{20}] = 0 \cdot P[x = 0] + 1 \cdot P[x = 1] + 2 \cdot P[x = 2] + 3 \cdot P[x = 3]$$
$$= 0 \cdot 0 + 1 \cdot 0 + 2 \cdot 0{,}077 + 3 \cdot 0{,}923 = 0{,}154 + 2{,}769 = 2{,}923.$$

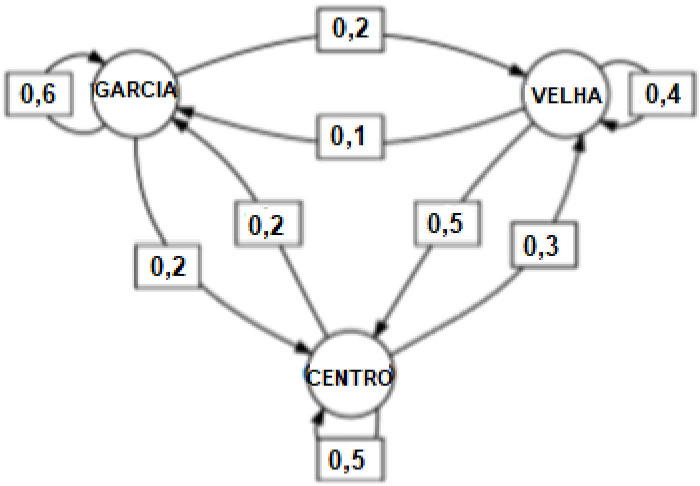
Conforme o esperado, $E[X_1] \leq E[X_2] \leq E[X_3] \leq \cdots \leq E[X_n] \leq \cdots$ Para dois lançamentos, a esperança aponta para a obtenção de, pelo menos, um número primo distinto. Para 10 lançamentos, a esperança aponta para que dois números primos distintos já tenham sido identificados.

Intuitivamente, identificamos que $\lim_{n \to \infty} E[X_n] = 3$. A expectativa é que com mais jogadas, maiores sejam as chances de obter o aparecimento dos três números primos distintos.

3.2 EXEMPLO II – MIGRAÇÃO EM REGIÕES DA CIDADE

Nesta seção, lidaremos com uma situação fictícia. Suponhamos que a cidade de Blumenau seja dividida em três regiões: Região 1 – Centro, Região 2 – Garcia, Região 3 – Velha. Os moradores de Blumenau, ao final de cada ano, permanecem na sua região ou migram para uma outra dentro da cidade. O diagrama abaixo apresenta as probabilidades de permanecer em uma mesma região ou de mudar para outra ao fim de um ano. A probabilidade de migração de moradores de Blumenau para outras cidades é desprezível.

Figura 3 – Diagrama das migrações em regiões da cidade de Blumenau



Fonte: Elaborada pela autora.

Respondamos a seguinte pergunta: Qual a probabilidade de um morador do Centro, após três anos, estar residindo na região do Garcia?

Resolução. Conforme o diagrama, após um ano, um morador que reside no Centro tem probabilidade de 0,2 de migrar para a região do Garcia, probabilidade de 0,3 de migrar para a região da Velha e probabilidade de 0,5 de permanecer no Centro. Um morador que reside no Garcia tem probabilidade de 0,2 de migrar para o Centro, probabilidade de 0,2 de migrar para a Velha e probabilidade de 0,6 de permanecer no Garcia. Já um morador da região da Velha tem probabilidade de 0,5 de migrar para o Centro, probabilidade de 0,1 de migrar para o Garcia e probabilidade de 0,4 de permanecer na Velha.

Consideremos $X_n\colon \Omega \to \mathbb{R}$, de forma que Ω seja o espaço amostral (comum), para todo $n \in \mathbb{N}$. $\{X_n\}_{n\in\mathbb{N}} = \{X_n\colon n \in \mathbb{N}\}$ é uma sequência aleatória, ou ainda, uma família de variáveis aleatórias tal que

$$\Gamma = \bigcup_{n\in\mathbb{N}} Im(X_n) = \{1, 2, 3\}.$$

Notemos que os elementos do conjunto Γ (espaço de estados) estão relacionados às regiões 1, 2 e 3, respectivamente, Centro, Garcia e Velha. E n representa o número de anos com migrações em análise.

Este exemplo configura uma Cadeia de Markov estacionária, pois, se considerarmos que $i, j \in \Gamma$ e que após a m-ésima verificação de migração o cidadão resida na região i, a probabilidade condicional de que na $(m+1)$-ésima migração o mesmo esteja na região j independe do número natural m caracterizado.

Sejam $i, j \in \Gamma$. Vamos determinar as probabilidades de transição, denotadas por p_{ij}:

$$p_{ij} = P\big[X_{m+1} = j \,\big|\, X_m = i\big].$$

O próprio diagrama permite determinar, por exemplo, que

$$p_{11} = P[X_{m+1} = 1 | X_m = 1] = 0{,}5.$$

A probabilidade condicional de estar na região 1 após a $(m+1)$-ésima permissão de migração, sendo que após a m-ésima estava na região 1 é de 0,5.

Para p_{12}, temos que

$$p_{12} = P[X_{m+1} = 2 | X_m = 1] = 0{,}2.$$

Isto é, a probabilidade de ir da região 1 para a região 2 numa abertura de migração é de 0,2.

Analogamente, identifica-se que

$$p_{13} = P[X_{m+1} = 3 | X_m = 1] = 0{,}3$$

e as demais probabilidades de transição são todas extraídas facilmente do diagrama proposto no enunciado.

Vamos construir a matriz de transição que caracteriza as probabilidades de migrações, ano a ano, dos moradores de Blumenau dentro das regiões que constituem a cidade. Lembrando que $\Gamma = \{1, 2, 3\}$, constrói-se a matriz de transição $P = (p_{ij})_{\substack{1\le i\le 3\\ 1\le j\le 3}}$:

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & a_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} 0{,}5 & 0{,}2 & 0{,}3 \\ 0{,}2 & 0{,}6 & 0{,}2 \\ 0{,}5 & 0{,}1 & 0{,}4 \end{pmatrix}.$$

Inicialmente, o morador enunciado reside no Centro que representa a Região 1. Daí, o vetor estado inicial para a resolução deste problema é:

$$x^{(0)} = (1 \quad 0 \quad 0).$$

Aplicando o Teorema 2.1, segue que:

$$x^{(1)} = x^{(0)} \cdot P = (1 \quad 0 \quad 0) \cdot \begin{pmatrix} 0,5 & 0,2 & 0,3 \\ 0,2 & 0,6 & 0,2 \\ 0,5 & 0,1 & 0,4 \end{pmatrix} = (0,5 \quad 0,2 \quad 0,3).$$

Analisando o vetor estado acima, como esperado, temos que após um ano, o morador que inicialmente reside no Centro tem probabilidade de 0,5 de continuar no Centro, probabilidade de 0,2 de mudar para a região do Garcia e probabilidade de 0,3 de mudar para a região da Velha.

Após mais um ano, vê-se que:

$$x^{(2)} = x^{(1)} \cdot P = (0,5 \quad 0,2 \quad 0,3) \cdot \begin{pmatrix} 0,5 & 0,2 & 0,3 \\ 0,2 & 0,6 & 0,2 \\ 0,5 & 0,1 & 0,4 \end{pmatrix} = (0,44 \quad 0,25 \quad 0,31).$$

Notemos que, após tal período, o morador do Centro tem probabilidade de 0,44 de estar residindo no Centro, probabilidade de 0,25 de estar na região do Garcia e probabilidade de 0,31 de residir na região da Velha. Finalmente, façamos

$$x^{(3)} = x^{(2)} \cdot P = (0,44 \quad 0,25 \quad 0,31) \cdot \begin{pmatrix} 0,5 & 0,2 & 0,3 \\ 0,2 & 0,6 & 0,2 \\ 0,5 & 0,1 & 0,4 \end{pmatrix} = (0,425 \quad 0,269 \quad 0,306)$$

Respondendo à pergunta: Espera-se que, após 3 anos, uma pessoa que mora inicialmente no Centro (Região 1) pode, com probabilidade de 0,269, estar residindo na região do Garcia (Região 2).

O mesmo tem uma probabilidade de 0,306 de residir na região da Velha. Finalmente, a probabilidade de estar residindo no Centro é de 0,425.

Considerando que o indivíduo reside inicialmente na região central, a tabela a seguir mostra as probabilidades para a residência dele ao longo dos próximos 20 anos.

Tabela 8 – Entradas dos vetores estado com n variando entre 0 e 20. Situação problema: Migrações em regiões da cidade de Blumenau

	$P[X_n = 1]$	$P[X_n = 2]$	$P[X_n = 3]$
$n = 0$	1	0	0
$n = 1$	0,500	0,200	0,300
$n = 2$	0,440	0,250	0,310
$n = 3$	0,425	0,269	0,306
$n = 4$	0,419	0,277	0,304
$n = 5$	0,417	0,280	0,303
$n = 6$	0,416	0,282	0,302
$n = 7$	0,415	0,283	0,302
$n = 8$	0,415	0,283	0,302
$n = 9$	0,415	0,283	0,302
$n = 10$	0,415	0,283	0,302
$n = 15$	0,415	0,283	0,302
$n = 20$	0,415	0,283	0,302

Fonte: Elaborada pela autora.

Todas as probabilidades listadas na Tabela 8 são valores aproximados com 3 casas decimais. Note que em 4 migrações (4 anos), o morador do Centro tem probabilidade de 0,419 de estar residindo no Centro, probabilidade de 0,277 de estar na região do Garcia e probabilidade de 0,304 de residir na região da Velha.

Para $n = 5$, vemos a probabilidade de 0,417 do cidadão estar residindo no Centro, probabilidade de 0,280 de estar na região do Garcia e probabilidade de 0,303 de residir na região da Velha. Percebe-se que, após 20 anos, a probabilidade de 0,415 de o indivíduo estar residindo na região do Centro, 0,283 é a probabilidade de estar residindo na região do Garcia e 0,302 é a probabilidade deste habitante estar localizado na região da Velha.

Notemos que, conforme aumenta o número de mudanças de estado, ou seja, a passagem dos anos, as probabilidades tendem a um valor fixo.

Este exemplo, na verdade, configura uma Cadeia de Markov Regular. Conforme tratado na Subseção 2.5.4, basta fazermos $n = 1$ para verificar que a matriz P^n apresenta todas as probabilidades de transição (elementos da matriz) estritamente positivas.

Conforme o Teorema 2.2, existe um único vetor estado q para o qual $q \cdot P = q$.

Vamos encontrar algebricamente o vetor estado q associado ao exemplo explorado nesta seção. Inicialmente, atribuímos as incógnitas x, y e z para os 3 elementos desse vetor estado q. Desejamos que

$$(x \quad y \quad z) \cdot \begin{pmatrix} 0{,}5 & 0{,}2 & 0{,}3 \\ 0{,}2 & 0{,}6 & 0{,}2 \\ 0{,}5 & 0{,}1 & 0{,}4 \end{pmatrix} = q \cdot P = q = (x \quad y \quad z).$$

Aplicando o produto matricial, nos deparamos com um sistema de três equações com três incógnitas. Segue que:

$$\begin{cases} 0{,}5x + 0{,}2y + 0{,}5z = x \\ 0{,}2x + 0{,}6y + 0{,}1z = y \\ 0{,}3x + 0{,}2y + 0{,}4z = z \end{cases} \Rightarrow \begin{cases} -0{,}5x + 0{,}2y + 0{,}5z = 0 \\ 0{,}2x - 0{,}4y + 0{,}1z = 0 \\ 0{,}3x + 0{,}2y - 0{,}6z = 0. \end{cases}$$

Resolvendo esse sistema, percebemos que ele possui uma infinidade de soluções. Apresentamos todas as soluções em função da incógnita z:

$$S = \{(1{,}375z, 0{,}9375z, z); \forall z \in \mathbb{R}\}.$$

Como o vetor estado envolve probabilidades, a soma de seus 3 elementos é 1. Sendo assim, $x + y + z = 1$. Logo, concluímos que

$$1{,}375z + 0{,}9375z + z = 1 \Rightarrow z = \frac{10000}{33125}$$

e, consequentemente,

$$y = \frac{9375}{33125} \text{ e } x = \frac{13750}{33125}.$$

Assim, identificamos o vetor estado $q = \begin{pmatrix} \frac{13750}{33125} & \frac{9375}{33125} & \frac{10000}{33125} \end{pmatrix}$, que é único.

Para o vetor estado inicial $x^{(0)} = (1 \quad 0 \quad 0)$, a sequência de vetores estado

$$x^{(0)}, x^{(0)} \cdot P, x^{(0)} \cdot P^2, x^{(0)} \cdot P^3, \ldots, x^{(0)} \cdot P^n, \ldots$$

tende a $q = \begin{pmatrix} \frac{13750}{33125} & \frac{9375}{33125} & \frac{10000}{33125} \end{pmatrix}$ como limite. Ou seja,

$$x^{(n)} = x^{(0)} \cdot P^n \to \begin{pmatrix} \frac{13750}{33125} & \frac{9375}{33125} & \frac{10000}{33125} \end{pmatrix}, \text{ quando } n \to \infty.$$

O vetor q é chamado de vetor de estado estacionário. É valido ressaltar que, independente de $x \geq 0$, $y \geq 0$ e $z \geq 0$, tais que $x + y + z = 1$, veremos

$$(x \quad y \quad z) \cdot P^n \to q, \text{ quando } n \to \infty.$$

Escrevendo os valores na forma decimal, podemos observar que até a terceira casa decimal, as entradas de q são as mesmas deduzidos na Tabela já no caso em que $n = 7$, pois, tem-se o vetor $q = (0{,}415 \quad 0{,}283 \quad 0{,}302)$.

4 O PROBLEMA DE COMPLETAR UMA COLEÇÃO DE FIGURINHAS

Neste capítulo, vamos tratar de um problema de completar uma coleção de figurinhas. O estudo é baseado no artigo "Álbuns de figurinhas: uma abordagem via Cadeias de Markov", dos professores Leandro Batista Morgado e Leonardo Silveira Borges (ver Ref. [15]). Neste artigo, os autores abordam o preenchimento de um álbum de figurinhas da copa do mundo de futebol de 2018. Dentre outros assuntos, eles estimam um número de figurinhas a serem compradas para garantir chances razoáveis de completar uma coleção, apesar das peças repetidas que eventualmente aparecem.

O nosso problema consiste em identificar a quantidade de compras necessárias para obter uma chance razoável de completar uma coleção de 30 figurinhas, adquiridas na compra dos chocolates Surpresa. Enfatizamos aqui que, apesar de se tratar de um produto inexistente nos dias atuais, o problema pode ser facilmente adaptado para qualquer outro tipo de colecionismo. Além de utilizarmos uma abordagem que envolve cadeias de Markov, usamos uma estimativa apresentada no artigo citado que evita o cálculo de uma considerável quantidade de produtos matriciais.

4.1 AS FIGURINHAS DO CHOCOLATE SURPRESA

Colecionar figurinhas é uma brincadeira que marca gerações. Crianças ou adultos, não importa a idade, o interesse por colecionar é compartilhado por muita gente. Para um colecionador, a sensação de adquirir uma nova figurinha para tentar completar uma coleção é indescritível. Já existiram diversas coleções de figurinhas em nossa sociedade, como figurinhas de jogadores de copas do mundo, de campeonatos nacionais, de animais, super-heróis, entre outros. As figurinhas podem ser compradas em bancas ou acompanham revistas ou produtos alimentícios. Em geral, o comprador não sabe se a figurinha adquirida é inédita ou repetida em sua coleção.

Aqui, vamos tratar de uma coleção de figurinhas que marcou a infância das crianças dos anos 1990: as figurinhas de dinossauros adquiridas na compra de chocolates. Mas antes de tratar matematicamente, vamos voltar no tempo. Os chocolates Surpresa eram vendidos pela empresa de alimentos Nestlé, na forma de barra de chocolate ao leite, com uma figurinha acompanhando o doce dentro da embalagem. Antes de comprar a barra, portanto, não era possível identificar qual a figurinha que acompanhava o produto.

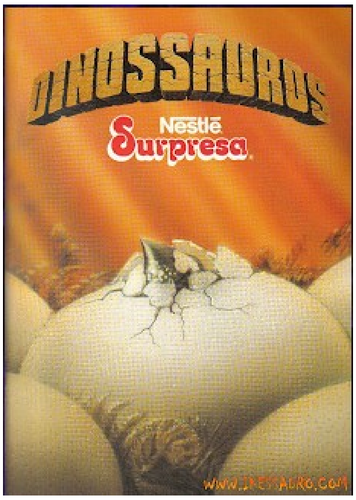
O chocolate Surpresa começou a ser vendido no Brasil nos anos 1980 e as figurinhas que o acompanhavam sempre foram um atrativo. Inicialmente, as figurinhas traziam imagens de animais do mundo, do Pantanal, da Amazônia, seres marinhos, entre outros. Com o sucesso dos dinossauros nos anos 1990 e o lançamento do filme Jurassic Park em 1993, foi comercializada a célebre coleção de figurinhas de dinossauros.

Figura 4 – Figurinha da coleção Dinossauros do Chocolate Surpresa

Fonte: Nestlé

A coleção dos dinossauros era constituída por 30 figurinhas de animais pré-históricos, ilustradas com imagens realistas e coloridas. Para os fãs mais entusiasmados, foi comercializado um álbum a ser preenchido que dividia os dinossauros em três grupos (Triássico, Jurássico e Cretássico), destacando o período geológico em que cada animal teria vivido. Além disso, no álbum eram mostradas outras informações e curiosidades sobre estes répteis.

Figura 5 – Capa do Álbum Chocolate Surpresa – Dinossauros



Fonte: Nestlé

Baseados nesta coleção que marcou uma geração, vamos desenvolver o nosso problema que utiliza cadeias de Markov. Uma questão importante a ser respondida aqui é: quantas compras deveriam ser efetuadas para que haja uma chance razoável de preencher o álbum?

Lembramos que para completar o álbum precisa-se encontrar as 30 figurinhas distintas que vêm com o chocolate Surpresa. Cada chocolate acompanha uma única figurinha. Ao comprar o seguinte chocolate, o colecionador pode ter uma figurinha nova ou uma figurinha repetida. Intuitivamente, entende-se que com menos figurinhas, maiores as chances de obter uma figurinha inédita. Com a coleção quase completa, maiores as chances da compra resultar em uma figurinha repetida.

Assumimos que a coleção é honesta, ou seja, que todas as figurinhas têm a mesma probabilidade de aparecer no ato da compra do chocolate. Considerando que é feita a compra de um chocolate por vez, nas próximas seções teremos subsídios para responder a questão aqui levantada.

4.2 DISCUSSÃO DO PROBLEMA VIA PROBLEMAS DE MARKOV

Continuamos a abordar o problema do preenchimento do álbum relativo às figurinhas do chocolate Surpresa. Como há 30 figurinhas diferentes, denotamos cada uma delas como F_k, ressaltando que $k \in \{1,2,\dots,30\}$.

Dizemos que Ω_n é o espaço amostral que lista as possíveis figurinhas adquiridas após n compras.

Para efeito de entendimento, se $n = 1$,

$$\Omega_1 = \{F_1, F_2, F_3, \dots, F_{29}, F_{30}\}.$$

Se $n = 2$, o espaço amostral tem $30^2 = 900$ elementos, pois,

$$\Omega_2 = \left\{ \begin{array}{c} (F_1,F_1), (F_1,F_2), (F_1,F_3), (F_1,F_4), \dots, (F_1,F_{29}), (F_1,F_{30}), \\ (F_2,F_1), (F_2,F_2), (F_2,F_3), (F_2,F_4), \dots, (F_2,F_{29}), (F_2,F_{30}), \\ (F_3,F_1), (F_3,F_2), (F_3,F_3), (F_3,F_4), \dots, (F_3,F_{29}), (F_3,F_{30}), \\ \vdots \\ (F_{30},F_1), (F_{30},F_2), (F_{30},F_3), (F_{30},F_4), \dots, (F_{30},F_{29}), (F_{30},F_{30}) \end{array} \right\}.$$

Se $n = 3$, o espaço amostral tem $30^3 = 27000$ elementos e

$$\Omega_3 = \begin{Bmatrix} (F_1,F_1,F_1),(F_1,F_1,F_2),(F_1,F_1,F_3),(F_1,F_1,F_4),\dots,(F_1,F_1,F_{29}),(F_1,F_1,F_{30}), \\ (F_1,F_2,F_1),(F_1,F_2,F_2),(F_1,F_2,F_3),(F_1,F_2,F_4),\dots,(F_1,F_2,F_{29}),(F_1,F_2,F_{30}) \\ (F_1,F_3,F_1),(F_1,F_3,F_2),(F_1,F_3,F_3),(F_1,F_3,F_4),\dots,(F_1,F_3,F_{29}),(F_1,F_3,F_{30}), \\ \vdots \\ (F_{30},F_{29},F_1),(F_{30},F_{29},F_2),(F_{30},F_{29},F_3),\dots,(F_{30},F_{29},F_{29}),(F_{30},F_{29},F_{30}), \\ (F_{30},F_{30},F_1),(F_{30},F_{30},F_2),(F_{30},F_{30},F_3),\dots,(F_{30},F_{30},F_{29}),(F_{30},F_{30},F_{30}) \end{Bmatrix}.$$

Seguindo esta linha de raciocínio, identifica-se que Ω_n possui 30^n elementos.

Definamos $X_n\colon \Omega_n \to \mathbb{R}$, a função variável aleatória que associa o espaço amostral ao número de figurinhas distintas adquiridas após n compras.

No caso da primeira compra, o colecionador adquire uma figurinha. Veja que pode ser qualquer uma das 30 da coleção. Então, $X_1\colon \Omega_1 \to \mathbb{R}$ é uma função tal que $Im(X_1) = \{1\}$. Em termos de probabilidade, temos que

$$P[X_1 = 1] = \frac{30}{30} = 1.$$

Obviamente, a chance do colecionador adquirir uma figurinha distinta na primeira compra é 100%.

No caso de duas compras seguidas, é possível aparecer uma figurinha repetida ou adquirir duas figurinhas distintas. Sendo assim, $X_2\colon \Omega_2 \to \mathbb{R}$ é uma função tal que $Im(X_2) = \{1, 2\}$.

Já citamos que o espaço amostral, nesse caso, tem $30^2 = 900$ elementos. A chance de ter duas figurinhas iguais (portanto, uma única figurinha distinta) é vista em 30 casos. Nos demais 870 casos, o colecionador teria duas figurinhas distintas. As probabilidades associadas a $n = 2$ são:

$$P[X_2 = 1] = \frac{30}{900} = \frac{1}{30} \text{ e } P[X_2 = 2] = \frac{870}{900} = \frac{29}{30}.$$

Em resumo, na compra de 2 chocolates, a chance de adquirir apenas uma figurinha distinta é $\frac{1}{30}$ e de obter duas figurinhas distintas é $\frac{29}{30}$.

Para três compras, pode-se ficar só com uma figurinha distinta, com duas figurinhas distintas ou com as três figurinhas diferentes. Sendo assim, $X_3\colon \Omega_3 \to \mathbb{R}$ é uma função tal que $Im(X_3) = \{1, 2, 3\}$. Seguindo o raciocínio detalhado anteriormente, as probabilidades relacionadas para $n = 3$ são:

$$P[X_3 = 1] = \frac{30}{27000} = \frac{1}{900},\ P[X_3 = 2] = \frac{2610}{27000} = \frac{29}{300} \text{ e } P[X_3 = 3] = \frac{24360}{27000} = \frac{203}{225}.$$

Portanto, na compra de 3 chocolates, as probabilidades de adquirir 1, 2 ou 3 figurinhas distintas são de $\frac{1}{900}$, $\frac{29}{300}$ e $\frac{203}{225}$, respectivamente.

Notamos que esse estudo pode ser feito para n natural qualquer. Obviamente, se o inteiro for $n > 30$, é possível que a imagem de X_n esteja no conjunto $\{1, 2, 3, \ldots, 29, 30\}$. À medida que n cresce, o cálculo das probabilidades fica mais trabalhoso. Usaremos, inicialmente, o Teorema 2.1 para um estudo efetivo.

À princípio, identificamos que $\{X_n; n \in \mathbb{N}\}$ constitui uma Cadeia de Markov. Como em uma primeira compra passa-se de zero figurinhas distintas para uma figurinha distinta, é preciso que 0 pertença ao espaço de estados. Nesta abordagem, identificamos que o conjunto $\Gamma = \{0, 1, 2, 3, \ldots, 29, 30\}$ é o espaço de estados do processo estocástico.

Em verdade, $\{X_n; n \in \mathbb{N}\}$ configura uma Cadeia de Markov estacionária. Sejam os índices $i, j \in \Gamma$. Suponha que até a m-ésima compra de figurinhas, temos i figurinhas distintas. A probabilidade condicional de que na $(m+1)$-ésima compra tenhamos j figurinhas distintas independe do número natural m considerado na caracterização.

A aplicação do Teorema 2.1 requer a dedução da matriz de transição $P = (p_{ij})_{\substack{0 \leq i \leq 30 \\ 0 \leq j \leq 30}}$, em que

$$p_{ij} = P\big[X_{m+1} = j \,\big|\, X_m = i\big] = P\big[X_1 = j \,\big|\, X_0 = i\big] \text{ e } i, j \in \Gamma.$$

A matriz P é de ordem $31x31$, pois, o conjunto $\Gamma = \{0, 1, 2, 3, \ldots, 29, 30\}$ tem 31 elementos.

Vamos refletir a identificação de algumas probabilidades de transição p_{ij}.

Vemos que:

$$p_{01} = P[X_{m+1} = 1 | X_m = 0] = 1.$$

Se não há figurinha até então, a probabilidade condicional de adquirir 1 figurinha distinta em uma compra de chocolate Surpresa é 100%. Tal observação é suficiente para garantir, por exemplo, que

$$p_{02} = P[X_{m+1} = 2 | X_m = 0] = 0.$$

Isto se deve ao fato de que $p_{00} + p_{01} + p_{02} + \cdots + p_{0u} = 1$, em que $u = 30$.

Além disso,

$$p_{10} = P[X_{m+1} = 0 | X_m = 1] = 0,$$

Pois, se já há uma figurinha, ao comprar um chocolate o número de figurinhas existentes não pode zerar. Para uma figurinha distinta já existente, uma compra pode levar a duas situações: uma figurinha distinta ou duas figurinhas distintas.

Sendo assim, identificamos que

$$p_{11} = P[X_{m+1} = 1|X_m = 1] = \frac{1}{30}.$$

Dada a existência de uma figurinha distinta, a compra de uma nova figurinha igual é dada na probabilidade de 1 em 30.

Para p_{12}, temos que

$$p_{12} = P[X_{m+1} = 2|X_m = 1] = \frac{29}{30}.$$

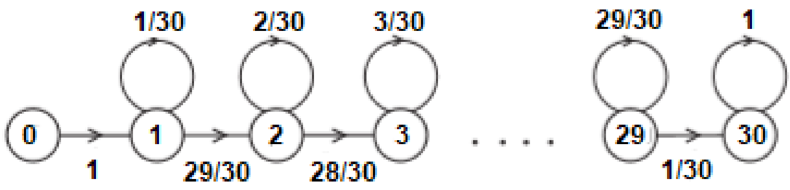
Isto é, a probabilidade condicional de termos 2 figurinhas distintas na $(m + 1)$-ésima compra do chocolate, sendo que na m-ésima havia apenas 1 figurinha distinta, é de $\frac{29}{30}$.

Por sua vez, identifica-se que

$$p_{13} = P[X_{m+1} = 3|X_m = 1] = 0.$$

Todas as probabilidades de transição, inclusive as detalhadas acima, aparecem indicadas na figura a seguir.

Figura 6 – Transição da Cadeia de Markov associada ao problema das figurinhas do Chocolate Surpresa



Fonte: Elaborada pela autora.

A estrutura mostra as probabilidades de transição na compra de uma nova figurinha, podendo o colecionador adquirir uma figurinha inédita ou repetida. As transições omitidas têm, obviamente, probabilidade nula de ocorrência.

Abaixo, a matriz de transição $P = \left(p_{ij}\right)_{\substack{0 \le i \le 30 \\ 0 \le j \le 30}}$:

$$P = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{30} & \frac{29}{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{15} & \frac{14}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{10} & \frac{9}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{2}{15} & \frac{13}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{5}{6} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{5} & \frac{4}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{7}{30} & \frac{23}{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{15} & \frac{11}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{10} & \frac{7}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{11}{30} & \frac{19}{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{5} & \frac{3}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{13}{30} & \frac{17}{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{7}{15} & \frac{8}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & \frac{7}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{17}{30} & \frac{13}{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{5} & \frac{2}{5} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{19}{30} & \frac{11}{30} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{7}{10} & \frac{3}{10} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{11}{15} & \frac{4}{15} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{23}{30} & \frac{7}{30} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{5} & \frac{1}{5} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{13}{15} & \frac{2}{15} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{9}{10} & \frac{1}{10} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{14}{15} & \frac{1}{15} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{29}{30} & \frac{1}{30} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}$$

Com a matriz de transição P e conhecendo o vetor estado $x^{(0)}$, podemos determinar os vetores estado $x^{(1)}$, $x^{(2)}$, $x^{(3)}, \dots, x^{(n)}, \dots$ nas observações seguintes, usando o já mencionado Teorema 2.1.

Como, inicialmente, não foi comprada figurinha alguma, temos o vetor estado inicial:

$$x^{(0)} = (1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0).$$

O primeiro elemento dessa matriz linha representa a probabilidade de ter adquirido zero figurinhas distintas, que é de 100% - já que antes das compras o colecionador não possui qualquer figurinha. A segunda entrada representa a probabilidade de ter adquirido 1 figurinha distinta, a terceira entrada mostra a probabilidade de ter adquirido 2 figurinhas distintas, a quarta entrada apresenta a probabilidade de ter adquirido 3 figurinhas distintas e, assim por diante, até a trigésima primeira entrada que mostra a probabilidade de conseguir 30 figurinhas distintas.

Pelo Teorema 2.1, se $n \in \mathbb{N}$,

$$x^{(n+1)} = x^{(n)} \cdot P.$$

Sendo assim,

$$x^{(1)} = (0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0),$$

$$x^{(2)} = \left(0\ \frac{1}{30}\ \frac{29}{30}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\right)$$

e

$$x^{(3)} = \left(0\ \frac{1}{900}\ \frac{29}{300}\ \frac{203}{225}\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\right).$$

Na compra de 3 figurinhas tem-se $\frac{1}{900}$ de chance de adquirir 1 figurinha distinta, $\frac{29}{300}$ de chance de comprar 2 figurinhas distintas e $\frac{203}{225}$ de chance de comprar 3 figurinhas distintas. Isto confere com o que observamos manualmente no início dessa seção.

Usando o programa Excel para efetuar o produto de matrizes, constrói-se a tabela a seguir em que se estudam as probabilidades de obtenção de figurinhas distintas em até 50 compras do chocolate.

Tabela 9 – Entradas dos vetores estado com n variando entre 0 e 50. Situação-problema: Completar uma coleção de figurinhas

	$P[X_n = 0]$	$P[X_n = 1]$	$P[X_n = 2]$	$P[X_n = 3]$	$P[X_n = 4]$	$P[X_n = 5]$	$P[X_n = 6]$	$P[X_n = 7]$	$P[X_n = 8]$	$P[X_n = 9]$	$P[X_n = 10]$	$P[X_n = 11]$	$P[X_n = 12]$	$P[X_n = 13]$	$P[X_n = 14]$	$P[X_n = 15]$	$P[X_n = 16]$	$P[X_n = 17]$	$P[X_n = 18]$	$P[X_n = 19]$	$P[X_n = 20]$	$P[X_n = 21]$	$P[X_n = 22]$	$P[X_n = 23]$	$P[X_n = 24]$	$P[X_n = 25]$	$P[X_n = 26]$	$P[X_n = 27]$	$P[X_n = 28]$	$P[X_n = 29]$	$P[X_n = 30]$
$n = 0$	100%	0	0	0	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 1$	0%	100%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 2$	0%	3,33%	96,67%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 3$	0%	0,11%	9,69%	90,2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 4$	0%	~0%	0,75%	18,07%	81,18%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 5$	0%	~0%	0,05%	2,51%	27,09%	70,35%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 6$	0%	~0%	~0%	0,30%	5,87%	35,20%	58,63%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 7$	0%	~0%	~0%	0,03%	1,05%	10,96%	41,06%	46,90%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 8$	0%	~0%	~0%	~0%	0,17%	2,74%	17,34%	43,79%	35,96%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 9$	0%	~0%	~0%	~0%	~0%	0,60%	5,75%	24,09%	43,16%	26,4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 10$	0%	~0%	~0%	~0%	~0%	0,10%	1,65%	10,22%	29,98%	39,57%	18,48%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 11$	0%	~0%	~0%	~0%	~0%	0,02%	0,42%	3,71%	15,83%	33,86%	33,86%	12,3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 12$	0%	~0%	~0%	~0%	~0%	~0%	0,10%	1,20%	7,07%	21,76%	34,98%	27,10%	7,79%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 13$	0%	~0%	~0%	~0%	~0%	~0%	~0%	0,36%	2,80%	11,71%	26,89%	33,30%	20,27%	4,67%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 14$	0%	~0%	~0%	~0%	~0%	~0%	~0%	0,09%	1,02%	5,57%	17,17%	30,1%	29,2%	14,19%	2,66%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 15$	0%	~0%	~0%	~0%	~0%	~0%	~0%	0,02%	0,34%	2,42%	9,62%	22,5%	30,74%	23,66%	9,28%	1,42%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 20$	0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	0,1%	1,0%	4,97%	13,22%	22,93%	26,32%	19,5%	9,07%	2,5%	0,36%	0,02%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
$n = 30$	0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	0,22%	1,24%	4,4%	10,65%	18,53%	23,13%	20,68%	13,15%	5,9%	1,81%	0,3%	~0%	~0%	~0%	~0%	~0%	~0%
$n = 50$	0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	~0%	0,01%	0,13%	0,8%	2,9%	7,8%	15,6%	22,4%	23%	16,6%	8%	2,5%	0,1%	~0%

Fonte: Elaborada pela autora.

Na tabela acima, todas as porcentagens são mostradas com valores aproximados em duas casas decimais.

Para $n = 4$, vemos que a probabilidade do colecionador ter apenas uma única figurinha distinta é ínfima. Já as probabilidades de adquirir 2, 3 ou 4 figurinhas distintas são de $0{,}75\%$, $18{,}07\%$ e $81{,}18\%$, respectivamente.

Para $n = 5$, notamos que a probabilidade de conseguir 2, 3, 4 ou 5 figurinhas distintas são de $0{,}05\%$, $2{,}51\%$, $27{,}09\%$ e $70{,}35\%$, respectivamente. Percebemos que é alta a chance de conseguir todas as figurinhas distintas em cinco compras.

Conforme os valores de n aumentam, a probabilidade de adquirir somente figurinhas distintas se reduz. Notamos que na compra de 6 chocolates, a chance de adquirir 6 figurinhas distintas é 58,63%. Na compra de 7 chocolates, a chance de conseguir todas as figurinhas distintas é 46,90%. Na compra de 8 chocolates é 35,96% a probabilidade de aquisição de 8 figurinhas distintas. Quando tratamos $n = 20$, a chance de conseguir nessas compras 20 figurinhas distintas é de 0,02%, ou seja, configura um evento pouco provável de acontecer.

O caso $n = 30$ é emblemático. Comprando 30 chocolates, a chance de vir as 30 figurinhas distintas, sem repetição portanto, existe, mas, é praticamente nula. Ao comprar 50 chocolates, a chance de vir 29 figurinhas distintas é de 0,1% e a probabilidade de completar a coleção é desprezível.

Este exemplo da aquisição de figurinhas configura uma Cadeia de Markov Absorvente. Isto pois, no caso em que $j = 30$, a probabilidade de transição é $p_{jj} = 1$. Em verdade, 30 é o único estado absorvente da cadeia. Quando o estado 30 for atingido não será mais possível sair dele. Notemos que nas compras dos chocolates Surpresa quando o colecionador adquirir as 30 figurinhas de dinossauros distintas, não será mais possível sair desse estado. Por mais compras que sejam feitas, as 30 figurinhas distintas já foram obtidas e não será permitido retroceder para uma quantidade inferior de figurinhas distintas.

Não identificamos na Tabela 9, mas a tendência é que, com mais compras de chocolates, a probabilidade de termos as 30 figurinhas distintas fica maior e tão próximo de 1 (100%) o quanto desejarmos.

Esperamos que dado o vetor estado inicial

$$x^{(0)} = (1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)$$

tenhamos

$$x^{(0)} \cdot P^n \to (0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1)$$

quando $n \to \infty$.

A situação acima conjecturada pode ser igualmente observada quando levamos em consideração a noção de esperança de uma variável aleatória.

Seja $E[X_n]$ a esperança da variável aleatória X_n. Na sequência, calculamos a esperança para algumas compras de figurinhas.

Para compra de 1 figurinha:

$$E[X_1] = 0 \cdot P[x = 0] + 1 \cdot P[x = 1] = 0 \cdot 0 + 1 \cdot 1 = 1.$$

Para compra de 2 figurinhas:

$$\begin{aligned} E[X_2] &= 0 \cdot P[x = 0] + 1 \cdot P[x = 1] + 2 \cdot P[x = 2] \\ &= 0 \cdot 0 + 1 \cdot \frac{1}{30} + 2 \cdot \frac{29}{30} = 0{,}033 + 1{,}93 = 1{,}963. \end{aligned}$$

Outros valores estão dispostos na Tabela a seguir:

Tabela 10 – Esperança da compra de n figurinhas.

Compra de n figurinhas	**Esperança**
$n = 3$	$E[X_3] = 2{,}9$
$n = 4$	$E[X_4] = 3{,}8043$
$n = 10$	$E[X_{10}] = 8{,}6271$
$n = 20$	$E[X_{20}] = 14{,}7775$
$n = 30$	$E[X_{30}] = 19{,}1452$
$n = 50$	$E[X_{50}] = 24{,}4375$

Fonte: Elaborada pela autora.

Conforme o esperado, $E[X_1] \leq E[X_2] \leq E[X_3] \leq \cdots \leq E[X_n] \leq \cdots$. Em 2 compras, a esperança aponta para a obtenção provável de 2 figurinhas distintas. Em 10 compras, a esperança aponta para algo entre 8 e 9 figurinhas distintas. Para 50 compras, a esperança sugere uma aquisição de ao menos 24 figurinhas diferentes. Intuitivamente, identificamos que $\lim_{n\to\infty} E[X_n] = 30$. A expectativa é que com mais compras de chocolates, maiores sejam as chances de obter as 30 figurinhas distintas e de completar, portanto, o álbum.

Agora, suponhamos que o colecionador esteja disposto a ter a coleção completa das figurinhas do álbum dos dinossauros. Ele pode comprar muitos chocolates e, infelizmente, não ter a sorte de completar sua coleção. Mas, ele também pode estimar uma quantidade de compras que permita que a probabilidade de preencher o álbum seja relativamente alta. Uma forma de fazer isso seria determinando os $x^{(0)} \cdot P^n$ com valores grandes para n. Isso, entretanto, sem auxílio de um programa ou software computacional, pode ser um trabalho consideravelmente árduo. Na seção abaixo, entretanto, é apresentada uma estimativa que requer o conhecimento de funções exponenciais e logarítmicas.

4.3 ESTIMATIVAS PARA COMPLETAR UMA COLEÇÃO

Nesta seção, vamos obter estimativas concretas para que seja razoável a probabilidade de completar o álbum de figurinhas do chocolate Surpresa no processo de compras honestas. A abordagem se baseia na Ref. [15].

Considere a seguinte variável aleatória que, para cada sequência possível de compras de figurinhas, associa o número mínimo necessário para completar o álbum:

$$\tau_{fig} = min\{n \geq 0; X_n = 30\}.$$

Desta forma, τ_{fig} representa a quantidade de tentativas mínimas para se obter as 30 figurinhas distintas levando em consideração uma sequência de compras. Por exemplo, se

$$E_1 = \{X_1 = 1, X_2 = 2, X_3 = 3, \dots, X_{56} = 29, X_{57} = 30\}$$

caracterizar uma sequência possível de compras de figurinhas, tem-se que $\tau_{fig} = 57$. Neste caso, foram requisitadas 57 compras para completar a coleção.

Outro exemplo: suponha que a sequência de compras para achar as trinta figurinhas distintas seja estabelecida em

$$E_2 = \{X_1 = 1, X_2 = 2, X_3 = 2, \dots, X_{97} = 29, X_{98} = 30\}.$$

Nesta sequência, o número mínimo de compras necessárias para completar o álbum é 98. Logo, $\tau_{fig} = \tau_{fig}(E_2) = 98$.

Além disso, para calcular a estimativa de completar o álbum de figurinhas é requisitado que trabalhemos com a noção de menor inteiro. Se $a \in \mathbb{R}$, $\lceil a \rceil$ representa o menor inteiro maior que a. Como exemplos,

$$\lceil 68 \rceil = 69,\ \lceil -3{,}24 \rceil = -3, \lceil 68{,}23 \rceil = 69 \ \text{ e } \ \lceil 124{,}5 \rceil = 125.$$

Seja $c > 0$ uma constante. Com a finalidade de estimar a quantidade de figurinhas que o colecionador deve comprar para ter uma chance razoável de completar o álbum, vamos provar o resultado a seguir:

$$P\left(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil\right) \leq e^{-c}.$$

O resultado estabelece que a probabilidade de não preencher o álbum usando até $\lceil 30 \cdot ln(30) + c \cdot 30 \rceil$ compras é menor ou igual a e^{-c}.

Lembrando que e^{-c} é um número entre 0 e 1, como consequência, estabelece-se outra relação: $P(\tau_{fig} \leq \lceil 30 \cdot ln(30) + c \cdot 30 \rceil) > 1 - \mathrm{e}^{-\mathrm{c}}$. Isto é, a probabilidade de preencher o álbum usando até $\lceil 30 \cdot ln(30) + c \cdot 30 \rceil$ compras é maior que o número real $1 - e^{-c}$>0.

Teorema 4.1 Se $c > 0$, então, $P(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil) \leq e^{-c}$.

Demonstração: Assumimos que todas as 30 figurinhas têm a mesma probabilidade de serem compradas. Sendo assim, a probabilidade de obter a k-ésima figurinha do álbum em uma determinada compra é $\frac{1}{30}$. Logo, a probabilidade de não obter a k-ésima figurinha nessa compra é $\left(1 - \frac{1}{30}\right) = \frac{29}{30}$.

Seja A_k o evento do colecionador não ter obtido a k-ésima figurinha nas primeiras $\lceil 30 \cdot ln(30) + c \cdot 30 \rceil$ compras. Temos que:

$$P(A_k) = \left(1 - \frac{1}{30}\right)^{\lceil 30 \cdot ln(30) + c \cdot 30 \rceil}.$$

Então, a probabilidade do álbum de figurinhas não estar completo nas primeiras $\lceil 30 \cdot ln(30) + c \cdot 30 \rceil$ compras é dada pela probabilidade da união dos eventos A_k, com k variando de 1 a 30. Portanto,

$$P(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil) = P\left(\bigcup_{k=1}^{30} A_k\right). \tag{4.1}$$

No Capítulo 2, foi citado que $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$. Em particular, $P(A_1 \cup A_2) \leq P(A_1) + P(A_2)$. Usando um raciocínio similar, mostra-se que

$$P\left(\bigcup_{k=1}^{30} A_k\right) \leq \sum_{k=1}^{30} P(A_k). \tag{4.2}$$

Substituindo (4.2) em (4.1), temos que:

$$P(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil) \leq \textstyle\sum_{k=1}^{30} P(A_k). \tag{4.3}$$

Mas,

$$\sum_{k=1}^{30} P(A_k) = 30 \cdot \left(1 - \frac{1}{30}\right)^{\lceil 30 \cdot ln(30) + c \cdot 30 \rceil} = 30 \cdot \left(1 - \frac{1}{30}\right)^{\frac{30 \cdot \lceil 30 \cdot ln(30) + c \cdot 30 \rceil}{30}} \tag{4.4}$$

Estimamos que $\left(1-\frac{1}{30}\right)^{30} = 0{,}3617$ e $\frac{1}{e} = 0{,}3679$. Sendo assim,

$$\left(1-\frac{1}{30}\right)^{30} < e^{-1}$$

e

$$30 \cdot \left(1-\frac{1}{30}\right)^{\frac{30 \cdot \lceil 30 \cdot ln(30) + c \cdot 30 \rceil}{30}} \leq 30 \cdot e^{\frac{-1 \cdot \lceil 30 \cdot ln(30) + c \cdot 30 \rceil}{30}}. \tag{4.5}$$

As relações (4.3), (4.4) e (4.5) implicam que

$$P\left(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil\right) \leq \sum_{k=1}^{30} P(A_k) \leq 30 \cdot e^{\frac{-1 \cdot \lceil 30 \cdot ln(30) + c \cdot 30 \rceil}{30}}. \tag{4.6}$$

Lembrando que

$$30 \cdot ln(30) + c \cdot 30 \leq \lceil 30 \cdot ln(30) + c \cdot 30 \rceil,$$

vê-se que

$$-\lceil 30 \cdot ln(30) + c \cdot 30 \rceil \leq -30 \cdot ln(30) - c \cdot 30.$$

Como exponencial é uma função crescente, já que $e > 1$, devido a (4.6),

$$P\left(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil\right) \leq 30 \cdot e^{\frac{-1 \cdot \lceil 30 \cdot ln(30) + c \cdot 30 \rceil}{30}} \leq 30 \cdot e^{\frac{-30 \cdot ln(30) - c \cdot 30}{30}}. \tag{4.7}$$

Finalmente,

$$e^{\frac{-30 \cdot ln(30) - c \cdot 30}{30}} = e^{\frac{30(-\ln(30) - c)}{30}} = e^{-(\ln(30) + c)} = \frac{1}{e^{\ln(30) + c}} = \frac{1}{e^{\ln(30)} \cdot e^{c}} = \frac{e^{-c}}{30}. \tag{4.8}$$

Substituindo (4.8) em (4.7), temos que:

$$P\left(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil\right) \leq 30 \cdot e^{\frac{-30 \cdot ln(30) - c \cdot 30}{30}} \leq 30 \cdot \frac{e^{-c}}{30} \leq e^{-c}.$$

■

É importante ressaltar que o colecionador não saberá a quantidade exata de figurinhas a serem compradas para preenchimento do álbum. Isso é impossível de ser determinado, pois, trata-se de uma situação que envolve experimentos aleatórios. O que pode ser feito, portanto, é estimar uma quantidade de compras em que a probabilidade de preenchimento do álbum seja adequada às expectativas do colecionador.

Agora, estudaremos algumas situações:

<u>Situação I</u>:

Queremos calcular quantos chocolates deverão ser comprados para que a probabilidade de obtenção de todas as figurinhas distintas seja de, pelo menos, 90%.

O estudo deve ser feito usando uma noção de complementar. Seguindo o resultado $P(\tau_{fig} > \lceil 30 \cdot ln(30) + c \cdot 30 \rceil) \leq e^{-c}$, vamos tratar que a probabilidade de não preencher o álbum usando até $\lceil 30 \cdot ln(30) + c \cdot 30 \rceil$ compras deva ser menor ou igual a 10%, ou seja,

$$P(\tau_{fig} > \lceil 30 \cdot ln(30) + 30 \cdot c \rceil) \leq 0{,}1.$$

Para isto, faça $e^{-c} = 0{,}1$. Tem-se que:

$$\ln(e^{-c}) = \ln(0{,}1) \Rightarrow -c = \ln(0{,}1) \Rightarrow c = 2{,}3.$$

Segue que o número de compras que interessa é estimado por

$$\lceil 30 \cdot ln(30) + 30 \cdot c \rceil = \lceil 30 \cdot ln(30) + 30 \cdot 2{,}3 \rceil = \lceil 171{,}04 \rceil.$$

Para que a probabilidade de ter todas as 30 figurinhas distintas seja, de pelo menos, 90% é preciso que se compre 172 chocolates.

Observação 5.1. O teste dá um indício de quantas compras serão feitas para que a probabilidade de 90% de chance de preencher o álbum seja atingida. O número estimado de compras pode, entretanto, não ser o melhor possível. Fazendo uso do Teorema 2.1, é possível deduzir que com 168 compras, $P[X = 30] = 0{,}9031$. Já no caso de 172 compras, $P[X = 30] = 0{,}9149$. Concluímos que a estimativa se mostra adequada, já que a diferença entre o caso estimado e o caso de interesse é 4.

<u>Situação II</u>:

Agora, queremos calcular quantos chocolates teriam que ser comprados para que a probabilidade de ter todas as figurinhas ser de, pelo menos, 95%. Com isso a chance de não ter completado o álbum é inferior ou igual a 5%.

Trabalhamos com

$$P(\tau_{fig} > \lceil 30 \cdot ln(30) + 30 \cdot c \rceil) \leq 0{,}05.$$

Portanto,

$$e^{-c} = 0{,}05 \Rightarrow \ln(e^{-c}) = \ln(0{,}05) \Rightarrow c = 3{,}00.$$

Segue que o número que nos interessa é dado por

$$\lceil 30 \cdot ln(30) + 30 \cdot c \rceil = \lceil 30 \cdot ln(30) + 30 \cdot 3 \rceil = \lceil 192{,}04 \rceil.$$

Para que a probabilidade de ter todas as 30 figurinhas distintas seja, de pelo menos, 95% é preciso que se compre 193 chocolates.

<u>Situação III</u>:

E para 99% temos,

$$P\left(\tau_{fig} > \lceil 30 \cdot ln(30) + 30 \cdot c \rceil\right) \leq 0{,}01.$$

Com isso,

$$e^{-c} = 0{,}01 \Rightarrow \ln(e^{-c}) = \ln(0{,}01) \Rightarrow c = 4{,}6.$$

Segue que

$$\lceil 30 \cdot ln(30) + 30 \cdot c \rceil = \lceil 30 \cdot ln(30) + 30 \cdot 4{,}6 \rceil = \lceil 240{,}04 \rceil.$$

E para que a probabilidade de ter todas as 30 figurinhas seja, de pelo menos, 99% é preciso que se compre 241 chocolates.

Conforme o número de compras aumenta, a probabilidade de o colecionador adquirir somente figurinhas inéditas se reduz. Prova disto é que na compra de 6 chocolates, a chance de adquirir 6 figurinhas distintas é 58,63%. Na compra de 7 chocolates, a chance de conseguir todas as figurinhas distintas é 46,90%. Na compra de 8 chocolates é 35,96% a probabilidade de aquisição de 8 figurinhas distintas. Na compra de 20 chocolates é 0,02% a probabilidade de aquisição de 20 figurinhas distintas.

Além disso, as probabilidades de transição

$$P[X_{m+1} = 28 | X_m = 27] = \frac{3}{30}, P[X_{m+1} = 29 | X_m = 28] = \frac{2}{30} \text{ e}$$

$P[X_{m+1} = 30 | X_m = 29] = \frac{1}{30}$ mostram que à medida que o colecionador fica mais perto de completar o álbum de figurinhas, a chance de ele adquirir uma nova figurinha se reduz.

Para que a probabilidade de ter todas as 30 figurinhas distintas ser, de pelo menos, 90%, 95% ou 99%, a estimativa é que se compre 172, 193 ou 241 chocolates, respectivamente. Fica claro que numericamente seria preciso comprar muitos chocolates para ter uma chance viável de fechar a coleção.

Por isso, uma alternativa interessante e muito utilizada por colecionadores é criar grupos onde se faz a troca de figurinhas repetidas, e assim consegue-se completar o álbum de uma forma bem mais barata e rápida.

A próxima seção corrobora que é viável esta possibilidade de trocas de figurinhas repetidas, visto que estuda o número de tentativas médias de achar a última figurinha que completa a coleção.

4.4 COMPRAS PARA OBTER A FIGURINHA QUE COMPLETA UMA COLEÇÃO

Como tratado nas seções anteriores, na compra do chocolate Surpresa podem aparecer 30 figurinhas distintas: $F_1, F_2, F_3, \dots, F_{29}, F_{30}$. Suponhamos que na m-ésima compra apareça a 29ª figurinha distinta da coleção.

Assumamos, portanto, que até a m-ésima compra o colecionador tenha em mãos 29 figurinhas ($F_1, F_2, F_3, \dots, F_{29}$) e que pelo menos uma destas não apareceu de forma repetida (sem perda de generalidade, F_{29} – a última a ser obtida). Desta forma, ao colecionador falta apenas a figurinha F_{30} para que ele complete sua coleção.

Gostaríamos de saber, em média, quantas compras, a partir de agora, serão requisitadas para que o colecionador seja contemplado com a figurinha F_{30}, ou seja, para que ele obtenha a figurinha faltante para completar seu álbum.

A figurinha F_{30} pode aparecer na $1^{\underline{a}}$, na $2^{\underline{a}}$, na $3^{\underline{a}}$ ou na $1000^{\underline{a}}$ compra. Por ser uma obra do acaso, não temos controle sobre a situação.

Para atingir o objetivo da seção, usamos a teoria da distribuição geométrica de probabilidade que encontramos na Ref. [7], pág. 131. Vamos descrever os eventos como "comprar a próxima figurinha". Se adquirirmos a última figurinha faltante temos um sucesso, caso contrário, temos um fracasso.

A probabilidade de obter a última figurinha faltante é de $\frac{1}{30}$ em cada nova compra. Consequentemente, a probabilidade de não conseguir é $1 - \frac{1}{30} = \frac{29}{30}$.

Calculamos a probabilidade de conseguir a figurinha F_{30} em exatamente k tentativas. Isso acontece, portanto, quando obtemos $(k-1)$ fracassos nas primeiras tentativas e sucesso na k-ésima tentativa.

$Y = k$ significa obter a figurinha F_{30} "apenas" na k-ésima compra.

Inicialmente,

$$P[Y = 1] = \frac{1}{30}.$$

A chance de obter a figurinha F_{30} na primeira compra é $\frac{1}{30}$.

Por sua vez,

$$P[Y = 2] = \frac{29}{30} \cdot \frac{1}{30} = \frac{29}{900}.$$

A probabilidade de obter a figurinha F_{30} apenas na segunda compra é $\frac{29}{900}$. Obtém-se, analogamente, que

$$P[Y=3]=\left(\frac{29}{30}\right)^2\cdot\frac{1}{30}=\frac{29^2}{30^3}$$

$$\vdots$$

$$P[Y=k]=\left(\frac{29}{30}\right)^{k-1}\cdot\frac{1}{30}=\frac{29^{k-1}}{30^k}=\frac{1}{30}\cdot\left(\frac{29}{30}\right)^{k-1}.$$

Para encontrar o número médio de compras de figurinhas para obter a faltante, calculamos a esperança da variável aleatória Y.

Assumamos que $q=\frac{29}{30}$. Deduzimos que:

$$E[Y]=\sum_{k=1}^{\infty}k\cdot P[Y=k]=\sum_{k=1}^{\infty}k\cdot\frac{1}{30}\cdot\left(\frac{29}{30}\right)^{k-1}=\frac{1}{30}\sum_{k=1}^{\infty}k\cdot\left(\frac{29}{30}\right)^{k-1}=\frac{1}{30}\sum_{k=1}^{\infty}k\cdot q^{k-1}$$

$$=\frac{1}{30}\sum_{k=1}^{\infty}\frac{d}{dq}(q^k)=\frac{1}{30}\cdot\frac{d}{dq}\left(\sum_{k=0}^{\infty}q^k\right)=\frac{1}{30}\cdot\frac{d}{dq}\left(\frac{1}{1-q}\right)=\frac{1}{30}\cdot\frac{1}{(1-q)^2}$$

$$=\frac{1}{30}\cdot\frac{1}{\left(1-\frac{29}{30}\right)^2}=\frac{\frac{1}{30}}{\left(\frac{1}{30}\right)^2}=30.$$

Espera-se pela teoria da distribuição geométrica de probabilidade que, em média, 30 compras sejam suficientes para aparecer a figurinha F_{30} faltante. O que mostra uma quantidade considerável de figurinhas a serem compradas para conseguir a última figurinha faltante da coleção. Assim, fica claro mais um motivo para incentivarmos a troca de figurinhas repetidas para aqueles que desejam completar uma coleção.

5 UMA PROPOSTA DE ABORDAGEM DE CADEIAS DE MARKOV NO ENSINO MÉDIO

Neste capítulo, desenvolvemos uma proposta de aplicação de variáveis aleatórias e cadeias de Markov no Ensino Médio. Os dois exemplos detalhados no Capítulo 3 serão transmitidos como conhecimentos em nível escolar.

Resumidamente, nossa proposta consiste em aplicar seis exercícios distribuídos em três atividades para estudantes de uma turma de 3º ano de Ensino Médio de uma escola pública estadual da cidade de Blumenau. É recomendado que a aplicação ocorra após a turma dominar assuntos prévios como probabilidade e probabilidade condicional. Além disso, é esperado que antes da abordagem seja feita uma revisão do tema Matrizes, normalmente trabalhado no 2º ano do Ensino Médio. Para tratar das Cadeias de Markov, requisita-se, portanto, aspectos como representação, classificação e operações com matrizes. A multiplicação de matrizes é primordial para o desenvolvimento da proposta.

A ideia é que os exercícios ajudem na introdução ao assunto e, por meio de cada um deles, os conceitos associados serão apresentados e aprofundados. O objetivo é que os exercícios sejam solucionados em classe, com participação efetiva dos alunos e intermediação por parte do docente. Os alunos serão instigados a utilizar calculadora e computador em momentos oportunos.

No **Apêndice A – Roteiro de Atividades** encontra-se o plano de trabalho descrevendo a aplicação. No **Apêndice B – Gabarito do Roteiro de Atividades** está descrito o gabarito dos exercícios propostos. Sugere-se a utilização de aproximadamente seis aulas de 45 minutos cada para a aplicação da proposta.

Inicialmente, entretanto, mostramos a fundamentação da nossa proposta de aplicação estruturada na Base Nacional Comum Curricular (BNCC) e nos Parâmetros Curriculares Nacionais (PCNs) do Ensino Médio.

Como já comentado, a proposta de aplicação consiste em exercícios com perguntas que induzem o estudante a criar, coletivamente, a aprendizagem. Questão a questão, o estudante aprofunda a compreensão dos conceitos. Nesta proposta, o aluno deve participar ativamente na construção do conhecimento. Ele explora a construção dos modelos e o docente vai guiando-o a partir do momento que ele ultrapassa cada etapa. Como o estudante atuará desta forma, cabe ao professor entender que a resolução pode não ser única e a socialização das estratégias é bastante oportuna e enriquecedora. Tal proposta vem ao encontro do texto da Base Nacional Comum Curricular. Conforme a BNCC,

> Para que esses propósitos se concretizem nessa área, os estudantes devem desenvolver habilidades relativas aos processos de investigação, de construção de modelos e de resolução de problemas. Para tanto, eles devem mobilizar seu modo próprio de raciocinar, representar, argumentar, comunicar e, com base em discussões e validações conjuntas, aprender conceitos e desenvolver representações e procedimentos cada vez mais sofisticados.

No decorrer das atividades propostas, os alunos tratam com situações que envolvem os conceitos de probabilidade, probabilidade condicional e matrizes. Desenvolvem procedimentos matemáticos para a resolução dos problemas, seguindo um caminho usual para encontrar as probabilidades, com a construção do espaço amostral e eventos. Gradativamente, um outro método de resolução através das variáveis aleatórias e Cadeias de Markov começa a ser explorado. A abordagem proposta não se preocupa em formalizar com detalhes a teoria (assim como feito em outros capítulos desse trabalho) – o estudante é instigado a entender como utilizar um método. De acordo com a BNCC:

> Os estudantes também precisam construir significados para os problemas próprios da Matemática. Para resolver esses problemas, eles devem, logo no início, identificar os conceitos e procedimentos matemáticos necessários ou os que possam ser utilizados, na chamada formulação matemática do problema. Depois disso, eles devem aplicar esses conceitos, executar procedimentos e, ao final, compatibilizar os resultados com o problema original, comunicando a solução aos colegas por meio de argumentação consistente.

O aluno vai investigar, explicar e justificar os problemas resolvidos durante todo o processo. A ideia é que no primeiro momento, ele desenvolva a resolução de exemplos usando conhecimentos previamente abordados (no caso, as noções usuais de probabilidade). Na sequência, ele será introduzido a uma nova estrutura (aqui, as Cadeias de Markov) e verá duas possibilidades distintas de resolução que o levam a mesma resposta. A construção de um novo modelo de resolução para determinados problemas de probabilidade permite ao aluno que encontre as respostas de forma mais eficaz e rápida. Ao estudante, entretanto, devem ficar claras as condições que permitem a utilização deste novo método e as eventuais impossibilidades em utilizá-lo.

Apesar dos problemas trabalhados não configurarem necessariamente situações práticas da vida do discente, eles abordam situações que podem indicar modelos e ferramentas úteis em outros contextos. Para a Base Nacional Comum Curricular:

> Devem-se considerar, também, os problemas cujas tarefas não estão explícitas e para as quais os estudantes deverão mobilizar seus conhecimentos e habilidades, a fim de

> identificar conceitos e conceber um processo de resolução. Em alguns desses problemas, os estudantes precisam identificar ou construir um modelo para que possam gerar respostas adequadas. Esse processo envolve analisar os fundamentos e propriedades de modelos existentes, avaliando seu alcance e validade para o problema em foco.

Nossa proposta aprofunda o que a legislação prevê para a abordagem da matemática no Ensino Médio. Ressaltamos que para algumas atividades envolvendo o cálculo de probabilidade será solicitada a descrição do espaço amostral, eventos e a contagem das possibilidades. A BNCC destaca os tópicos importantes no conteúdo de probabilidades no Ensino Médio:

> (EM13MAT311) Resolver e elaborar problemas que envolvem o cálculo da probabilidade de eventos aleatórios, identificando e descrevendo o espaço amostral e realizando contagem das possibilidades.
> (EM13MAT312) Resolver e elaborar problemas que envolvem o cálculo de probabilidade de eventos em experimentos aleatórios sucessivos.

Em geral, os assuntos probabilidades e matrizes são tratados de forma independente no Ensino Médio. A proposta de aplicação visa estabelecer conexões entre esses dois conteúdos, atendendo, de alguma forma, às noções de relação e inter-relação esperadas na BNCC, Ref. [23], pág. 520. De acordo com a referência, o aluno realmente demonstra que compreendeu os conceitos trabalhados quando transita serenamente entre diferentes conteúdos, relacionando-os logicamente. Os PCNs, por sua vez, apontam a importância de estabelecer relações entre diferentes temas matemáticos:

> O critério central é o da contextualização e da interdisciplinaridade, ou seja, é o potencial de um tema permitir conexões entre diversos conceitos matemáticos e entre diferentes formas de pensamento matemático, ou, ainda, a relevância cultural do tema, tanto no que diz respeito às suas aplicações dentro ou fora da Matemática, como à sua importância histórica no desenvolvimento da própria ciência.

Finalmente, a utilização de recursos tecnológicos nessa proposta de aplicação é uma alternativa facilitadora na aprendizagem do estudante. A calculadora para transformar números fracionários em números decimais. E o computador, com o programa de Excel, para efetuar multiplicações de matrizes. A BNCC mostra a importância do uso desses recursos:

> O uso de tecnologias possibilita aos estudantes aprofundar sua participação ativa nesse processo de resolução de problemas. São alternativas de experiências variadas e facilitadoras de aprendizagens que reforçam a capacidade de raciocinar logicamente, formular e testar conjecturas, avaliar a validade de raciocínios e construir argumentações.

A seguir, temos as três atividades propostas para os alunos do 3º ano do Ensino Médio, com as devidas reflexões.

5.1 REFLEXÕES SOBRE A PROPOSTA DE APLICAÇÃO

A primeira atividade proposta consiste em retiradas de bolas de dentro de uma urna, a segunda está relacionada a lançamentos de um dado e a terceira atividade relata a migração de pessoas em regiões da cidade de Blumenau. As duas últimas atividades estão detalhadas no Capítulo 3. A seguir, relata-se as 6 aulas utilizadas para aplicação do projeto.

Inicia-se a primeira aula da aplicação, propondo que os estudantes leiam e busquem solucionar a primeira questão enunciada no roteiro:

1) *Uma urna contém 3 bolas de cores diferentes: 1 bola AZUL, 1 bola VERMELHA e 1 bola PRETA.*

Retira-se, ao acaso, uma bola da urna e identifica-se sua cor. Essa bola será reposta na urna e o processo de retirada de bola será efetuado novamente. Em matemática, entende-se que são feitas ***duas retiradas com reposição.***

Baseado no enunciado:

a) *Apresente o espaço amostral que descreve todas as possibilidades de cores de bolas após duas retiradas com reposição.*
b) *Qual a probabilidade de aparecer a mesma cor de bola nas duas retiradas?*
c) *Qual a probabilidade de retirada de duas bolas de cores diferentes?*

A questão 1 revê o assunto probabilidade – servindo como uma revisão e também introdução ao novo tema trabalhado. Nessa questão, os alunos descreverão manualmente todas as possibilidades de cores de bolas após duas retiradas.

O professor pode identificar as maneiras como os estudantes deduzem o espaço amostral e fazê-los socializar com a turma. Pode aparecer diagramas em árvore ou simples listas, por exemplo. Mas, é importante que os alunos denotem corretamente o espaço amostral usando uma linguagem de conjuntos.

Deve ser enfatizado que as retiradas são com reposição. Se fosse considerada uma situação sem reposição, o espaço amostral seria outro. Esse aspecto pode ser comparado, por exemplo, quando for feita a correção do item (a).

Com o item (a) solucionado, a ideia é que se identifique os eventos que interessam aos itens (b) e (c). E que, consequentemente, os alunos sejam capazes de calcular as probabilidades solicitadas.

Claro que dúvidas podem surgir e o professor deve estar atento para respondê-las. A ideia é dar tempo para os alunos resolverem a questão, que pode ser feita individualmente ou em grupo. Realizada a questão 1, um ou dois alunos podem apresentar sua resolução para a turma. É importante escrever as respostas no quadro, pois, isto pode minimizar as dúvidas dos demais discentes.

Na sequência, a questão 2, cuja abordagem pode ser iniciada na primeira aula e concluída ao fim da segunda:

2) *Uma urna contém 3 bolas de cores diferentes: 1 bola AZUL, 1 bola VERMELHA e 1 bola PRETA. Seguindo o enunciado da atividade anterior, agora serão feitas* ***três retiradas com reposição****.*
   a) *Apresente o espaço amostral que descreve todas as possiblidades de cores de bolas após três retiradas com reposição.*
   b) *Qual a probabilidade de aparecer a mesma cor de bola nas três retiradas?*
   c) *Qual a probabilidade de retirada de três bolas de cores diferentes?*

Na questão 2, os alunos não deverão ter grandes dificuldades em solucionar os itens, pois, eles deverão perceber que há uma similaridade com a questão anterior. Obviamente, os alunos tenderão a notar que o trabalho será maior, pois, há mais retiradas envolvidas e, portanto, mais elementos no espaço amostral. Orienta-se a fazer uma resolução repetindo as mesmas estratégias da questão anterior.

Feita a correção da questão 2, algumas reflexões podem ser propostas: pode-se calcular o número de elementos presentes no espaço amostral usando princípio de contagem. Veja que o número de elementos do espaço amostral é $3 \cdot 3 \cdot 3 = 27$, já que há 3 possibilidades de cores diferentes para as 1ª, 2ª e 3ª retiradas. Além disso, pode-se determinar por contagem também o número de elementos dos eventos que interessam aos itens (b) e (c). Na letra (b), o número de possibilidades de aparecer a mesma cor de bola nas três retiradas é 3: todas azuis, vermelhas ou pretas. Já na letra (c), existe $3 \cdot 2 \cdot 1 = 6$ possibilidades para retiradas de 3 bolas com cores distintas, pois há 3, 2 e 1 possibilidades de cores de bolas diferentes para a 1ª, 2ª e 3ª retirada, respectivamente.

Perceba que as probabilidades dos itens (b) e (c) da questão 1 são complementares. Na questão 2, as probabilidades dos itens (b) e (c) não são complementares, pois são 3 retiradas, e existem outros elementos no espaço amostral que não são três bolas com a mesma cor ou três bolas de cores diferentes. Essa reflexão pode ser pontuada aos estudantes.

Ao finalizar a questão 2, o professor pode questionar os alunos sobre quais resultados obteríamos se fizéssemos quatro retiradas de bola da urna. E discutir as dificuldades que poderiam ser enfrentadas para o cálculo de probabilidades. Pode-se refletir que a probabilidade de encontrar uma única cor de bola pode ser facilmente estabelecida usando os princípios de contagem (como comentado acima). Mas, os demais casos poderão requerer uma análise mais detalhada. Uma das estratégias para facilitação no estudo, entretanto, é usar a teoria de Cadeias de Markov.

Na terceira aula, antes de apresentar a questão 3, o docente pode embasar teoricamente o que são as cadeias de Markov. Aqui, o professor pode comentar, usando uma linguagem simples e acessível ao Ensino Médio, que a Cadeia de Markov é um processo associado às probabilidades condicionais, em que o seu estado futuro depende apenas do seu estado no presente, não importando aquilo que ocorreu no passado. A ideia aqui, portanto, não é se aprofundar no conceito de variáveis aleatórias e Cadeias de Markov (assim como feito no texto deste trabalho), mas, citar que há maneiras mais simples de abordar a situação eventualmente instigada ao fim da segunda aula.

É importante evidenciar que o problema apresentado nas questões anteriores se adequa à proposta de uma cadeia de Markov, no sentido que com as retiradas sucessivas de bolas, podem aparecer uma, duas ou três bolas de cores diferentes. O aparecimento ou não de uma nova cor de bola após uma retirada depende apenas da quantidade de cores identificadas até o momento daquela retirada.

O professor pode alegar que o estudo de probabilidades usando cadeia de Markov requer a identificação de uma matriz e esse processo é explicado no texto da questão 3:

3) *Os exemplos que envolvem a reposição de elementos podem ter propriedades de probabilidade estudadas com auxílio de matrizes.*

Voltamos às questões 1 e 2. Em um urna, estão 3 bolas de cores diferentes: 1 bola AZUL, 1 bola VERMELHA e 1 bola PRETA. Serão retiradas, com reposição, bolas desta urna. Antes de iniciar o processo, se observa zero cor de bola – uma vez que nenhuma bola *foi retirada da urna para análise de cor. Com o passar das retiradas, dependendo do acaso, poderão ser observadas bolas de 1, 2 ou 3 cores distintas.*

O número a_{ij} *descreve a probabilidade condicional de passarmos de* i *cores vistas para j cores em uma determinada retirada de bola da urna.*

Por exemplo,

$a_{00} = 0$ *e* $a_{01} = 1$*, pois, em uma retirada, passaremos obrigatoriamente do caso 0 cor observada para 1 cor observada. De forma similar, conclui-se que* $a_{02} = 0$ *e* $a_{03} = 0$*, pois, é impossível passar de 0 cor para duas ou três cores em uma única retirada da bola da urna.*

$a_{10} = 0$*, pois, se já vemos uma cor de bola, após uma retirada, poderemos ver 1 ou 2 (nunca zero). Aqui aparece uma situação interessante. Já se conhece a existência de uma cor de bola. Na próxima retirada, há uma chance em três de aparecer a mesma cor de bola. Há duas chances em três de aparecer uma segunda cor de bola. Conclui-se, portanto, que* $a_{11} = \frac{1}{3}$ *e* $a_{12} = \frac{2}{3}$*. Destaca-se também que* $a_{13} = 0$*.*

Seguindo o procedimento, considera-se $A = \left(a_{ij}\right)_{4x4}$*, uma matriz de transição:*

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Essa matriz é útil para o cálculo de probabilidades relativas ao número de cores distintas que aparecem após n *retiradas, com reposição, de bolas da urna.*

A questão 3 traz uma explicação referente a construção da matriz de transição, mostrando passo a passo a dedução de algumas probabilidades de transição. Não há nada a ser completado pelo estudante. Trata-se apenas da exposição do tema.

Após a leitura, espera-se que o professor, usando o quadro, ajude o estudante a compreender a construção da matriz A detalhada no texto.

Neste momento, certamente os alunos ficarão surpresos, pois terão que o número a_{ij}, visto como a localização de um elemento na linha i e coluna j, agora descreve a probabilidade condicional de passarmos de i cores vistas para j cores em uma determinada retirada de bola da urna. E há a formação de uma matriz quadrada com todos os elementos representados por probabilidades condicionais. O docente pode reforçar que os assuntos estudados em matemática têm aplicabilidade em diversos contextos e o uso de matrizes não se resume à resolução de sistemas lineares.

Na quarta aula, sugere-se que o estudante resolva o exercício 4 que fecha a Atividade 1. Nesse exercício, é apresentada uma parte que explica a utilização do método das Cadeias de Markov, uma forma didática de apresentar o Teorema 2.1 deste trabalho para a compreensão dos alunos do terceiro ano do Ensino Médio. Na questão 4, há 5 itens a serem feitos em conjunto com os alunos:

4) *Na matriz linha* $1X4$: $\begin{pmatrix} 1^{\underline{o}} & 2^{\underline{o}} & 3^{\underline{o}} & 4^{\underline{o}} \end{pmatrix}$ *os elementos representam:*
   - *1º elemento: probabilidade de aparecer zero cores na retirada das bolas.*
   - *2º elemento: probabilidade de aparecer apenas uma única cor na retirada das bolas.*
   - *3º elemento: probabilidade de aparecer duas cores distintas na retirada das bolas.*
   - *4º elemento: probabilidade de aparecer três cores distintas na retirada das bolas.*

Antes de iniciar o processo de retirada de bolas da urna há, certamente 0 cor vista. Então, a probabilidade de se observar 0 quantidade de cores é 100%. $\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$ *é a matriz que caracteriza as probabilidades antes de iniciar o processo de retirada de bolas da urna.*

O produto da matriz linha $\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$ *com a matriz quadrada de transição, destacada acima, gera uma nova matriz linha. Tal matriz caracteriza as probabilidades de encontrarmos as quantidades distintas de cores* $0, 1, 2, ou\ 3$.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix}$$

Analisando os elementos dessa nova matriz linha, percebe-se que em <u>uma retirada</u> a chance de se verificar uma única cor é de 100% (independente se a cor for azul, vermelha ou preta, é certo que exatamente uma cor será destacada após a primeira retirada de bola da urna).

O produto da matriz linha encontrada com a matriz de transição quadrada gera uma nova matriz linha. Essa nova matriz descreve a probabilidade de aparecerem cores distintas de bolas, após duas retiradas, com reposição. Segue abaixo:

$$\begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}$$

Analisando os elementos dessa nova matriz linha, temos que em duas retiradas, com reposição, a chance de aparecer uma única cor é de $\frac{1}{3}$ e a chance de aparecer duas cores diferentes é de $\frac{2}{3}$. Baseado nessa leitura, faça o que se pede:

a) *Determine o seguinte produto matricial:*

$$\begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} = (\frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad})$$

b) *O produto do item a) resulta uma matriz linha. Qual o significado de cada um dos elementos da matriz?*

c) *Faça o produto da matriz linha, encontrada no item a), com a matriz de transição, conforme a seguinte estrutura:*

$$(\frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad}) \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} = (\frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad})$$

d) *O produto do item c) resulta uma matriz linha. Qual o significado de cada um dos elementos da matriz?*

e) *Suponha que sejam feitas as retiradas de cinco bolas da urna, com reposição. Determine a probabilidade de, após as cinco retiradas, aparecer um única cor de bola (obviamente, não importando qual). Determine a probabilidade de, após cinco retiradas, aparecer só a bola vermelha.*

A questão 4 traz uma explicação sobre os elementos da matriz linha, que conhecemos como vetor estado. É importante que o professor retome adição e multiplicação de frações com denominadores diferentes, caso os alunos tenham dúvidas. E permita que os estudantes utilizem a calculadora como um recurso facilitador, para encontrar os resultados na sua forma decimal e em porcentagem.

Veja que dois produtos matriciais foram feitos como modelo e o aluno deve solucionar outros dois nos itens (a) e (c). Aqui, seria interessante o professor comentar que as probabilidades da matriz linha $\begin{pmatrix} 0 & \frac{1}{3} & \frac{2}{3} & 0 \end{pmatrix}$ trazem informações identificadas manualmente na questão 1. Ao resolver os itens (a) e (b), o aluno deve compreender que as entradas da matriz linha $\begin{pmatrix} 0 & \frac{1}{9} & \frac{6}{9} & \frac{2}{9} \end{pmatrix}$ correspondem às probabilidades calculadas manualmente na questão 2. Neste ponto, inclusive, aparece a informação adicional de que $\frac{6}{9}$ é a probabilidade referente a obter duas cores distintas na retirada de 3 bolas de dentro da urna. Essa retomada das questões 1 e 2, como verificação de resultados, reforça com os alunos a validação do processo de Markov.

Ao abordar os itens (c) e (d), o docente pode refletir que as entradas do novo vetor linha descrevem as probabilidades de identificar 0, 1, 2 ou 3 cores distintas após quatro retiradas de bolas com reposição. Deve-se relatar que descrever o espaço amostral e os eventos manualmente para encontrar as probabilidades para 4, 5, 6 ou mais retiradas de bolas de dentro da urna é um caminho trabalhoso e inviável. E que o método descrito na atividade é eficiente e prático para encontrar as probabilidades solicitadas quando trata-se de um número considerável de retiradas.

Na sequência, o professor pode mostrar e explicar que no programa do Excel é possível fazer multiplicações de matrizes, o que torna o software um recurso facilitador nesse estudo. Apesar de outros programas também permitirem o cálculo de produtos matriciais, o Excel é um dos mais acessíveis, inclusive, nos laboratórios de informática presentes nas escolas. A parte inicial do item (e) da questão 4, portanto, pode ser feita com auxílio do computador.

O último tópico proposto no item (e) da questão 4 não deve ser solucionado usando o processo de Markov. É válido para que os alunos reflitam os casos em que tal abordagem deve ser evitada.

Ao finalizar a questão 4, o professor pode rever os resultados de probabilidade associados a retirada de quatro ou cinco bolas da urna e instigar os discentes a encontrar informações relativas às probabilidades ao se fazer 6 retiradas de bolas ou mais. Por isso, deve ser reforçado que os 4 elementos da matriz linha correspondem as probabilidades de aparecerem 0, 1, 2 ou 3 cores distintas na retirada de bolas e, que a 1ª, 2ª, 3ª, 4ª, 5ª,... matriz linha, geradas nas multiplicações matriciais a partir de $\begin{pmatrix}1 & 0 & 0 & 0\end{pmatrix}$, estão relacionadas a $1,\ 2,\ 3,\ 4,\ 5,$ ... número de bolas retiradas de dentro da urna, respectivamente.

Na quinta aula, a ideia é trabalharmos uma nova atividade. A Atividade 2 é baseada na Seção 3.1 deste trabalho. Iniciamos, sugerindo que o estudante leia e busque solucionar à questão 5:

5) Um dado não viciado tem seis faces: 1, 2, 3, 4, 5 e 6.

Sabe-se que, 2, 3 e 5 são números primos e 1, 4 e 6 não são primos. Após uma quantidade de lançamentos, queremos estudar a probabilidade de que apareçam primos distintos na face voltada para cima.

Para isso, faça o que se pede:

a) Suponha que sejam feitos dois lançamentos desse dado. Determine o espaço amostral que descreve as maneiras possíveis de vermos os pares de faces voltadas para cima.

b) Determine, após duas jogadas, a probabilidade de não aparecer face com número primo voltada para cima.

c) Determine, após duas jogadas, a probabilidade de aparecer duas faces com números primos distintos voltadas para cima.

d) Construa a matriz de transição associada ao modelo de lançamentos de dados e a observação de 0, 1, 2 ou 3 números primos distintos voltadas para cima.

e) Use a matriz determinada no item d), e complete os espaços destacados.

Em três lançamentos do dado, a probabilidade de não sair número primo é de _________ e a probabilidade de sair três números primos distintos é______________.

Em quatro lançamentos do dado, a probabilidade de não sair número primo é de ___________ e a probabilidade de sair três números primos distintos é______________.

Na Atividade 2, os estudantes terão liberdade e tempo para criar e desenvolver suas estratégias de resolução, e assim, compartilhar ideias e soluções com os colegas da turma.

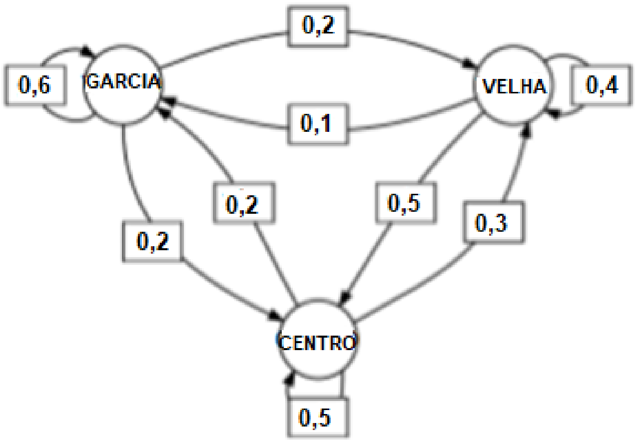
O professor pode retomar com os alunos o conceito de números primos. Nos itens (a), (b) e (c), o aluno se depara com a resolução tradicional do problema para dois lançamentos do dado. No item (d), ele precisa construir a matriz de transição $4x4$ para esse problema, que consiste nos elementos a_{ij}, que retratam a probabilidade condicional de passarmos de i números primos distintos vistos para j números primos distintos em um determinado lançamento do dado. Lembrando que $i, j \in \{0,1,2,3\}$.

O docente pode interagir levando em consideração os seguintes questionamentos: Quais os elementos da matriz linha para zero lançamentos e o que cada um deles representa? Quais serão as matrizes linha associadas a 1, 2, 3 e 4 lançamentos de um dado? Que processo deve ser feito para obtenção dessas matrizes? Tais questionamentos podem indicar, portanto, a dinâmica para o fechamento desta atividade. Feitos esses esclarecimentos, o item (e) pode ser completado.

Ao fim da resolução do Exercício 5, outras duas perguntas podem ser levantadas: É possível responder as probabilidades solicitadas no item (e) construindo os espaços amostrais e os eventos de forma tradicional (como aprendido anteriormente à aplicação)? Se sim, por que não optamos por este caminho? Espera-se que o aluno compreenda que o método tradicional de resolução é trabalhoso e pode se tornar inviável. Ao término da Atividade 2, objetiva-se que os alunos compreendam que o método da cadeia de Markov é eficaz para a resolução de uma classe de problemas que envolvem experimentos aleatórios sucessivos.

Por fim, na sexta aula, trabalhamos uma nova atividade, com uma abordagem distinta das anteriores. A Atividade 3 é baseada na Seção 3.2 deste trabalho. Iniciamos, sugerindo que o estudante leia e busque solucionar à questão 6:

6) *A cidade de Blumenau é dividida em três regiões: Região 1 – Centro, Região 2 – Garcia , Região 3 – Velha. Os moradores de Blumenau, ao final de cada ano, permanecem na sua região ou migram para uma outra dentro da cidade. O diagrama abaixo apresenta as probabilidades de permanecer em um mesmo bairro ou de mudar para outro. A probabilidade de migração de moradores de Blumenau para outras cidades é desprezível.*



a) Determine a matriz de transição que caracteriza as migrações, ano a ano, dos moradores de Blumenau dentro das regiões que constituem a cidade.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}$$

b) Encontre a probabilidade de um morador do Centro, após três anos, estar residindo no bairro Garcia.

A Atividade 3 retrata uma situação diferente das anteriores, pois, há um esboço em que as probabilidades de transição são dadas sem a necessidade de dedução – apesar de abordar uma situação hipotética. A cidade de Blumenau foi usada como referência, pois, é o local em que a aplicação está sendo planejada. Recomenda-se que os alunos sejam alertados se tratar de uma situação fictícia, mas, que a ideia pode ser adaptada para o estudo de outras situações, inclusive realistas.

No item (a), portanto, o número a_{ij} descreve a probabilidade condicional de uma pessoa migrar da região i para a região j em um determinado período de observação. Os alunos não terão dificuldades em construir a matriz de transição, pois, as probabilidades de transição são todas extraídas facilmente do diagrama proposto no enunciado.

Para solucionar o item (b), o estudante deve lembrar que é preciso um vetor linha, com três entradas – que representa o estado inicial. O aluno pode ser instigado por questionamentos como: Quais os elementos da matriz linha para o momento atual, em que ainda não houve migração? Quais os elementos das matrizes linha que caracterizam cada uma das regiões da cidade? Quais as matrizes linha que descrevem as probabilidades após as 1ª, 2ª e 3ª migrações? Qual o processo para deduzirmos essas matrizes? Se assim desejarem, os estudantes podem usar ferramentas como calculadora e Excel para a obtenção das respostas.

Antes de finalizar a aplicação, o estudante pode ser instigado a pensar ou a propor situações reais de estudo em que as cadeias de Markov poderiam ser viáveis para a resolução de problemas.

Em caso de dificuldades de ideias por parte dos discentes, o professor pode citar a Ref. [4], pág. 28, que aborda a seguinte situação: há três usinas hidrelétricas em funcionamento e uma quarta usina que está desativada com a possibilidade de ser reativada. Todas as usinas possuem a mesma capacidade de produção. Pretende-se descobrir se a reativação da quarta usina é uma decisão sustentável e qual a melhor forma de se fazer a reativação. Para isso, se utiliza as Cadeias de Markov para realizarem simulações a fim de encontrar a melhor maneira de realizar a transferência da produção das três usinas para a quarta usina. Na caracterização, cria-se um tabela com a porcentagem de energia transferida entre as usinas e a quantidade de energia que permanece em cada uma delas. Esta tabela atua como a matriz de transição. A matriz linha inicial possui como elementos a quantidade de energia gerada inicialmente por cada usina. E assim, realizando os produtos matriciais, sabe-se ao longo dos meses qual a produção de energia de cada usina. Partindo das matrizes fornecidas, pode-se realizar várias outras simulações, a fim de alcançar o melhor resultado possível no âmbito socioambiental.

A depender da disponibilidade e interesse dos estudantes, desdobramentos da atividade poderiam ser aplicados a fim de aprofundar o estudo do tema Variáveis Aleatórias e Cadeias de Markov.

6 CONSIDERAÇÕES FINAIS

O objetivo desse trabalho foi realizar um estudo sobre cadeias de Markov e algumas aplicações. Três problemas bastante distintos entre si foram apresentados ao leitor.

As situações-problema estudadas ajudaram no entendimento das cadeias de Markov e também mostraram sua aplicabilidade. Em especial, com o problema de "completar o álbum de figurinhas" foi possível encontrar uma quantidade de compras necessárias para obter uma chance razoável de completar uma coleção de 30 figurinhas, adquiridas na compra dos chocolates Surpresa – o que resulta uma importante ferramenta para previsões. Também chamou a atenção a estimativa obtida através de inúmeras desigualdades que prevê uma eventual quantidade de compras usando uma estratégia simples e eficaz. Tais estratégias são pouco exploradas em nível de ensino básico e podem aproximar a matemática à realidade do estudante.

Como professora de Matemática, sempre gostei de trabalhar com o assunto de probabilidades, e por meio das cadeias de Markov foi possível ter um conhecimento mais amplo sobre este conteúdo. O estudo sobre o processo de Markov sanou dúvidas que tinha sobre probabilidade condicional. E mostrou uma relação entre os assuntos probabilidades e matrizes que não conhecia.

Acredito que a abordagem de Cadeias de Markov seja viável no Ensino Médio. Com conhecimento prévio de matrizes e probabilidades, o docente pode aplicar a proposta desenvolvida no Capítulo 5 sem grandes dificuldades. Destacamos, entretanto, que o assunto de matrizes não tem sido reforçado na recente Base Nacional Comum Curricular (BNCC) e isso pode ser um dificultador para futuras aplicações.

Como docente, terei um olhar diferenciado sobre o conteúdo de probabilidades, e quando retornar as aulas presenciais, pretendo aplicar a proposta com meus alunos do 3°ano do Ensino Médio. Ressalto que a ideia inicial era de fazer a aplicação a uma turma que leciono e relatar a experiência dos alunos com o tema. Devido a pandemia da Covid-19, entretanto, decidiu-se que a apresentação e detalhamento da proposta seriam suficientes para refletir o estudo.

Esperamos que esse trabalho seja uma opção de consulta para outras pessoas que forem estudar Variáveis Aleatórias, Cadeias de Markov e aplicações. E que seja utilizado por professores do Ensino Médio que queiram introduzir o conceito das Cadeias de Markov em suas aulas. Que ele sirva como inspiração!

REFERÊNCIAS

[1] ALENCAR, Marcelo Sampaio. **Probabilidade e processos estocásticos**. Editora Érica: São Paulo, 2008.

[2] ANTON, Howard; RORRES, Chris. **Álgebra linear com aplicações**. 10. ed. Bookman: Porto Alegre, 2012.

[3] ATUNCAR, Gregório S. **Conceitos Básicos de Processos Estocásticos.** Material Didático – Universidade Federal de Minas Gerais, Belo Horizonte, 2011.

[4] BARROS NETO, Ávido S. **Cadeias de Markov no Ensino Médio.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Federal de Sergipe, São Cristóvão, 2013.

[5] CASTRO, Diogo M. S. **Cadeias de Markov: Uma aplicação para o Ensino de Matrizes e Probabilidades.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Federal de Alagoas, Maceió, 2015.

[6] COSTA, Fábio S. **Cadeias de Markov Regulares: Uma abordagem para alunos e professores do Ensino Médio.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Federal do Amazonas, Manaus, 2017.

[7] LOESCH, Claudio; STEIN, Carlos Efrain. **Estatística Descritiva e Teoria das Probabilidades.** 2. Ed. EDIFURB: Blumenau, 2011.

[8] MAGELA, Mateus M. **Teoria Básica das Cadeias de Markov.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) - Universidade Federal do Espírito Santo, Vitória, 2015.

[9] MAIA, Leonardo P. **Uma introdução à dinâmica estocásticas de populações**. SBMAC: São Carlos, 2008 – (Notas em Matemática Aplicada; v. 35).

[10] MANOEL, Marcelo R. **Cadeias de Markov: Uma abordagem voltada para o Ensino Médio.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) - Universidade Estadual de Campinas, 2016.

[11] MARCHESINI, André L. S. **Probabilidade e Cadeias de Markov: Uma proposta para os Ensinos Fundamental e Médio.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Federal do ABC, Santo André, 2017.

[12] MARIOTTO, Rafael P. **Introdução às Variáveis Aleatórias e Cadeias de Markov.** Trabalho de Conclusão de Curso (Licenciatura em Matemática) – Universidade Federal de Santa Catarina, Florianópolis, 2009.

[13] MEDEIROS, Sérgio S. **Cadeias de Markov Ocultas.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Federal do ABC, Santo André, 2017.

[14] MORETTIN, Pedro A.; BUSSAB, Wilton de O. **Estatística Básica.** 9. ed. Saraiva: São Paulo, 2017.

[15] MORGADO, Leandro B.; BORGES, Leonardo S. **Álbuns de figurinhas: uma abordagem via Cadeias de Markov**. REMAT: Revista Eletrônica de Matemática, v. 5, n.1, p. 101-113, 2019.

[16] OLIVEIRA, José C. **Noções de Grafos dirigidos, Cadeias de Markov e as buscas do Google.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Federal de Sergipe, São Cristóvão, 2014.

[17] RAMOS, Yuri T. **Aplicações de Cadeias de Markov no Ensino Médio.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Estadual de Campinas, 2017.

[18] RIBEIRO, Thaís S. G. **Processos de Markov discretos: exemplos voltados para o Ensino Médio.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) - Universidade Estadual Paulista, São José do Rio Preto, 2017.

[19] RODRIGUES, Welton C. **Cadeias de Markov: Uma aula para alunos do Ensino Médio.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) - Universidade Federal de Juiz de Fora, 2013.

[20] SILVA, Carlos E. V. **Aplicações da Álgebra Linear nas Cadeias de Markov.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) - Universidade Federal de Goiás, Goiânia, 2013.

[21] SOUZA JUNIOR, Fernando L. **Cadeias de Markov e o jogo Monopoly.** Dissertação (Mestrado Profissional em Matemática – PROFMAT) – Universidade Federal do ABC, Santo André, 2016.

[22] TAHA, Hamdy A. **Pesquisa Operacional.** Pearson Prentice Hall, São Paulo, 2008.

[23] **Base Nacional Comum Curricular Ensino Médio.** Disponível em: <http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=85121-bncc-ensino-medio&category_slug=abril-2018-pdf&Itemid=30192>. Acesso em: 13 de julho de 2020.

[24] **Imagem de Andrei Markov**. Disponível em: <https://commons.wikimedia.org/wiki/File:Andrei_Markov.jpg >. Acesso em: 05 de agosto de 2020.

[25] **Imagem do álbum de figurinhas**. Disponível em: <http://www.ikessauro.com/2008/04/album-marcante_8659.html >. Acesso em: 06 de outubro de 2020.

[26] **Parâmetros Curriculares Nacionais Ensino Médio.** Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/ciencian.pdf>. Acesso em: 13 de julho de 2020.

[27] **Saudades? Relembre todas as coleções do chocolate Surpresa.** Disponível em: <https://super.abril.com.br/mundo-estranho/saudades-relembre-todas-as-colecoes-do-chocolate-surpresa>. Acesso em: 06 de outubro de 2020.

APÊNDICE A – Roteiro de Atividades

ATIVIDADE 1

1) Uma urna contém 3 bolas de cores diferentes: 1 bola AZUL, 1 bola VERMELHA e 1 bola PRETA.

   Retira-se, ao acaso, uma bola da urna e identifica-se sua cor. Essa bola será reposta na urna e o processo de retirada de bola será efetuado novamente. Em matemática, entende-se que são feitas **duas retiradas com reposição.**

   Baseado no enunciado:

   a) Apresente o espaço amostral que descreve todas as possibilidades de cores de bolas após duas retiradas com reposição.
   b) Qual a probabilidade de aparecer a mesma cor de bola nas duas retiradas?
   c) Qual a probabilidade de retirada de duas bolas de cores diferentes?

2) Uma urna contém 3 bolas de cores diferentes: 1 bola AZUL, 1 bola VERMELHA e 1 bola PRETA. Seguindo o enunciado da atividade anterior, agora serão feitas **três retiradas com reposição**.

   a) Apresente o espaço amostral que descreve todas as possiblidades de cores de bolas após três retiradas com reposição.
   b) Qual a probabilidade de aparecer a mesma cor de bola nas três retiradas?
   c) Qual a probabilidade de retirada de três bolas de cores diferentes?

3) Os exemplos que envolvem a reposição de elementos podem ter propriedades de probabilidade estudadas com auxílio de MATRIZES.

   Voltamos às questões 1 e 2. Em um urna, estão 3 bolas de cores diferentes: 1 bola AZUL, 1 bola VERMELHA e 1 bola PRETA. Serão retiradas, com reposição, bolas desta urna. Antes de iniciar o processo, se observa zero cor de bola – uma vez que nenhuma bola foi retirada da urna para análise de cor. Com o passar das retiradas, dependendo do acaso, poderão ser observadas bolas de 1, 2 ou 3 cores distintas.

   O número a_{ij} descreve a probabilidade condicional de passarmos de i cores vistas para j cores em uma determinada retirada de bola da urna.

   Por exemplo,

$a_{00} = 0$ e $a_{01} = 1$, pois, em uma retirada, passaremos obrigatoriamente do caso 0 cor observada para 1 cor observada. De forma similar, conclui-se que $a_{02} = 0$ e $a_{03} = 0$, pois, é impossível passar de 0 cor para duas ou três cores em uma única retirada da bola da urna.

$a_{10} = 0$, pois, se já vemos uma cor de bola, após uma retirada, poderemos ver 1 ou 2 (nunca zero). Aqui aparece uma situação interessante. Já se conhece a existência de uma cor de bola. Na próxima retirada, há uma chance em três de aparecer a mesma cor de bola. Há duas chances em três de aparecer uma segunda cor de bola. Conclui-se, portanto, que $a_{11} = \frac{1}{3}$ e $a_{12} = \frac{2}{3}$. Destaca-se também que $a_{13} = 0$.

Seguindo o procedimento, considera-se $A = \left(a_{ij}\right)_{4x4}$, uma matriz de transição:

$$\begin{pmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & a_{21} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Essa matriz é útil para o cálculo de probabilidades relativas ao número de cores que aparecem após n retiradas, com reposição, de bolas da urna. O método será explicado abaixo.

4) Na matriz linha $1x4$: $(1^{\underline{o}} \quad 2^{\underline{o}} \quad 3^{\underline{o}} \quad 4^{\underline{o}})$ os elementos representam:
   - 1º elemento: probabilidade de aparecer zero cores na retirada das bolas.
   - 2º elemento: probabilidade de aparecer apenas uma única cor na retirada das bolas.
   - 3º elemento: probabilidade de aparecer duas cores distintas na retirada das bolas.
   - 4º elemento: probabilidade de aparecer três cores distintas na retirada das bolas.

Antes de iniciar o processo de retirada de bolas da urna há, certamente 0 cor vista. Então, a probabilidade de se observar 0 quantidade de cores é 100%. $(1 \quad 0 \quad 0 \quad 0)$ é a matriz que caracteriza as probabilidades antes de iniciar o processo de retirada de bolas da urna.

O produto da matriz linha $(1 \quad 0 \quad 0 \quad 0)$ com a matriz quadrada de transição, destacada acima, gera uma nova matriz linha. Tal matriz caracteriza a probabilidades de encontrarmos as quantidades distintas de cores $0, 1, 2$ ou 3.

$$(1 \quad 0 \quad 0 \quad 0)\cdot\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & \frac{1}{3} & \frac{2}{3} & 0\\ 0 & 0 & \frac{2}{3} & \frac{1}{3}\\ 0 & 0 & 0 & 1\end{pmatrix} = (0 \quad 1 \quad 0 \quad 0)$$

Analisando os elementos dessa nova matriz linha, percebe-se que em uma retirada a chance de se verificar uma única cor é de 100% (independente se a cor for azul, vermelha ou preta, é certo que exatamente uma cor será destacada após a primeira retirada de bola da urna).

O produto da matriz linha encontrada com a matriz de transição quadrada gera uma nova matriz linha. Essa nova matriz descreve a probabilidade de aparecerem cores distintas de bolas, após duas retiradas, com reposição. Segue abaixo:

$$(0 \quad 1 \quad 0 \quad 0)\cdot\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & \frac{1}{3} & \frac{2}{3} & 0\\ 0 & 0 & \frac{2}{3} & \frac{1}{3}\\ 0 & 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}0 & \frac{1}{3} & \frac{2}{3} & 0\end{pmatrix}$$

Analisando os elementos dessa nova matriz linha, temos que em duas retiradas, com reposição, a chance de aparecer uma única cor é de $\frac{1}{3}$ e a chance de aparecer duas cores diferentes é de $\frac{2}{3}$. Baseado nessa leitura, faça o que se pede:

a) Determine o seguinte produto matricial:

$$\begin{pmatrix}0 & \frac{1}{3} & \frac{2}{3} & 0\end{pmatrix}\cdot\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & \frac{1}{3} & \frac{2}{3} & 0\\ 0 & 0 & \frac{2}{3} & \frac{1}{3}\\ 0 & 0 & 0 & 1\end{pmatrix} = (\frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad})$$

b) O produto do item a) resulta uma matriz linha. Qual o significado de cada um dos elementos da matriz?

c) Faça o produto da matriz linha, encontrada no item a), com a matriz de transição, conforme a seguinte estrutura:

$$(\frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad})\cdot\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & \frac{1}{3} & \frac{2}{3} & 0\\ 0 & 0 & \frac{2}{3} & \frac{1}{3}\\ 0 & 0 & 0 & 1\end{pmatrix} = (\frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad} \quad \frac{\quad}{\quad})$$

d) O produto do item c) resulta uma matriz linha . Qual o significado de cada um dos elementos da matriz?

e) Suponha que sejam feitas as retiradas de cinco bolas da urna, com reposição. Determine a probabilidade de, após as cinco retiradas, aparecer um única cor de bola (obviamente, não importando qual). Determine a probabilidade de, após cinco retiradas, aparecer só a bola vermelha.

ATIVIDADE 2

5) Um dado não viciado tem seis faces: 1, 2, 3, 4, 5 e 6.

Sabe-se que, 2, 3 e 5 são números primos e 1, 4 e 6 não são primos. Após uma quantidade de lançamentos, queremos estudar a probabilidade de que apareçam primos distintos na face voltada para cima.

Para isso, faça o que se pede:

a) Suponha que sejam feitos dois lançamentos desse dado. Determine o espaço amostral que descreve as maneiras possíveis de vermos os pares de faces voltadas para cima.

b) Determine, após duas jogadas, a probabilidade de não aparecer face com número primo voltada para cima.

c) Determine, após duas jogadas, a probabilidade de aparecer duas faces com números primos distintos voltadas para cima.

d) Construa a matriz de transição associada ao modelo de lançamentos de dados e a observação de 0, 1, 2 ou 3 faces distintas voltadas para cima.

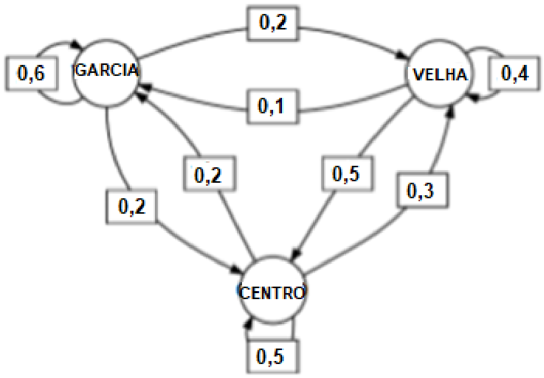
Use a matriz determinada no item d), complete os espaços destacados.

Em três lançamentos do dado, a probabilidade de não sair número primo é de _________ e a probabilidade de sair três números primos distintos é _______________.

Em quatro lançamentos do dado, a probabilidade de não sair número primo é de ___________ e a probabilidade de sair três números primos distintos é_______________.

ATIVIDADE 3 - DESAFIO

6) A cidade de Blumenau é dividida em três regiões: Região 1 – Centro, Região 2 – Garcia, Região 3 – Velha. Os moradores de Blumenau, ao final de cada ano, permanecem na sua região ou migram para uma outra dentro da cidade. O diagrama abaixo apresenta as probabilidades de permanecer em um mesmo bairro ou de mudar para outro. A probabilidade de migração de moradores de Blumenau para outras cidades é desprezível.



a) Determine a matriz de transição que caracteriza as migrações, ano a ano, dos moradores de Blumenau dentro das regiões que constituem a cidade.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow \begin{pmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{pmatrix}$$

b) Encontre a probabilidade de um morador do Centro, após três anos, estar residindo no bairro Garcia.

APÊNDICE B – Gabarito do Roteiro de Atividades

1)

a) $S = \{(A,V);(A,P);(A,A);(V,A);(V,P);(V,V);(P,A);(P,V);(P,P)\}$.

b) $P(A) = \frac{3}{9} = \frac{1}{3}$

c) $\mathrm{P}(\bar{A}) = 1 - \frac{1}{3} = \frac{2}{3}$

2)

a) $$S = \begin{Bmatrix} (A,V,A);(A,V,V);(A,V,P);(A,P,A);(A,P,V);(A,P,P);(A,A,A); \\ (A,A,V);(A,A,P);(V,A,A);(V,A,V);(V,A,P);(V,P,A);(V,P,V); \\ (V,P,P);(V,V,A);(V,V,V);(V,V,P);(P,A,A);(P,A,V);(P,A,P); \\ (P,V,A);(P,V,V);(P,V,P);(P,P,A);(P,P,V);(P,P,P). \end{Bmatrix}$$

b) $P(B) = \frac{3}{27} = \frac{1}{9}$

c) $P(C) = \frac{6}{27} = \frac{2}{9}$

3) Apenas leitura.

4)

a) $\begin{pmatrix} 0 & \frac{1}{9} & \frac{6}{9} & \frac{2}{9} \end{pmatrix}$

b) Em três retiradas de bolas de dentro da urna, a probabilidade de sair 0 cores é 0, de sair uma única cor é $\frac{1}{9}$, sair duas cores distintas é $\frac{6}{9}$ e sair três cores distintas é $\frac{2}{9}$.

c) $$\begin{pmatrix} 0 & \frac{1}{9} & \frac{6}{9} & \frac{2}{9} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} & 0 \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{27} & \frac{14}{27} & \frac{12}{27} \end{pmatrix}$$

d) Em quatro retiradas de bolas de dentro da urna, a probabilidade de sair 0 cores é 0, de sair uma única cor é $\frac{1}{27}$, sair duas cores distintas é $\frac{14}{27}$ e sair três cores distintas é $\frac{12}{27}$.

e) A probabilidade de, após as cinco retiradas, aparecer um única cor de bola é $\frac{1}{81}$. E a probabilidade de, após cinco retiradas, aparecer só a bola vermelha é $\frac{1}{243}$.

5)

a) $$S = \begin{Bmatrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{Bmatrix}$$

b) Após duas jogadas a probabilidade de não aparecer face com número primo voltada para cima é $\frac{1}{4}$.

c) Após duas jogadas, a probabilidade de aparecer duas faces com números primos distintos voltadas para cima é $\frac{1}{6}$.

d) $$P = \begin{pmatrix} p_{00} & p_{01} & p_{02} & p_{03} \\ p_{10} & p_{11} & p_{12} & p_{13} \\ p_{20} & p_{21} & p_{22} & p_{23} \\ p_{30} & p_{31} & p_{32} & p_{33} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 2/3 & 1/3 & 0 \\ 0 & 0 & 5/6 & 1/6 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

e) Em três lançamentos do dado, a probabilidade de não sair número primo é de $\frac{1}{8}$ e a probabilidade de sair três números primos distintos é $\frac{1}{36}$. Em quatro lançamentos do dado, a probabilidade de não sair número primo é de $\frac{1}{16}$ e a probabilidade de sair três números primos distintos é $\frac{1}{12}$.

6)

a) $$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \rightarrow \begin{pmatrix} 0{,}5 & 0{,}2 & 0{,}3 \\ 0{,}2 & 0{,}6 & 0{,}2 \\ 0{,}5 & 0{,}1 & 0{,}4 \end{pmatrix}$$

b) A probabilidade de um morador do Centro, após três anos, estar residindo no bairro Garcia é 0,269.