Fernando Luiz de Souza Junior

Cadeias de Markov e o Jogo *Monopoly*

Santo André, 2016

Universidade Federal do ABC

Centro de Matemática, Computação e Cognição

Fernando Luiz de Souza Junior

Cadeias de Markov e o Jogo *Monopoly*

Orientador: Prof. Dr. Rafael de Mattos Grisi

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

Este exemplar corresponde a versão final da dissertação defendida pelo aluno Fernando Luiz de Souza Junior, e orientada pelo Prof. Dr. Rafael de Mattos Grisi.

Santo André, 2016

Ficha Catalográfica elaborada pela Biblioteca da Universidade Federal do ABC

SOUZA JUNIOR, Fernando Luiz de

Cadeias de Markov e o Jogo Monopoly / Fernando Luiz de Souza Junior – Santo André: Universidade Federal do ABC, 2016.

94 fls. : il

Orientador: Prof. Dr. Rafael de Mattos Grisi

Dissertação (Mestrado) – Universidade Federal do ABC. Mestrado Profissional em Matemática em Rede Nacional – PROFMAT, Santo André, 2016.

1. Princípios de probabilidade 2. Cadeias de Markov 3. Medida invariante 4. Ergodicidade 5. Jogo Monopoly I. SOUZA JUNIOR, Fernando Luiz de II. Mestrado Profissional em Matemática em Rede Nacional – PROFMAT, 2016 III. Cadeias de Markov e o Jogo Monopoly

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

FOLHA DE ASSINATURAS

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Fernando Luiz de Souza Junior, realizada em 21 de setembro de 2016:

Prof.(a) Dr.(a) **Rafael de Mattos Grisi** (UFABC) – Presidente

Daniel Miranda Machado

Prof.(a) Dr.(a) **Daniel Miranda Machado** (UFABC) – Membro Titular

Paola A. Kassama

Prof.(a) Dr.(a) **Paola Andrea Gaviria Kassama** (UNIFESP) – Membro Titular

Prof.(a) Dr.(a) **Jeferson Cassiano** (UFABC) – Membro Suplente

Prof.(a) Dr.(a) **Gleiciane da Silva Aragão** (UNIFESP) – Membro Suplente

Dedico este trabalho à minha esposa Jânia, que sempre esteve ao meu lado, lembrando-me que nada é impossível para aquele que crê em si mesmo.

AGRADECIMENTOS

Agradeço a Deus, por ter conservado minhas forças nesta caminhada.

Agradeço a meus pais, Elza e Fernando, por não medirem esforços em me ajudar na busca pelo conhecimento.

Agradeço à Universidade Federal do ABC por ter me proporcionado tal riquíssima experiência de voltar a estudar Matemática, aprendendo com os melhores.

Agradeço aos caríssimos amigos Erik, Jairo e Kléber, pelo auxílio nas horas difíceis, pelo companheirismo, pela união.

Agradeço a todos os professores do PROFMAT, pelas excelentes aulas e pela paciência com os alunos.

E agradeço especialmente ao Professor Rafael Grisi, pela sua competência, seriedade e sabedoria.

"All we have to decide is what to do with the time that is given us."

(J.R.R. Tolkien, *The Fellowship of the Ring*)

RESUMO

Neste trabalho analisamos uma versão simplificada do jogo Monopoly utilizando um modelo de Cadeia de Markov com parâmetro de tempo discreto. No primeiro capítulo discorremos sobre a Teoria Clássica das Probabilidades, trazendo os resultados mais importantes para este estudo, precedida por uma breve introdução acerca das ideias sobre o acaso ao longo da história da humanidade e os principais pensadores envolvidos no desenvolvimento dessa Teoria. No segundo capítulo fazemos uma introdução histórica aos processos estocásticos e às Cadeias de Markov; em seguida, explicamos os conceitos fundamentais sobre Cadeias de Markov, colocando alguns exemplos e por fim discutindo a ergodicidade de uma Cadeia de Markov. No terceiro capítulo, após uma breve explicação sobre o surgimento e posterior evolução do jogo Monopoly ao longo do século XX, analisamos a dinâmica do jogo pelo modelo de uma Cadeia de Markov, utilizando como objeto de estudo uma versão mais simples do jogo em questão.

Palavras-chave: Cadeias de Markov, medida invariante, ergodicidade

ABSTRACT

In this work we analyze a simplified version of the Monopoly game using a Markov chain model with discrete time parameter. In the first chapter we discuss on the Classical Theory of Probability, bringing the most important results for this study, preceded by a brief introduction about the ideas of chance throughout the history of mankind and leading thinkers involved in the development of this theory. In the second chapter we make a historical introduction to stochastic processes and Markov chains; then we explain the fundamental concepts of Markov Chains, putting some examples and finally discussing the ergodicity of a Markov chain. In the third chapter, after a brief explanation of the emergence and subsequent evolution of the Monopoly game throughout the twentieth century, we analyze the dynamics of the game by the model of a Markov chain, using as an object of study a simpler version of the game in question.

Keywords: Markov chains, invariant measure, ergodicity

CONTEÚDO

INTRODUÇÃO

> *"A humanidade precisou de centenas de anos para se acostumar com um mundo onde alguns eventos não tinham causa...ou eram determinados por causas tão remotas que somente podiam ser razoavelmente representados por modelos probabilísticos."*
>
> — Maurice g. Kendall

O que é o acaso? Diante de pergunta tão simples, nossa mente imediatamente é convidada a analisar todos os acontecimentos que acontecem à sua volta, e de novo questionar: de tudo o que se vê, se sente ou se percebe, quais são as coisas realmente regidas pelo acaso? Haverá uma lei geral de tudo o que está em nosso redor?

Supondo que exista o acaso e que existam situações que são caracterizadas pela sua presença, é possível, mesmo aceitando o acaso como fato, avançar na compreensão desse acaso, caminhando pelas trilhas da investigação e método científico? Será que o acaso permite algum grau de determinação?

Antes de tudo é necessário especificar ou conceituar o que se entende por "acaso". Ao lançarmos uma moeda, por exemplo, é senso comum que os possíveis resultados, exceto por alguma extravagância da natureza, são "cara" e "coroa". No entanto, antes de realizada a experiência não é possível antecipar com certeza qual dos dois possíveis resultados irá ocorrer. Isto acontece porque os fatores que determinam um destes particulares resultados não podem ser identificados e caso isto ocorra não são passíveis de controle.

A ideia de acaso é quase tão antiga quanto as primeiras civilizações, só que a percepção de que isto é um fenômeno natural e passível de explicações científicas, veio ocorrer bem mais tarde. Inicialmente o acaso era percebido como fruto ou obra da divindade e atualmente, devemos admitir, a ideia de existirem fenômenos sem causa definida ainda encontra dificuldade para ser aceita; basta notar a popularidade alcançada por pessoas que dizem prever o futuro, dentro de determinados grupos da sociedade contemporânea.

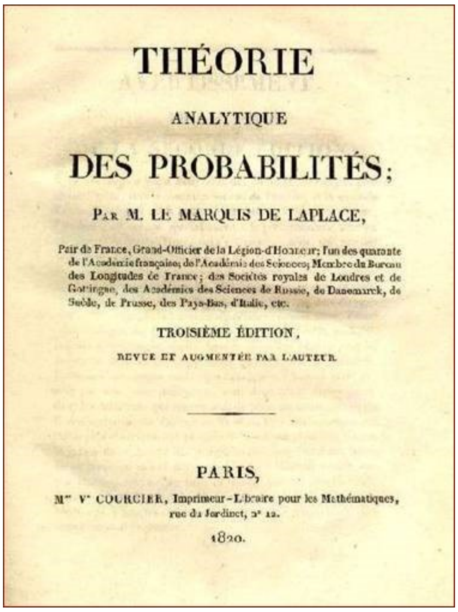
Para escaparmos ao mero senso comum, que ora tende para a extrema simplificação, ora tende para o místico, podemos entender o acaso, e por conseguinte a probabilidade, sob o ponto de vista formal. Desse modo, de acordo com Viali em [22], devemos entender o termo "acaso" como "[...] um conjunto de forças, em geral, não determinadas ou controladas, que exercem individualmente ou coletivamente papel preponderante na ocorrência de diferentes resultados de um experimento ou fenômeno. "

Para exemplificar, tomemos o lançamento de um dado, no qual se obtém qualquer uma das seis faces e o "acaso" está na impossibilidade de prever quais faces serão observadas ao longo de novos lançamentos, não podendo identificar seus resultados previamente. O "acaso" relaciona-se também com muitos outros exemplos, como o lançamento de uma moeda, a determinação da vida útil de um equipamento eletrônico, as previsões do tempo, etc. Desta maneira, observamos como o "acaso" pode estar associado aos jogos de azar (nos quais a chance de ganhar não depende da habilidade do jogador), aos fenômenos naturais ou aos eventos ocorridos no cotidiano.

No entanto, não podemos nos afastar por completo das abordagens mais intuitivas e até ingênuas do conceito de probabilidade; tais abordagens, do ponto de vista de uma pedagogia da matemática que entenda ser o objetivo último da ciência que ela possa ser acessível e entendida por todos, têm importante valor. Considerar o esmero didático que deve ser empreendido para que essas várias discussões sobre aleatoriedade, chance e percepção do universo que nos cerca cheguem ao entendimento dos estudantes, através de eficazes estratégias de ensino, é absolutamente fundamental.

As visões divergentes sobre a compreensão do que é "probabilidade" e "evento aleatório" remontam à Antiguidade; diversos físicos, filósofos e matemáticos dedicaram-se nessa tarefa de dar sentido a esse debate, e quantas correntes distintas de pensamento foram originadas! Ingenuidade seria a nossa pensar que, no ato de ensinar tais conceitos, sobrepujasse de maneira natural e incontestável na mente dos alunos os conceitos de probabilidade clássica tais como os colocados na obra *Teoria Analítica das Probabilidades*, de Pierre S. Laplace.

Com efeito, na Matemática certas suposições fundamentais são feitas e delas são deduzidas as conclusões, segundo o raciocínio lógico formal, garantindo assim validade ao nosso pensamento. Entretanto, nem todos os pensamentos são matemáticos, grande parte das crenças não é certa, apenas provável. Mas dizer que um acontecimento é determinado pelo acaso é declarar que não se sabe como ele é determinado. Mesmo assim, no "reino do acaso", percebe-se certa regularidade, certa "ordem dentro



THÉORIE
ANALYTIQUE
DES PROBABILITÉS;

PAR M. LE MARQUIS DE LAPLACE,

Pair de France, Grand-Officier de la Légion-d'Honneur; l'un des quarante de l'Académie française; de l'Académie des Sciences; Membre du Bureau des Longitudes de France; des Sociétés royales de Londres et de Gottingue, des Académies des Sciences de Russie, de Danemarck, de Suède, de Prusse, des Pays-Bas, d'Italie, etc.

TROISIÈME ÉDITION,
REVUE ET AUGMENTÉE PAR L'AUTEUR.

PARIS,
M^me V^e COURCIER, Imprimeur-Libraire pour les Mathématiques, rue du Jardinet, n° 12.
1820.

Figura 1: Teoria Analítica das Probabilidades (1ª edição em 1812), que traz a definição de probabilidade matemática tal qual a encontramos atualmente nos livros didáticos de matemática do Ensino Básico

da desordem" e acerca de acontecimentos atribuídos ao acaso formamos uma gradação de crença racional.

Em geral, os julgamentos na tomada de decisão são feitos por dedução provável, nos negócios, na mesa de jogos, em um processo de júri, até mesmo em experimentos. Costumamos dizer em dias quentes e nebulosos que provavelmente choverá. Um meteorologista pode precisar de maiores evidências, mas não o cidadão comum. Costuma-se raciocinar assim, no senso comum, em assuntos que vão do mais trivial ao mais importante, com o uso frequente das palavras "provável" e "probabilidade", sem precisar seu significado.

De fato, segundo Bennett em [3], se a incerteza em um processo aleatório é fruto da nossa ignorância a respeito das forças que determinam o seu desfecho ou da existência de mecanismos inacessíveis inerentes aos recursos e condições que o circunscrevem e que determinam o seu desfecho, estas são questões que estão no centro da discussão filosófica envolvendo as noções de aleatoriedade e chance e que continuam em debate até os dias de hoje.

Em qual seara deveremos então caminhar? Como investigar a verdade que subjaz oculta no âmago dos acontecimentos aleatórios? Qual o equilíbrio que devemos encontrar entre a formalização matemática, a observação dos fatos, a intuição geral e as concepções filosóficas?

Claro está que esta lide ainda se encontra em processo, mas dentro das ciências exatas, mais especificamente dentro da Matemática, deveremos optar por um caminho que leve à construção lógica da verdade, que seja naturalmente sistematizável e traduzível para uma linguagem simbólica. A tentativa de se estabelecer um elo entre a Matemática e a intuição geral, falando de Probabilidades, passa por voltar às origens de todas essas discussões: retornar ao elemento lúdico, ao jogo, à percepção que existe uma chance de ganhar e uma chance de perder em um jogo.

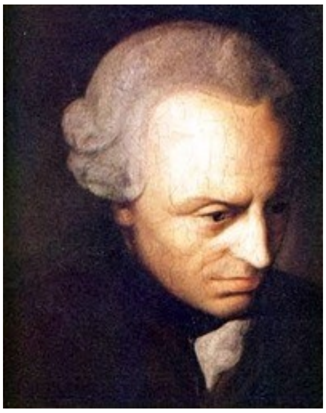


Figura 2: "*Todo o conhecimento humano começou com intuições, passou daí aos conceitos e terminou com ideias*" - Immanuel Kant (1724 – 1804)

Portanto, à luz dessas considerações iniciais, a proposta desse trabalho será, ao analisar um objeto lúdico, poder discorrer sobre os conceitos matemáticos da Teoria Clássica da Probabilidade, e ir além, fazendo uso de uma ferramenta um pouco mais moderna do ponto de vista teórico e histórico para aprofundar tal análise, adentrando nos domínios da Estatística e da Teoria Moderna da Probabilidade.

No Capítulo 1 traremos uma breve introdução histórica sobre a Teoria da Probabilidade, destacando os principais nomes de vulto na construção científica nesta área. Em seguida, faremos uma introdução clássica às Probabilidades, trazendo situações-exemplo e formalizando os elementos básicos desse sistema teórico, chegando às Probabilidades da União e da Intersecção de eventos.

No Capítulo 2 abordamos os processos estocásticos, que descrevem sistemas que variam em algum grau, de forma imprevisível, à medida que o tempo passa; é apresentada ao leitor uma rápida digressão sobre a natureza desses processos e colocados alguns exemplos. Dentre os vários processos estocásticos, uma Cadeia de Markov é aquele no qual o conjunto de estados do sistema é discreto, onde a probabilidade futura de um estado depende unicamente de seu estado corrente. Estudamos, então, as principais características de uma Cadeia de Markov. Compreenderemos algumas aplicações básicas, formalizaremos os elementos fundamentais e sistematizaremos toda a parte inicial dessa teoria, chegando à discussão sobre a ergodicidade de uma Cadeia.

No Capítulo 3 trazemos uma aplicação do estudo feito nos capítulos anteriores, ao analisar a dinâmica de uma versão simplificada do internacionalmente famoso jogo *Monopoly*. É possível enxergar o desenrolar do jogo como um processo iterativo do tipo estocástico, mais precisamente como a progressão temporal de uma Cadeia de Markov. Oferecemos ao leitor uma sucinta visão sobre o desenvolvimento do jogo, desde os seus primórdios até os dias atuais. Fazemos então algumas considerações sobre características do jogo que o tornam interessante para análise, estudando algumas iterações para n razoavelmente grande, dentro de uma versão mais simples desse jogo; chegamos, por fim, a algumas conclusões parciais.

1

PRINCÍPIOS DE PROBABILIDADE

"Se o ensino de Matemática se deve ocupar mais de uma forma de pensar do que de uma forma de escrever fórmulas ou numerais, se o ensino da Matemática se deve ocupar mais da tomada consciente de decisões do que do estrito cálculo, então a teoria das probabilidades é fundamental."

— Bernardes, 1987

1.1 PREÂMBULO

1.1.1 *Sobre as Ideias*

O desenvolvimento matemático da teoria das probabilidades foi uma invenção do homem que permitiu a ele fazer afirmações definitivas acerca de acontecimentos cujos resultados não podiam ser determinados com antecedência. Em outras palavras, a teoria das probabilidades forneceu um método quantitativo para avaliar fenômenos randômicos. Todos aqueles que já jogaram cartas, lançaram dados ou moedas, investem na bolsa ou dirigem numa autoestrada percebem que a aleatoriedade desempenha um papel fundamental nas experiências do dia-a-dia.

A física clássica lida com eventos de resultado final bem determinado, ou pelo menos com eventos cujo desfecho é razoavelmente previsto. No entanto, a teoria das probabilidades abarca aqueles eventos nos quais o desfecho é incerto. Qualquer pessoa pode afirmar, impunemente, que o sol se levantará amanhã, que objetos próximos à superfície da terra serão atraídos gravitacionalmente por ela, que a água vai ferver após aquecida até uma certa temperatura, que a corrente elétrica existirá em um fio condutor se existirem locais nesse condutor com potenciais elétricos distintos. As leis da física inventadas pelo homem que explicam esses fatos parecem funcionar muito bem, e seria surpreendente se elas fossem violadas.

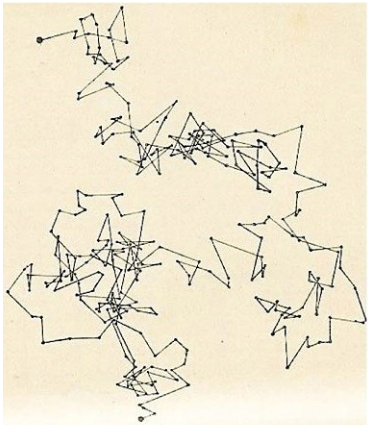


Figura 3: Partículas em Movimento Browniano, uma das aplicações da moderna teoria da probabilidade em Física

Em contraponto com esses acontecimentos com resultados bem determinados, existem acontecimentos nos quais o resultado final é completamente desconhecido antes que o evento ocorra de fato. Se uma moeda honesta é lançada, ninguém pode dizer com certeza antecipada se sairá cara ou coroa. Antes de viajar, nenhum de nós poderá afirmar com certeza que não ocorrerá um acidente com o veículo que nos transporta. Ninguém poderá afirmar com toda a certeza qual será o tempo atmosférico na cidade daqui a uma semana.

Vamos supor que estamos interessados em determinar a probabilidade de sair cara quando uma moeda honesta é lançada. Se tivermos paciência, podemos jogar a moeda um número muito grande de vezes. Se após n lançamentos nós tivermos obtido r_n caras, podemos então formar a razão r_n/n, que indica a proporção de caras em relação ao número total de lançamentos. Definimos então

$$p = \lim_{n \to \infty} \frac{r_n}{n}$$

como sendo a probabilidade de ocorrer cara no lançamento dessa moeda. Infelizmente, nós não podemos de fato definir esse limite precisamente, pois r_n *não* é uma função determinada por n. Essa é a grande questão básica subjacente à teoria de probabilidades.

E já que estamos falando em ideias, algumas ideias aparentemente triviais podem "pregar peças" em gente gabaritada...

Vamos pensar em uma questão aparentemente simples: qual resultado é mais provável no lançamento simultâneo de duas moedas: duas caras, duas coroas ou uma de cada? Para esta pergunta, algumas pessoas poderiam pensar (equivocadamente) que os três resultados são equiprováveis, cada um com probabilidade de $1/3$. É comum deixarem de considerar que sair faces distintas nesse tipo de lançamento de moedas tem probabilidade dobrada do que faces iguais.

Entretanto, mesmo o conceituado matemático do século XVIII, Jean Le Rond D'Alembert, à época um dos mais influentes cientistas franceses, sustentou que a probabilidade de se conseguir uma cara em dois lançamentos consecutivos era de $2/3$, por entender que haveria apenas três casos possíveis equiprováveis, ou seja, $C,\ KC$ e KK (K representa o resultado coroa e C, cara). Na verdade, ao pensar num primeiro lançamento sendo cara, nem cogitou o segundo lançamento neste caso, por já ter obtido o resultado esperado...

Em Teoria da Probabilidade, talvez por essa aproximação histórica com o senso comum e o pensamento do cidadão médio, todo o cuidado é pouco - inclusive nos conceitos mais basilares.

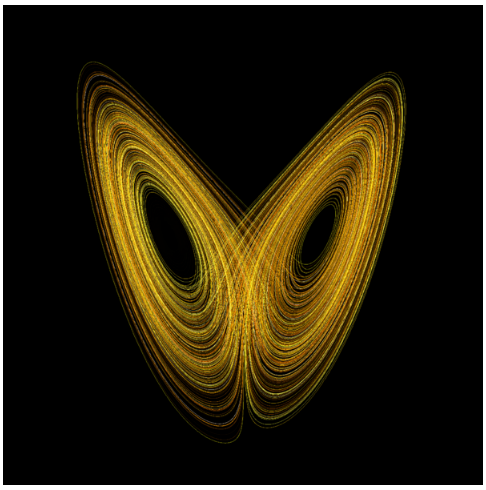


Figura 4: Representação gráfica das Equações de Edward Lorenz (1963) – o famoso "Efeito Borboleta"

Por fim, é importante lembrar que com o desenvolvimento da teoria moderna das probabilidades, a aproximação entre Probabilidades e Estatística foi se estreitando cada vez mais. Teorias sobre regularidades estatísticas estudadas em fenômenos aleatórios fortaleceram a concepção de que há uma *teoria da desordenação ordenada* que cerca o imprevisível e o define de certa forma. No Capítulo 2 deste trabalho estudaremos os rudimentos da teoria das *Cadeias de Markov*, que é historicamente um dos primeiros esforços dos estatísticos e probabilistas em compreender essa "arrumação no caos" e atribuir significado a ela.

1.1.2 *Um pouco de História*

A humanidade tem lidado com a incerteza desde épocas as mais remotas na tentativa de obter vantagens em disputas e evitar perdas advindas de fatores imprevisíveis. Há milhares de anos jogos de azar têm sido parte de nossa civilização. Pinturas em tumbas egípcias feitas em 3500 a.C. mostram pessoas jogando uma forma primitiva de dados feitos de um osso do calcanhar de nome *astragalus*.

Dados de 6 faces datados de 3000 a.C. foram encontrados no norte do Iraque. Durante as Cruzadas vários jogos de dados foram trazidos para o Ocidente (a palavra "azar" sendo derivada de *al zahr*, que significa "dado" em árabe). O baralho moderno surgiu na França no século XIV. Por outro lado, levantamentos de dados estatísticos para censos populacionais e avaliação de produções agrícolas foram realizados na Europa a partir do século XI; apólices de seguros navais baseados em conceitos de risco foram emitidos na Itália e Holanda já no século XIV. No período dos séculos XV e XVI, além de considerações filosóficas sobre causalidade e acaso, várias investigações de problemas relativos a jogos de azar, ou, mais geralmente, a eventos sujeitos ao acaso, foram realizadas de forma esparsa; maiores detalhes podem ser encontrados em [7].

O frei **Luca Pacioli (1445 – 1517)**, embora não tenha publicado nada de original é reconhecido pela obra *Summa de arithmetica, geometria, proportioni e proportionalità* (Veneza, 1494), que o tornou conhecido para a história do desenvolvimento da probabilidade, se firmando como um tratado grandioso para a época. Pacioli apresentava na *Summa* conteúdos de Aritmética, Álgebra, Geometria e Trigonometria e é o primeiro autor conhecido que estudou os jogos de azar. Estudou ainda o problema dos pontos (divisão da aposta), embora sua solução tenha sido incorreta, conforme lemos em [4, 22].

O autodidata **Niccolo de Fontana de Brescia (1499 – 1557)**, também conhecido como *Tartaglia*, publicou sua obra *General Trattato*, em 1556, que dedicava algumas páginas aos problemas propostos por Pacioli, dentre eles o problema dos pontos. Se coloca assim tal problema: um jogo equitativo termina quando um dos jogadores vencer seis partidas. Suponha-se que por algum motivo o jogo tenha que ser interrompido quando o primeiro jogador tenha vencido cinco partidas e o segundo apenas três. Como as apostas devem ser repartidas? Niccolo argumentou que a divisão deveria ser 5:3, que não é correta; a solução correta para esse problema foi dada mais tarde por Fermat e Pascal, de acordo com o que se encontra em [4, 22].

O médico **Girolamo Cardano (1501 – 1576)**, que também era um inveterado jogador de cartas, dados e xadrez, publicou o livro *Liber de Ludo Aleae*, que na realidade era um manual de jogos de azar. Contudo, ele foi o primeiro a estudar o lançamento de dados, baseado na hipótese de que existia um princípio científico governando as probabilidades de se obter um par de "seis", além de mera sorte. Foi também o primeiro a introduzir técnicas de combinatória no cálculo dos casos possíveis de um evento e considerar a probabilidade de um evento como a razão entre o número de casos favoráveis e o número de casos possíveis, vide [4, 22].

O destacado cientista italiano **Galileo Galilei (1564 – 1642)** também publicou um manual sobre jogos, o *Considerações sobre o Jogo de Dados*. Nesta obra ele fornece a famosa explicação da razão pela qual, embora sejam seis as somas que fornecem nove pontos no lançamento de três dados e que também sejam seis as que fornecem dez pontos, a experiência mostra que a soma dez é mais comum de ocorrer que a soma nove.

O filósofo, matemático e teólogo francês **Blaise Pascal (1623 – 1662)** teve conhecimento do problema dos pontos (divisão da aposta) através de Chevalier de Mèrè, um intelectual francês apaixonado por jogos e influente na corte de Luís XIV, e que ficou imortalizado por essa sua participação na gênese da teoria da probabilidade. Para resolver o problema dos pontos era preciso técnicas mais apuradas envolvendo um grande número de possibilidades, que foi o que Pascal fez. Esse e outros problemas motivaram a correspondência entre Pascal e **Pierre de Fermat (1601 – 1665)**, um advogado de grande erudição e talvez um dos mais brilhantes matemáticos amadores que já existiu. As correspondências entre Pascal e Fermat foram publicadas em 1679, e desde então têm sido consideradas a origem do desenvolvimento da teoria matemática da probabilidade. Pascal e Fermat foram os pioneiros na solução de problemas genéricos em probabilidade, e a solução correta do problema dos pontos (divisão da aposta) foi o marco inicial desse tipo de abordagem. Confirmamos essas informações em [4, 7, 22].

O desenvolvimento da probabilidade como área da matemática teve grande impulso em 1657, com a publicação de um pequeno livro, *De Ratiociniis in Ludo Aleae*, escrito pelo astrônomo e físico **Christiaan Huygens (1629 – 1695)**. Huygens era de uma importante família holandesa, e ficou muito mais conhecido pelas suas importantes contribuições à Astronomia e à Física Ondulatória. No entanto, em uma visita à Paris em 1655, tomou conhecimento da correspondência entre Pascal e Fermat e dos problemas de probabilidade nela investigados, e isto foi um impulso importante para a escrita do *Ludo Aleae*. Escrevendo a van Schooten (que era seu professor de Geometria), Huygens justifica a publicação do *Ludo Aleae* afirmando que "...não estamos tratando apenas com jogos mas com os fundamentos de uma nova teoria, tanto profunda como interessante". Foi o primeiro a utilizar o conceito de esperança matemática; outros detalhes são encontrados em [7, 22].

Trabalhando sobre a abordagem (combinatória) que Fermat utilizava na resolução dos problemas em probabilidade, o matemático suíço **Jacob Bernoulli (1654 – 1705)** iniciou o processo de sistematização da probabilidade deixando de lado os seguros e os jogos de azar. Em torno de 1689 ele publicou um trabalho sobre séries dando a conhecer um dos primeiros resultados fundamentais da Teoria Clássica de Probabilidades: que a frequência relativa de um evento tende para a probabilidade matemática deste evento, quando o número de repetições do experimento tende ao infinito. Foi devido a J. Bernoulli, então, o primeiro grande tratado de probabilidade, *Ars Conjectandi*, publicado postumamente em 1713. Neste seu livro, J. Bernoulli provou a *Lei dos Grandes Números*, que é a formalização do resultado fundamental citado acima; esta demonstração marcou o início de uma nova era na teoria da probabilidade. Ele provou tal Lei usando cálculos com coeficientes binomiais, sem usar a aproximação de Stirling (abordagem feita por De Moivre), e sem evidentemente utilizar ferramentas da "ainda não nascida" análise (abordagem feita por matemáticos russos ao final do século XIX, culminando com o trabalho de Komolgorov), conforme o que se lê em [7, 12, 22].

Logo após, em 1718, foi publicado pelo matemático franco-inglês **Abraham de Moivre (1667 – 1754)** a obra *Doctrine of Chances*, onde são exibidas técnicas de reduzir problemas de probabilidade a equações de diferenças, e de usar funções geratrizes para solucionar essas equações, que foram mais tarde aperfeiçoadas por Laplace. Ainda nesta primeira edição, De Moivre apresenta a definição de independência de eventos, além de investigar taxas de mortalidade e os fundamentos da teoria das anuidades. A segunda edição do *Doctrine* apareceu em 1738; nesta obra é introduzida pela primeira vez a distribuição normal, que usou como uma aproximação para a distribuição binomial (apresentando a distribuição normal em forma de série). De Moivre nesta obra ainda provou e usou em seus resultados a aproximação de Stirling; mais detalhes são encontrados em [7, 22].

No entanto, foi devido ao matemático francês **Pierre Simon de Laplace (1749 – 1827)** o grande salto qualitativo para consolidar a Teoria Clássica da Probabilidade. Até Laplace, a probabilidade ainda mantinha sua essência no cálculo de jogos de azar. Laplace ampliou o campo de aplicações da teoria para outras áreas, como a teoria dos erros, a matemática atuarial e a mecânica estatística. De fato, em 1812 publica a obra fundamental desta fase, *Théorie Analytique des Probabilités*. Os fundamentos da teoria de probabilidade foram colocados por Laplace nesta obra em uma forma que se manteve praticamente inalterada até o início do século XX. Nesse tratado, Laplace fez novas contribuições e reuniu, sistematizou e ampliou resultados desenvolvidos por seus predecessores: estabeleceu os métodos de equações de diferenças e de funções geratrizes e deu uma nova formulação e uma prova heurística do teorema central do limite. Foi a partir da obra de Laplace que os estudos na área realmente cresceram e tiveram a atenção de outros grandes matemáticos. Informações complementares verificar [7, 21, 22].

No final do século XIX, o russo **Pafnuty Lvovich Chebyshev (1821–1884)** fundou a denominada escola de São Petersburgo que produziu grandes matemáticos russos com contribuições fundamentais à teoria de probabilidade. Chebyshev foi influenciado pelos matemáticos russos **V. Ya. Bunyakovskii**, que publicou o importante livro *Fundamentos da Teoria Matemática de Probabilidades*, e **M. V. Ostrogradskii**, que fazem parte da linhagem inaugurada por Daniel Bernoulli e Euler em São Petersburgo. Chebyshev foi o primeiro a raciocinar sistematicamente em termos de variáveis aleatórias e seus momentos. Usando esses conceitos, Chebyshev estabeleceu uma simples desigualdade que permitiu uma prova trivial da Lei dos grandes números. O conceito de momentos foi utilizado por ele e, em seguida, por seu aluno **Andrei Andreyevich Markov (1856–1922)** para dar uma prova rigorosa do teorema central do limite. Um outro de seus famosos estudantes, Alexander Mikhailovich Lyapunov (1857–1918), posteriormente usou o conceito de funções características para dar uma prova mais simples desse importante teorema. Markov fez estudos sobre dependência de variáveis aleatórias analisando as hoje denominadas cadeias de Markov em tempo discreto. O trabalho fundamental sobre as cadeias de Markov em tempo contínuo foi desenvolvido posteriormente por Kolmogorov, como lemos em [7].

Andrei Nikolaevich Kolmogorov (1903–1987) foi um dos mais importantes matemáticos do século XX. Komolgorov axiomatizou a teoria da probabilidade da mesma forma que a Geometria foi axiomatizada por Euclides. Seu interesse por probabilidades iniciou-se em 1924 e em 1925 publicou seu primeiro trabalho nesta área com **Aleksandr Ya. Khintchine (1894–1959)** contendo o importante "teorema das três séries" e resultados em inequações de somas parciais de variáveis aleatórias que se tornaram a base para as desigualdades de martingales e do cálculo estocástico. Em 1928 ele determinou condições necessárias e suficientes para a Lei forte dos grandes números e provou, sob condições bastante gerais, a lei do logaritmo iterado para somas de variáveis aleatórias independentes. Em 1929, havia sido publicado seu trabalho *Teoria Geral de Medidas e Teoria de Probabilidade*, onde era apresentada a primeira descrição de uma construção axiomática baseada na teoria de medidas que havia sido criada em torno de 1901 por **Henri Lebesgue (1875–1941)** e **Èmile Borel (1871–1956)**. Em 1933 publicou em alemão um pequeno livro intitulado *Grundbegriffe der Wahrscheinlichkeitsrechnung* (traduzido para o inglês sob o título *Foundations of the Calculus of Probability*) onde desenvolve a teoria de probabilidade de forma matematicamente rigorosa a partir dos fundamentos axiomáticos baseados na teoria de medidas. A axiomatização de Kolmogorov marcou o início do desenvolvimento da teoria moderna de probabilidade. Em 1931 Kolmogorov publicou um importante artigo, *Métodos Analíticos na Teoria da Probabilidade*, no qual estabelece os fundamentos da teoria moderna de processos estocásticos, dando início à teoria dos processos de difusão. Em 1938 ele publicou mais um artigo na área da teoria probabilística que se tornou a base dos processos aleatórios de Markov. Maiores informações são encontradas em [7, 22].

1.2 ELEMENTOS DA TEORIA DA PROBABILIDADE

A seguir vamos exibir alguns conceitos básicos da chamada teoria da probabilidade. Nos aprofundaremos apenas o necessário para a compreensão dos temas tratados neste trabalho. O leitor interessado em mais detalhes pode encontrar em [10, 13, 19].

Há vários acontecimentos que, embora sejam repetidos muitas vezes e sob as mesmas condições, não apresentam os mesmos resultados – há uma variabilidade intrínseca na gama de resultados finais, mesmo se garantirmos que os parâmetros iniciais sejam os mesmos ou pelo menos muito próximos. Chamamos de *experimentos aleatórios* ou *fenômenos aleatórios* esses acontecimentos dotados de tal variabilidade: o resultado final é imprevisível, não se pode determiná-lo antes de ser realizado. Por exemplo, são aleatórios os fenômenos:

a) Lançar um dado não viciado e registrar seu resultado;

b) Lançar duas moedas e observar as faces voltadas para cima;

c) Retirar 1 carta de um baralho com 52 cartas e observar o seu naipe;

d) Número de pessoas que ganharão o próximo sorteio da Megasena;

e) Número de gotas de água que atingirão o telhado de uma casa durante uma chuva forte.

A matemática busca compreender esses fenômenos, buscando mensurar as chances (ou probabilidades) de, em um certo experimento aleatório, um determinado resultado ocorrer, já que não há maneiras de se ter exatidão na previsão do resultado. A Teoria das Probabilidades é o ramo da matemática que se encarrega de compreender mais a fundo esses fenômenos aleatórios e seus resultados, criando, elaborando e pesquisando modelos para seu estudo.

Estudando um experimento aleatório, nos interessaremos naturalmente por conhecer todos os resultados que podem acontecer naquele experimento, ao mesmo passo que atentaremos para resultados específicos dentro daquele experimento.

Desse modo, definimos primeiramente **Espaço Amostral** de um experimento como sendo o conjunto de todos os seus possíveis resultados de um experimento aleatório. Assim, para os experimentos exemplificados acima, teremos respectivamente

a) $\Omega_a = \{1, 2, 3, 4, 5, 6\}$;

b) $\Omega_b = \{KK, KC, CK, CC\}$ onde K representa *cara* e C representa *coroa*;

c) $\Omega_c = \{Copas,\ Espadas,\ Ouros,\ Paus\}$;

d) $\Omega_d = \{1, 2, 3, 4, \cdots\}$;

e) $\Omega_e = \{1, 2, 3, 4, \cdots\}$;

Também definiremos por **Evento** como sendo qualquer subconjunto do Espaço Amostral daquele experimento. Um evento pode ser caracterizado por um fato que se quer observar, uma condição matemática, ou uma outra escolha específica qualquer que focalize o olhar do experimentador para um certo tipo de resultado desejado. No exemplo a) acima descrito, poderíamos estar interessados em observar a ocorrência de um número que seja divisor de 6. Neste caso, o evento de interesse seria

$$D = \{1, 2, 3, 6\}.$$

Tomando o exemplo e), poderíamos querer estudar quando uma quantidade maior que 10000 gotas e menor que 15000 gotas atinge o telhado, e assim teríamos

$$E = \{10001, 10002, 10003, \ldots, 14999\}.$$

Como subconjuntos do espaço amostral, podemos manipular eventos através das operações tradicionais de conjuntos, como união, interseção e complemento.

Neste sentido, o evento complementar de um dado evento E, notado por E^c, será exatamente o conjunto dos resultados do experimento que não estão em E. Voltando nos exemplos que estamos explorando, poderíamos colocar para o exemplo a), sobre a rolagem dos dados, o evento

$$A = \{\text{Número par}\} = \{2, 4, 6\},$$

e portanto

$$A^c = \{1, 3, 5\}.$$

E se no exemplo b) estivéssemos interessados no evento

$$B = \{\text{Ao menos um resultado coroa}\} = \{KC, CK, KK\},$$

teríamos

$$B^c = \{CC\}.$$

Como comentamos anteriormente, ao olhar para eventos estamos olhando para conjuntos, e portanto pode-se falar em união de dois ou mais eventos, bem como em intersecção de dois ou mais eventos. Voltando a clássica rolagem de um dado, já citado várias vezes, podemos considerar, por exemplo os eventos

$$A = \{1, 2, 3\}, \; B = \{3, 6\} \; \text{ e } \; C = \{2, 4, 6\},$$

e assim

$$A \cup B = \{1, 2, 3, 6\},$$

enquanto

$$A \cap C = \{2\}.$$

Como comentamos acima é comum descrevermos os eventos de interesse de maneira textual, descrevendo uma propriedade ou característica que desejamos observar no resultado. Deste modo, os eventos A, B e C acima poderiam ser descritos por

$$\begin{aligned} A &= \{\text{resultado menor que } 4\}; \\ B &= \{\text{resultado é multiplo de } 3\}; \\ C &= \{\text{resultado é par}\}. \end{aligned}$$

Para entender o que é a união ou a intersecção de dois eventos, lembre-mos primeiro como são definidas tais operação entre dois conjuntos A e B.

Lembrando, diremos que um elemento $a \in A \cup B$ se, e somente se, $a \in A$ **ou** $A \in B$. Ou seja, a união é formada por todos os elementos que fazem parte de **ao menos** um dos conjuntos.

Da mesma forma, diremos que um elemento $a \in A \cap B$ se, e somente se, $a \in A$ **e** $A \in B$. Ou seja, a intersecção é formada por todos os elementos que fazem parte dos dois conjuntos **simultaneamente**.

Quando falamos de eventos, podemos usar estas mesmas definições para descrever a união ou intersecção de eventos.

Ao realizarmos um dada experimento aleatório, podemos dizer que observamos $A \cup B$ quando o observamos o evento A **ou** o evento B.

Já para observarmos a intersecção de dois eventos teremos de observar a ocorrência de A e de B **simultaneamente**.

Assim, para os eventos descritos acima teríamos

$$A \cup B = \{\text{resultado é menor que 4 OU múltiplo de 3}\},$$

e

$$A \cap C = \{\text{resultado é menor que 4 E par}\}.$$

Para melhor ilustrar todos os conceitos acima, consideremos um novo exemplo.

Exemplo 1.1. Colocamos 3 bolas numeradas aleatoriamente em 3 urnas, identificadas pelas letras a, b e c. Estamos interessados em observar o evento

$$A = \{\text{nenhuma urna está vazia}\}.$$

Gostaríamos também de entender a relação entre A e os eventos

$$\begin{aligned} A_1 &= \{\text{a urna } a \text{ está vazia}\} \\ A_2 &= \{\text{a urna } b \text{ está vazia}\} \\ A_3 &= \{\text{a urna } c \text{ está vazia}\}. \end{aligned}$$

O primeiro passo aqui é descrever o espaço amostral que descreve tal experimento. Para isso precisamos primeiro decidir como descreveremos resultados individuais do experimento, que neste caso são uma relação de bolas e suas respectivas urnas. Se valendo da numeração das bolas, podemos representar um resultado por uma sequência de três letras, indicando as urnas onde estão as bolas 1, 2 e 3 respectivamente.

Assim, (a, b, a) indica que as bolas 1 e 3 foram colocadas na urna a e a bola 2 na urna b.

Deste modo, o espaço amostral é

$$\begin{aligned} \Omega = \{&(a,a,b);(a,b,a);(b,a,a);(a,b,b);(b,a,b);(b,b,a);(a,a,c);(a,c,a);(c,a,a); \\ &(a,c,c);(c,a,c);(c,c,a);(b,b,c);(b,c,b);(c,b,b);(b,c,c);(c,b,c);(c,c,b); \\ &(a,b,c);(a,c,b);(b,a,c);(b,c,a);(c,a,b);(c,b,a);(a,a,a);(b,b,b);(c,c,c)\}. \end{aligned}$$

Para descrever os eventos A, A_1, A_2 e A_3 note se um trio em Ω não possui uma letra, então naquele resultado a caixa ausente está vazia. Assim,

$$\begin{aligned} A &= \{(a,b,c);(a,c,b);(b,a,c);(b,c,a);(c,a,b);(c,b,a)\} \\ A_1 &= \{(b,b,c);(b,c,b);(c,b,b);(b,c,c);(c,b,c);(c,c,b);(b,b,b);(c,c,c)\} \\ A_2 &= \{(a,a,c);(a,c,a);(c,a,a);(a,c,c);(c,a,c);(c,c,a);(a,a,a);(c,c,c)\} \\ A_3 &= \{(a,a,b);(a,b,a);(b,a,a);(a,b,b);(b,a,b);(b,b,a);(a,a,a);(b,b,b)\}. \end{aligned}$$

Agora, para escrever A em função de A_1, A_2 e A_3, volte à descrição do evento A e observe que podemos descrever A^c por

$$\begin{aligned} A^c &= \{\textbf{ao menos} \text{ uma caixa está vazia}\} \\ &= \{\text{a caixa } a \text{ está vazia } \textbf{ou} \text{ a caixa } b \text{ está vazia } \textbf{ou} \text{ a caixa } c \text{ está vazia}\}, \end{aligned}$$

e portanto

$$A^c = A_1 \cup A_2 \cup A_3,$$

ou ainda

$$A = (A_1 \cup A_2 \cup A_3)^c.$$

Tal igualdade pode ser facilmente verificada simplesmente listando os elementos dos conjuntos envolvidos.

Outra maneira de descrever tal relação é olhar para os eventos A_1, A_2 e A_3 e notar que

$$\begin{aligned} A_1^c &= \{\text{a caixa } a \text{ não está vazia}\} \\ A_2^c &= \{\text{a caixa } b \text{ não está vazia}\} \\ A_3^c &= \{\text{a caixa } c \text{ não está vazia}\}. \end{aligned}$$

e além disso

$$A = \{\text{a caixa } a \text{ não está vazia } \mathbf{e} \text{ a caixa } b \text{ não está vazia } \mathbf{e} \text{ a caixa } c \text{ não está vazia}\}.$$

Segue daí que

$$A = A_1^c \cap A_2^c \cap A_3^c.$$

1.3 DEFININDO PROBABILIDADE

Nosso objetivo agora é tentar encontrar uma boa definição para o que conhecemos por probabilidade. É importante observar que este é um processo de *modelagem*. Ou seja, queremos tentar representar matematicamente a imprevisibilidade do resultado em um certo experimento. Queremos então associar a cada resultado do experimento, ou mais geralmente a cada evento, um número que representará *probabilidade* de observarmos tal evento.

Algumas situações são bastante simples, e foram historicamente as primeiras a serem tratadas. Tais situações aparecem quando o espaço amostral é finito com n elementos, e não existe nada que indique que um dado resultado deva ser mais provável que os demais. Neste caso simplesmente dizemos que cada resultado possui a mesma probabilidade, nominalmente $1/n$.

Mais formalmente, dado um espaço amostral Ω e um evento $E \subseteq \Omega$, a probabilidade de ocorrer o evento E pode ser definida como o número real $\mathbf{P}(E)$ dado por

$$\mathbf{P}(E) = \frac{n(E)}{n(\Omega)},$$

onde $n(E)$ representa o número de elementos do evento E e $n(\Omega)$ o número de elementos do espaço amostral Ω.

No escopo dessa definição, fica claro que $0 \leq \mathbf{P}(E) \leq 1$, pois $n(E) \leq n(\Omega)$, qualquer que seja o evento $E \subseteq \Omega$.

Exemplo 1.2. Considere o experimento de lançarmos duas moedas honestas e observarmos o resultado obtido. Queremos então medir as probabilidades da ocorrência de dois resultados iguais, e da ocorrência de no mínimo uma cara.

Para começar, o espaço amostral deve ser

$$\Omega = \{KK, KC, CK, CC\}.$$

Os eventos que considerados são

$$E_1 = \{\text{dois resultados iguais}\} = \{KK, CC\}$$

e

$$E_2 = \{\text{ao menos uma cara}\} = \{KK, CK, KC\}.$$

Assim, considerando que todos os resultados possíveis tem a mesma probabilidade, temos

$$\mathbf{P}(E_1) = \frac{n(E_1)}{n(\Omega)} = \frac{2}{4} = \frac{1}{2}$$

e

$$\mathbf{P}(E_2) = \frac{n(E_2)}{n(\Omega)} = \frac{3}{4}.$$

Essa forma de definir probabilidade funciona bem para conjuntos simples, experimentos estáticos e situações pouco complexas. Mas alguns exemplos simples não podem ser descritos desta forma. Considere, por exemplo, o experimento onde jogamos uma moeda até observar "cara" pela primeira vez. O total de resultados possíveis neste experimento é infinito, e portanto a definição acima não contempla este caso.

Mas mesmo em experimentos com finitos resultados possíveis, não podemos garantir que todos eles tenham a mesma probabilidade de ocorrência. Podemos, por exemplo, rolar um dado feito de material não homogêneo, ou que durante sua fabricação tenha adquirido bolhas internas próximas a uma das faces. Este dado seria então "viciado" e dizer que todos os valores tem a mesma probabilidade de ocorrer seria incoerente.

Observação 1.3.1. *Hoje sabemos que existe um valor limite e ele é exatamente a probabilidade do evento. Este resultado é conhecido Lei Forte dos Grandes Números de Bernoulli, e é uma das ferramentas utilizadas pela estatística para testar a validade de modelos.*

Mas como decidir então qual probabilidade deveremos associar a cada evento? Uma possibilidade é proceder de forma *experimental*. Para isso considere um experimento aleatório, cujo espaço amostral é Ω. Realize repetidas vezes esse experimento, sob as mesmas condições iniciais. Para cada evento $E \subseteq \Omega$, seja $n(E)$ o número de vezes que acontece o evento E ao longo de n repetições do experimento. Dessa forma, podemos definir $\mathbf{P}(E)$, a probabilidade de ocorrência do evento E, por

$$\mathbf{P}(E) = \lim_{n\to\infty} \frac{n(E)}{n}.$$

Isto é, $\mathbf{P}(E)$ é definida como a frequência-limite da ocorrência de E, ao considerar um número arbitrariamente grande (tão grande quanto se queira) de repetições do experimento em questão.

Não é difícil perceber que esta definição não é satisfatória. Antes de mais nada, ela depende de um limite que em geral não temos como calcular ou tão pouco garantir a existência. Além disso, na prática deveríamos repetir o experimento um total muito grande de vezes, e esperar que este resultado esteja próximo suficiente do limite.

Mas este não é o caminho natural de se fazer modelagem matemática. A ideia central da modelagem é encontrar um modelo matemático baseado no que conhecemos sobre o que queremos medir, para depois tentar corroborar o modelo com experimentos ou dados.

Assim, cada experimento aleatório trará um desafio a parte, e precisaremos alocar a cada evento um valor, baseado no que conhecemos do problema.

Exemplo 1.3. Consideremos o experimento de lançar uma moeda não viciada até atingir cara pela primeira vez.

Cada resultado deste experimento pode ser representado por uma sequência de C's terminando em um K, representando as coroas lançadas e o último lançamento onde observamos cara. Algo como

$$\underbrace{CC\cdots C}_{n-1 \text{ vezes}}K$$

é um resultado típico deste experimento, onde foram necessárias n lançamentos para se observar cara pela primeira vez.

Assim

$$\Omega = \{K, CK, CCK, CCCK, \cdots\}.$$

Para determinar as probabilidades dos diferentes eventos, vamos inicialmente definir as probabilidade de um resultado individual. Ou seja, vamos definir a probabilidade dos eventos

$$E_n = \{\underbrace{CC\cdots C}_{n-1 \text{ vezes}}K\}.$$

Do fato da moeda ser não viciada, podemos concluir que em cada lançamento temos probabilidade $1/2$ de observar cara e $1/2$ de observar coroa. É razoável supor, portanto, que

$$\mathbf{P}(E_1) = \mathbf{P}(\{K\}) = \frac{1}{2}.$$

Para observar E_2 precisamos que em um lançamento de 2 moedas aconteça exatamente o resultado CK, que é um em 4 possíveis. Portanto, colocamos

$$\mathbf{P}(E_2) = \frac{1}{4},$$

e generalizando esta ideia fazemos

$$\mathbf{P}(E_n) = \frac{1}{2^n}.$$

Agora, para definir a probabilidade de um evento qualquer podemos simplesmente somar as probabilidades dos resultados individuais que formam aquele evento. Assim, por exemplo, a probabilidade do evento

$$A = \{\text{no máximo dois lançamentos foram feitos}\} = \{K, CK\}$$

seria

$$\mathbf{P}(A) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}.$$

Assim como no exemplo acima, para definir um modelo de probabilidades devemos associar valores para a probabilidade de cada evento. Mas temos que tomar cuidado com os valores que associamos, para que o modelo não fique inconsistente.

As regras que devem ser observadas são conhecidas como axiomas das Probabilidades, e são parte da seguinte definição.

Definição 1.1. Considere um experimento aleatório, cujo espaço amostral é Ω. Diremos que uma função $\mathbf{P}$ que associa a cada evento E desse espaço amostral um número real $\mathbf{P}(E)$, é uma **distribuição de probabilidade em** Ω (ou simplesmente uma probabilidade em Ω) se satisfaz os seguintes três axiomas:

Axioma 1 - Para todo evento E, $0 \leq \mathbf{P}(E) \leq 1$;

Axioma 2 - $\mathbf{P}(\Omega) = 1$;

Axioma 3 - Para toda sequência de eventos $E_1, E_2, \ldots, E_n$ tais que $E_i \cap E_j = \emptyset$ sempre que $i \neq j$, vale que

$$\mathbf{P}(E_1 \cup E_2 \cup \cdots \cup E_n) = \mathbf{P}(E_1) + \mathbf{P}(E_2) + \cdots + \mathbf{P}(E_n).$$

Os axiomas acima formam um conjunto mínimo de exigências para que a associação de probabilidades feitas seja consistente. Os axiomas 1 e 2 são bastante intuitivos e concordam com tudo o que fizemos até este ponto. O axioma 3 pode ser mais complicado, e talvez seja melhor entendido se olharmos no caso particular de 2 eventos. Neste caso ele diz que a probabilidade de 2 eventos que não podem ocorrer simultaneamente é a soma das probabilidades individuais dos eventos, o que está de acordo com a noção de probabilidade como frequência.

Para entender melhor os axiomas, voltemos ao Exemplo 1.3. Lembrando, neste exemplo consideramos o lançamento de uma moeda não viciada repetidamente até observarmos a primeira cara. Definimos $\mathbf{P}$ associando a probabilidade $1/2^n$ ao resultado onde são necessários exatamente n lançamentos. Ou seja,

$$\mathbf{P}(\{\underbrace{CC\cdots C}_{n-1 \text{ vezes}}K\}) = \frac{1}{2^n}.$$

Para E um evento qualquer, definimos então $\mathbf{P}(E)$ como a soma das probabilidades de seus elementos, o que é imediatamente compatível com o axioma 3.

A seguir verifiquemos o axioma 2. Para isso observe que pela construção de $\mathbf{P}$ temos

$$\begin{aligned}
\mathbf{P}(\Omega) &= \mathbf{P}(\{K\}) + \mathbf{P}(\{CK\}) + \mathbf{P}(\{CCK\}) + \mathbf{P}(\{CCCK\}) + \cdots \\
&= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \cdots \\
&= \left(\frac{1}{2}\right)^1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \cdots \\
&= \frac{1}{2}\frac{1}{1-1/2} \\
&= 1,
\end{aligned}$$

o que segue direto da expressão para a soma de uma PG infinita.

O axioma 1 agora segue diretamente do fato de que cada evento, por ser subconjunto de Ω, tem probabilidade dada pela soma de parte dos valores da soma acima, e portanto não pode ultrapassar 1. Além disso, como os valores somados são positivos, a probabilidade também deve ser.

1.3.1 *Propriedades de uma Probabilidade*

Vamos agora explorar algumas propriedades comuns a toda distribuição de probabilidade. Estas propriedades nos permitem estudar probabilidades de nosso interesse conhecendo quase nada sobre elas, e estão por detrás de quase todos os cálculos feitos no próximo capítulo.

PROPRIEDADES DE UMA PROBABILIDADE

1. $\mathbf{P}(\emptyset) = 0$

   Para perceber isso, basta notar que $\Omega \cap \emptyset = \emptyset$, de modo que Ω e $\emptyset$ são disjuntos. Assim, pelos Axiomas 2 e 3, vale que

   $$1 = \mathbf{P}(\Omega) = \mathbf{P}(\Omega \cup \emptyset) = \mathbf{P}(\Omega) + \mathbf{P}(\emptyset) = 1 + \mathbf{P}(\emptyset),$$

   e a propriedade segue.

2. **Probabilidade do Complementar**

   Para qualquer evento $A \subseteq \Omega$ vale que

   $$\mathbf{P}(A^c) = 1 - \mathbf{P}(A).$$

Para perceber isso, note que $A \cap A^c = \emptyset$ (A e A^c são disjuntos) e que $A \cup A^c = \Omega$. Assim, novamente pelos Axiomas 2 e 3, temos que

$$1 = \mathbf{P}(\Omega) = \mathbf{P}(A \cup A^c) = \mathbf{P}(A) + \mathbf{P}(A^c),$$

e o resultado segue.

3. **Monotonicidade**

   Para quaisquer eventos $A, B \subseteq \Omega$ tais que $A \subseteq B$, vale que

   $$\mathbf{P}(A) \leq \mathbf{P}(B).$$

   Para isso observe que, como ilustramos na Figura 5, vale que $B = A \cup (A^c \cap B)$. Além disso, A e $A^c \cap B$ são disjuntos. Assim, pelos Axiomas 1 e 3,

   $$\mathbf{P}(B) = \mathbf{P}(A) + \mathbf{P}(A^c \cap B) \geq \mathbf{P}(A).$$

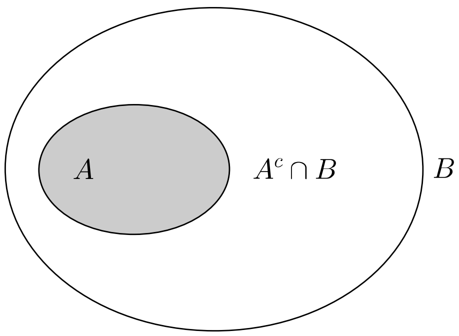


Figura 5: Escrevendo B como união de dois eventos disjuntos

4. **Inclusão-Exclusão**

   Para quaisquer eventos $A, B \subseteq \Omega$ vale que

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

   Para obtermos a expressão para $\mathbf{P}(A \cup B)$, notemos inicialmente que $A \cup B$ pode ser escrito como a união de dois eventos disjuntos (ver Figura 6). Nominalmente

   $$A \cup B = A \cup (A^c \cap B).$$

   Do Axioma 3, podemos escrever

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

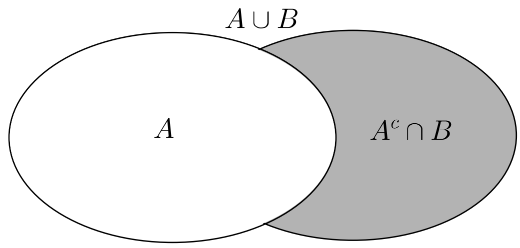


Figura 6: Escrevendo $A \cup B$ como união de dois eventos disjuntos

Por outro lado, é fácil ver que $B = (A \cap B) \cup (A^c \cap B)$. Obtemos novamente do do Axioma 3 que $P(B) = P(A \cap B) + P(A^c \cap B)$.

Substituindo a última igualdade obtida na expressão de $P(A \cup B)$, chegamos à expressão final

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

Tal expressão pode ser estendida para os casos de três, quatro ou mais eventos. Como exemplo, uma expressão para a situação envolvendo três eventos A, B e C quaisquer seria

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

Exemplo 1.4. Criando uma aplicação prática: se uma pessoa retirar uma carta ao acaso de um baralho comum de 52 cartas, e considerar os eventos não-disjuntos

$$A = \{\text{Observa-se uma carta de espadas}\}$$

e

$$B = \{\text{Observa-se um valete}\}.$$

ela poderá calcular a probabilidade de a carta retirada ser de espadas ou ser um valete da seguinte maneira

$$P(A \cup B) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52}.$$

1.4 PROBABILIDADE CONDICIONAL E INDEPENDÊNCIA

Para finalizar este capítulo vamos tratar de um conceito fundamental na teoria das probabilidades: a medida de quanto o conhecimento da ocorrência de um evento altera a probabilidade de ocorrência de outro evento.

Para melhor entender o problema, comecemos por analisar um exemplo simples para posteriormente tentar generalizar o que aprendemos.

Exemplo 1.5. Em um certo momento durante um jogo de dados Breno aposta que em uma rolagem de 2 dados distintos um deles mostrará o número 6. Alberto, o responsável pelo jogo, lança os dados e diz para Breno que a soma dos valores sorteados foi igual à 8, dando a ele a possibilidade de alterar sua aposta. Como Breno pode decidir o que fazer?

Para começar, vamos antes de mais nada explicitar o espaço amostral deste experimento. Aqui cada resultado pode ser representado por um par de valores, (a, b), com $a, b \in \{1, 2, 3, 4, 5, 6\}$, representado a face sorteada nos dados 1 e 2 respectivamente. Assim

$$\Omega = \{(1, 1); (1, 2); \ldots; (6, 5); (6, 6)\},$$

e $n(\Omega) = 6 \cdot 6 = 36$.

Observe agora que Breno está interessado no evento

$$\begin{aligned} A &= \{\text{em um dos dados a face sorteada é } 6\} \\ &= \{(1, 6); (2, 6); (3, 6); (4, 6); (5, 6); (6, 6); (6, 5); (6, 4); (6, 3); (6, 2); (6, 1)\}, \end{aligned}$$

que a princípio, considerando que todos os resultados são igualmente prováveis, tem probabilidade

$$\mathbf{P}(A) = \frac{n(A)}{n(\Omega)} = \frac{11}{36}.$$

Mas Alberto dá uma nova informação sobre o sorteio, dizendo que ocorreu o evento

$$\begin{aligned} B &= \{\text{a soma dos resultados foi } 8\} \\ &= \{(2, 6); (3, 5); (4, 4); (5, 3); (6, 2)\}. \end{aligned}$$

Ao receber esta informação Breno pode imediatamente descartar qualquer resultado fora do evento B. Assim, o evento $\{(6,1);(1,6)\}$, por exemplo, passa a ter probabilidade 0, uma vez que neste evento a soma dos resultados é sempre 7.

Assim ele "reduz seu espaço amostral" à B, que possui apenas 5 elementos. A seguir ele "descarta" todos os elementos de A que não estão em B, ficando apenas com

$$A \cap B = \{(2,6);(6,2)\}.$$

Assim, considerando que todos os resultados em B sejam ainda equiprováveis, Breno calcula nova probabilidade de vitória fazendo

$$\frac{n(A \cap B)}{n(B)} = \frac{2}{5} \geq \frac{11}{36}, \tag{1.1}$$

e decide aumentar sua aposta.

O valor calculado acima por Breno é conhecido na matemática como *probabilidade condicional de A dado B*, e denotado por $\mathbf{P}(A|B)$.

No caso de Breno foi razoavelmente simples calcular $\mathbf{P}(A|B)$, mas o calculo final dependeu fortemente do fato dos resultados individuais em Ω terem todos a mesma probabilidade. Surge então a dúvida: como calcular esta quantidade de modo geral?

Vamos para isso analisar passo a passo o raciocínio de Breno. Antes de mais nada ele "reduziu" o espaço amostral para B, o que na prática levou-o a considerar apenas o evento $A \cap B$ no lugar de todo o evento A. Isso por que a informação sobre a ocorrência de B tem o efeito primário de levar a probabilidade de todo resultado em B^c a 0.

Matematicamente, isso significa que $\mathbf{P}(A|B)$ deve ser proporcional à $\mathbf{P}(A \cap B)$. Ou seja

$$\mathbf{P}(A|B) = \alpha_A \cdot \mathbf{P}(A \cap B),$$

para algum α_A.

É importante observar, no entanto, que a princípio α_A pode depender tanto de A quanto de B, de modo que até agora a expressão diz apenas que $\mathbf{P}(A|B) = 0$ sempre que $\mathbf{P}(A \cap B) = 0$.

A seguir Breno ponderou que se os resultados eram equiprováveis a princípio, a nova informação não deveria mudar tal fato. Em um espaço onde os resultados não são todos igualmente prováveis, poderíamos argumentar que para dois eventos A e C a razão entre as probabilidades de $A \cap B$ e $C \cap B$ deveria ser mantida após a informação da ocorrência de B. Ou seja, se $C \cap B$ era, digamos, 2 vezes mais provável que $A \cap B$ antes, a informação da ocorrência de B não deveria alterar isso.

Escrevendo isso, encontramos que

$$\frac{\alpha_C \cdot \mathbf{P}(C \cap B)}{\alpha_A \cdot \mathbf{P}(A \cap B)} = \frac{\mathbf{P}(C|B)}{\mathbf{P}(A|B)} = \frac{\mathbf{P}(C \cap B)}{\mathbf{P}(A \cap B)},$$

e portanto $\alpha_A = \alpha_C$, mostrando que α_A deve depender apenas de B e

$$\mathbf{P}(A|B) = \alpha \cdot \mathbf{P}(A \cap B).$$

Para encontrar o valor de α basta notar que a informação sobre a ocorrência de B leva a probabilidade de B para 1! Ou seja,

$$1 = \mathbf{P}(B|B) = \alpha \cdot \mathbf{P}(B \cap B) = \alpha \cdot \mathbf{P}(B),$$

e portanto

$$\alpha = \frac{1}{\mathbf{P}(B)}.$$

Com isso chegamos à seguinte definição.

Definição 1.2. Dada uma probabilidade $\mathbf{P}$ em um espaço amostral Ω e eventos $A, B \subseteq \Omega$, com $\mathbf{P}(B) > 0$, definimos a **probabilidade condicional de** A **dado** B por

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}. \tag{1.2}$$

É interessante observar que no caso de espaços onde os resultados são igualmente prováveis, a equação (1.2) se reduz à

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

que é exatamente a expressão usada por Breno em (1.1).

Outra forma de escrever (1.2) é dizer que

$$\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\mathbf{P}(B),$$

ou ainda

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

Expressões que estão "escondidas" no currículo do ensino médio, normalmente em exercícios do tipo "sorteio sem reposição", como mostramos no exemplo abaixo, bastante comum em livros de matemática de ensino médio.

Exemplo 1.6. Uma caixa contém 8 bolas, sendo 5 brancas e 3 pretas. Uma bola é sorteada ao acaso desta caixa e a seguir, sem que a primeira bola seja devolvida à caixa, uma nova bola é sorteada. Neste caso, qual a probabilidade de sortearmos uma bola branca seguida de uma bola preta?

Para resolver o problema considere os eventos

$$A = \{\text{a primeira bola sorteada é branca}\}$$

e

$$B = \{\text{a segunda bola sorteada é preta}\}.$$

Queremos calcular $\mathbf{P}(A \cap B)$, mas antes disso observe que

$$\mathbf{P}(A) = \frac{5}{8},$$

pois existem 5 bolas brancas dentre as 8 bolas na caixa. E além disso

$$\mathbf{P}(B|A) = \frac{3}{7},$$

pois ao sortearmos uma bola branca no primeiro sorteio, restaram 7 bolas na caixa, das quais 3 são pretas.

Assim

$$\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B|A) = \frac{5}{8} \cdot \frac{3}{7} = \frac{15}{54}.$$

É importante comentar que podemos usar a probabilidade condicional para calcular a probabilidade da intersecção de 3 eventos, ou mais. Para isso basta fazer

$$\mathbf{P}(A \cap B \cap C) = \mathbf{P}(C|A \cap B)\mathbf{P}(A \cap B) = \mathbf{P}(C|A \cap B)\mathbf{P}(B|A)\mathbf{P}(A),$$

ou seja

$$\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\mathbf{P}(B|A)\mathbf{P}(C|A \cap B).$$

O mesmo raciocínio no caso de n eventos $A_1, A_2, \ldots, A_n$ nos dá

$$\mathbf{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbf{P}(A_1)\mathbf{P}(A_2|A_1)\mathbf{P}(A_3|A_1 \cap A_2) \cdots \mathbf{P}(A_n|A_1 \cap \cdots \cap A_{n-1}).$$

1.4.1 *Independência*

Vamos agora tentar entender um dos conceitos centrais na teoria das probabilidades, a independência de eventos. Em linhas gerais, diremos que dois eventos são independentes quando a ocorrência de um dos eventos não trás nenhuma informação sobre o outro. Antes de colocar esta definição de forma mais clara, vamos tentar deixar as coisas mais claras com um exemplo.

Exemplo 1.7. Voltando ao jogo de dados de Breno, considere que os dados são distintos, sendo um deles azul e o outro vermelho. Nesta jogada Breno apostou que o dado vermelho mostraria a face 6. Alberto, por sua vez, revela a Breno após a jogada que o dado azul mostrou um resultado menor que 3.

Breno se recorda então que o espaço amostral Ω deste experimento é composto pelos 36 pares de número (a, b) onde a representa o resultado do dado azul e o b o do vermelho. Os eventos em jogo desta vez são

$$\begin{aligned} E &= \{\text{o dado vermelho mostrou } 6\} \\ &= \{(1,6);(2,6);(3,6);(4,6);(5,6);(6,6)\}, \end{aligned}$$

e

$$\begin{aligned} F &= \{\text{o dado azul mostrou um número menor que } 3\} \\ &= \{(1,1);(1,2);(1,3);(1,4);(1,5);(1,6);(2,1);(2,2);(2,3);(2,4);(2,5);(2,6)\}, \end{aligned}$$

e portanto

$$\mathbf{P}(E) = \frac{6}{36} = \frac{1}{6} \qquad \text{e} \qquad \mathbf{P}(F) = \frac{12}{36} = \frac{1}{3}.$$

Para calcular $\mathbf{P}(E|F)$ precisamos ainda de $\mathbf{P}(E \cap F)$, mas como

$$E \cap F = \{(1,6);(2,6)\},$$

temos que $\mathbf{P}(E \cap F) = 1/18$, e

$$\mathbf{P}(E|F) = \frac{\mathbf{P}(E \cap F)}{\mathbf{P}(F)} = \frac{1/18}{1/3} = \frac{1}{6} = \mathbf{P}(E).$$

Assim Breno conclui que a informação de Alberto não ajudou em nada, e decide manter a aposta inicial.

No exemplo acima $\mathbf{P}(E|F) = \mathbf{P}(E)$, mostrando que a informação sobre a ocorrência de F não interferiu na probabilidade de E, e neste sentido podemos dizer que E é independente de F.

A primeira questão que aparece é se E ser independente de F também implica que F é independente de E. Ou ainda se

$$\mathbf{P}(E|F) = \mathbf{P}(E) \Longleftrightarrow \mathbf{P}(F|E) = \mathbf{P}(F).$$

Antes de responder a questão, podemos voltar ao exemplo e ver que

$$\mathbf{P}(F|E) = \frac{\mathbf{P}(E \cap F)}{\mathbf{P}(E)} = \frac{1/18}{1/6} = \frac{1}{3} = \mathbf{P}(F),$$

e portanto neste caso podemos também dizer que F é independente de E.

De modo geral, tomando dois eventos A e B de probabilidade positiva, note que

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

Mostramos assim que independência é uma característica do par de eventos, e olhando na segunda linha do argumento acima encontramos uma maneira mais simples de caracterizar independência, que explicitamos na definição a seguir.

Definição 1.3. Dada uma probabilidade $\mathbf{P}$ em um espaço amostral Ω, diremos que dois eventos A e B são **independentes** se

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

Como ilustrado no exemplo 1.7, independência é uma propriedade esperada quando um mesmo experimento é repetido sem que o resultado de um influa diretamente no outro, como seguidos sorteios da loteria, repetidos lançamentos de uma moeda ou algo similar.

Mas não é apenas nestes casos que esperamos encontrar independência. Para ilustrar, voltemos por um instante ao exemplo 1.7, e suponha que a informação dada por Alberto seja que a soma dos resultados dos dois dados é 7. Neste caso, a informação de Alberto leva em conta o resultado dos dois dados, então é razoável supor que a probabilidade de vitória de Breno seja alterada.

Para verificar esta hipótese defina o evento

$$\begin{aligned} G &= \{\text{a soma dos resultados é } 7\} \\ &= \{(1,6);(2,5);(3,4);(4,3);(5,2);(6,1)\}, \end{aligned}$$

lembre que

$$\begin{aligned} E &= \{\text{o dado vermelho mostrou } 6\} \\ &= \{(1,6);(2,6);(3,6);(4,6);(5,6);(6,6)\}, \end{aligned}$$

e observe que

$$\begin{aligned} E \cap G &= \{\text{o dado vermelho mostrou 6 e a soma dos resultados foi } 7\} \\ &= \{(1,6)\}. \end{aligned}$$

Segue assim que

$$\mathbf{P}(E \cap G) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \mathbf{P}(E) \cdot \mathbf{P}(G),$$

de onde concluímos que E e G são de fato independentes, contrariando nossa hipótese inicial!

É interessante observar que o evento F do exemplo 1.7 também é independente de G, uma vez que

$$F \cap G = \{(1,6);(2,5)\},$$

e

$$\mathbf{P}(F \cap G) = \frac{2}{36} = \frac{2}{6} \cdot \frac{1}{6} = \mathbf{P}(F) \cdot \mathbf{P}(G).$$

Isso mostra que E, F são independentes, F, G são independentes, assim como E, G também são. Mas o trio E, F, G não pode ser independente! Isso por que, se a soma dos resultados é 7 e o dado vermelho é 6, então o dado azul é obrigatoriamente menor que 3. Ou seja, $\mathbf{P}(F | E \cap G) = 1$.

Isso mostra que para definir a independência de diversos eventos precisamos pedir mais que apenas independência dois a dois, como se verifica na definição a seguir.

Definição 1.4. Dada uma probabilidade $\mathbf{P}$ em um espaço amostral Ω, diremos que eventos $A_1, A_2, \ldots, A_n$ em Ω são **independentes** se para quaisquer índices $i_1, i_2, \ldots, i_k \in \{1, 2, \ldots, n\}$, distintos 2 a 2, com $1 \leq k \leq n$, vale que

$$\mathbf{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \cdot \mathbf{P}(A_{i_2}) \cdots \mathbf{P}(A_{i_k}).$$

O exemplo a seguir descreve uma importante distribuição de probabilidade que faz uso do conceito de independência em sua definição.

Exemplo 1.8 (*Distribuição Geométrica*)**.** Para este exemplo considere um experimento aleatório qualquer de sua preferência. Pode ser a rolagem de um dado, o lançamento de uma moeda, o sorteio de carta, um sorteio da loteria, ou mesmo a escolha de uma pessoa na rua para responder uma pesquisa eleitoral.

Considere também um evento A no espaço amostral de tal experimento, com probabilidade p.

Imagine agora a possamos repetir tal experimento quantas vezes forem necessárias, de modo que cada repetição seja independente das demais.

Vamos agora repetir este experimento até que o evento A seja observado. Feito isso, queremos associar uma probabilidade $\mathbf{P}$ aos elementos de $\Omega = \{1, 2, 3, 4, \ldots\}$ que indicam o total de repetições necessárias.

A chave para a definição de $\mathbf{P}$ está na palavra *independente* usada para descrever as repetições do experimento. Matematicamente, podemos dizer que os eventos $A_1, A_2, \ldots,$ dados por

$$A_k = \{\text{o evento } A \text{ foi observado na } k\text{-ésima repetição do experimento}\},$$

são independentes, no sentido da definição 1.4.

Para definir $\mathbf{P}$ em Ω precisamos atribuir uma probabilidade para cada valor $k \in \Omega$, o que formalmente é o mesmo que atribuir uma probabilidade aos eventos

$$B_k = \{\text{foram necessárias } k \text{ repetições}\} = \{k\}.$$

Para isso, observe que

$$B_k = A_1^c \cap \cdots \cap A_{k-1}^c A_k,$$

e como, por hipótese, $\mathbf{P}(A_i) = p$ para todo $i \geq 1$ e $A_1, \cdots, A_k$ são independentes, então

$$\mathbf{P}(B_k) = \mathbf{P}(A_1^c \cap \cdots \cap A_{k-1}^c A_k) = (1-p)^{k-1} p.$$

2

CADEIAS DE MARKOV

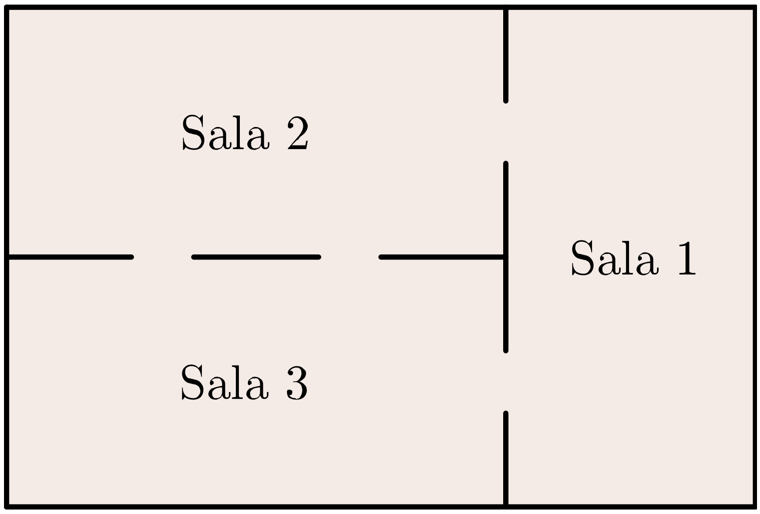
> *"O raciocínio estatístico será um dia tão necessário à cidadania eficiente quanto a capacidade de ler e escrever"*
>
> — H.G.Wells

2.1 INTRODUÇÃO

Já vimos que um grande número de matemáticos contribuiu para a gênese da Teoria Moderna das Probabilidades, em especial a escola russa do final do século XIX e início do século XX, iniciada por Chebyshev e coroada por Kolmogorov com brilhantes e fundamentais estudos. Era uma época em que se buscava, além da axiomatização da Teoria, um modo de compreender padrões de distribuição de probabilidades emergentes em sistemas simples. Começava, de maneira mais programática, a busca pela *ordenação do caos* a que fizemos referência no capítulo anterior.

Para tentar ilustrar o que vamos estudar neste capítulo, pensemos no seguinte experimento aleatório, encontrado em [9] ou [18].

Exemplo 2.1. Em um laboratório de testes comportamentais, um ratinho é colocado numa gaiola com três espaços distintos, todos comunicáveis entre si. Cada espaço tem aproximadamente a mesma área. Ele se encontra inicialmente no espaço 1 e já foi treinado para mudar de espaço atravessando uma porta sempre que soa um alarme (de minuto em minuto). Cada vez que soa o alarme o ratinho escolhe qualquer uma das portas disponíveis, com igual probabilidade e sem ser afetado por escolhas anteriores.



Imaginemos agora esse experimento se repetindo continuamente, minuto após minuto, por um intervalo arbitrariamente grande de tempo. Será que o ratinho ficou mais ou menos a mesma proporção de tempo em cada um dos espaços?

Dado o que estudamos até o momento, o problema acima não possui uma resposta trivial, e à primeira vista é possível pensarmos sim. Mas uma análise mais cuidadosa rapidamente coloca dúvidas em tal hipótese. Note que entre as salas 2 e 3 existem 2 portas, enquanto a sala 1 se comunica com as demais através de uma única porta para cada sala. Isso faz com que o rato tenha uma maior probabilidade de ir da sala 2 para a 3 (ou vice-versa) do que retornar para a sala inicial, fazendo com que ele fique mais tempo nestas salas.

Ao longo deste capítulo vamos estudar as ferramentas necessárias para responder tal pergunta. Para tal modelaremos este experimento usando um tipo especial de processo temporal aleatório conhecido como *Cadeias de Markov*, que levam este nome em homenagem à **Andrei Andreyevich Markov** (1856 - 1922), um matemático russo que fez importantes contribuições no estudo da perda de memória em processos estocásticos.

Acompanhando o desenvolvimento contemporâneo da Teoria de Probabilidades, o estudo das Cadeias de Markov também tem experimentado notáveis avanços, desde sua sistematização, realizada em 1938 por Kolmogorov das bases dos processos aleatórios que as caracterizam. Os trabalhos do próprio Kolmogorov e de **Aleksandr Yakovlevich Khintchine** (1894 – 1959) permitiram o surgimento desigualdades martingales e do cálculo estocástico mais rigorosamente estabelecido, ao lado das contribuições dos matemáticos **Joseph Leo Doob** (1910 – 2004) e **Paul Lévy** (1886 – 1971); ao longo do século XX, a Teoria das Cadeias de Markov vem ganhando mais espaço e se expandindo para a aplicação em várias áreas: física, química, biologia, ciências sociais, Teoria dos Jogos, música, linguística, processos do tipo Monte-Carlo, neurociências, bioinformática, reconhecimento de imagens, reconhecimento de assinaturas, sistemas de comunicação e internet, ciências da computação, teorias econômicas, Teoria das filas, estudo de processos migratórios, processos de modelagem de evolução populacional, modelos de difusão, dentre muitos outros.

2.2 CONCEITOS INICIAIS

A seguir introduziremos alguns conceitos e resultados envolvendo cadeias de Markov. Vários detalhes serão omitidos, pois fogem ao escopo deste trabalho. Maiores detalhes poderão ser encontrados em [9, 18].

Antes de introduzir mais formalmente o que entendemos por uma Cadeia de Markov, voltemos ao exemplo 2.1, e tentemos descrever seus elementos principais. Antes de mais nada, denote por $X_0, X_1, X_2, X_3, \ldots$ a sequência aleatória de quartos visitados pelo rato durante seu passeio. Assim, X_0 denota o quarto onde o rato começa seu passeio ($X_0 = 1$), X_1 o primeiro quarto visitado depois de iniciar o passeio, X_2 o segundo quarto visitado, e assim por diante. Deste modo $X_0, X_1, X_2, \ldots$ é uma sequência aleatória de números do conjunto $\{1, 2, 3\}$, formada de acordo com algumas regras, que tentaremos descrever a seguir.

Antes de prosseguir, vamos estabelecer algumas notações, que deixaram nossa vida mais simples. Para este problema, e ao longo de todo o capítulo, estaremos interessados principalmente nos eventos $\{X_n = i\}$. Para deixar a notação mais limpa faremos duas pequenas alterações na notação que usamos até aqui.

1. Deixaremos de lado as chaves em $\{X_n = i\}$ e escreveremos simplesmente $\mathbf{P}(X_n = i)$ ou $\mathbf{P}(X_{n+1} = j | X_n = i)$;

2. Quando estivermos falando da intersecção de eventos deste tipo, substituiremos o símbolo $\cap$ por ponto e vírgula (;). Assim, ao invés de $\mathbf{P}(\{X_n = 1\} \cap \{X_k = 2\})$, escreveremos $\mathbf{P}(X_n = 1; X_k = 2)$.

Voltando ao problema, considere que sempre que soa a campainha o rato escolhe aleatoriamente uma das portas disponíveis no quarto onde ele está e passa por ela, mudando de quarto. Isso pode ser representado pelas seguintes equações, validas para qualquer instante de tempo $n \geq 0$.

$$
\begin{aligned}
P(X_{n+1} = 2|X_n = 1) &= \frac{1}{2} \\
P(X_{n+1} = 3|X_n = 1) &= \frac{1}{2} \\
P(X_{n+1} = 1|X_n = 2) &= \frac{1}{3} \\
P(X_{n+1} = 3|X_n = 2) &= \frac{2}{3} \\
P(X_{n+1} = 1|X_n = 3) &= \frac{1}{3} \\
P(X_{n+1} = 2|X_n = 3) &= \frac{2}{3}.
\end{aligned}
$$

Como o rato começa na sala 1, temos que $\mathbf{P}(X_0 = 1) = 1$ e $\mathbf{P}(X_0 = 2) = \mathbf{P}(X_0 = 3) = 0$. Com estas informações em mãos podemos começar a entender o problema.

Queremos tentar calcular a proporção de tempo que o rato passa em cada sala após uma grande número de movimentos, o que mais formalmente poderia ser encontrado como

$$\lim_{n\to\infty} \frac{n(i)}{n},$$

onde $n(i)$ denota o total de vezes que o rato visitou a sala i após n minutos.

Para entender um pouco melhor a expressão acima, pensamos em um caso mais simples. Suponha que jogamos um dado n vezes, e contamos a proporção de vezes que o número 1 foi observado. Como a probabilidade do valor 1 ser observado é $1/6$ em cada uma das n jogadas, esperamos que quando n for suficientemente grande, a fração $n(1)/n$ esteja próxima de $1/6$.

Um elemento essencial neste exemplo é o fato de que a probabilidade de observarmos o valor 1 é constante e igual a $1/6$. Mas isso não ocorre no caso do nosso ratinho. No nosso caso $\mathbf{P}(X_n = i)$ parece depender de n para qualquer i. Mais a frente mostraremos como calcular tais valores, mas por enquanto observe, por exemplo que

$$\mathbf{P}(X_0 = 2) = 0$$

e como da sala 1 o rato pode ir para as salas 2 ou 3 com a mesma probabilidade,

$$\mathbf{P}(X_1 = 2) = \frac{1}{2} = \mathbf{P}(X_1 = 3).$$

Além disso

$$\begin{aligned}\mathbf{P}(X_2 = 2) &= \mathbf{P}(X_2 = 2; X_1 = 1) + \mathbf{P}(X_2 = 2; X_1 = 3)\\ &= \mathbf{P}(X_2 = 2|X_1 = 1)\mathbf{P}(X_1 = 1) + \mathbf{P}(X_2 = 2|X_1 = 3)\mathbf{P}(X_1 = 3)\\ &= \frac{1}{2}\cdot 0 + \frac{2}{3}\cdot\frac{1}{2}\\ &= \frac{1}{3}.\end{aligned}$$

Assim $\mathbf{P}(X_0 = 2) = 0$, $\mathbf{P}(X_1 = 2) = 1/2$ e $\mathbf{P}(X_2 = 2) = 1/3$, o que ilustra a dependência de $\mathbf{P}(X_n = 2)$ em n. Isso deixa claro que se quisermos encontrar a proporção de tempo que o rato passa em cada sala, precisamos antes entender como se comporta $\mathbf{P}(X_n = i)$ para cada sala i.

Neste sentido, observe que neste exemplo, sempre que o rato decide para que sala ele vai a única informação necessária é a sala onde ele se encontra, sem precisar saber como ele chegou até ali ou mesmo a quanto tempo ele está andando pelo labirinto. Sequências aleatórias com este tipo de propriedade são conhecidas como *Cadeias de Markov*, e serão nosso objeto de estudo pelo resto deste capítulo.

Definição 2.1. Uma **Cadeia de Markov** é uma sequência aleatória $X_0, X_1, X_2, \dots$ de elementos em um conjunto $S = \{a_1, a_2, \dots, a_l\}$ tal que

$$\mathbf{P}(X_{n+1} = j|X_n = i; X_{n-1} = i_{n-1}; \dots; X_0 = i_0) = \mathbf{P}(X_{n+1} = j|X_n = i), \qquad (2.1)$$

para qualquer $n \geq 0$ e quaisquer $i_0, i_1, \dots, i_{n-1}, i, j \in S$.

O conjunto S é chamado **espaço de estados** da cadeia, e a probabilidade $\mathbf{P}(X_{n+1} = j|X_n = i)$ é conhecida como **probabilidade de transição** do estado i para o estado j, e representada por P_{ij}. Ou seja,

$$P_{ij} = \mathbf{P}(X_{n+1} = j|X_n = i).$$

A equação (2.1) é central na definição de Cadeias de Markov e nos diz que, assim como no caso do rato no exemplo 2.1, a única informação necessária na decisão do estado da cadeia no instante n é o estado no instante anterior, sendo portanto independente de todo o resto da trajetória até aquele momento.

Para cada sítio i, as entradas $P_{ia_1}, P_{ia_2}, \ldots, P_{ia_l}$ representam as probabilidades de cada caminho disponível quando a cadeia sai do estado i, e devem portanto formar por si só uma distribuição de probabilidade. Ou seja, $P_i a_j \geq 0$ e

$$P_{ia_1} + P_{ia_2} + \cdots + P_{ia_l} = 1.$$

MATRIZ DE TRANSIÇÃO Agora que definimos os elementos que caracterizam uma cadeia de Markov, precisamos organizar estas informações de modo a melhor visualizar todos os elementos envolvidos. Assim, dada uma cadeia de Markov $(X_n)_{n\geq 0}$ com espaço de estados $S = \{a_1, \ldots, a_m\}$ e probabilidades de transição $P_{ij},\ i, j \in S$, definimos a **matriz de transição** $P = (P_{a_ia_j})_{m\times m}$ colocando na $i-$ésima linha e $j-$ésima coluna de P a probabilidade de transição $P_{a_ia_j}$ do estado a_i para o estado a_j.

A matriz P não apenas ajuda na organização dos elementos da cadeia, como tem um papel central nos cálculos que faremos mais a frente.

A seguir damos alguns exemplos para ilustrar melhor tudo o que vimos até aqui.

Exemplo 2.2. Voltando ao exemplo 2.1, o passeio do rato entre as 3 salas pode ser representado por uma cadeia de Markov $X_1, X_2, \ldots$, com espaço de estados $S = \{1, 2, 3\}$ e matriz de transição dada por

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

Exemplo 2.3. Considere um sistema de comunicação que transmite os dígitos 0 e 1. Cada dígito passa por várias etapas, e em cada passagem há uma probabilidade p que o dígito que entrou permaneça inalterado quando sair.

Se denotarmos por X_n o digito transmitido após a n-ésima passagem temos que

$$\mathbf{P}(X_{n+1} = 1|X_n = 1) = \mathbf{P}(X_{n+1} = 0|X_n = 0) = p$$

e

$$\mathbf{P}(X_{n+1} = 1|X_n = 0) = \mathbf{P}(X_{n+1} = 0|X_n = 1) = 1 - p$$

Assim, este problema pode ser modelado por uma Cadeia de Markov de dois estados, $\{0,1\}$, com matriz de transição dada por

$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}.$$

Exemplo 2.4. Sabemos que o humor de uma pessoa varia com o tempo. Para tentar modelar estas mudanças, suponha que uma pessoa pode estar bem humorada (B), indiferente (I) ou mal humorada (M). Se ela está bem humorada hoje, então ela estará bem humorada (B), indiferente (I) ou mal humorada (M) amanhã com probabilidades $0,5$, $0,4$ e $0,1$ respectivamente. Se ela está indiferente hoje, então ela estará B, I ou M amanhã com probabilidades $0,3$, $0,4$ e $0,3$ respectivamente. Finalmente, se ela estiver mal humorada hoje, então ela estará B, I ou M amanhã com probabilidades $0,2$, $0,3$ e $0,5$ respectivamente. Deste modo modelamos o humor de uma pessoa dia a dia como uma Cadeia de Markov de três estados ($S = \{B, M, I\}$), e matriz de transição

$$P = \begin{bmatrix} 0,5 & 0,4 & 0,1 \\ 0,3 & 0,4 & 0,3 \\ 0,2 & 0,3 & 0,5 \end{bmatrix}$$

Exemplo 2.5 (Bêbado no trânsito)**.** Consideremos uma pessoa guiando um automóvel por uma pista que tem três faixas de rolamento (estados $i = 1, i = 2$ e $i = 3$); tal pista não tem acostamentos, contando com um abismo à esquerda (estado $i = 0$) e à direita (estado $i = 4$). A permanência ou não do carro na faixa de rolamento atual é completamente aleatória, e a cada minuto (digamos) existe a possibilidade de mudança de faixa. Coloquemos probabilidade de $1/3$ do carro ir para a faixa à esquerda no minuto seguinte, probabilidade de $1/3$ dele permanecer na mesma faixa no minuto seguinte, e probabilidade de $1/3$ do carro ir para a faixa à direita no minuto seguinte. Podemos modelar tal situação por uma Cadeia de Markov, dado que o fato do carro mudar ou não de faixa no minuto seguinte é determinado apenas pelo estado atual em que está o carro. A matriz de transição para um passo (mudança do primeiro minuto para o segundo minuto) será:

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Essa fatídica viagem terminará se o carro cair em qualquer um dos dois abismos, o que significa alcançar o estado $i = 0$ ou o estado $i = 4$. Alcançado esse estado, o carro não sai mais de lá. Isso está representado nas linhas associadas aos estados 0 e 4 da matriz de transição, nas quais a probabilidade de permanência é 1 e as demais são 0. Assim , ao chegar nesses estados, o processo é "absorvido" e permanece ali para sempre. Por esta razão, chamamos tais estados de **absorventes**.

2.3 A TRANSIÇÃO EM n PASSOS E A EQUAÇÃO DE CHAPMAN-KOLMOGOROV

Dando sequência ao nosso estudo, lembre-se que queremos tentar entender como se comportam as probabilidades $\mathbf{P}(X_n = a)$ da cadeia estar em cada estado a no instante n. É importante observar que $\mathbf{P}(X_n = a)$ depende não apenas do estado a e do instante n, como também do estado em que iniciamos o processo. Deste modo, o que estamos interessados de fato são as probabilidades $\mathbf{P}(X_n = a | X_0 = b)$ para cada par de estados $a, b \in S$.

Tais valores para $n = 1$ são representados na matriz de transição da cadeia, mas gostaríamos de entender como calcular tais valores para qualquer instante n. Antes de mais nada, lembre-se que a cada passo a cadeia "esquece" sua trajetória até aquele instante, de modo que $\mathbf{P}(X_{k+n} = i | X_k = j)$ não depende de k. Isso nos permite definir $P_{ij}^{(n)}$ como a probabilidade do processo sair do estado a e chegar no estado b em n passos, isto é

$$P_{ij}^{(n)} = \mathbf{P}(X_{k+n} = j | X_k = i).$$

Assim podemos definir a **matriz de transição de n passos** para a cadeia com estados $\{a_1, \ldots, a_m\}$ como a matriz $P^{(n)}$ com entradas $P_{a_i a_j}^{(n)}$.

Deste ponto em diante notaremos $P_{a_i a_j}$ simplesmente por P_{ij}, e desta forma i poderá ser um estado em S ou representará o estado $a_i \in S$.

Para entender como calcular $\mathbf{P}^{(n)}$ precisamos antes entender como as matrizes $\mathbf{P}^{(n)}$ se relacionam entre si para distintos valores de n. Neste sentido, vamos primeiro tentar entender a probabilidade de que o processo, estando no estado i, chegar ao estado k em n passos e de, na sequência, sair do estado k e alcançar o estado j em mais m passos adicionais. Ou seja, queremos calcular

$$\mathbf{P}(X_n = k; X_{n+m} = j | X_0 = i).$$

Mas aplicando diretamente as definições de probabilidade condicional e de cadeia de Markov, encontramos que

$$\begin{aligned}
\mathbf{P}(X_n = k; X_{n+m} = j | X_0 = i) &= \frac{\mathbf{P}(X_{n+m} = j; X_n = k; X_0 = i)}{\mathbf{P}(X_0 = i)} \\
&= \mathbf{P}(X_{n+m} = j | X_n = k; X_0 = i) \cdot \frac{\mathbf{P}(X_n = k; X_0 = i)}{\mathbf{P}(X_0 = i)} \\
&= \mathbf{P}(X_{n+m} = j | X_n = k; X_0 = i)\mathbf{P}(X_n = k | X_0 = i).
\end{aligned}$$

Ou seja,

$$\mathbf{P}(X_n = k; X_{n+m} = j | X_0 = i) = P_{ik}^{(n)} \cdot P_{kj}^{(m)}.$$

Com isso podemos relacionar as entradas de $\mathbf{P}^{(n+m)}$, $\mathbf{P}^{(n)}$ e $\mathbf{P}^{(m)}$, bastando para isso perceber que para chegar de i até j em $n + m$ passos, a cadeia precisará passar por algum estado $k \in S$ no instante n, antes de seguir até j após mais m passos. Assim,

$$\begin{aligned}
P_{ij}^{(n+m)} &= \mathbf{P}(X_{n+m} = j | X_0 = i) \\
&= \mathbf{P}(X_n = 1; X_{n+m} = j | X_0 = i) + \cdots + \mathbf{P}(X_n = l; X_{n+m} = j | X_0 = i) \\
&= P_{i1}^{(n)} \cdot P_{1j}^{(m)} + P_{i2}^{(n)} \cdot P_{2j}^{(m)} + \cdots + P_{il}^{(n)} \cdot P_{lj}^{(m)}.
\end{aligned}$$

Mostramos assim o seguinte resultado.

Proposição 2.2 (Equação de Chapman-Kolmogorov)**.** *Dada uma cadeia de Markov com matriz de transição de n passos $P^{(n)} = (P_{ij}^{(n)})_{l \times l}$, vale que*

$$P_{ij}^{(n+m)} = \sum_{k=1}^{l} P_{ik}^{(n)} \cdot P_{kj}^{(m)}. \tag{2.2}$$

A equação (2.2) vem de encontro a interpretação matricial das equações. De fato, lembre-se que para calcular o elemento da linha i coluna j em $P^{(n)} \cdot P^{(m)}$ devemos

$$\begin{bmatrix} P_{11}^{(n)} & P_{12}^{(n)} & \cdots & P_{1l}^{(n)} \\ P_{21}^{(n)} & P_{21}^{(n)} & \cdots & P_{2l}^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{l1}^{(n)} & P_{l2}^{(n)} & \cdots & P_{ll}^{(n)} \end{bmatrix} \cdot \begin{bmatrix} P_{11}^{(m)} & P_{12}^{(m)} & \cdots & P_{1l}^{(m)} \\ P_{21}^{(m)} & P_{21}^{(m)} & \cdots & P_{2l}^{(m)} \\ \vdots & \vdots & \ddots & \vdots \\ P_{l1}^{(m)} & P_{l2}^{(m)} & \cdots & P_{ll}^{(m)} \end{bmatrix}$$

Figura 7: Esquema de cálculo de $P_{21}^{(n+m)}$

tomar os elementos da linha i de $P^{(n)}$ e multiplicar pelos respectivos elementos da coluna j de $P^{(m)}$ e somar, exatamente como é feito na equação (2.2). E portanto

$$P^{(n+m)} = P^{(n)} \cdot P^{(m)},$$

como ilustrado na figura 7.

Conseguiremos perceber melhor esta afirmação utilizando um exemplo simples. Seja $(P^{(n)})_{2\times 2}$ a matriz que traz a evolução de n passos de uma cadeia de dois estados, e seja $(P^{(m)})_{2\times 2}$ a matriz que traz a evolução de m passos da mesma cadeia. Tendo duas matrizes de mesma ordem, lançamos mão da multiplicação usual obtendo

$$\begin{bmatrix} P_{11}^{(n)} & P_{12}^{(n)} \\ P_{21}^{(n)} & P_{22}^{(n)} \end{bmatrix} \cdot \begin{bmatrix} P_{11}^{(m)} & P_{12}^{(m)} \\ P_{21}^{(m)} & P_{22}^{(m)} \end{bmatrix} = \begin{bmatrix} P_{11}^{(n)}P_{11}^{(m)} + P_{12}^{(n)}P_{21}^{(m)} & P_{11}^{(n)}P_{12}^{(m)} + P_{12}^{(n)}P_{22}^{(m)} \\ P_{21}^{(n)}P_{11}^{(m)} + P_{22}^{(n)}P_{21}^{(m)} & P_{21}^{(n)}P_{12}^{(m)} + P_{22}^{(n)}P_{22}^{(m)} \end{bmatrix},$$

que pela equação (2.2) é equivalente à

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

É interessante observar também que

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

Outra consequência importante da equação de Chapman-Kolmogorov é que $P^{(n)} = P^n$. Para ver isso, comece percebendo que por definição

$$P^{(1)} = P = P^1.$$

Calculando para $n = 2$ vemos que, por Chapman-Kolmogorov

$$P^{(2)} = P^{(1+1)} = P^{(1)} \cdot P^{(1)} = P \cdot P = P^2.$$

Seguindo, para $n = 3$ vale

$$P^{(3)} = P^{(2+1)} = P^{(2)} \cdot P^{(1)} = P^2 \cdot P = P^3.$$

Assim, seguindo o mesmo raciocínio para os demais valores de n concluímos que, de fato, $P^{(n)}$ é a n-ésima potência de P.

Por esta razão, deste ponto em diante, passaremos a usar P^n para indicar a matriz $P^{(n)}$, e $\mathbf{P}_{ij}^n$ no lugar de $P_{ij}^{(n)}$. É importante apontar, no entanto, que apesar de P^n ser a n-ésima potência de P, em geral **não podemos afirmar** que $P_{ij}^n = \left(P_{ij}\right)^n$, de modo que P_{ij}^n não deve ser visto como uma potência.

Exemplo 2.6. Voltando mais uma vez ao exemplo 2.1, queremos agora calcular a probabilidade do rato ir da sala 1 para a sala 3 em 4 passos. Ou seja, queremos encontrar $P_{13}^4 = \mathbf{P}(X_4 = 3|X_0 = 1)$.

Para isso lembramos que a matriz de transição P é dada por

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

e calculando P^4 encontramos

$$P = \begin{bmatrix} \frac{7}{27} & \frac{10}{27} & \frac{10}{27} \\ \frac{20}{81} & \frac{77}{182} & \frac{5}{18} \\ \frac{20}{81} & \frac{5}{18} & \frac{77}{182} \end{bmatrix}.$$

Assim

$$\mathbf{P}(X_4 = 3|X_0 = 1) = \frac{10}{27} \approx 0,3703.$$

2.3.1 *Tentando Interpretar P^n*

Como comentamos no início do capítulo, estamos interessados em calcular a proporção de tempo que o rato passa em cada um das salas 1, 2 e 3, e para isso concluímos que precisamos antes entender o comportamento de $\mathbf{P}(X_n = i|X_0 = 1)$ para n grande.

Calculando então a matriz de transição de n passos para $n = 8$ e $n = 16$ encontramos

$$P^8 \approx \begin{bmatrix} 0,25012 & 0,37494 & 0,37494 \\ 0,24996 & 0,39452 & 0,35552 \\ 0,24996 & 0,35552 & 0,39452 \end{bmatrix}$$

e

$$P^{16} \approx \begin{bmatrix} 0,25 & 0,375 & 0,375 \\ 0,24999 & 0,37576 & 0,37425 \\ 0,24999 & 0,37425 & 0,37576 \end{bmatrix}.$$

Apesar dos valores acima serem aproximados, eles parecem indicar que algo bastante interessante está acontecendo: as distribuições de probabilidade de cada linha parecem estar se aproximando rapidamente de $[0,25 \quad 0,375 \quad 0,375]$.

Se isso for verdade podemos concluir duas coisas bastante importantes. A primeira é que a proporção de tempo que o rato passa em cada sala após um longo período de caminhada, deve ser aproximadamente igual à $[0,25 \quad 0,375 \quad 0,375]$, respondendo nossa questão inicial!

Mas além disso, este comportamento parece ser o mesmo para cada linha da matriz, o que reforçaria a ideia de que o rato realmente esquece a sala de onde saiu.

Mas surgem agora algumas novas questões. Este comportamento está realmente acontecendo, ou foi apenas uma feliz escolha de potências? Ele acontece para qualquer cadeia de Markov? Se não, sob quais condições ele acontece?

Antes de continuar explorando estas questões, vamos colocar mais alguns exemplos.

Exemplo 2.7. Um laboratório faz testes para medir a memória recente de camundongos, e para isso os submete a um labirinto simples, que pode levar a três caminhos: recompensa (R), choque elétrico (C) e encontro com o espelho (E). Um camundongo é exposto a uma sequência de passeios pelo labirinto, com um intervalo de algumas horas entre um passeio e outro. Através de levantamento das frequências relativas de ocorrência dos diversos comportamentos da cobaia, chegou-se a algumas conclusões:

a) O camundongo nunca vai para o espelho duas vezes seguidas;

b) Se ele encontra a recompensa, a chance de encontrar novamente a recompensa no próximo passeio é de 50%, e a probabilidade de encontrar o espelho (bem como o choque elétrico) é de 25%;

c) Se ele encontra o espelho, as probabilidades de ele ir para a recompensa ou para o choque elétrico no próximo passeio são iguais;

d) Se ele encontra o choque, a chance dele no próximo passeio encontrar novamente o choque é de 50%, e a probabilidade de encontrar a recompensa (bem como o espelho) é de 25%.

Tem-se então uma matriz de transição da forma:

$$P = \begin{bmatrix} 0,5 & 0,25 & 0,25 \\ 0,5 & 0 & 0,5 \\ 0,25 & 0,25 & 0,5 \end{bmatrix}$$

Podemos agora calcular as probabilidades de transição para tempos maiores.

$$P^2 = \begin{bmatrix} 0,4375 & 0,1875 & 0,375 \\ 0,375 & 0,25 & 0,375 \\ 0,375 & 0,1875 & 0,4375 \end{bmatrix}$$

$$P^4 \approx \begin{bmatrix} 0,40234 & 0,19922 & 0,39844 \\ 0,39844 & 0,20312 & 0,39844 \\ 0,39844 & 0,19922 & 0,40234 \end{bmatrix}$$

e

$$P^8 \approx \begin{bmatrix} 0,40001 & 0,2 & 0,39999 \\ 0,39999 & 0,20002 & 0,3999 \\ 0,39999 & 0,2 & 0,40001 \end{bmatrix}$$

Assim como no exemplo 2.1, P^n parece estar se aproximando de uma matriz de transição com todas as linhas iguais. Isso significa que após algumas tentativas, a probabilidade de que o camundongo encontre o choque elétrico será 0,40 e isso **independe de qualquer resultado de passeios anteriores**.

Considere agora que, por qualquer razão, nós saibamos que P^n se aproxima de uma dada matriz P^* quando n cresce, e que todas as linhas de P^* são iguais. Será que é possível encontrar tal matriz sem precisar calcular P^n exaustivamente?

Para responder esta pergunta, vamos primeiro lembrar uma importante propriedade da multiplicação de matrizes: ao multiplicarmos duas matrizes A e B, a i-ésima linha do produto $A \cdot B$ é de fato o resultado do produto da i-ésima linha de A, vista como uma matriz com 1 linha, por B.

Para ilustrar melhor esta propriedade considere o produto

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & -1 \\ -1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 2 \\ -5 & 0 & 4 \\ 7 & 5 & 0 \end{bmatrix},$$

e observe que

$$\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 2 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -5 & 0 & 4 \end{bmatrix},$$

e

$$\begin{bmatrix} -1 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 3 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 7 & 5 & 0 \end{bmatrix}.$$

Posto isso, voltemos à nossa matriz P^*. Sabemos que P^n se aproxima de P^* quando n é muito grande, e para simplificar notaremos isso por $P^n \to P^*$. Com isso podemos dizer que, mas se $P^n \to P^*$, então deveríamos ter que

$$P^n = P^{n-1} \cdot P \approx P^* \cdot P,$$

mostrando que P^* deve ser tal que

$$P^* = P^* \cdot P.$$

Agora, como todas as linhas de P^* são iguais, se chamarmos esta linha comum de v, segue da propriedade acima que

$$vP = v.$$

Além disso, como para todo n as entradas de cada linha de P^n são positivas, e somam 1, o mesmo acontece com os elementos de v.

Em resumo, se a cadeia possui m estados, estamos buscando $v = [x_1\ x_2\ \cdots\ x_m]$, com $x_i \geq 0,\ i = 1, \dots m,$ e $x_1 + \cdots + x_m = 1$, tal que $vP = P$.

Voltando ao exemplo anterior, onde

$$P = \begin{bmatrix} 0,5 & 0,25 & 0,25 \\ 0,5 & 0 & 0,5 \\ 0,25 & 0,25 & 0,5 \end{bmatrix},$$

Estamos buscando $v = [x\ y\ z]$, com $x + y + z = 1$ tal que $vP = P$. Ou seja,

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} 0,5 & 0,25 & 0,25 \\ 0,5 & 0 & 0,5 \\ 0,25 & 0,25 & 0,5 \end{bmatrix} = \begin{bmatrix} x & y & z \end{bmatrix}.$$

Isso nos leva ao sistema

$$\begin{cases} 0,5x & + & 0,5y & + & 0,25z & = & x \\ 0,25x & + & & + & 0,25z & = & y \\ 0,25x & + & 0,5y & + & 0,5z & = & z \\ x & + & y & + & z & = & 1 \end{cases},$$

e consequentemente a

$$v = \begin{bmatrix} 0,4 & 0,2 & 0,4 \end{bmatrix},$$

como esperávamos.

Como comentamos antes, v é uma distribuição de probabilidade, que representa o comportamento das probabilidades $\mathbf{P}(X_n = j|X_0 = 1)$ para n suficientemente grande, e não depende do estado i. Pois estas razões, ele é conhecido como **distribuição de equilíbrio da cadeia**. Voltaremos a isso mais tarde, ainda neste capítulo.

2.4 CLASSIFICANDO ESTADOS DE UMA CADEIA

Em uma cadeia de Markov, tanto a análise das características de cada estado como o estudo da relação entre dois estados quaisquer têm grande relevância.

Com efeito, se consideramos o conjunto dos estados possíveis em uma Cadeia de Markov, podemos partir de um ponto arbitrário para iniciar o processo interativo? Se isso estiver garantido, é possível passar por todos os estados da Cadeia ou existirão estados que são inacessíveis? E se damos início ao processo partindo desses estados inacessíveis, o que acontece no decorrer do processo? Existe caminho de ida e de volta entre quaisquer dois estados escolhidos ao acaso? Haverá na Cadeia um estado (ou um grupo de estados) com a característica peculiar de absorver o processo para ele? Se o processo passa por um dado estado, é possível retornar a ele? Com qual probabilidade? Existe algum tipo de periodicidade no processo interativo? Como detectá-la?

Sob este enfoque, traremos aqui algumas definições centrais para a construção da teoria; utilizaremos, além do tratamento matricial das cadeias, uma explicitação de cada definição por meio de um desenho esquemático do espaço de estados e das probabilidades de transição, que auxiliará a visualização de algumas destas propriedades.

Definição 2.3. Dada uma cadeia de Markov com espaço de estados S e matriz de transição P, diremos que o estado $j \in S$ é **acessível** a partir do estado $i \in S$ $(i \to j)$, se existir $n \geq 0$ tal que $P_{ij}^n > 0$. Ou seja, se partindo de i temos uma probabilidade positiva de atingir j em algum instante.

Se $i \to j$ e $j \to i$, diremos que i e j são **comunicáveis**, e notaremos por $i \leftrightarrow j$.

Verificar se $i \to j$ calculando potências da matriz de transição pode ser um trabalho árduo e pouco frutífero. Uma maneira mais prática é usar a equação de Chapman-Kolmogorov. Para isso, lembre-se que de (2.2) temos que

$$P_{ij}^n = \sum_{k=1}^{l} P_{ik} P_{kj}^{n-1} \geq P_{ii_1} P_{i_1 j}^{n-1}.$$

Assim, se encontrarmos um caminho $i = i_o, i_1, i_2 \ldots, i_n = j$ com

$$P_{i_k i_{k+1}} > 0,$$

para todo $k = 0, 1, \ldots, n-1$, temos que

$$P_{ij}^n \geq P_{ii_1} P_{i_1 j}^{n-1} \geq P_{ii_1} P_{i_1 i_2} P_{i_2 j}^{n-2} \geq P_{ii_1} P_{i_1 i_2} \cdots P_{i_{n-2} i_{n-1}} P_{i_{n-1} j} > 0.$$

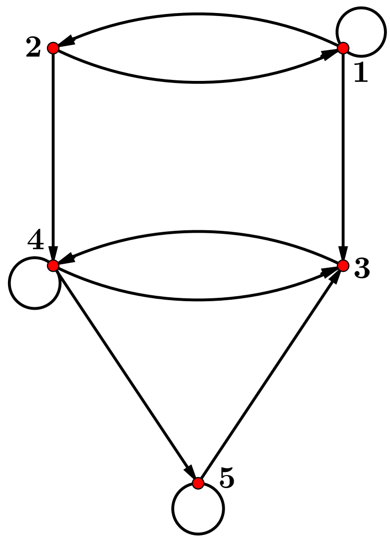
Uma forma que pode auxiliar na decisão de acessibilidade entre estados é fazer uma representação pictórica da cadeia usando um *grafo*. Para cada estado da cadeia, colocamos um ponto no plano. Se $P_{ij} > 0$ colocamos uma seta de i para j. Se $P_{ii} > 0$ desenhamos um laço em i.

Para ilustrar melhor o processo e a definição de acessibilidade, vamos a uma exemplo.

Exemplo 2.8. Considere uma Cadeia de Markov com espaço de estados $S = \{1, 2, 3, 4, 5\}$ e matriz de transição

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & 0 & \frac{1}{4} & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Graficamente, a cadeia acima pode ser representada por



Pela figura é fácil ver que $1 \to 5$, usando o caminho $1, 2, 4, 5$ ou o caminho $1, 3, 4, 5$. No entanto $5 \not\to 1$.

Do mesmo modo $2 \to 3$, mas $3 \not\to 2$.

Por outro lado, $3 \to 5$, pelo caminho $3, 4, 5$, e $5 \to 3$, de modo que $3 \leftrightarrow 5$.

De fato, uma análise mais cuidadosa mostra que

$$1 \leftrightarrow 2$$

e

$$3 \leftrightarrow 4, \;\; 4 \leftrightarrow 5 \text{ e } 3 \leftrightarrow 5,$$

e qualquer outra relação de acessibilidade tem apenas uma direção. Com isso podemos dividir os estados desta cadeia em dois grupos, ou classes, formados por elementos que se comunicam entre si.

Esta divisão é possível pelo seguinte resultado.

Proposição 2.4. *A relação de comunicabilidade ($\leftrightarrow$) entre estados de uma cadeia de Markov é uma* ***relação de equivalência.*** *Ou seja, para quaisquer $i, j, k \in S$ vale que*

1. *$i \leftrightarrow i$;*
2. *Se $i \leftrightarrow j$, então $j \leftrightarrow i$;*
3. *Se $i \leftrightarrow j$ e $j \leftrightarrow k$, então $i \leftrightarrow k$.*

Demonstração. Note que, pela definição de probabilidade condicional, $\mathbf{P}(X_0 = i|X_0 = i) = 1$ para qualquer $i \in S$, e portanto $P_{ii}^0 = 1 > 0$ e $i \leftrightarrow i$.

A propriedade *2* segue direto da definição de comunicabilidade.

Para terminar, note que se $i \leftrightarrow j$ e $j \leftrightarrow k$, então existem n e m tais que $P_{ij}^n > 0$ e $P_{jk}^m > 0$, e portanto por Chapman-Kolmogorov

$$P_{ik}^{n+m} \geq P_{ij}^n \cdot P_{jk}^m > 0,$$

e portanto $i \leftrightarrow k$. □

Como comentamos acima, este resultado nos permite separar os estados da cadeia em conjuntos ou **classes** disjuntas, formadas apenas por estados que se comunicam entre si.

Para isso, dado um estado $i \in S$ defina a **classe** de i por

$$C_i = \{j \in S : i \leftrightarrow j\}.$$

Ou seja, C_i é o conjunto de todos os estados que se comunicam com i. A Classe C_i é **irredutível** se j percorre todo S.

É interessante notar que se $C_i \cap C_j \neq \emptyset$, então $C_i = C_j$. Para ver isso tome $a \in C_i \cap C_j$, e note que um estado k está em C_j se, se só se, $j \leftrightarrow k$ ou ainda $k \leftrightarrow j$. Assim, como $a \in C_j$, temos que $j \leftrightarrow a$, e portanto $k \leftrightarrow a$. Agora, como $a \in C_i$ temos que $i \leftrightarrow a$, de onde segue que $i \leftrightarrow k$, de onde concluímos que $k \in C_i$. Assim

$$k \in C_j \Leftrightarrow k \in C_i,$$

e portanto $C_i = C_j$.

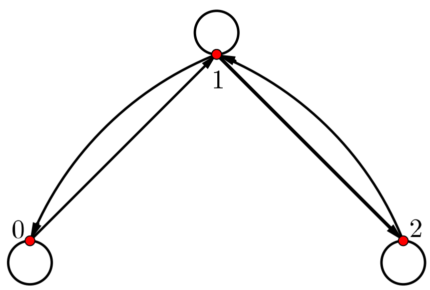
Outra forma de colocar a propriedade acima é que duas classes irredutíveis em uma Cadeia de Markov, ou são idênticas ou são disjuntas.

Definição 2.5. Uma cadeia de Markov é dita **irredutível** se a relação de comunicabilidade divide o espaço de estados em apenas uma classe irredutível. Em outras palavras, a cadeia é irredutível se para qualquer par de estados $i, j \in S$ temos que $i \leftrightarrow j$.

Exemplo 2.9. A Cadeia de três estados $\{0, 1, 2\}$ representada pela matriz de transição

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix},$$

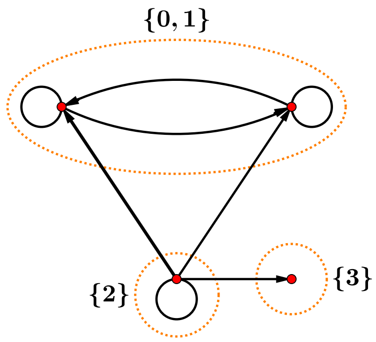
é claramente irredutível, pois se transita do estado 0 para o estado 1 com probabilidade $1/2$, do estado 1 para o estado 2 com probabilidade $1/4$, do estado 2 para o estado 1 com probabilidade $1/3$ e finalmente do estado 1 para o estado 0 com probabilidade $1/2$, completando o ciclo.



Exemplo 2.10. Uma Cadeia, de quatro estados $\{0, 1, 2, 3\}$ e que tenha matriz de transição

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

revela três classes, a saber: $\{0, 1\}$, $\{3\}$ e $\{2\}$. Colocando mais explicitamente, $\{0, 1\}$ é uma classe pois, entre os estados 0 e 1 há acessibilidade de ida e volta independentemente se o processo está em 0 ou em 1. Se configura $\{3\}$ como uma classe pois é um estado absorvente da Cadeia e isolado do processo, inacessível a partir de qualquer outro estado e que não acessa nenhum outro ponto da Cadeia, a não ser ele mesmo. Temos que $\{2\}$ é uma classe pois não há acessibilidade para ele a partir dos estados 0 e 1. Pelo exposto acima, a Cadeia apresentada não é irredutível, o que também fica mostrado na representação esquemática da figura abaixo



Continuando nosso estudo, vamos agora olhar para o comportamento dos estados da cadeia ao longo do tempo. A próxima definição diz respeito a possibilidade da cadeia retornar a um dado estado ao longo da sua trajetória. Em outras palavras, estamos interessados aqui em saber se ao sair de um certo estado i, a cadeia voltará ou não a visitar tal estado.

Definição 2.6. Para um estado $i \in S$ qualquer, denotemos por f_i a probabilidade que o processo iniciado no estado i retorne novamente ao estado i. Diremos que o estado i é recorrente se $f_i = 1$ e transiente se $f_i < 1$.

Além disso, se todos os estados da cadeia forem recorrentes (resp. transientes), diremos simplesmente que a cadeia é recorrente (resp. transiente).

Esta é um conceito mais complicado de entender, então gastaremos algum tempo nele. Verificar a recorrência ou transiência de um estado nem sempre é fácil de determinar olhando apenas para a matriz, mas vamos tentar colocar algumas ideias de como fazê-lo.

Para isso, voltemos ao exemplo 2.10, e analisemos estado por estado.

- **Estado 3** - Este é o mais simples, pois a $\mathbf{P}(X_{n+1} = 3|X_n = 3) = 1$ e assim, se a cadeia iniciar em 3, ela nunca mais sairá de lá, e portanto $f_3 = 1$ e 3 é recorrente;
- **Estados 0 e 1** - Começando em qualquer um destes estados, o próximo poderá ser qualquer um deles com probabilidade $1/2$ cada. Ou seja,

$$\mathbf{P}(X_1 = |X_0 = 1) = \mathbf{P}(X_1 = 2|X_0 = 1) = \mathbf{P}(X_1 = 1|X_0 = 2) = \mathbf{P}(X_1 = 2|X_0 = 2) = \frac{1}{2}.$$

Isso se repetirá a cada instante de tempo, e portanto os estados se alterarão como no lançamento de uma moeda, e portanto $f_1 = f_2 = 1$.

- **Estado 2** - Este estado é um pouco mais complicado. Observe que se a cadeia iniciar em 4, para retornar à 4 ela, de fato, não pode sair! Isso por que se ele for para 3, a cadeia permanecerá lá, como já vimos. E o mesmo acontecerá se for para 1 ou 2. Com isso,

$$f_4 = \mathbf{P}(X_1 = 4 | X_0 = 4) = \frac{1}{4} < 1,$$

e portanto 4 é transiente.

Para entender melhor o conceito de recorrência e transiência, suponha que um certo estado i seja recorrente ($f_i = 1$), e que a cadeia partiu deste mesmo estado. Como a probabilidade de retorno é 1, certamente a cadeia visitará novamente o estado i. Quando isso acontecer, como os próximos passos da cadeia não dependem da trajetória dela até este momento, a probabilidade de retorno é novamente 1, e cadeia certamente retornará uma terceira vez ao estado i. Este argumento pode ser repetido a cada visita da cadeia à i, e portanto o estado i será visitado **infinitas vezes** durante a trajetória da cadeia.

Por outro lado, se o estado i for transiente há uma probabilidade $f_i^c = 1 - f_i > 0$ de que o processo não retorne mais ao estado i, sempre que passar por ele. Ou seja, iniciado o processo no estado i, a probabilidade do processo retornar ao menos k vezes à i tendo começar a i, antes de abandonar definitivamente o estado, é f_i^k.

Chame de N_i o total de vezes que a cadeia visita i ao longo de sua trajetória. Dizer que N_i é infinito é o mesmo que dizer que $N_i \geq n$ **para todo** $n > 0$, e portanto

$$\mathbf{P}(N_i = \infty) = \mathbf{P}(N_i \geq n \textbf{ para todo } n > 0) \leq \mathbf{P}(N_i \geq k) = f_i^k, \qquad \text{para qualquer } k.$$

Assim, como $f_i < 1$ e

$$\mathbf{P}(N_i = \infty) \leq \mathbf{P}(N_i \geq k) f_i^k, \qquad \text{para qualquer } k,$$

podemos tomar k tão grande quanto quisermos para concluir que $\mathbf{P}(N_i = \infty) = 0$.

Logo, podemos dizer que i é transiente se o total de visitas à i é finita.

Dentre outras coisas, que estudaremos posteriormente, esta relação nos revela uma propriedade interessante, descrita na seguinte proposição.

Proposição 2.7. *Em uma Cadeia de Markov com finitos estados, é impossível que todos sejam transientes.*

Demonstração. Para ver isso, suponha que $S = \{1, 2, \cdots, M\}$ e denote por a_1 o estado no qual a cadeia tem início. Como a_1 é transiente, vai existir um instante t_1 que marcará a última visita da cadeia à a_1. Chame o estado que a cadeia se encontra em $t_1 + 1$ de a_2, e observe que $a_2 \neq a_1$ e como a_2 é transiente, existirá um tempo t_2 tal que em $t_1 + t_2$ a cadeia visitará a_2 pela última vez. Chamando o próximo elemento de a_3, temos $a_3 \neq a_1$, $a_3 \neq a_2$, e seguindo o raciocínio até a_M, encontraremos um tempo $t_1 + \cdots + t_M$ para o qual a cadeia visitou a_M pela última vez. Como a cadeia não para, no tempo $t_1 + \cdots + t_M + 1$ ela deve visitar um novo sítio x, que deverá ser igual a a_k para algum k, contrariando o fato de a_k já ter sido visitado pela última vez.

Concluímos assim que existe ao menos um estado recorrente. □

Suponha que i é recorrente e que $i \to j$. Isso significa que existe n tal que $P_{ij}^n > 0$, e portanto a probabilidade q_{ij} da cadeia chegar à j após passar por i é positiva. De fato $q_{ij} \geq P_{ij}^n > 0$.

Assim, podemos dizer que após n visitas ao estado i (n grande), em aproximadamente $q_{ij} \cdot n$ destas visitas, a cadeia também visitou j, e portando o total de visitas à j é infinita, e j é recorrente.

Mostramos assim o seguinte resultado.

Proposição 2.8. *Em uma Cadeia de Markov, tanto recorrência quanto transiência são propriedades de classe.*

Juntando agora as Proposições 2.7 e 2.8, concluímos o seguinte.

Corolário 2.9. *Toda cadeia irredutível com espaço de estados finito é recorrente.*

A proposição 2.8 começa a simplificar bastante nossa busca por estados transientes e recorrentes em uma cadeia, permitindo que analisemos apenas um elemento em cada classe.

Para facilitar um pouco mais, considere agora uma cadeia de Markov qualquer, e tome uma classe irredutível $C \subset S$. Suponha que exista nesta classe um estado j com probabilidade $q_j > 0$ da cadeia não mais retornar à classe C uma vez que passe por j. Na representação gráfica, isso significaria a existência de uma seta de j para fora da classe C, sem nenhum outra de um estado externo para dentro de C.

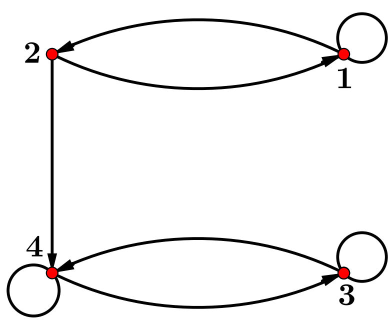
Isso significa que j é transiente, pois $1 - f_j \geq q_j > 0$, e pela Proposição 2.8 todo elemento em C são também transientes.

Por outro lado, se não existir tal elemento em C, então ao iniciar a trajetória em C a cadeia não deixará a classe, e pelos mesmos argumentos que levam à Proposição 2.7, podemos afirmar que existe um estado recorrente em C, e portanto todo estado em C é recorrente.

Exemplo 2.11. Considere uma cadeia de Markov com matriz de transição dada por

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & 0 & \frac{3}{4} & \frac{1}{4} \end{bmatrix}.$$

Podemos representar graficamente esta cadeia da seguinte forma.



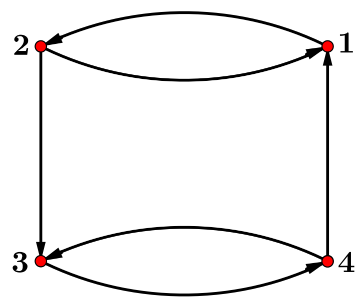
Antes de analisar a recorrência, vamos localizar as classes desta cadeia. Pela representação gráfica, é fácil ver que $1 \leftrightarrow 2$ e $3 \leftrightarrow 4$. Também é fácil de ver que existe apenas uma seta do conjunto $\{1,2\}$ para o conjunto $\{3,4\}$, mas não o contrário. Assim, podemos concluir que elementos de $\{3,4\}$ são acessíveis a partir de elementos de $\{1,2\}$, mas não o contrário. Portanto, as classes são $\{1,2\}$ e $\{3,4\}$.

Observando a cadeia com cuidado vemos que existe uma probabilidade positiva da cadeia ir de 2 para 4 e deixar a classe $\{1,2\}$. Segue daí que 1 e 2 são estados transientes.

Por outro lado, se a cadeia tiver início na classe $\{3,4\}$, ela permanecerá na classe, mostrando que os estados 3 e 4 são recorrentes.

A próxima propriedade que estudaremos está relacionado com a existência de algum "padrão" na sequência de visitas que uma cadeia faz a um determinado estado.

Para tentar ilustrar a propriedade que buscamos, considere uma cadeia de Markov cuja representação gráfica é dada abaixo.



Considere que a cadeia tem início no estado 3, e observe que para retornar à 3 existem, a princípio, duas possibilidades:

1. Seguir para 4 e imediatamente voltar para 3, o que tomaria 2 movimentos;
2. Ir para 4 (1 movimento), seguir para 1 (1 movimento), ficar alternando entre 1 e 2, terminando em 1 ($2k$ movimentos, ir de 1 para 2 e em seguida para 3 (mais 2 movimentos). Isso tomaria um total de $2k + 4$ movimentos.

Concluímos assim que a cadeia precisa de um total par de passos para retornar ao estado 3, e por esta razão diremos que 3 é periódico de período 2. Sendo um pouco mais formal podemos afirmar que $P_{33}^n > 0$ apenas para n par, ou ainda que o maior inteiro que divide todo $n > 0$ tal que $P_{33}^n > 0$ é 2.

Generalizando esta ideia, temos a seguinte definição.

Definição 2.10. Dada uma cadeia de Markov com matriz de transição P e espaço de estados S, definimos o **período** de i por

$$d(i) = mdc\{n \geq 0 : P_{ii}^n > 0\}.$$

Se $d(i) \geq 2$ diremos que i é periódico com **período** $d(i)$, e se $d(i) = 1$ diremos que i é **aperiódico**.

Se todos os estados da cadeia tiverem o mesmo período d diremos que a cadeia é periódica, se $d \geq 2$, ou aperiódica, se $d = 1$.

A relação de acessibilidade entre dois estados quaisquer da Cadeia também pode ser objeto de considerações sobre a periodicidade desses estados. Dizendo de uma forma mais direta, se dois estados i, j arbitrários são comunicáveis, que tipo de relação existe entre os respectivos períodos?

Conduziremos essa discussão supondo que $i \leftrightarrow j$. Assim, existem k, m tais que $P_{ij}^k > 0$ e $P_{ji}^m > 0$. Pelas equações de Chapman-Komolgorov segue então

$$P_{ii}^{(k+m)} \geq P_{ij}^k \cdot P_{ji}^m > 0,$$

e como o período do estado i é $d(i) = mdc\{n \geq 0 | P_{ii}^n > 0\}$, concluímos que $d(i)|k+m$.

Damos continuidade à análise tomando $n > 0$ tal que $P_j j^n > 0$. Nos valemos novamente de Chapman-Komolgorov para colocar que

$$P_{ii}^{k+n+m} \geq P_{ij}^k \cdot P_{jj}^n \cdot P_{ji}^m > 0,$$

trazendo-nos a conclusão que $d(i)|k+m+n$. Mas vimos que $d(i)|k+m$, e portanto $d(i)|n$.

Em resumo, vimos que $d(i)$ é divisor de todo elemento em $\{n \geq 0 : P_{jj}^n > 0\}$, e portanto $d(i) \leq d(j)$.

Usando exatamente os mesmos argumentos, apenas invertendo os papéis de i e de j, mostramos que $d(j) \leq d(i)$, e concluímos finalmente que $d(i) = d(j)$.

Acabamos de demonstrar que a periodicidade é uma **propriedade de classe**, ou seja, entre estados comunicáveis os períodos se transmitem e são numericamente idênticos. Nessa altura, é fundamental perceber que, se estivermos lidando com uma Cadeia de Markov irredutível, então teremos então um único padrão de periodicidade que regerá toda a Cadeia.

2.5 CONSIDERAÇÕES SOBRE A CONVERGÊNCIA DE P^n

Nesta sessão vamos voltar ao problema que permeou todo este capítulo. Ou seja, o comportamento de P^n para n suficientemente grande. Já vimos exemplos que nos indicaram a possibilidade de P^n convergir uma matriz P^*, com todas as linhas iguais.

Esta matriz-linha ou vetor, que chamamos de v, representa a chamada **distribuição de equilíbrio** da cadeia, e representa a probabilidade da cadeia se encontrar em cada estado do sistema depois de um total razoável de movimentos.

Mais claramente, se a cadeia tem estados $S = \{1, 2, 3 \ldots, m\}$ então a distribuição de equilíbrio tem forma

$$v = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix},$$

onde v_j representa a probabilidade de equilíbrio do estado j, ou seja, independente que estado a cadeia tenha início, a probabilidade dela estar em j se aproxima de v_j depois de um tempo suficientemente grande. Matematicamente,

$$P_{ij}^n = \mathbf{P}(X_n = j | X_0 = i) \to v_j,$$

para qualquer i.

Outra interpretação, como já colocamos no início do capítulo, é ver estas entradas como a proporção de tempo que a cadeia passa em cada estado ao longo de sua trajetória.

Nós já vimos como calcular a distribuição de equilíbrio, caso ela exista, mas ainda não discutimos quando ela realmente existe.

Dessa maneira, queremos saber quais serão as condições que garantiriam essa convergência de P^n. Vimos conceitos como: comunicabilidade entre estados, transiência e recorrência, periodicidade, e tudo isso estará em jogo para decidirmos se uma Cadeia de Markov convergirá.

2.5.1 *Convergência, Recorrência e Irredutibilidade*

Comecemos tratando da relação entre recorrência/transiência e o comportamento de P^n.

Lembre-se que estamos estudando apenas cadeias com espaço de estados finito, e portanto teremos sempre algum estado recorrente.

Suponha agora que i é um estado transiente de uma certa cadeia de Markov. Como i é transiente, o total de visitas a ele durante a cadeia é certamente 0, e assim a proporção de visitas à i durante a trajetória da cadeia deverá ser 0. Em outras palavras, se i é transiente $v_i = 0$.

Isso significa apenas que estados transientes não são importantes no estudo da convergência de $\mathbf{P}^n$, e portanto podemos considerar apenas cadeias recorrentes.

Considere inicialmente uma cadeia de Markov, com matriz P, e com mais de uma classe irredutível de estados recorrentes.

Tome C_1 e C_2 duas classes distintas, e considere $i \in C_1$ e $j \in C_2$. Como ambas as classes são recorrentes, sabemos que não existe comunicação entre estados destas classes. Em particular, sabemos que $i \not\to j$ e $j \not\to i$. Isso significa que

$$P^n_{ji} = P^n_{ij} = 0, \quad \text{para todo } n \geq 0,$$

e deveríamos ter $v_i = v_j = 0$, o que claramente é um absurdo.

Isso mostra que para garantir a convergência de P^n nas condições que estamos buscando, é importante que a cadeia tenha apenas uma classe de estados recorrentes.

Isso não significa que não exista um vetor de probabilidades que satisfaça $vP = v$, ou mesmo que não ocorra a convergência de P^n para uma matriz com linhas distintas uma das outras.

De fato, se $i, j \in C_1$, por exemplo, então $P^n_{ij} > 0$ para algum $n > 0$, de modo que P^n_{ij} ainda pode se aproximar de algum valor, mas este valor agora dependeria tanto de i quanto de j.

Vamos tentar esclarecer isso com um exemplo.

Exemplo 2.12. Considere uma cadeia de Markov com estados $\{1, 2, 3, 4\}$ e matriz de transição

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{3}{4} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}.$$

Fica claro que a cadeia tem classes $\{1, 2\}$ e $\{3, 4\}$ e que ambas são compostas de estados recorrentes.

Calculando P^2 encontramos

$$\begin{bmatrix} \frac{5}{12} & \frac{7}{12} & 0 & 0 \\ \frac{7}{18} & \frac{11}{18} & 0 & 0 \\ 0 & 0 & \frac{5}{16} & \frac{11}{16} \\ 0 & 0 & \frac{11}{36} & \frac{25}{36} \end{bmatrix}.$$

Observe que os elementos de P^2 que representam transição entre classes permanecem 0, como deveriam. O mesmo acontece nas demais potências de P.

O que vemos no exemplo anterior é que matrizes de transição não-irredutíveis podem ser escritas em "blocos matriciais", que representarão as distintas classes irredutíveis da Cadeia. Assim, ao passo em que n cresce, os blocos continuarão presentes na matriz P^n, e cada um desses conjuntos disjuntos poderá convergir para um vetor distinto dos outros.

Para tentar entender um pouco mais sobre o problema, tentemos resolver o sistema que levaria à distribuição de equilíbrio. Neste caso teríamos

$$\begin{cases} \frac{1}{2}x + \frac{1}{3}y = x \\ \frac{1}{2}x + \frac{2}{3}y = y \\ \frac{1}{4}z + \frac{1}{3}w = z \\ \frac{3}{4}z + \frac{2}{3}w = w \\ x + y + z + w = 1 \end{cases},$$

que é um sistema com infinitas soluções. Para tentar encontrar algumas destas soluções, vamos separar os equações por classes irredutíveis. Para isso, vamos separar a equação $x+y+z+w=1$ em $x+y=\lambda$ e $z+w=1-\lambda$, para algum $\lambda > 0$, e encontrar assim os sistemas

$$\begin{cases} \frac{1}{2}x + \frac{1}{3}y = x \\ \frac{1}{2}x + \frac{2}{3}y = y \\ x + y = \lambda \end{cases} \quad \text{e} \quad \begin{cases} \frac{1}{4}z + \frac{1}{3}w = z \\ \frac{3}{4}z + \frac{2}{3}w = w \\ z + w = 1-\lambda \end{cases}.$$

Resolvendo os sistemas acima, encontramos

$$\begin{aligned} \begin{bmatrix} x & y & z & w \end{bmatrix} &= \begin{bmatrix} \frac{2}{5}\lambda & \frac{3}{5}\lambda & \frac{4}{13}(1-\lambda) & \frac{9}{13}(1-\lambda) \end{bmatrix} \\ &= \lambda \begin{bmatrix} \frac{2}{5} & \frac{3}{5} & 0 & 0 \end{bmatrix} + (1-\lambda) \begin{bmatrix} 0 & 0 & \frac{4}{13} & \frac{9}{13} \end{bmatrix} \end{aligned}$$

Com isso encontramos infinitas "distribuições de equilíbrio" para a cadeia, em função de dois vetores, cada um deles assumindo valores positivos para uma classe diferente. Podemos ver tais vetores como as probabilidades de equilíbrio de cada classe. Isso porque, iniciando dentro da classe $\{1,2\}$, a probabilidade da cadeia estar em 1 se aproxima de $2/5$ quando n cresce.

Percebemos, concluindo essa análise, que é fundamental que a Cadeia seja recorrente e irredutível para que possamos tentar obter a convergência que desejamos.

2.5.2 *Convergência e Aperiodicidade*

Para entender como a periodicidade de elementos da cadeia pode influenciar no comportamento das matrizes P^n, vamos primeiro analisar uma cadeia específica.

Considere então a cadeia de Markov $(X_n)_{n\geq 0}$ com estados $\{0,1\}$ e matriz de transição

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

Esta é uma cadeia periódica, de período 2, que consiste apenas da inversão ao longo do tempo entre os estados 0 e 1.

Assim, se $X_0 = 0$ teremos $X_n = 1$ nos instantes ímpares e $X_n = 0$ para n par. O contrário acontece se $X_0 = 1$.

Estamos buscando uma distribuição de equilíbrio, e isso é conseguido se resolvermos a equação matricial

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix},$$

com $x+y=1$. Ou seja, queremos x,y com $x=y$ e $x+y=1$. Isso nos leva a $x=y=1/2$.

Portanto, a periodicidade da cadeia não nos impediu de calcular o candidato à distribuição de equilibrio, mas isso ainda não garante que P^n converge para o vetor encontrado.

De fato, como os estados se alternam, garantimos que, por exemplo, $P^n_{0,1} = \mathbf{P}(X_n = 1|X_0 = 0) = 0$ sempre que n for par, e do mesmo modo $P^n_{0,1} = 1$ para n impar. Isso mostra que, neste exemplo ao menos, os valores de P^n_{01} ficam se alternando, sem se aproximar de nenhum valor.

Matricialmente, calculando P^n vemos facilmente que

$$P^n = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

para n par, e

$$P^n = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

para n ímpar.

Este simples exemplo já nos mostra que se quisermos que a matriz P^n se aproxime de uma matriz P^* com linhas iguais, é necessário que a cadeia seja aperiódica.

2.6 O TEOREMA ERGÓDICO

Nas duas últimas seções entendemos melhor o significado das linhas da matriz de transição de uma Cadeia de Markov, bem como conseguimos classificar seus estados, chamando atenção especial para as características de recorrência/transiência e periodicidade de um estado.

Conseguimos também entender a relação de tais propriedades com a possível convergência de P^n. Para concluir este capítulo vamos apresentar um teorema central no estudo de cadeias de Markov. Este é o teorema que cuida da convergência da sequência P^n, dando condições suficientes para que ela ocorra.

Infelizmente, a demonstração de tal resultado foge muito do escopo deste trabalho, de modo que o enunciaremos sem provar.

Teorema 2.11 (Teorema Ergódico)**.** *Seja $(X_n)_{n\geq 0}$ uma cadeia de Markov recorrente, irredutível e aperiódica, com matriz de transição P e espaço de estados $S = \{1, 2, \cdots, m\}$.*

Nestas condições, para cada $j \in S$, existe $v_j \geq 0$ tal que

$$P_{ij}^n \to v_j,$$

para qualquer estado inicial $i \in S$.

Além disso, os valores v_j, $j \in S$ formam a única solução não-negativa do sistema de equações

$$\begin{cases} v_j = \sum_{i=0}^m v_i P_{ij}, \quad j \in S \\ \sum_{j=0}^m v_j = 1 \end{cases}.$$

Por esta razão, cadeias recorrentes, irredutíveis e aperiódicas são chamadas de cadeias **ergódicas**.

Para finalizar o capítulo, estudemos um último exemplo.

Exemplo 2.13. Consideremos três hotéis à beira-mar, em uma das movimentadas praias do nordeste brasileiro, que recebem milhares de turistas todo o ano. Os hotéis Areado (A), Bromélias (B) e Ceará (C), todos da mesma rede hoteleira, estão dispostos em sequência ao longo da faixa de areia; devido à estreiteza dessa praia, é impossível ir diretamente do hotel Areado ao hotel Ceará. As distâncias entre os hotéis Areado e Bromélias e entre os hotéis Bromélias e Ceará são aproximadamente as mesmas.

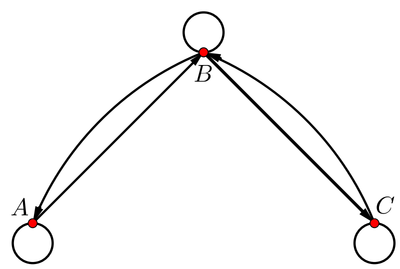
Alguns turistas gostam de ir trocando de hotel durante a estada nesta praia, outros não. Qualquer que seja o hotel em que o turista esteja, seja de 50% a chance de ele permanecer nesse mesmo hotel no dia seguinte. Consideremos também que quem está no Areado tem 50% de se dirigir ao Bromélias no dia seguinte; em contrapartida, quem está no Bromélias tem apenas 25% de chance de optar pelo Areado no dia seguinte. Além disso, os hóspedes do Bromélias tem 25% de chance de querer trocar de hospedagem e se dirigirem ao Ceará no outro dia, enquanto 50% das pessoas que estão no Ceará acabam indo para o Bromélias no dia subsequente.

Para poder controlar melhor os custos de cada hotel, a rede que os gerencia gostaria de saber qual a proporção do total de hospedes da rede que se hospeda em cada hotel ao longo de uma temporada agitada.

Para isso, vamos modelar o problema com uma matriz de transição de 3 estados $A,\ B$ e C, cuja matriz de transição será dada por

$$\begin{bmatrix} 0,5 & 0,5 & 0 \\ 0,25 & 0,5 & 0,25 \\ 0 & 0,5 & 0,5 \end{bmatrix}.$$

Representando a cadeia graficamente, encontramos



e fica claro que a matriz é irredutível, e portanto recorrente. Para ver que é aperiódica, basta notar que $P_{ii} > 0$ para todo estado da cadeia.

Por ser uma cadeia irredutível, com estados todos recorrentes e aperiódicos, sua distribuição de equilíbrio, denotada por $v = (v_1, v_2, v_3)$, pode ser obtida resolvendo o sistema linear associado à cadeia. Assim temos

$$\begin{cases} 0,5x & + & 0,25y & & & = & x \\ 0,5x & + & 0,5y & + & 0,5z & = & y \\ & + & 0,25y & + & 0,5z & = & z \\ x & + & y & + & z & = & 1 \end{cases},$$

o que nos dá $x = 0,25$, $y = 0,5$ e $z = 0,25$.

Pelo teorema ergódico sabemos que $P_{ij}^n \to v_j$ para todo i, j, e assim podemos concluir que na alta temporada a cada momento, o hotel Areado fica com 25% dos hóspedes, o hotel Bromélias com 50% e o hotel Ceará com os outros 25%.

3

CADEIAS DE MARKOV E O JOGO *MONOPOLY*

3.1 UM POUCO DE HISTÓRIA

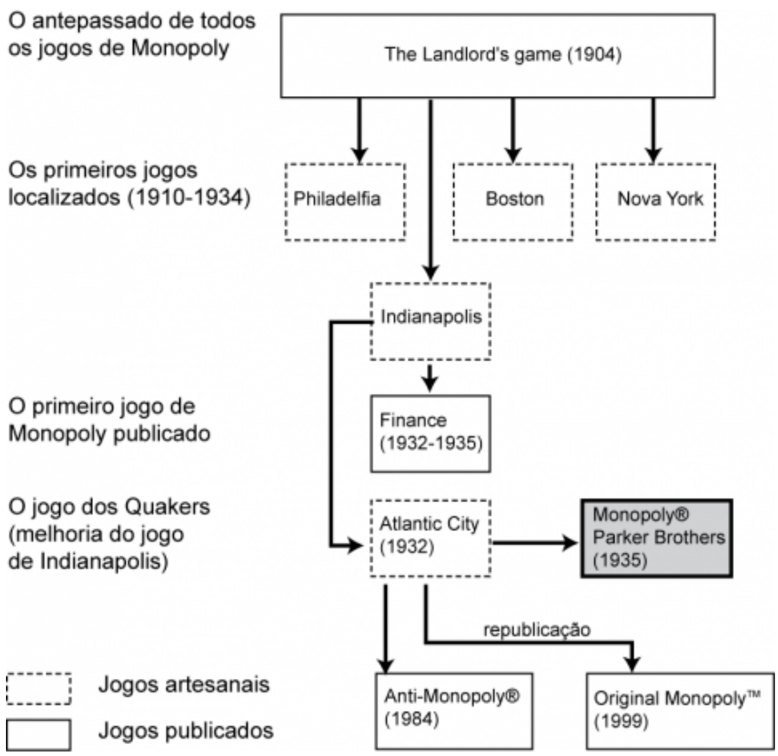
O *Monopoly* é um jogo mundialmente conhecido, do qual existem múltiplas versões nacionais; no Brasil, conhecemos este jogo pelo nome *Banco Imobiliário*. Nestas diversas versões mudam o nome das cidades e das propriedades a ser comercializadas, mas a dinâmica do jogo permanece razoavelmente inalterada desde a primeira versão conhecida. A primeira versão foi feita em uma toalha de mesa e desde o início o passatempo dava ao jogador a possibilidade de comprar e vender terrenos em locais valorizados.

Figura 8: Tabuleiro piloto original de 1934. Neste jogo cada uma das peças foi moldada em madeira à mão, e o tabuleiro foi todo desenhado com caneta e tinta.

Em 1934 Charles B. Darrow, americano da Pensilvânia, mostrou o jogo aos executivos da Parker Brothers. Mas o jogo foi rejeitado por conter “52 erros específicos". Darrow, desempregado naquela época de recessão, decidiu produzir o jogo sozinho.

Com a ajuda de um amigo, Darrow vendeu 5.000 unidades feitas à mão para lojas de departamentos da Filadélfia, Boston e Nova York. As pessoas adoraram o jogo e a demanda cresceu rapidamente, ao ponto que Darrow não podia mais supri-la. Mas o sucesso instantâneo fez com que a Parker Brothers revisse sua posição e lançasse o jogo em 1935.

Mas na verdade, ele não levou à Parker Brothers uma ideia original. Um jogo muito parecido havia sido patenteado por Lizzie J. Magie em 1904, com o nome de *The Landlord's Game*. A diferença principal é que no jogo de Magie, o conceito básico era o de mostrar como os monopólios são injustos e como as cobranças de aluguel por grandes proprietários inescrupulosos podiam ser exorbitantes.



Por seu conteúdo pró-capitalismo, o jogo foi proibido na União Soviética até 1987.

3.2 REGRAS E VERSÕES

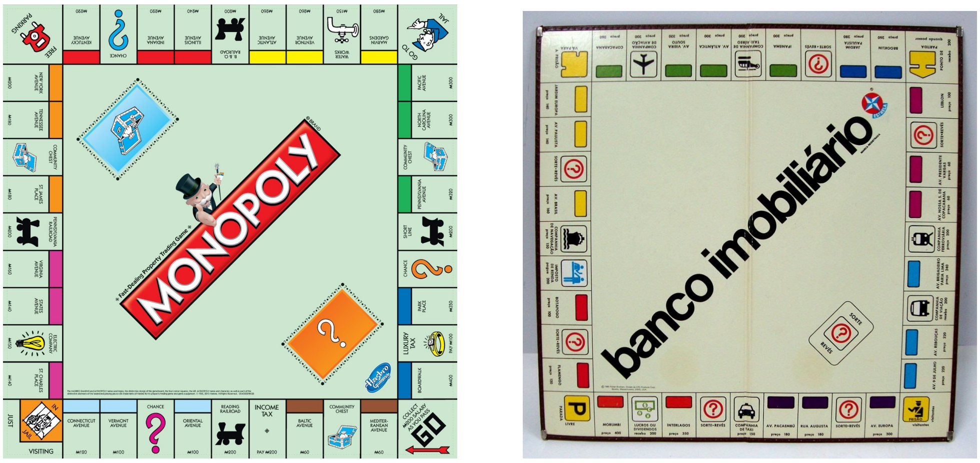
A partir de 1935, numerosos *Monopoly* floresceram na Europa, marcando o início de uma imensa difusão, que continua até nossos dias. Mais de 200 milhões de jogos foram vendidos no mundo. O *Monopoly* permanece um dos jogos mais populares: mais de 80 países têm a sua versão e o jogo existe já em mais de 30 línguas. Existem igualmente versões sobre temas inesperados, em função das modas do momento ou dos acontecimentos da atualidade.

No entanto, a dinâmica do jogo permanece a mesma nas versões comerciais espalhadas pelo mundo. As regras podem ser resumidas da seguinte forma:

- os jogadores (geralmente de 2 a 6) alternam turnos, nos quais rolam dois dados que definirão quantas casas devem andar;
- ao cair em uma casa que representa uma propriedade ou uma empresa, esta poderá ser comprada pelo jogador (caso ainda não tenha dono) ou então o jogador deverá pagar um valor de aluguel ao dono da propriedade, sempre de acordo com o valor atual daquela propriedade;
- ao longo do jogo cada jogador pode desenvolver suas propriedades, construindo casas e hoteis, aumentando assim o valor a ser cobrado de seus adversários;
- algumas casas exigem que o jogador pague algo ao banco, como taxas;
- uma casa no tabuleiro envia o jogador para a prisão, cujas regras para sair serão resumidas mais adiante;
- algumas casas dão ao jogador a chance de sortear cartas aleatórias que podem ter diversos efeitos, como enviar para a prisão, ou mesmo dar uma chance do jogador deixar a prisão quando entrar nela.

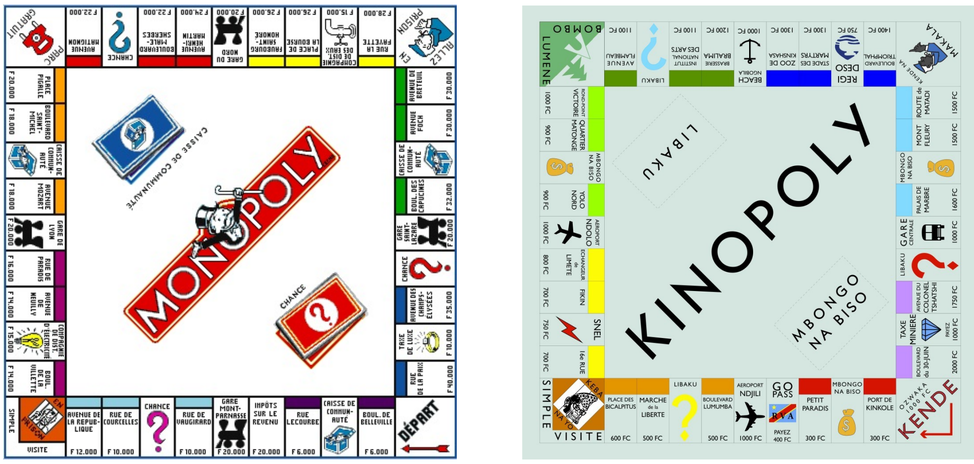
A ideia original do jogo é a mesma desde a sua aparição: um jogo de base aleatória (pois cada jogada consiste no lançamento de dois dados idênticos), com peças se deslocando através de casas no caminho de um tabuleiro cíclico, onde os jogadores ganham dinheiro com suas propriedades e perdem dinheiro dependendo da casa em que estão, e o vencedor sendo aquele jogador que permaneceu com dinheiro enquanto todos os outros foram à falência. De acordo com as regras oficiais, “a idéia do jogo é vender, comprar ou alugar propriedades de maneira vantajosa, de tal maneira que um dos jogadores fique mais rico e chegue ao monopólio’.

As inúmeras versões encontradas por todo o planeta diferem basicamente no tabuleiro; o nome das propriedades ofertadas em cada versão e obviamente a língua em que estão escritas as nominações são distintas. Em todas as versões há notas de dinheiro-fantasia (em exemplos mais modernos, há cartão de crédito), ícones representando casas/hotéis (ou similares), dados e peões para movimentação dos jogadores.



(a) Versão Norte Americana (b) Versão alternativa brasileira

Figura 9: Versões de *Monopoly* ao redor do mundo



(a) Versão Francesa (b) Versão do Congo (cidade de Kinshasa)

Figura 10: Versões de *Monopoly* ao redor do mundo

Percebe-se no tabuleiro a mesma configuração nestas e nas outras tantas versões ao redor do mundo: são 40 casas no total, sendo

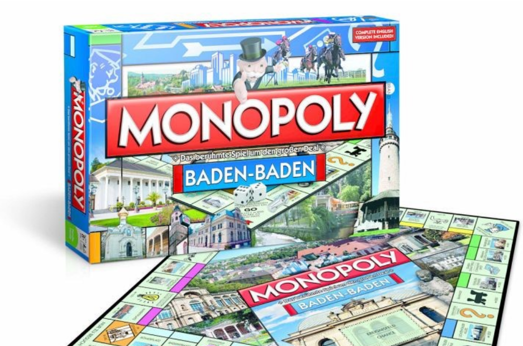
- Uma casa para o início;
- Uma casa de “Parada Livre”;

- Uma casa de “Prisão”;
- Uma casa “Vá para a Prisão”;
- Três ou seis casas do tipo “sorte ou revés”, que na versão original, em inglês, se chamam “chance” ou “comunity chest”.

(a) Versão Japonesa

(b) Versão Iraquiana

Figura 11: Versões de *Monopoly* ao redor do mundo



(a) Versão Alemã (cidade de Baden-Baden)

(b) Versão Israelense

Figura 12: Versões de *Monopoly* ao redor do mundo

A PRISÃO Esta é uma casa bastante importante para o jogo, e suas regras são alvo de diversas alterações em versões alternativas.

No jogo oficial um jogador pode ser enviado para a prisão de duas formas: sorteando uma carta que o envie para lá, ou terminando seu movimento na casa "Vá para a Prisão". Quando em uma jogada regular o jogador termina seu movimento na casa "Prisão" ela funcionará exatamente como a casa de "Parada Livre".

Para sair da prisão existem algumas possibilidades:

- Pagar uma quantia fixa ($50 na versão original);
- Usar uma carta "sair da prisão";
- Sortear uma dupla de números iguais nos dados. Situação na qual o jogador deverá andar exatamente a quantidade sorteada;
- Na terceira rodada após a prisão o jogador sai sem pagar nada;

Por ser um jogo essencialmente *probabilístico*, vem o natural questionamento sobre a existência ou não de uma *estratégia vencedora*, bem como a dúvida sobre a *finitude* do tempo de disputa de uma partida. Para responder parcialmente à questão sobre a melhor estratégia utilizaremos as ferramentas desenvolvidas ao longo deste trabalho.

3.3 A MATEMÁTICA DO JOGO

3.3.1 *Uma visão geral*

Queremos agora tentar modelar matematicamente o jogo de *Monopoly*, e para isso vamos usar cadeias de Markov. A escolha de cadeias de Markov é natural, uma vez que a cada jogada um jogador salta de uma casa para outra aleatoriamente, de acordo com o resultado da rolagem de 2 dados.

Para simplificar as análises que faremos, vamos antes montar uma versão reduzida do jogo, com regras similares, mas em tabuleiros menores.

Considerar uma versão do *Monopoly* como uma Cadeia de Markov implica construir uma matriz de transição P que nos dê a probabilidade de, estando na casa i, ir para a casa j imediatamente em um passo, na próxima jogada, independentemente das jogadas anteriores. Nesse ponto, importa saber quais as características do tabuleiro que vamos construir.

Como vimos nas várias versões do jogo ao redor do mundo, o tabuleiro de Monopoly é sempre de formato quadrado, e o número total de casas do tabuleiro (C) dado por

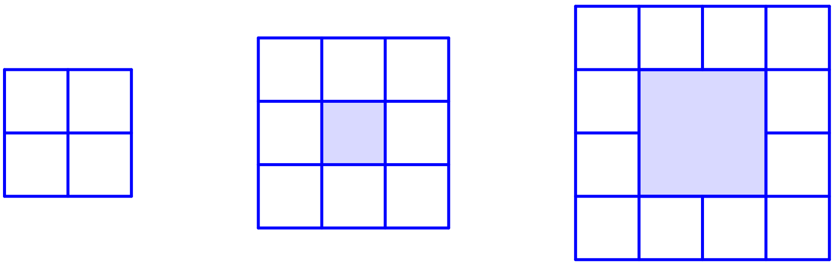


Figura 13: Tabuleiros alternativos com $k = 1, 2$ e 3 respectivamente.

$C = 4.k,\ k > 0$. Além disso, os quatro cantos do tabuleiro estão sempre definidos na mesma ordem: um início na casa 1, uma casa de prisão (que funciona tanto como visita livre para quem passa, como para saída para quem está preso) na casa $k+1$, uma casa de parada livre na casa $2k+1$, e "vá para a prisão" na casa $3k+1$. As demais casas são as "propriedades' que devem ser compradas e manipuladas ao longo da partida, juntamente com casas de sorte ou revés.

Nessa perspectiva, ponderamos que:

- Para obter uma dinâmica similar ao jogo real, nosso modelo deverá comportar o jogo com dois dados, de modo que em uma única jogada não será possível dar a volta completa no tabuleiro, garantindo a inexistência de *loops* espúrios que atrapalhariam a análise do jogo. Descartamos então os valores de $k = 1, 2$ ou 3;

- A prisão é uma casa complicada para a modelagem. Existem diversas regras diferentes de como sair da prisão, dificultando muito a modelagem. Em [17], o autor considerou que a saída da prisão se daria em uma jogada, mediante pagamento de uma multa. Neste caso a ida para a prisão é apenas um transporte da casa $3k+1$ para a casa $k+1$, sem nenhum ônus adicional. Tornaremos nosso modelo mais realístico, usando a regra que para a saída da prisão o jogador deve tirar uma dupla de números iguais na rolagem dos dados, senão permanece na prisão;

- Para $k = 4, 5$ e 6 surgem problemas com a regra de saída da prisão: o jogador está na prisão, tira uma dupla e portanto sai da prisão...e dependendo da rolagem retorna à prisão imediatamente! Isso configura um "loop" que altera demasiadamente a dinâmica do modelo e a similaridade com o jogo *Monopoly*. O menor valor para o qual consegue-se resolver esse problema é para $k = 7$. De todo modo, apresentaremos uma versão com $k = 4$ para esclarecer melhor estes pontos;

- O objetivo deste trabalho é analisar apenas a dinâmica probabilística do jogo. Questões como casas de sorte-revés, preço das propriedades, colocação de casas/hotéis, salário recebido a cada rodada etc. não serão consideradas aqui.

3.3.2 *Modelando o jogo*

O ponto central desse trabalho é analisar o jogo *Monopoly* (e seus similares mais simples) como sendo uma cadeia de Markov. Nos deteremos brevemente para entender como isso pode ser feito.

a) Identificamos as casas do tabuleiro como estados de uma cadeia: Existem $4.k$ casas (estados) no tabuleiro que compõem o jogo, $k > 0$. A cada jogada, uma e apenas uma casa é alcançada. A interpretação fundamental que permite classificar o processo como cadeia de Markov é que a casa que o jogador estará na jogada $n+1$ depende exclusivamente da casa em que o jogador está no instante n. De fato, a quantidades de casas a andar depende apenas do lançamento de dois dados, de modo que a chance de ir para uma certa casa é completamente independente das casas anteriormente visitadas, dependendo apenas da casa onde o jogador está.

b) Montamos a matriz de transição: Para isso, como explicamos anteriormente, numeramos as casas de 1 a $4.k$, fazendo 1 a casa de início e seguindo direção de movimento das peças. Para determinar a probabilidade P_{ij} lembre-se que, de acordo com a regra, estando em na casa i o jogador moverá sua peça para a casa $i+s$ sempre que a soma dos resultados da rolagem dos dados for igual à s. Assim temos que $P_{13} = 1/36$ e $P_{18} = 1/6$, ao mesmo tempo que $P_{12} = 0$, assim como $P_{1,1+j} = 0$ para todo $j \geq 13$.

   Como o tabuleiro é cíclico, devemos reiniciar a contagem quando passarmos da casa $4k$. Assim, a casa $4k+5$ será considerada como a casa 5. Para exemplificar, considere $4k = 40$ e note que se o jogador está na casa 36 e rola 8 na soma dos dados, ele cai então na casa 4, de modo que $P_{36,4} = 1/6$.

   Para a casa $k+1$, onde fica a prisão, funciona exatamente como as demais. Isso por que, de acordo com as regras, se o jogador apenas passar por esta casa com uma rolagem de dados, ela funciona exatamente como as demais. Assim, as probabilidades de transição a partir de $k+1$ serão feitas normalmente, como nas demais casas.

O jogador só vai "preso" de fato quando cai na casa $3k+1$, "vá para a prisão". Uma vez nesta casa, a regra diz que o jogador vai para a casa $k+1$, onde ele ficará preso até algo acontecer (varia de regra para regra). O fato do jogador ficar "preso" na casa $k+1$ não tem nenhum efeito prático, a não ser o fato de que a contagem de para onde o jogador vai ao ser liberado é feita a partir desta casa. Podemos portanto considerar que o jogador fica parado nesta casa até conseguir sair, e quando isso acontecer contamos as transições a partir da casa $k+1$.

As transições da casa "vá para a prisão" dependem portanto de que regra usaremos para sair, e explicaremos cada caso separadamente, quando ocorrer.

Montada a matriz de transição, queremos ter certeza que podemos seguir nossa análise sem problemas. Para considerar algo sobre a estratégia precisamos descobrir a proporção de tempo que um jogador passa em cada casa ao longo do jogo, o que para nós significa determinar se a cadeia que acabamos de montar é ou não ergódica. Precisamos então analisar duas características principais: irredutibilidade e aperiodicidade.

O JOGO É IRREDUTÍVEL Por simplicidade, desconsideremos momentaneamente as casas "vá para a prisão" e "prisão", localizadas em $3k+1$ e $k+1$ respectivamente, tratando-as como casas de passagem livre.

O jogo tem início sempre na casa 1, e saindo dela podemos ir para qualquer casa entre 3 e 13 (lembrando que passando da casa $4k$ devemos reiniciar a contagem). Como $P_{i,i+2} > 0$ para todo i, então a partir de 1 podemos acessar qualquer outra casa ímpar. De modo que $1 \to m$ para todo m ímpar. Em particular, temos que após alguns passos podemos retornar à 1, fechando o ciclo, e mostrando que $1 \leftrightarrow m$ para todo m ímpar.

Considere agora que o jogador se encontra na casa 4. A mesma análise que fizemos acima mostra que $4 \leftrightarrow m$ para toda casa m par. Já sabemos que $1 \to 4$, pois o jogador pode tirar 3 na soma dos dados. Da mesma forma, temos que $4k - 2 \to 1$, e como $4k_2$ é par sabemos que $4 \to 4k - 2$. Chapman-Kolmogorov nos garante agora que $4 \to 1$ ligando pares e impares, e garantindo a irredutibilidade.

A análise acima só poderia ser atrapalhada pela casa "vá para a prisão", cuja transição é feita de forma distinta.

Para incorporar esta casa ao jogo, considere primeiro que k é par, como no tabuleiro oficial. Neste caso, a casa "vá para a prisão" é impar, pois está na posição $3k+1$. Desta forma, considerando apenas movimentos de dois passos, vemos facilmente que ela é acessível a partir de qualquer casa i impar, e que podemos nos mover entre casas pares sem passar pela casa $3k+1$.

É importante notar também que também é possível o jogador se mover entre duas casas impares sem passar por $3k+1$. De fato, fazendo movimentos de dois passos o jogador chegará primeiro chegar à casa $3k-1$. Neste momento, fazendo um movimento de 4 casas, ele vai diretamente para $3k+3$ sem passar por $3k+1$. Ilustramos isso na figura 14.

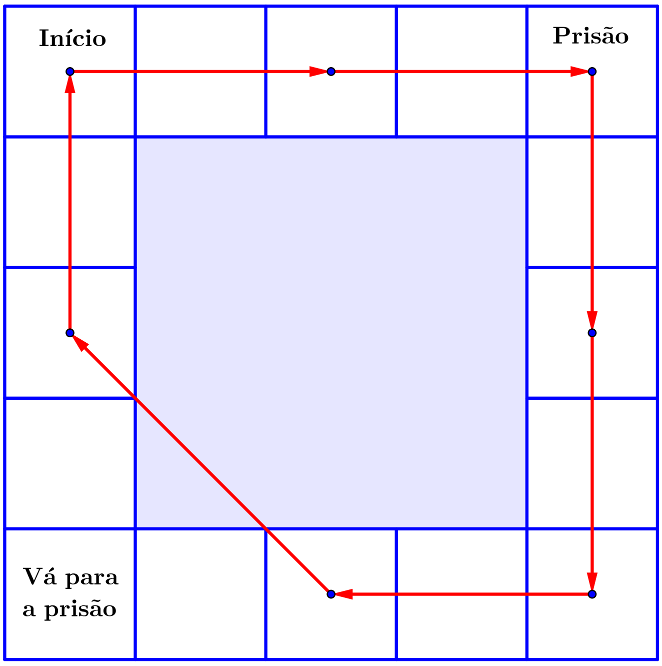


Figura 14: Um ciclo fechado, passando por todas as casas ímpares, evitando a casa "Vá Para a Prisão", no tabuleiro $k = 4$.

Isso tudo mostra que é qualquer par de casas se comunicam ($i \leftrightarrow j$) por caminhos que não incluam $3k+1$.

Com isso, para mostrar que a cadeia é irredutível, nos resta apenas mostrar que $3k+1 \leftrightarrow i$ para algum $i \neq 3k+1$.

Já vimos que se i é impar, então $i \rightarrow 3k+1$. Agora note que, independente da regra usada, ao sair da prisão o jogador seguirá para alguma casa $j \neq 3k+1$. Sabemos assim que existe j tal que $3k+1 \to j$, e como certamente $j \to i$ sem voltar a $3k+1$, temos que $3k+1 \to 1$, concluindo que $i \leftrightarrow 3k+1$.

Concluímos assim que a cadeia é irredutível, e pelo Corolário 2.9 é também recorrente.

O JOGO É APERIÓDICO A aperiodicidade, neste caso, é um problema bem mais simples de ser abordado.

Já sabemos que a cadeia é irredutível, e como periodicidade é propriedade de classe, só precisamos encontrar um sítio aperiódico e concluiremos a aperiodicidade de toda a cadeia.

Supondo que k é impar, tome então o sítio 1 (se k for par, basta tomar o sítio 2 e fazer a mesma análise). Já sabemos que fazendo apenas movimentos de comprimento 2, é possível voltar para a casa 1 sem passar pela casa $3k+1$.

O total de movimentos de tamanho 2 necessários para dar a volta em um tabuleiro de $4k$ casas é exatamente $2k$. Se apenas no primeiro movimento o jogador andar 4 casas, e depois seguir com movimentos de tamanho 2, ele demorará $2k-1$ movimentos para retornar à 1.

Mas $mdc(2k-1, 2k) = 1$ para todo k, e portanto o período $d(1) = 1$, e 1 é aperiódico.

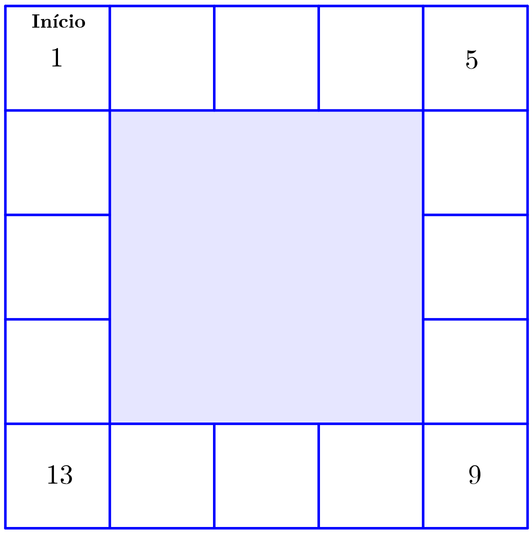
Com isso concluímos que a cadeia é recorrente, irredutível e aperiódica, e portanto ergódica. Podemos assim seguir normalmente a análise de proporção de visitas que queremos fazer.

Na sequência, começaremos a analisar as características de tabuleiros menores, similares ao jogo original, e que obedecem à lei de formação de tabuleiro $C = 4k, k \geq 1$, como já explicado antes.

O trabalho computacional que será visto nas próximas seções foi realizado com o auxílio do Software *Excel 2013* operacional em uma máquina com processador *IntelCore-i5* a 2,6GHz.

3.3.3 *Considerações para o caso $k = 4$*

Tomemos por base o seguinte tabuleiro.



Considere agora as seguintes regras alternativas

Valor de k	Total de Casas	Casas Especiais	Regra para Saída de Prisão
4	16	Casa 1 - Início Casa 2 - parada livre Casa 3 - parada livre Casa 4 - parada livre	Não existe prisão

Esse é o modelo mais simples que podemos exibir, no qual o jogo com dois dados é possível sem existência de *loops* indesejados. A matriz de transição nesse caso será:

$$P = \begin{bmatrix}
0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} \\
\frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} \\
\frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} \\
\frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} \\
\frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} \\
\frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} \\
\frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} \\
\frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} \\
\frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} \\
\frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0
\end{bmatrix}$$

Aqui obtemos uma convergência razoavelmente rápida: em 16 passos já conseguimos uma boa aproximação para o equilíbrio, que neste caso é dado por

$$v = \begin{bmatrix} \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} & \frac{1}{16} \end{bmatrix}.$$

Esta uniformidade pode, a princípio, parecer estranha. Mas sem a prisão, o jogo ganha uma grande simetria. De fato, sem a prisão, podemos começar o jogo em qualquer casa, e ele se comportará exatamente da mesma forma, apenas “rotacionado”! Esta simetria se exprime pela uniformidade da distribuição no equilíbrio.

Mas este modelo está longe da realidade, então propomos uma alteração para deixa-lo mais próximo do que queremos estudar. Tome então o seguinte conjunto de regras especiais.

Valor de k	Total de Casas	Casas Especiais	Regra para Saída de Prisão
4	16	Casa 1 - Início Casa 2 - parada livre Casa 3 - parada livre Casa 4 - parada livre	Uma jogada após chegar na casa 13, o jogador sai partir da casa 5.

Neste caso é introduzido uma casa de desequilíbrio da distribuição de probabilidades: a casa 13. A perturbação causada neste modelo será evidenciada, mesmo que a prisão só se mantenha por uma jogada. Cair na prisão, neste modelo, é simplesmente ser transportado de volta para a casa 5 na jogada seguinte. A matriz de transição será

$$P = \begin{bmatrix}
0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} \\
\frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} \\
\frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} \\
\frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} \\
\frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} \\
\frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} \\
\frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} \\
\frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0
\end{bmatrix},$$

com distribuição de equilíbrio dada, aproximadamente, pelos valores abaixo.

Início	Casa 2	Casa 3	Casa 4	**Prisão**	Casa 6	Casa 7	Casa 8
$0,057$	$0,054$	$0,053$	$0,051$	$0,116$	$0,053$	$0,055$	$0,057$

Casa 9	Casa 10	Casa 11	Casa 12	**V.P.P.**	Casa 14	Casa 15	Casa 16
$0,059$	$0,062$	$0,063$	$0,065$	$0,064$	$0,063$	$0,061$	$0,059$

É interessante notar que a prisão é a casa mais visitada do tabuleiro! Isso faz com que as casas subsequentes, especialmente as que estão à distância 5, 6, 7, 8 e 9 da prisão, sejam mais visitadas que as demais. Ao mesmo tempo, as casas mais distantes da prisão são menos frequentes. Principalmente aquelas que não são atingidas em um só passo, como as casas 2, 3, 4 e 6.

Isso acontece por conta da não uniformidade dos resultados no lançamento de dois dados, como vemos no histograma da figura 15.

Esta é também a principal alteração na distribuição de equilíbrio causada pela presença da prisão, e estará presente em todas as análises que faremos a seguir.

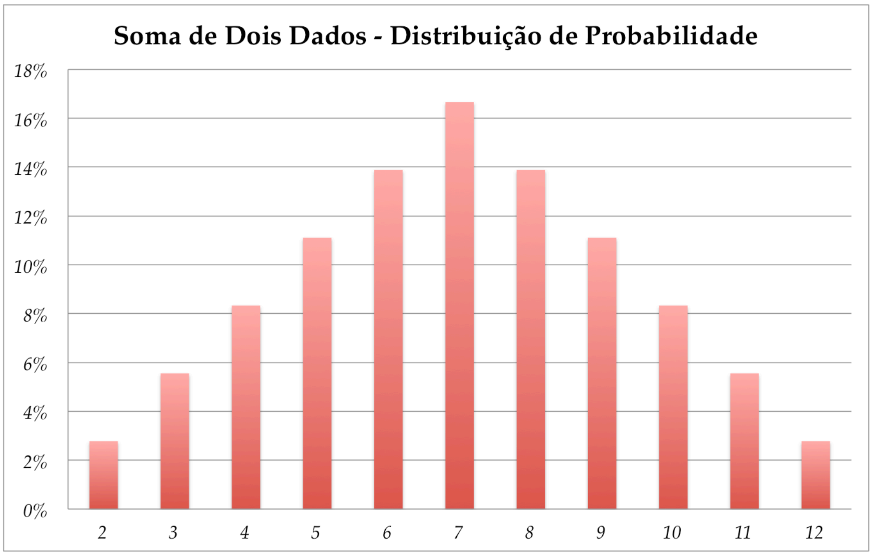


Figura 15: Histograma mostrando a distribuição de probabilidade da soma de dois dados.

Analisemos agora um próximo modelo, um pouco mais próximo do modelo real, com as seguintes regras.

Valor de k	Total de Casas	Casas Especiais	Regra para Saída de Prisão
4	16	Casa 1 - Início Casa 2 - parada livre Casa 3 - parada livre Casa 4 - parada livre	Sai da prisão ao conseguir rolar uma dupla de 1, 3 ou 5.

Aqui consideramos uma regra alternativa onde, para sair da cadeia o jogador deve rolar uma dupla de 1, de 3 ou de 5 nos dados, andando o total indicado pelo dado. Isso inclui no modelo uma alta probabilidade (11/12) do jogador permanecer na casa "*vá para a prisão*", antes de sair pela casa da prisão, movendo de acordo com o resultado dos dados.

Observação 3.3.1. *Alteramos a regra original neste caso, considerando apenas duplas de números ímpares, pois uma dupla de 4's levaria o jogador de volta para a casa 13.*

A matriz de transição será então

$$P = \begin{bmatrix}
0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} \\
\frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} \\
\frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} \\
\frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} \\
\frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} \\
\frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} \\
\frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & 0 & 0 & 0 & \frac{1}{36} & 0 & \frac{11}{12} & 0 & \frac{1}{36} & 0 \\
\frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} \\
\frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0
\end{bmatrix},$$

com distribuição de equilíbrio dada, aproximadamente, pelos valores abaixo.

Início	Casa 2	Casa 3	Casa 4	**Prisão**	Casa 6	Casa 7	Casa 8
$0,036$	$0,035$	$0,034$	$0,033$	$0,034$	$0,035$	$0,047$	$0,035$

Casa 9	Casa 10	Casa 11	Casa 12	**V.P.P.**	Casa 14	Casa 15	Casa 16
$0,035$	$0,036$	$0,048$	$0,036$	$0,437$	$0,038$	$0,049$	$0,036$

Vemos agora que a maior probabilidade pulou agora para a casa 13. Isso por que, do modo como modelamos, após chegar à casa "*vá para a prisão*", o jogador não chega a ir para a casa prisão. Esta serve apenas como referência para determinarmos que casa o jogador vai quando sair.

Assim, as casas mais prováveis depois da casa 13 são aquelas que estão à distância 2 (jogador rola dois 1's), distância 6 (dois 3's) ou distância 10 (dois 5's), sempre contando da casa 5 (prisão).

Mas perceba que neste modelo, cada jogador passa próximo a 43% do jogo na prisão, o que está bastante longe da realidade. Temos então que alterar tal modelo de alguma forma.

Pelas regras originais o jogador sai automaticamente da prisão após três jogadas, ou pode sair antes mediante o pagamento de alguma quantia.

Para aproximar estas regras, vamos considerar que cada jogador sempre tenta rolar o dado na primeira jogada preso, e se não conseguir sair ele paga para sair na segunda rodada. Ou seja, na primeira rodada após a chegada na casa "*vá para a prisão*" o jogador terá uma probabilidade $11/12$ de ir a casa 5, de onde sairá na próxima jogada. Com probabilidade $1/12$ ele se moverá de acordo com o resultado do dado, como na regra anterior.

Valor de k	Total de Casas	Casas Especiais	Regra para Saída de Prisão
4	16	Casa 1 - Início Casa 2 - parada livre Casa 3 - parada livre Casa 4 - parada livre	rolar uma dupla de 1, 3 ou 5 ou sair automaticamente na jogada seguinte

Isso nos leva à seguinte matriz de transição.

$$P = \begin{bmatrix}
0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 \\
0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} \\
\frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} \\
\frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} \\
\frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} \\
\frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} \\
\frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} \\
\frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} \\
\frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} \\
0 & 0 & 0 & 0 & \frac{11}{12} & 0 & \frac{1}{36} & 0 & 0 & 0 & \frac{1}{36} & 0 & 0 & 0 & \frac{1}{36} & 0 \\
\frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 & \frac{1}{36} \\
\frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{1}{36} & \frac{1}{18} & \frac{1}{12} & \frac{1}{9} & \frac{5}{36} & \frac{1}{6} & \frac{5}{36} & \frac{1}{9} & \frac{1}{12} & \frac{1}{18} & \frac{1}{36} & 0 & 0 & 0 & 0
\end{bmatrix},$$

com distribuição de equilíbrio dada, aproximadamente, pelos valores abaixo.

Início	Casa 2	Casa 3	Casa 4	**Prisão**	Casa 6	Casa 7	Casa 8
0,057	0,055	0,054	0,052	0,112	0,054	0,058	0,058

Casa 9	Casa 10	Casa 11	Casa 12	**V.P.P.**	Casa 14	Casa 15	Casa 16
0,060	0,062	0,065	0,065	0,064	0,063	0,063	0,059

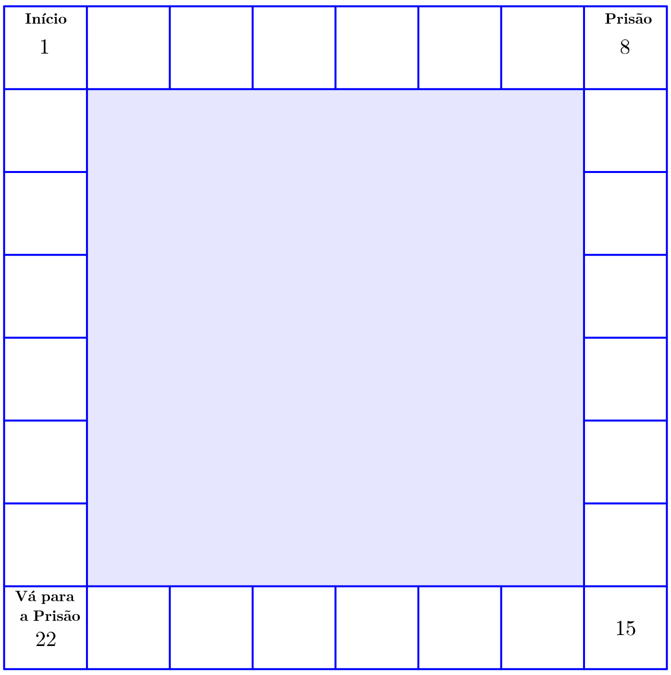
Observe que a prisão é novamente a mais visitada, mas agora apenas 11% do tempo de jogo. Novamente, as casas mais visitadas são as casas de 10 à 14, com uma pequena influência (quase imperceptível) da probabilidade de sair com a rolagem de dois dados iguais.

Este é o modelo mais próximo do original que encontramos, e portanto vamos ficar com ele. A única coisa ainda problemática é o fato de que ao sair da casa 5, podemos voltar à casa 13 (vá para a prisão) imediatamente. Mas isso ocorre devido ao tamanho do tabuleiro.

Antes de passar para a análise do jogo original, façamos um tabuleiro com $k = 7$.

3.3.4 *Considerações para o caso* $k = 7$

Seguindo a análise para $k = 7$, o tabuleiro será agora da seguinte forma.



Aqui a "prisão" está na casa 8, enquanto a casa "vá para a prisão" está na posição 22.

Como já apontado, este é o primeiro valor de k para o qual o tamanho do tabuleiro não traz problemas à anállise do jogo, como proposto em *Monopoly*. Fazemos, portanto, diretamente a análise do modelo mais fiel ao desenrolar do jogo clássico. Ou seja, na jogada seguinte à chegada na casa 22, o jogador deve rolar um dado. Ao tirar qualquer dupla de resultados iguais ele sai para a casa indicada, contando a partir da casa 8. Caso contrário ele segue para a casa 8, de onde sairá normalmente na jogada seguinte.

Valor de k	Total de Casas	Casas Especiais	Regra para Saída de Prisão
7	28	Casa 1 - Início Casa 8 - prisão Casa 15 - parada livre Casa 22 - vá para a prisão	rolar dupla de valores iguais ou sair automaticamente na jogada seguinte

A matriz de transição neste caso tem dimensão 28×28 e portanto sua representação aqui é muito complicada. De todo modo, as linhas se comportam muito similarmente às do caso $k = 4$, mudando apenas a linha relativa à casa 22 (vá para a prisão), que apresentamos abaixo.

$$\begin{array}{cc} & \begin{array}{ccccccccccccccccccc} 1 & 2 & 3 & \cdots & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & \cdots & 28 \end{array} \\ 22 & \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & \frac{5}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \cdots & 0 \end{bmatrix} \end{array}$$

A distribuição de equilíbrio da cadeia é dada, aproximadamente, pelos valores da tabela abaixo.

3,04%	3,11%	3,15%	3,18%	3,20%	3,23%	3,27%	6,29%
3,18%							3,19%
3,31%							3,36%
3,43%							3,35%
3,54%							3,55%
3,64%							3,55%
3,73%							3,77%
3,70%	3,65%	3,77%	3,68%	3,80%	3,71%	3,84%	3,78%

Tabela 1: Distribuição de equilíbrio para tabuleiro com $k = 7$. Os valores estão dispostos de acordo com a posição correspondente no tabuleiro, a casa mais frequente (6,29%) corresponde à prisão.

Notamos algumas coisas interessantes. A primeira é que à medida que aumenta o tabuleiro, as porcentagens vão diminuindo, como que se "diluindo" no conjunto total de estados, e vemos que as diferenças são pequenas para um tempo de jogo muito curto (por exemplo, 0,8% é a diferença entre a casa mais visitada e a casa menos visitada, excetuando obviamente a prisão, que continua a ser a casa mais visitada).

A segunda coisa é que padrão semelhante ao do caso $k = 4$ continua acontecendo. Ou seja, as casas seguintes à do prisão são as mais visitadas (excetuando-se a prisão). Aqui, no entanto, não são apenas as probabilidades das somas dos dados que determinam as casas mais prováveis. Parece existir uma contribuição da probabilidade de saída rolando duplas. Observe que as casas 6, 8, 10 e 12 são as mais prováveis após a prisão.

3.4 O *monopoly* E SEU TABULEIRO CLÁSSICO

Para os tabuleiros anteriores, versões simplificadas do jogo, observamos que:

- Se não existe a casa "Prisão", o jogo tem uma distribuição simétrica;
- Existindo a casa "Prisão" nota-se que é introduzido um desequilíbrio no jogo, de forma que passam a existir casas que serão mais visitadas ao longo do tempo e casas que serão menos visitadas. A casa "Prisão" aparece como a mais visitada nos dois tabuleiros analisados;
- As casas com maior probabilidade de visita são as que seguem a casa da "Prisão", em especial as casas que estão à distância 6, 8, 10 e 12 da prisão.

A análise no tabuleiro oficial se dará da mesma forma. Utilizaremos as mesmas regras de saída da prisão do caso $k = 7$ para aproximar as regras originais.

Valor de k	Total de Casas	Casas Especiais	Regra para Saída de Prisão
10	40	Casa 1 - Início Casa 8 - prisão Casa 21 - parada livre Casa 31 - vá para a prisão	rolar dupla de valores iguais ou sair automaticamente na jogada seguinte

Aqui a matriz de transição tem dimensão 40×40, e assim como no tabuleiro anterior, apresentaremos a única linha que se comporta de forma distinta, que neste caso é a linha 31.

$$\begin{array}{cc} & \begin{matrix} 1 & 2 & 3 & \cdots & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 & 19 & 20 & 21 & 22 & \cdots & 40 \end{matrix} \\ 31 & \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & \frac{5}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \frac{1}{6} & 0 & \cdots & 0 \end{bmatrix} \end{array}$$

A distribuição de equilíbrio está representada na tabela abaixo, também representando a disposição de casas da tabuleiro original.

2,24	2,26	2,28	2,30	2,27	2,25	2,24	2,24	2,25	2,26	4,45
2,22										2,26
2,19										2,40
2,15										2,38
2,25										2,52
2,33										2,51
2,48										2,66
2,41										2,67
2,56										2,72
2,63										2,63
2,63	2,64	2,64	2,65	2,64	2,63	2,61	2,58	2,67	2,61	2,69

Tabela 2: Distribuição de equilíbrio para tabuleiro oficial. Os valores estão dispostos de acordo com a posição correspondente no tabuleiro e em percentagem, a casa mais frequente (4,45%) corresponde à prisão. A casa superior à esquerda marca o início.

Apresentamos também um histograma comparando as diversas probabilidades na figura.

É interessante observar que os padrões apresentados no tabuleiro $k = 7$ se repetem, mas agora os valores são mais próximos uns dos outros. A casa mais visitada após a prisão é a 19 (8 casas após a prisão), com 2,72% de probabilidade, enquanto a menos visitada é a 38, com 2,15%. A diferença entre as duas é de apenas 0,57%.

As casas que se encontram à distância 6, 8, 10 e 12 depois da prisão, continuam dentre as mais visitadas, assim como no caso anterior, enquanto as casas menos visitadas ainda estão concentradas próximas do início.

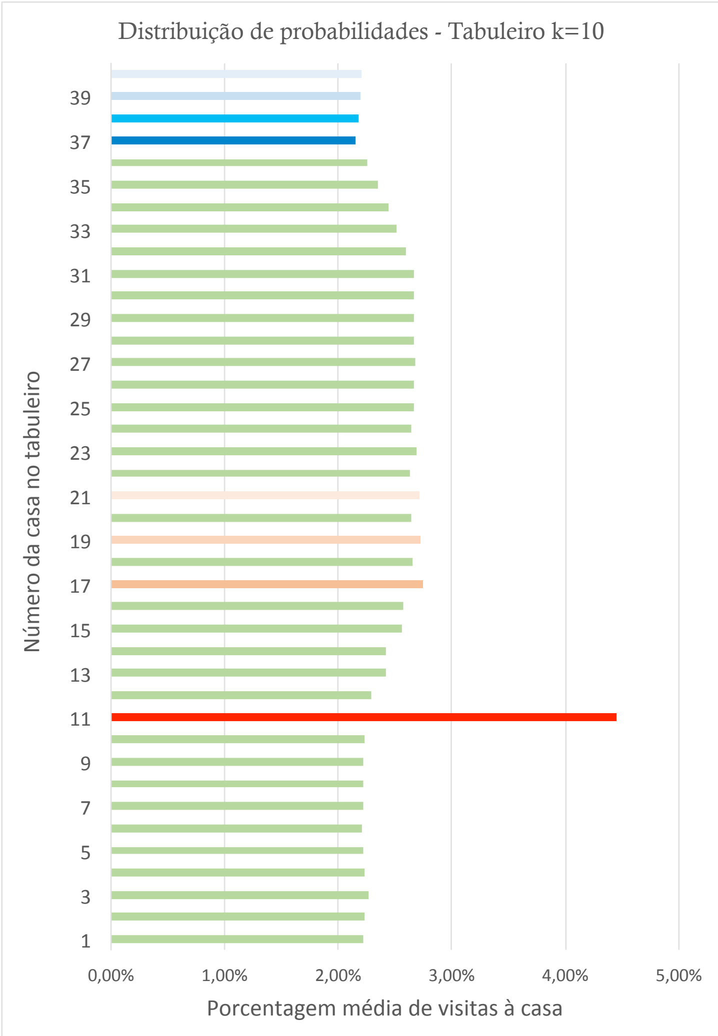


Figura 16: Histograma os as probabilidades de equilíbrio no tabuleiro oficial.

3.4.1 *Breve Análise Estratégica*

Os dados calculados acima dizem muito pouco sobre uma possível estratégia de jogo. Lembre-se que a maneira de ganhar dinheiro (e o jogo) é comprar terrenos e cobrar aluguel daqueles que parem no mesmo espaço. A primeira vista, a distribuição de equilíbrio parece indicar que a melhor estratégia é concentrar negócios entre as casas 17 e 30, especialmente nas casas 17, 18, 19, 21 e 23. Assim o jogador recolheria aluguel com mais frequência, aumentando a chance de ganhar o jogo.

No entanto existem diversas outros fatores que devem ser levados em consideração, e a pequena diferença entre as probabilidades de cada casa podem arruinar esta primeira análise.

Dentre os fatores mais importante, e fácil de ser incorporado em nossa análise, está a preço da propriedade. O valor em dinheiro marcado em cada casa define não apenas o custo de compra daquele terreno, mas também o preço do aluguel que será cobrado de quem parar ali.

Assim, denotando por v_1 a probabilidade de equilíbrio da casa i, e por M_i o valor da propriedade, os valores de $v_i \cdot M_i$ para as diferentes casas são mais importantes para a definição de estratégia que apenas os valores de v_i.

Considere então os valores de propriedades abaixo, como encontrados no tabuleiro oficial de *Monopoly*.

0	60	0	60	200	200	100	0	100	120	0
400										140
0										150
350										140
0										160
200										200
320										180
0										0
300										180
300										200
0	280	150	260	260	200	240	220	0	220	0

Tabela 3: Valores das propriedades no tabuleiro oficial. Deixamos marcadas as casas mais frequentes.

De cara notamos que 3 das casas mais frequentes não tem valor, enquanto as demais tem valor baixo se comparados com os das casas seguintes. Precisamos saber agora se a maior probabilidade destas casas compensa o seu baixo valor, e isso calculamos na próxima tabela.

Vemos agora que as casas que melhor pagam se concentrar no final do tabuleiro, logo antes da casa de início, justamente onde estavam as casas de menor frequência de visita.

0	1,36	0	1,38	4,53	4,50	2,24	0	2,25	2,71	0
8,88										3,17
0,00										3,59
7,53										3,33
0										4,03
4,66										5,02
7,95										4,79
0										0
7,67										4,89
7,88										5,25
0	7,38	3,96	6,88	6,87	5,27	6,27	5,67	0	5,73	0

Tabela 4: Valores das propriedades no tabuleiro oficial pesados pela frequência de visita.

Isso ainda não garante que a melhor estratégia seria comprar as casas finais. O jogo possui outros elementos importantes, como o desenvolvimento de propriedades que aumenta o valor do aluguel, e pode compensar esta diferença. Além disso as propriedades no tabuleiro são divididas em *tipos*, e possuir várias propriedades de um mesmo tipo pode aumentar o valor cobrado dos demais jogadores.

Ainda que não consigamos estabelecer uma estratégia clara para vencer o jogo, conseguimos fazer uma análise inicial importante, que nos mune de informações preciosas na hora de decidir como queremos jogar.

BIBLIOGRAFIA

[1] *Monopoly Wiki*. Fandom, 2016. `http://monopoly.wikia.com/wiki/Main_Page`.

[2] Almeida Nogueira, F. M. de: *Modelagem e Simulação: Cadeias de Markov*. Notas de aula da Disciplina EPD-042-Pesquisa Operacional II, UFJF, 2015.

[3] Bennett, D. J.: *Aleatoriedade*. Martins Fontes, São Paulo, 2003.

[4] Calabria, A. R. e M. F. Cavalari: *Coleção História da Matemática para Professores*. Em *Um passeio histórico pelo início da Teoria das Probabilidades*. Sociedade Brasileira de História da Matemática, Campinas, 2013.

[5] Dimuro, G. P., R. H. S. Reiser, A. C. R. Costa e P. L. R. Sousa: *Modelos de Markov e Aplicações (tutorial)*. Em Educat (ed.): *VI Oficina de Inteligência Artificial, Universidade Católica de Pelotas*, pp. 37–59, Pelotas, 2002.

[6] Eves, H.: *Introdução à História da Matemática*. Editora da Unicamp, 3ª ed., 2002.

[7] Gadelha, A.: *Uma Pequena História da Probabiidade*. Notas de aula da disciplina Teoria de Probabilidade I, Curso de Pós-Graduação em Estatística, UFRJ, 2004.

[8] Gray, J. R.: *Probability*. Oliver & Boyd, 1ª ed., 1967.

[9] Hinojosa, A. e A. Milanés: *Uma Introdução aos Processos Estocásticos com Aplicações*. UFMG-Dep. de Estatística, 2011.

[10] Hoel, P. G., S. C. Port e C. J. Stone: *Introdução à Teoria da Probabilidade*. Editora Interciência, 4ª ed., 1978.

[11] Júnior, D. P. F. e V. V. Júnior: *Conceitos e Simulação de Cadeias de Markov*. XIX Seminário de Iniciação Científica da UFG – PIVIC, 2011.

[12] Junqueira, A. L. N.: *A probabilidade que a história nos conta*. XIV Conferência Interamericana de Educação Matemática, México, 2015.

[13] Lass, H. e P. Gottlieb: *Probability and Statistics*. Addison-Wesley, 1ª ed., 1971.

[14] Lipschutz, S.: *Matemática Finita*. Editora McGraw-Hill do Brasil, 1ª ed., 1972.

[15] Lipschutz, S.: *Probabilidade*. Editora McGraw-Hill do Brasil, 2ª ed., 1978.

[16] Lopes, C. E. e E. Meirelles: *O Desenvolvimento da Probabilidade e da Estatística*. XVIII Encontro Regional de Professores de Matemática, LEM/IMECC/UNICAMP, 2005.

[17] Marino, R.: *Banco Imobiliário - Post do Blog Todas as Configurações Possíveis*. `http://www.todasasconfiguracoes.com/2012/10/19/banco-imobiliario/`.

[18] Ross, S. M.: *Introduction to Probability Models*. Probability and Statistics. Academic Press, 2007, ISBN 9780123736352.

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

[20] Santos, T. A. R. dos: *A Matemática por trás do Google*. Tese de Mestrado, Profmat, CMCC-UFABC, 2014.

[21] Silva, C. B. da e C. de Queiroz e Silva Coutinho: *O nascimento da Estatística e sua relação com o surgimento da Teoria de Probabilidade*. Integração, (41):191–196, 2005.

[22] Viali, L.: *Algumas Considerações sobre a Origem da Teoria da Probabilidade*. Revista Brasileira de História da Matemática, 8(16):143–153, 2008.