

Editado por

Célia A. Zorzo Barcelos

Universidade Federal de Uberlândia - UFU
Uberlândia, MG, Brasil

Eliana X.L. de Andrade

Universidade Estadual Paulista - UNESP
São José do Rio Preto, SP, Brasil

Maurílio Boaventura

Universidade Estadual Paulista - UNESP
São José do Rio Preto, SP, Brasil



A Sociedade Brasileira de Matemática Aplicada e Computacional - SBMAC publica, desde as primeiras edições do evento, monografias dos cursos que são ministrados nos CNMAC.

Para a comemoração dos 25 anos da SBMAC, que ocorreu durante o XXVI CNMAC em 2003, foi criada a série **Notas em Matemática Aplicada** para publicar as monografias dos minicursos ministrados nos CNMAC, o que permaneceu até o XXXIII CNMAC em 2010.

A partir de 2011, a série passa a publicar, também, livros nas áreas de interesse da SBMAC. Os autores que submeterem textos à série Notas em Matemática Aplicada devem estar cientes de que poderão ser convidados a ministrarem minicursos nos eventos patrocinados pela SBMAC, em especial nos CNMAC, sobre assunto a que se refere o texto.

O livro deve ser preparado em **Latex (compatível com o Miktex versão 2.7)**, as figuras em **eps** e deve ter entre **80 e 150 páginas**. O texto deve ser redigido de forma clara, acompanhado de uma excelente revisão bibliográfica e de **exercícios de verificação de aprendizagem** ao final de cada capítulo.

Veja todos os títulos publicados nesta série na página
<http://www.sbmac.org.br/notas.php>

UMA INTRODUÇÃO À DINÂMICA ESTOCÁSTICA DE POPULAÇÕES

Leonardo Paulo Maia
leonardo.maia@ufabc.edu.br
lpmaia@gmail.com

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC



Sociedade Brasileira de Matemática Aplicada e Computacional

São Carlos - SP, Brasil
2012

Coordenação Editorial: Sandra Mara Cardoso Malta

Coordenação Editorial da Série: Eliana Xavier Linhares de Andrade

Editora: SBMAC

Capa: Matheus Botossi Trindade

Patrocínio: SBMAC

Copyright ©2012 by Leonardo Paulo Maia.

Direitos reservados, 2012 pela SBMAC. A publicação nesta série não impede o autor de publicar parte ou a totalidade da obra por outra editora, em qualquer meio, desde que faça citação à edição original.

Catálogo elaborado pela Biblioteca do IBILCE/UNESP
Bibliotecária: Maria Luiza Fernandes Jardim Froner

Maia, Leonardo Paulo

Uma Introdução à Dinâmica Estocástica de Populações

- São Carlos, SP : SBMAC, 2012, 76 p.; 20,5cm

- (Notas em Matemática Aplicada; v. 35)

e-ISBN 978-85-86883-96-5

1. Dinâmica de Populações. 2. Cadeias de Markov.
3. Processos de Nascimento e Morte. 4. Funções Geradoras
I. Maia, Leonardo Paulo. II. Título. III. Série.

CDD - 51

Esta é uma republicação em formato de e-book do livro original do mesmo título publicado em 2008 nesta mesma série pela SBMAC.

Conteúdo

Prefácio	7
1 Conceitos Básicos	11
1.1 Processos Estocásticos	11
1.2 Populações e Evolução	13
1.2.1 Introdução	13
1.2.2 Um pouco de Biologia	14
1.2.3 Forças evolucionárias	15
1.3 A Modelagem Matemática	17
2 Cadeias de Markov de Tempo Discreto e os Modelos de Wright-Fisher e de Moran	19
2.1 Introdução	19
2.2 O Modelo de Wright-Fisher	22
2.2.1 A caracterização do modelo	22
2.2.2 Relações e classes de equivalência	23
2.2.3 Estados absorventes, fixação e genealogias	25
2.2.4 Variações no tema: mutações e seleção	30
2.3 Simulações	32
2.4 O Modelo de Moran	34
3 Cadeias de Markov de Tempo Contínuo e os Processos de Nascimento e Morte	41
3.1 Introdução	41
3.1.1 Um pouco de Cálculo	42
3.2 Taxas e o Processo de Poisson	43
3.2.1 Relembrando a dinâmica populacional determinística	43
3.2.2 Taxas probabilísticas e o processo de Poisson	45
3.2.3 Funções geradoras	49
3.3 Nascimentos e Mortes	50

3.4	Algoritmo de Gillespie	52
4	Aplicações: Evolução do Câncer e Jogos em Populações Finitas	57
4.1	Um Modelo Matemático para o Câncer	57
4.1.1	Mais um pouco de Biologia	57
4.1.2	Um modelo estocástico simples	60
4.2	Um Jogo Evolucionário	61
A	Teoria Axiomática de Probabilidade	65
	Bibliografia	69

Prefácio

A Teoria da Probabilidade e sua fiel companheira no mundo discreto, a Análise Combinatória, são as matérias cujo ensino é considerado o mais difícil entre os professores do nível médio. Independentemente das sutilezas do seu processo de ensino/aprendizagem, que estariam presentes mesmo que essas disciplinas fossem bem dominadas pelos instrutores, elas são, para sermos honestos, verdadeiras desconhecidas daqueles docentes. Não coincidentemente, mas sem que seja possível discernir claramente o que é causa e o que é efeito, a maioria dos Licenciados e Bacharéis em Matemática e Bacharéis em Matemática Aplicada brasileiros terminam suas graduações cursando, no máximo, duas disciplinas nas áreas de Probabilidade e/ou Estatística. Portanto, aprendendo apenas alguns aspectos superficiais dos princípios básicos dessas importantíssimas áreas, e tendo absolutamente nenhum contato com suas infindáveis aplicações, esses profissionais graduam-se com uma formação indiscutivelmente incompleta.

Tenho a firme convicção de que, muitas vezes, essa superficialidade atrasa o progresso até mesmo dos que se denominariam “matemáticos puros”, pois estou entre aqueles que acredita que a Matemática e as aplicações iluminam-se mutuamente. É só ver a importância crescente da Teoria Ergódica em Sistemas Dinâmicos, para ficar em apenas um exemplo. Porém, sem sombra de dúvida, essa lacuna de formação é particularmente escandalosa no caso dos Bacharéis em Matemática Aplicada. Incertezas e flutuações são inerentes ao mundo natural. Muitas vezes, esses aspectos estocásticos são irrelevantes para a construção de um modelo que descreva apropriadamente um sistema real, mas esses casos são a exceção e não a regra, como fazem parecer muitos currículos tupiniquins.

Tentando nadar contra essa corrente, elaborei a proposta de um minicurso que foi aceito em dois eventos, o I Encontro Regional de Matemática Aplicada e Computacional de Bauru-SP (ERMAC), que será realizado em junho de 2008, e o XXXI Congresso Nacional de Matemática Aplicada e Computacional (CNMAC), que ocorrerá em setembro do mesmo ano, na

cidade de Belém do Pará. Estas notas foram elaboradas para servirem como material de referência desse minicurso, que terá duração de 4 horas no evento regional e de 6 horas no nacional. Elas são uma introdução a um mundo ao qual eu já me dedico há alguns anos, o da Modelagem Estocástica em Evolução, mediante a discussão de técnicas analíticas e computacionais simples.

Dado o cenário que descrevi no início deste Prefácio, nota-se facilmente que me meti em uma bela encrenca. Como ensinar aplicações motivadoras da Teoria da Probabilidade (incluindo processos estocásticos simples como cadeias de Markov) a alunos provavelmente não muito familiarizados com as noções básicas do tema em apenas um punhado de horas?! Bem, tive que admitir que meus ouvintes/leitores conhecem os aspectos intuitivos da Teoria da Probabilidade e mantive a discussão nesse nível, embora em alguns trechos e exercícios eu tenha abordado alguns tópicos mais elementares. Além disso, fiel ao meu credo que as aplicações e a teoria se complementam, recorri a modelos de Dinâmica de Populações e apresentei técnicas simples de simulação computacional, além de disponibilizar alguns exercícios para ajudar os estudantes a avaliarem sua compreensão dos conceitos discutidos.

Dessa forma, estas notas também podem (e devem) ser vistas como uma introdução à Biomatemática complementar às tradicionais abordagens baseadas em equações diferenciais e com um nítido direcionamento à Genética de Populações. Apresento também modelos estocásticos para descrever a evolução do câncer e jogos em populações finitas para ressaltar a importância das técnicas discutidas e soprar um “ar de pesquisa” sobre meus leitores. Embora a proposta tenha sido concebida para alunos de graduação, creio que alunos de pós-graduação interessados em Modelagem Matemática e outros curiosos também poderão apreciar as discussões.

Finalmente, ressalto que não tenho absolutamente nenhuma pretensão de originalidade acerca dos temas aqui discutidos. É claro que a escolha e a organização dos tópicos refletem minha linha de pensamento, mas utilizei explicitamente material que conheci em diversos textos, como [9], [17] e [19], nesta empreitada. Este texto, bem como eventuais erratas e material adicional, como códigos-fonte para simulações, estará disponível na forma de um arquivo PDF no sítio

<http://lpmaia.googlepages.com>

e eu ficaria muito grato se meus leitores me avisassem sobre quaisquer erros que sejam encontrados enviando uma mensagem para

lpmaia@gmail.com.

Santo André, 08 de abril de 2008.
Leonardo Paulo Maia

Capítulo 1

Conceitos Básicos

1.1 Processos Estocásticos

Nestas notas, adotaremos uma abordagem heurística no que se refere à apresentação dos processos estocásticos, seguindo o clássico livro [9]. Como naquela referência, espera-se que o leitor tenha familiaridade com o conceito de uma variável aleatória (v.a., para abreviar) X definida como uma função que atribui um valor real $X(\omega)$ a cada evento elementar ω de um espaço amostral Ω , bem como com as idéias de uma função distribuição de probabilidade $F_X(x) = \mathbb{P}(X \leq x)$ da v.a. X (incluindo casos multidimensionais) e de esperança de X . Uma referência padrão é [11].

Pois bem, podemos dizer com precisão bem razoável (para um pouco mais de rigor, ver o apêndice A) que um **processo estocástico** com **conjunto indicial** \mathcal{T} é uma coleção $\{X_t : t \in \mathcal{T}\}$ de variáveis aleatórias indexadas por \mathcal{T} . O parâmetro $t \in \mathcal{T}$ é usualmente interpretado como o tempo físico, e esse será o significado dessa grandeza nas aplicações que veremos neste minicurso. Então, como podemos interpretar um processo X_t se pensamos em t como tempo?

Certamente, os leitores têm consciência de que é importantíssimo distinguir uma v.a. X (uma função) do número $X(\omega)$ que ela associa a um evento ω . Como outros autores, chamaremos aquele número de uma **realização** da v.a. X . Ora, enquanto uma realização de uma v.a. é simplesmente “um ponto”, um processo estocástico, sendo uma coleção de v.a.’s, tem como realização uma família de pontos ao longo do tempo, um tipo de trajetória, na interpretação física. Contudo, enquanto um problema de valor inicial que satisfaz as condições usuais de regularidade (no sentido usual da teoria de equações diferenciais ordinárias) admite uma única solução, um processo

estocástico admite diversas realizações, toda uma família de trajetórias ou caminhos amostrais (*sample paths*), como veremos na figura 2.3, no próximo capítulo.

Nestas notas, tipicamente estaremos interessados em descrever como o tamanho de alguma população (no sentido mais amplo possível, poderíamos pensar em pessoas, ratos, bactérias, genes ou até mesmo idéias!) varia no tempo quando sua dinâmica é regida por regras probabilísticas. Porém, na modelagem com equações diferenciais, já aprendemos que, embora o tamanho de uma população real seja um inteiro não negativo, pode ser muito conveniente admitir que essa grandeza possa ser qualquer número real não negativo. Além disso, ninguém esperaria que o tamanho de uma população se alterasse radicalmente em pouco tempo; em algum sentido, seria bastante razoável que houvesse algum tipo de dependência entre os tamanhos (aleatórios) que uma população poderia apresentar em diferentes instantes. Apresentamos esses argumentos para mostrar que é muito natural que classifiquemos os diferentes processos estocásticos de acordo com alguns critérios, e alguns deles revelar-se-ão muito mais úteis do que outros.

Um processo estocástico pode ser caracterizado pelo seu **espaço de estados** \mathcal{S} (conjunto de possíveis valores das variáveis X_t), pela natureza do seu conjunto indicial \mathcal{T} e pelas relações de dependência entre as v.a.'s X_t . Vamos descrever esses critérios com mais cuidado.

1. Em nível básico, uma v.a. é usualmente classificada como discreta ou contínua, dependendo dos valores que pode apresentar. Analogamente, se cada X_t apresentar valores em um conjunto enumerável (tipicamente $\mathcal{S} = \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$ ou $\mathcal{S} = \mathbb{Z}$), diremos que o processo em questão é de estado discreto (ou que ele é uma **cadeia**). Por outro lado, se \mathcal{S} for um conjunto como \mathbb{R} ou $[0, \infty)$ (usualmente, uma união de intervalos), diremos que se trata de um processo de estado contínuo. Neste caso, usaremos a notação $X(t)$ ao invés de X_t . Como no caso de uma única v.a., o contexto da aplicação não determina univocamente a escolha de \mathcal{S} , embora algum caso particular geralmente se mostre mais conveniente.
2. A interpretação do parâmetro t como o tempo físico é tão pervasiva no estudo dos processos estocásticos que as expressões “processo de tempo discreto” e “processo de tempo contínuo” são as usuais para classificar um processo cujo conjunto indicial \mathcal{T} é discreto (como $\mathcal{T} = \{0, 1, 2, \dots\}$) ou contínuo (como $\mathcal{T} = [0, \infty)$), respectivamente.
3. Dados \mathcal{S} e \mathcal{T} , um processo estocástico é completamente caracterizado pela distribuição de probabilidade conjunta de qualquer família finita $\mathbf{X} = (X_{t_1}, \dots, X_{t_n})$ de v.a.'s do processo, ou seja, precisamos conhecer

$$F_{\mathbf{X}}(\mathbf{x}) = \mathbb{P}[X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n] \text{ para quaisquer } n \in \mathbb{N}, \mathbf{x} = (x_1, \dots, x_n) \text{ e } \mathbf{t} = (t_1, \dots, t_n), \text{ com } x_i \in \mathcal{S} \text{ e } t_i \in \mathcal{T}, i = 1, \dots, n.$$

Em geral, especificar todas as condições do critério 3 poderia ser um desafio fabuloso. Na prática, só são considerados casos em que a dependência entre as v.a.'s do processo pode ser descrita de uma forma bastante simples. Especificamente, talvez os mais populares processos estocásticos sejam os **processos de Markov**, aqueles em que, dado X_t , a distribuição de probabilidade de X_s , com $s > t$, não depende dos valores de X_u , com $u < t$. Em outras palavras, dado o estado presente, as expectativas acerca do futuro independem do comportamento exibido no passado. Em símbolos, para todo $n \in \mathbb{N}$,

$$\mathbb{P}(X_t \leq x | X_{t_1} = x_1, \dots, X_{t_n} = x_n) = \mathbb{P}(X_t \leq x | X_{t_n} = x_n) \quad (1.1.1)$$

sempre que $t_1 < t_2 < \dots < t_n < t$. Portanto, processos de Markov têm “memória curta”.

Nestas notas, no entanto, não consideraremos os processos de Markov em toda sua generalidade. Vamos nos restringir a cadeias de Markov de tempo discreto ou contínuo. Nos dois próximos capítulos, descreveremos as propriedades adicionais desses processos e iremos aplicá-los em diversos modelos populacionais. Alguns destes modelos são oriundos da Genética de Populações e, portanto, é conveniente que neste capítulo introdutório nós nos dediquemos também a esclarecer alguns conceitos biológicos que serão empregados posteriormente.

1.2 Populações e Evolução

1.2.1 Introdução

A enorme maioria dos estudantes se inicia na Biomatemática mediante o estudo de modelos simples de crescimento exponencial, logístico e presa-predador (Lotka-Volterra) [15, 16]. Nos dois primeiros casos, só há um tipo de “indivíduo” na população, cujo comportamento (dada alguma condição inicial) só depende de alguma regra dinâmica estabelecida previamente (crescimento linear irrestrito e restrito segundo uma função quadrática, respectivamente) e uma propriedade intrínseca daqueles indivíduos, que é sua taxa de crescimento. O modelo presa-predador costuma ser o primeiro contato desses estudantes com o efeito de competição, que se dá naquele caso mediante encontros entre representantes de duas diferentes espécies, uma das quais eliminando a outra. Contudo, ressaltamos que indivíduos de um tipo só deixam descendentes do mesmo tipo. Estudaremos em breve efeitos

análogos (competição e crescimento livre ou restrito) em modelos estocásticos, mas já adiantamos que a competição pode ocorrer de uma forma indireta (sem encontros entre diferentes espécies).

Em última instância, o significado do termo **evolução** é simplesmente mudança. Nesse sentido, os modelos determinísticos mencionados logo acima descrevem a evolução das populações idealizadas a que eles se referem. No entanto, o conceito de evolução é usado mais freqüentemente para descrever populações expostas a mecanismos de variabilidade, no sentido que indivíduos de uma certa categoria podem dar origem a descendentes que se enquadrem em outras categorias, por exemplo. Para sermos mais precisos, é conveniente introduzirmos uma terminologia apropriada a organismos biológicos reais, embora as aplicações subseqüentes sejam descritas abstratamente, sem que se faça menção explícita a organismos verdadeiros.

1.2.2 Um pouco de Biologia

Muito provavelmente, os leitores estão familiarizados com a noção que o material genético presente nas células de um organismo vivo contém toda a informação necessária para gerá-lo. Em particular, esse material genético apresenta-se como filamentos de um dos dois ácidos nucleicos, o DNA (sigla em inglês do ácido desoxirribonucleico) e o RNA (ácido ribonucleico). Somente mencionaremos o DNA no que se segue. Apesar de toda a complexidade desses compostos químicos, é possível caracterizar um filamento de DNA simplesmente descrevendo a seqüência de nucleotídeos ou bases (a unidade variável dessas moléculas) que o compõem, como também deve ser de conhecimento dos leitores.

Sabe-se que certos trechos de uma seqüência de DNA contêm as informações necessárias para a realização de funções específicas (como gerar uma certa proteína, por exemplo). Um trecho como esse (não uma particular seqüência de bases, mas a região em si) é denominado um *locus* (termo do Latim cujo plural é *loci*) gênico. Dizemos que um *locus* gênico é ocupado por um gene, que (agora, sim) é um segmento específico de DNA que contém informação para desempenhar a função correspondente ao *locus* que ele ocupa. Porém, um mesmo gene pode ser encontrado em diversas formas, semelhantes, mas não idênticas, que diferem umas das outras por alguns nucleotídeos. Cada uma dessas variantes é denominada um **alelo** e uma propriedade (ou conjunto de propriedades) manifestada pela ação de um alelo em um certo ambiente é conhecida como um **fenótipo**.

Um cromossomo é uma longa seqüência de DNA que contém diversos genes, entre outros segmentos cuja natureza ainda é sujeita a controvérsias. Cada organismo tem sua identidade genética registrada em uma família de

cromossomos (os seres humanos apresentam 46 cromossomos, por exemplo) que é denominada o **genótipo** daquele organismo. Portanto, o termo genótipo refere-se ao conjunto de genes que um organismo particular possui, embora freqüentemente essa nomenclatura também seja usada para denotar a composição genética em um *locus* ou conjunto de *loci* específico.

Pois bem, se concentrarmos nossa atenção em um particular *locus* gênico do genótipo de todos os membros de uma população, podemos dizer, na verdade, que estamos estudando uma população de genes, aqueles presentes no *locus* em questão. Portanto, de forma muito simplificada, a Genética de Populações é simplesmente o estudo de populações “moleculares”, cujos integrantes são genes ou conjuntos de genes, levando em conta em princípios da Genética. O objetivo dessa Ciência é descrever o comportamento das diversas freqüências alélicas, ou seja, as proporções observadas de cada alelo.

Os diferentes organismos podem ter uma única cópia do seu material genético (os chamados haplóides) ou podem tê-lo armazenado aos pares (diplóides). Os seres humanos, por exemplo, têm seus 46 cromossomos organizados em 23 pares. Conseqüentemente, quando se observa um certo *locus* gênico, é possível que um indivíduo seja homocigoto (alelos iguais nos dois *loci* correspondentes, um em cada cromossomo) ou heterocigoto (alelos distintos) e, neste caso, um dos alelos, que contém informações distintas para exercer a função que cabe àquele *locus*, pode vir a sobrepujar o outro e expressar o mesmo fenótipo que se manifestaria caso o indivíduo fosse homocigoto para aquele alelo, o que constitui o fenômeno da dominância.

1.2.3 Forças evolucionárias

Na subseção anterior, conhecemos uma série de conceitos que podem ser úteis para visualizarmos quem são os membros de uma população “molecular”, mas nada aprendemos sobre o que afeta o comportamento dinâmico dessas populações e há várias perguntas importantes ainda sem resposta. Quais são as origens dos diversos alelos que podemos ver em um *locus*? Se diferentes alelos exercem uma mesma função de maneiras diferentes, isso deve afetar o desempenho do organismo em seu ambiente. Como?

O estado de uma população é o resultado da interação de várias forças evolucionárias. De fato, a diversidade genética observada em um sistema deve-se aos efeitos conjuntos de processos elementares como mutações, seleção natural e deriva genética.

Genericamente, denomina-se **mutação** qualquer alteração em um trecho de material genético e que, portanto, pode ser transmitida a um descendente. Tais alterações podem ocorrer, por exemplo, pela remoção ou inserção de segmentos em um gene. Mas as mutações conceitualmente mais

simples são as pontuais, aquelas em que, embora o número total de nucleotídeos permaneça constante, alguns deles são inseridos incorretamente quando uma nova seqüência é criada a partir de uma já existente. O processo de mutação é o mecanismo básico de geração de variabilidade em uma população e leva ao surgimento de diferentes alelos.

Alelos distintos podem, em princípio, fazer com que seus portadores apresentem diferentes capacidades de sobrevivência em um certo ambiente e/ou tenham variadas taxas de fecundidade. Nesse sentido, dizemos que alguns alelos podem conferir melhor capacidade de adaptação do que outros, da mesma forma que algumas espécies podem ser mais bem adaptadas do que outras, em um nível ecológico. Uma forma conveniente de visualizar essas diferenças é imaginar que seria possível construir uma medida dessa adaptabilidade de um indivíduo a um certo ambiente, um **valor adaptativo** (correspondente ao termo *fitness*, em inglês) que associasse um número real a cada genótipo. Em outras palavras, o valor adaptativo de um organismo (ou até de um alelo) seria uma medida da capacidade conjunta de sobrevivência (viabilidade) e reprodução (fecundidade) do organismo em dadas condições, um “sucesso relativo esperado”. Em modelos de tempo contínuo, o valor adaptativo normalmente é associado à diferença entre as taxas de nascimento e morte do indivíduo. Quando o tempo é discreto e não há superposição de gerações, muitas vezes o valor adaptativo é dado pelo tamanho médio da prole que um organismo deixa para a geração seguinte.

Dessa forma, a famosa **seleção natural** de Charles Darwin é simplesmente o crescimento da frequência relativa dos organismos mais bem adaptados ao meio em que se encontram, com o passar do tempo. É importante perceber que, ao contrário da pressão mutacional, a seleção atua no sentido de reduzir a variabilidade de uma população. Contudo, é igualmente importante destacar, embora não tenhamos espaço para justificar essa afirmação [3], que o mecanismo biológico da hereditariedade não somente permite que a variabilidade genética seja mantida na ausência de seleção como também faz com que a escala de tempo envolvida na eliminação de alelos menos aptos possa ser da ordem de centenas de gerações. Portanto, a seleção não é capaz de eliminar a diversidade populacional.

Algumas expressões das passagens acima merecem novamente a nossa atenção: “sucesso relativo esperado”, “tamanho médio da prole”, “crescimento da frequência relativa”. Há uma tendência da frequência dos organismos mais aptos crescer, mas isso não quer dizer que todos os integrantes dessa subpopulação realizarão toda sua potencialidade. Alguns deixarão menos descendentes do que deles se esperaria, outros deixarão mais. Essas flutuações são ainda mais fortes se a população não for muito grande, em algum sentido. Imagine um dado usual, com 6 faces equiprováveis. Pelas

leis dos grandes números, sabemos que a proporção dos resultados correspondente a cada uma das faces será bem próxima de $1/6$ se o número de lançamentos for suficientemente grande. Por outro lado, um número pequeno de lançamentos pode resultar em proporções bem distintas. Analogamente, frequências alélicas podem sofrer fortes flutuações devido a erros de amostragem em populações finitas. Esse fenômeno chama-se **deriva genética** e pode levar à perda de algumas classes de indivíduos, especialmente aquelas menos numerosas, reduzindo a variabilidade populacional. Em suma, a evolução é um fenômeno probabilístico e essa é a razão pela qual o seu estudo requer o conhecimento de processos estocásticos.

Na tese de doutorado [10], que está disponível no sítio indicado no Prefácio destas notas, podem ser encontrados um panorama da Teoria da Evolução e da Genética de Populações e uma introdução à Evolução Molecular. Embora curtos e direcionados a objetivos mais restritos, mesmo assim talvez alguns leitores considerem esses textos úteis, especialmente porque eles indicam diversas referências adicionais.

1.3 A Modelagem Matemática

Antes de passarmos ao estudo de modelos matemáticos específicos, vamos tecer algumas considerações acerca da modelagem matemática. Em alguns casos que estudaremos adiante, consideraremos populações que seguem uma dinâmica de tempo discreto. Como essa hipótese pode ser justificada? Poderíamos estar considerando uma população real, que obviamente viveria em um mundo físico com tempo contínuo ¹, mas só a observaríamos em intervalos regulares de tempo e, mesmo assim, tentaríamos descrever seu comportamento. Mas esse argumento “concreto” não é estritamente necessário para justificar aquela escolha.

O fato é que uma distorção da realidade aparentemente tão grave quanto imaginar que uma população real vive em um mundo “imaginário” onde o tempo é discreto, embora certamente não possa descrever fielmente muitos aspectos do sistema em estudo, pode se revelar, com uma frequência muitíssimo maior do que um leigo em modelagem poderia conceber, uma excelente ferramenta para descrever uma determinada característica daquele sistema. Por que isso é possível? Talvez ninguém no mundo tenha uma resposta direta para essa pergunta e isso é o que torna o ensino da modelagem tão difícil; aparentemente, a única recomendação coerente para a aprendizagem de técnicas de modelagem é estudar muitos exemplos. Em última instância,

¹ Talvez isso não seja tão óbvio assim... Você sabia que há cientistas que cogitam a possibilidade do tempo físico ser discreto?

só o confronto com a experiência é que pode dizer se um modelo matemático é útil ou não como uma representação da realidade, embora um modelo ineficiente possa, mesmo assim, envolver Matemática de bom nível e/ou inspirar novos desenvolvimentos puramente teóricos.

Capítulo 2

Cadeias de Markov de Tempo Discreto e os Modelos de Wright-Fisher e de Moran

2.1 Introdução

Neste capítulo, estudaremos cadeias de Markov de tempo discreto e algumas das suas aplicações em Genética de Populações. Adotaremos sempre o espaço indicial $\mathcal{T} = \{0, 1, \dots\}$ e é usual falarmos que X_t está no estado i se $X_t = i$. Devido ao caráter Markoviano do processo, expresso na eq. (1.1.1), podemos caracterizá-lo apresentando apenas as probabilidades de transição $\mathbb{P}(X_t = j | X_{t-1} = i)$, o que equivale a caracterizar a matriz $\mathbf{P}^{(t)} = \left(p_{i,j}^{(t)}\right)$ em que

$$p_{i,j}^{(t)} = \mathbb{P}(X_{t+1} = j | X_t = i).$$

Essa matriz, em geral, depende do instante em que se considera a transição. O uso iterado da definição de probabilidade condicional e da propriedade Markoviana revela que a probabilidade $\pi_i^{(t)} = \mathbb{P}(X_t = i)$ da cadeia estar no estado i no instante t é completamente determinada pela matriz de transição e pela distribuição de probabilidade do estado inicial. De fato, adotando o espaço de estados $\mathcal{S} = \{0, 1, \dots, N\}$, onde N é um inteiro pos-

itivo, e definindo o vetor $\boldsymbol{\pi}^{(t)} = (\pi_0^{(t)}, \dots, \pi_N^{(t)})$, é fácil ver que

$$\begin{aligned}\pi_j^{(t)} &= \mathbb{P}(X_t = j) = \sum_{i \in \mathcal{S}} \mathbb{P}(X_t = j, X_{t-1} = i) \\ &= \sum_{i \in \mathcal{S}} \mathbb{P}(X_t = j | X_{t-1} = i) \mathbb{P}(X_{t-1} = i) \\ &= \sum_{i \in \mathcal{S}} p_{i,j}^{(t-1)} \pi_i^{(t-1)}\end{aligned}$$

que, interpretando o vetor $\boldsymbol{\pi}^{(t)}$ como uma matriz linha, equivale à equação matricial

$$\boldsymbol{\pi}^{(t)} = \boldsymbol{\pi}^{(t-1)} \mathbf{P}^{(t-1)}, \quad (2.1.1)$$

cujas iterações levam à equação

$$\boldsymbol{\pi}^{(t)} = \boldsymbol{\pi}^{(0)} \mathbf{P}^{(0)} \mathbf{P}^{(1)} \dots \mathbf{P}^{(t-1)}, \quad (2.1.2)$$

onde a ordem de multiplicação das matrizes deve ser respeitada, naturalmente.

Em alguns casos, a probabilidade da cadeia passar de um estado a outro em um passo temporal não depende de t . Nesse caso, podemos definir univocamente uma **matriz de transição** $\mathbf{P} = (p_{i,j})$ onde $p_{i,j} = \mathbb{P}(X_t = j | X_{t-1} = i)$, omitindo a dependência temporal em t sem incorrer em qualquer imprecisão. Tais cadeias de Markov são denominadas homogêneas. Nessas condições, a eq. (2.1.2) adquire a forma

$$\boldsymbol{\pi}^{(t)} = \boldsymbol{\pi}^{(0)} \mathbf{P}^t, \quad (2.1.3)$$

em que \mathbf{P}^t é a t -ésima potência de \mathbf{P} .

Exercício: Verifique a validade da eq. (2.1.2).

Uma cadeia de Markov homogênea de tempo discreto pode ser representada por um **diagrama de estados**, que é simplesmente um grafo em que os nós são indexados pelos possíveis estados da cadeia de Markov e os arcos são indexados pelas correspondentes probabilidades de transição entre os estados. Esse conceito está ilustrado na figura 2.1, onde $\mathcal{S} = \{1, 2\}$ e $p_{1,2} = 0, 1$.

A eq. (2.1.3) suscita várias questões importantes. A mais direta de todas seria “como calcular essa potência?”. Bem, um resultado bem conhecido da Álgebra Linear garante que, se uma matriz for diagonalizável e pudermos efetivamente determinar seus autovalores e autovetores para diagonalizá-la mediante uma transformação de semelhança, o cálculo de uma potência qualquer dessa matriz torna-se trivial. Porém, mesmo quando essas

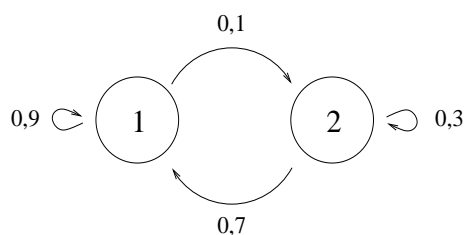


Figura 2.1: Exemplo de diagrama de estados de uma cadeia de Markov homogênea de tempo discreto.

condições são satisfeitas, essa trivialidade pode se manifestar apenas na estrutura conceitual das operações a serem realizadas, pois a complexidade dos cálculos pode inviabilizar qualquer tentativa de solução analítica. Nessas situações, o conhecimento de Álgebra Linear Numérica pode ser essencial para a obtenção de resultados confiáveis.

Exercício: Determine completamente a cadeia de Markov (ou seja, determine explicitamente $\pi^{(t)}$, resolvendo a eq. (2.1.3)) representada na figura 2.1 mediante a diagonalização da matriz de transição.

Uma outra pergunta bem pertinente, que os iniciados em Análise devem ter feito a si próprios assim que viram a eq. (2.1.3), é “essa recorrência tem um limite?”. Em muitos casos, sim, existe um único ponto fixo estável π tal que $\pi = \pi P$ e $\pi^{(t)} \rightarrow \pi$ quando $t \rightarrow \infty$ independentemente do estado inicial $\pi^{(0)}$. Nesses casos, o cálculo de π equivale à resolução de um sistema linear e métodos computacionais podem ser necessários, novamente. Porém, há diversas situações em que não existe uma distribuição estacionária no sentido descrito acima¹, como veremos na próxima seção.

Exercício: Mostre que, na cadeia de Markov representada na figura 2.1, independentemente do estado inicial $\pi^{(0)}$, $\pi^{(t)} \rightarrow \pi$ quando $t \rightarrow \infty$, onde $\pi = \pi P$.

Agora, vamos começar a ver diversos exemplos de populações idealizadas cujas dinâmicas podem ser representadas como cadeias de Markov, para descobriremos as dificuldades que podem surgir.

¹Podem haver um ou vários pontos fixos, mas nenhum deles ser alcançado a partir de certas condições iniciais; pode ser que todas as condições iniciais levem a um ponto fixo, mas nem sempre ao mesmo; ou pontos fixos podem simplesmente não existir.

2.2 O Modelo de Wright-Fisher

2.2.1 A caracterização do modelo

Vamos usar um modelo de tempo discreto para descrever uma população, que será composta por N unidades de dois tipos de alelos, por exemplo. A hipótese de tamanho constante da população tem o propósito de evidenciar os efeitos da deriva genética, como definida no capítulo 1. Em cada geração, a população é completamente caracterizada pelo número de um dos tipos alélicos. Indexemos essas categorias como 1 e 2 e vamos dizer que o estado do sistema é dado por X_t , o número de alelos do tipo 1 na t -ésima geração. Todos os membros dessa população só vivem durante uma única geração, o que nos leva a dizer que estudaremos um modelo sem superposição de gerações. A composição do sistema no instante seguinte é determinada por N repetições do seguinte procedimento: um representante da geração t é escolhido ao acaso e dá origem a um novo integrante idêntico ao seu ancestral, que participará normalmente do próximo sorteio.

Em outras palavras, pensando em termos de um tradicional modelo probabilístico de urnas, é como se tivéssemos à nossa disposição

- i. uma urna vazia, que representa a próxima geração a ser construída,
- ii. para representar a geração atual, uma urna cheia com N bolas, das quais i seriam brancas e $N - i$ seriam pretas, caso haja i alelos do tipo 1 e $N - i$ do tipo 2, e
- iii. um “reservatório” de bolas, com tantas bolas das duas cores quantas fossem necessárias,

e sorteássemos uma bola da urna cheia que, embora fosse devolvida à sua urna de origem, determinaria a cor de uma bola a ser retirada do reservatório e colocada na urna a ser preenchida, iterando esse procedimento N vezes. Trata-se, portanto, de um simples esquema de amostragem com reposição.

Esse esquema é conhecido como o **modelo de Wright-Fisher** por ter sido concebido independentemente por dois dos 3 cientistas responsáveis pela estruturação da Genética de Populações, Sewall Wright e Ronald Fisher, este também amplamente reconhecido como um dos pais da Estatística moderna. Apesar de parecer excessivamente simples, esse modelo é o ponto de partida para a construção de inúmeros modelos teóricos mais avançados e é discutido em todo os bons livros sobre Genética de Populações, até mesmo naqueles com enfoque mais qualitativo.

Denotemos por p a probabilidade de que um alelo do tipo 1 seja escolhido para reproduzir-se. Como essa amostragem é realizada com reposição,

aquela probabilidade é constante durante todas as N iterações necessárias para a formação da geração seguinte e, pela uniformidade no sorteio, é evidente que $p = \frac{i}{N}$. Se denominarmos “sucesso” a chegada de um indivíduo do tipo 1 à próxima geração e “fracasso” o evento complementar (neste caso, sorteio de um alelo do tipo 2 para reproduzir-se), a transição $X_t = i \rightarrow X_{t+1} = j$ corresponde à observação de j sucessos e $N - j$ fracassos. Ora, a ocorrência de j sucessos em N eventos que só podem resultar em sucesso ou fracasso quando a probabilidade de sucesso é constante é um evento descrito pela bem conhecida distribuição binomial de probabilidade. Portanto, é claro que o modelo de Wright-Fisher é uma cadeia de Markov homogênea de tempo discreto cuja matriz de transição é determinada por

$$p_{i,j} = \binom{N}{j} \left(\frac{i}{N}\right)^j \left(1 - \frac{i}{N}\right)^{N-j}. \quad (2.2.4)$$

Exercício: Calcule a média e a variância do número de sucessos em N eventos que podem resultar em sucesso com probabilidade p ou em fracasso com probabilidade $1 - p$. Em uma população de Wright-Fisher com i alelos do tipo 1 em uma certa geração, qual é o número médio de alelos do mesmo tipo na geração seguinte?

Porém, nada garante que o estudo desse modelo seja trivial. Não há uma forma fechada para \mathbf{P}^t que “ilumine” o comportamento dinâmico dessa cadeia. Além disso, o modelo de Wright-Fisher não admite uma distribuição de equilíbrio $\boldsymbol{\pi}$ no sentido da seção anterior, pois $\boldsymbol{\pi}^{(t)}$ pode convergir para mais de um vetor. De fato, imaginemos que $\boldsymbol{\pi}^{(0)} = (0, 0, \dots, 0, 1)$, ou seja, $\mathbb{P}(X_0 = N) = 1$ e a população é composta inicialmente apenas por alelos do tipo 1. Ora, é óbvio que essa condição nunca mudará e que $\boldsymbol{\pi}^{(t)} = (0, 0, \dots, 0, 1)$ para todo $t \in \mathcal{T}$. Analogamente, $(1, 0, 0, \dots, 0)$ também é um ponto fixo da recorrência dada pela eq. (2.1.1). Portanto, ao invés de empregarmos “força bruta”, devemos abrir mão da caracterização completa desse modelo e tentar identificar quais resultados relevantes poderiam ser obtidos se trilhássemos caminhos alternativos, baseados em conceitos diferentes dos que vimos até agora.

2.2.2 Relações e classes de equivalência

Alguns aspectos do comportamento de uma cadeia de Markov podem ser compreendidos se tivermos alguma idéia sobre como o espaço de estados pode ser “explorado” pela cadeia. Nesse contexto, é claro que é importante trabalharmos com um critério preciso de acessibilidade entre os estados do processo estocástico. Conseqüentemente, precisamos, nesta etapa dos

nostros estudos, esquecer momentaneamente a modelagem e nos voltarmos um pouco mais à Matemática.

Dada uma cadeia de Markov homogênea de tempo discreto, diremos que um estado j é acessível a partir do estado i (o que denotaremos por $i \hookrightarrow j$) se existe $t \in \mathcal{T}$ tal que $(\mathbf{P}^t)_{ij} = \mathbb{P}(X_{t+s} = j | X_s = i) > 0$, ou seja, se, partindo ² do estado i , a cadeia puder atingir o estado j após um número t suficientemente grande, mas finito, de transições. Se $i \hookrightarrow j$ e $j \hookrightarrow i$, diremos que i e j se comunicam, o que representaremos por $i \leftrightarrow j$.

Embora o conceito de comunicabilidade já tenha sido definido de forma precisa, podemos abordar nosso problema a partir de um novo ponto de vista se pensarmos em termos ainda mais abstratos. Para facilitar a compreensão dos conceitos a serem introduzidos, vamos reelaborar o conceito de função. Muitos estudantes têm a noção de uma função como uma “regra f ” que “gera um número $f(x)$ ” a partir de um “número x ”. Mas bons livros e bons professores enfatizam que, na verdade, uma função é um trio formado por dois conjuntos (não necessariamente de números) A (o domínio) e B (o contradomínio), e uma “regra” que associa um elemento de B a cada elemento de A . Esse conceito já é bem mais adequado do que o primeiro, mas ainda mantém um componente excessivamente impreciso, embora muitos profissionais das ciências exatas passem suas vidas inteiras satisfeitas com ele: o que é uma “regra”? Na verdade, podemos definir uma função sem recorrer a esse termo, usando apenas a tradicional Teoria de Conjuntos.

Dados dois conjuntos, A e B , uma relação R entre A e B é qualquer subconjunto do produto cartesiano $A \times B$ e denotamos por xRy o fato de (x, y) pertencer a R , com $x \in A$ e $y \in B$. Dessa forma, uma função com domínio A e contradomínio B é simplesmente uma relação R entre A e B tal que, para todo $x \in A$ (para garantir que todo elemento do domínio tenha sua imagem), existe um único $y \in B$ (a unicidade faz com que a imagem de um elemento do domínio esteja bem definida) de modo que $(x, y) \in R$. Portanto, uma função pode ser vista como um tipo particular de conjunto de pares ordenados.

Assim como uma função é um tipo particular (e importantíssimo) de relação, outros casos especiais de relações merecem destaque. Mas só é pertinente a este texto definirmos o que é uma **relação de equivalência** em um conjunto A . Trata-se de uma relação $R \subset A \times A$ em que, para quaisquer $x, y, z \in A$, valem as propriedades

- i. reflexiva, ou seja, $(x, x) \in R$ (ou xRx),
- ii. simétrica (se xRy , então yRx) e

²Note que $s \in \mathcal{T}$ é arbitrário.

iii. transitiva (se xRy e yRz , então xRz).

Embora pudéssemos dizer precisamente se dois estados se comunicavam ou não, não saberíamos dizer precisamente o que é a propriedade de comunicabilidade entre dois estados quando introduzimos aquele critério. Agora, os leitores já devem ter percebido que a comunicabilidade é uma relação de equivalência \leftrightarrow (um símbolo análogo a R na discussão acima) no espaço de estados \mathcal{S} . De fato, vê-se claramente que ela é reflexiva, no sentido que $i \leftrightarrow i$ (trivialmente, com $t = 0$ na definição da acessibilidade). É evidente também que a relação entre dois estados que se comunicam é simétrica. Além disso, para quaisquer estados i, j e k tais que $i \leftrightarrow j$ e $j \leftrightarrow k$, necessariamente $i \leftrightarrow k$ pois, como existem t_1 que garante a veracidade de $i \hookrightarrow j$ e t_2 que garante a veracidade de $j \hookrightarrow k$, $t_1 + t_2$ garante que $i \hookrightarrow k$ e um argumento análogo aplica-se ao caminho de volta entre os estados.

Dada uma relação de equivalência R em A , este conjunto pode ser particionado como a união de uma família ³ de conjuntos não vazios ⁴ dois a dois disjuntos, cada um dos quais denominado uma **classe de equivalência** C e caracterizado por qualquer um dos seus elementos u como $C = \{x \in A : xRu\}$. Veremos que a relação de equivalência comunicabilidade particiona o espaço de estados do modelo de Wright-Fisher como a união de 3 classes de equivalência que admitem uma importante interpretação biológica.

2.2.3 Estados absorventes, fixação e genealogias

No que se segue, vamos considerar a comunicabilidade \leftrightarrow como uma relação de equivalência no espaço de estados $\mathcal{S} = \{0, 1, \dots, N\}$ do modelo de Wright-Fisher. Pela eq. (2.2.4), é claro que, no modelo de Wright-Fisher, $p_{i,j} > 0$ se $i, j \in C_{12} = \{1, \dots, N-1\}$. Portanto, considerando $t = 1$ na definição de acessibilidade, conclui-se que $i \leftrightarrow j$ se $i, j \in C_{12}$. No entanto, como $p_{0,0} = 1 = p_{1,1}$, nenhum $i \in C_{12}$ é acessível a partir dos estados 0 ou N , de modo que C_{12} é uma classe de equivalência bem definida. Os estados i tais que $p_{i,i} = 1$, como 0 e N , são denominados **estados absorventes**, pois, como já tínhamos visto, a cadeia permanecerá indefinidamente naquele estado se eventualmente atingi-lo. Claramente, cada um deles não se comunica com qualquer estado diferente de si próprio e origina uma classe de equivalência, $C_1 = \{N\}$ e $C_2 = \{0\}$, de modo que $\mathcal{S} = C_{12} \cup C_1 \cup C_2$.

Portanto, quando existem, estados absorventes constituem classes de equivalência que são singletos, e sempre são pontos fixos naturais das cadeias de Markov homogêneas, embora não necessariamente os únicos. Como uma

³Que pode ter um único integrante.

⁴Mas que podem ser singletos, ou seja, ter um único elemento.

cadeia que se inicie em um estado absorvente nunca mudará de estado, uma cadeia que contenha dois ou mais estados absorventes certamente não atinge um estado estacionário independentemente das condições iniciais, embora essa “falta de convergência universal” possa ocorrer mesmo que haja um ou nenhum estado absorvente. Os leitores interessados em descrições mais completas desses temas devem procurar outras referências, como [9].

Mas veremos que, dada qualquer distribuição de probabilidade para o estado inicial de uma população que segue a dinâmica de Wright-Fisher, necessariamente um dos dois estados absorventes será alcançado. Para demonstrar isso, precisamos mostrar que a cadeia de Markov não pode permanecer indefinidamente em C_{12} e, quando visitar C_1 ou C_2 , nada mais interessante ocorrerá.

Vamos denotar por q a probabilidade condicional da cadeia de Wright-Fisher fazer uma transição para C_{12} dado que ela já se encontra nesse conjunto. Trata-se, simplesmente, de uma probabilidade de permanência naquela classe. Com base na definição de probabilidade condicional e na lei da probabilidade total, segue que

$$\begin{aligned}
q &= \mathbb{P}(X_t \in C_{12} | X_{t-1} \in C_{12}) \\
&= \sum_{i \in C_{12}} \mathbb{P}(X_t \in C_{12}, X_{t-1} = i | X_{t-1} \in C_{12}) \\
&= \sum_{i \in C_{12}} \mathbb{P}(X_t \in C_{12} | X_{t-1} = i, X_{t-1} \in C_{12}) \cdot \mathbb{P}(X_{t-1} = i | X_{t-1} \in C_{12}) \\
&= \sum_{i \in C_{12}} \mathbb{P}(X_t \in C_{12} | X_{t-1} = i) \cdot \mathbb{P}(X_{t-1} = i | X_{t-1} \in C_{12}) \\
&= \sum_{i \in C_{12}} [1 - \mathbb{P}(X_t \notin C_{12} | X_{t-1} = i)] \cdot \mathbb{P}(X_{t-1} = i | X_{t-1} \in C_{12}) \\
&= \sum_{i \in C_{12}} [1 - p_{i,0} - p_{i,N}] \cdot \mathbb{P}(X_{t-1} = i | X_{t-1} \in C_{12}) \\
&< \sum_{i \in C_{12}} 1 \cdot \mathbb{P}(X_{t-1} = i | X_{t-1} \in C_{12}) = 1,
\end{aligned} \tag{2.2.5}$$

pois podemos ver pela eq. (2.2.4) que $p_{i,0} > 0$ e $p_{i,N} > 0$ para qualquer $i \in C_{12}$. Portanto, $q < 1$. A probabilidade $q^{(n)}$ da cadeia de Markov permanecer em C_{12} por $n - 1$ transições seguidas e sair dessa classe no n -ésimo passo ($n = 1, 2, \dots$) é dada por uma distribuição geométrica,

$$q^{(n)} = q^{n-1}(1 - q), \tag{2.2.6}$$

de modo que a probabilidade de eventualmente a cadeia deixar C_{12} é

$$\begin{aligned} \sum_{n=1}^{\infty} q^{(n)} &= \sum_{n=1}^{\infty} q^{n-1}(1-q) \\ &= (1-q) \sum_{n=1}^{\infty} q^{n-1} = (1-q) \frac{1}{1-q} = 1, \end{aligned} \quad (2.2.7)$$

após calcularmos a soma de uma série geométrica.

Exercício: Determine a média e a variância do número de passos até o primeiro sucesso, no sentido da distribuição geométrica dada na eq. (2.2.6).

Dessa forma, sabemos que a população eventualmente será constituída por apenas um tipo de alelo, 1 ou 2. Porém, como um modelo para um problema biológico verdadeiro, precisamos nos lembrar que a população verdadeira seria constituída por organismos de alguma natureza e que estaríamos interessados em estudar o que acontece em um particular *locus* gênico. Em Genética de Populações, dizemos que um alelo alcançou **fixação** se todos os organismos da população analisada apresentam-no no *locus* em questão. A fixação é um fenômeno biológico extremamente importante, pois, na ausência de mutações, todos os descendentes daquela população terão aquele mesmo trecho de material genético.

Assim, seria desejável conhecer a probabilidade de fixação de cada um dos tipos de alelo, dado um estado inicial da população. Na verdade, dados i alelos do tipo 1 no instante inicial, a probabilidade de fixação dessa classe é i/N e, conseqüentemente, a do tipo 2 é $1 - i/N$. Esse resultado é simples, correto, e quase levemente intuitivo. Por que, afinal, as freqüências alélicas iniciais contribuem linearmente, digamos assim, para as probabilidades de fixação? Isso requer uma explicação que nos levará a pensar na dinâmica dessa população de uma forma completamente diferente do que fizemos até agora.

Para responder a essa pergunta, é conveniente considerarmos primeiro como o modelo de Wright-Fisher poderia ser generalizado para descrever uma população com um número arbitrário l de alelos, o que por si só já será interessante. Ora, precisamos apenas considerar que l resultados, e não apenas 2, podem ocorrer em cada um dos N sorteios. Como a quantidade de configurações possíveis é evidentemente finita ⁵, embora o estado da população não possa mais ser determinado pela quantidade de um único tipo alélico, ele ainda pode ser determinado por um único número, que seria

⁵Ela é igual ao número de soluções da equação $\sum_{i=1}^l x_i = N$ com variáveis inteiras e não negativas, e vale $\binom{N+l-1}{N}$. Esse é um problema clássico em Análise Combinatória, como pode ser visto em [13] ou [20].

o índice de qualquer estado particular em alguma enumeração de todas as configurações. No entanto, é muito mais transparente descrever as probabilidades de transição usando uma notação vetorial para caracterizar cada estado. Sejam U_k e V_k as quantidades de alelos do tipo k nos instantes $t + 1$ e t , respectivamente. A neutralidade no sorteio dos alelos que contribuirão para a geração seguinte faz com que a probabilidade de um alelo do tipo k ser sorteado no instante t para reproduzir-se seja $p_k = V_k/N$ e com que a probabilidade de transição entre os estados $\mathbf{U} = (U_1, \dots, U_l)$ e $\mathbf{V} = (V_1, \dots, V_l)$, onde $\sum_{k=1}^l U_k = N = \sum_{k=1}^l V_k$, seja dada por uma distribuição multinomial,

$$\mathbb{P}(X_{t+1} = \mathbf{U} | X_t = \mathbf{V}) = \left(\frac{N!}{\prod_{k=1}^l U_k!} \right) \prod_{k=1}^l p_k^{U_k}, \quad (2.2.8)$$

que os leitores devem comparar com a eq. (2.2.4).

No caso com 2 alelos, a notação escolhida para representar as classes de equivalência do espaço de estados indicava que todos os estados de uma mesma classe continham os mesmos tipos de alelos. Explicitamente, embora o estado da cadeia de Wright-Fisher só levasse em conta a quantidade de alelos do tipo 1, é claro que C_{12} , por exemplo, era o conjunto de todos os estados da população em que os dois tipos de alelos estavam presentes. Um pouco de reflexão fará com que os leitores percebam que, no caso geral, haverá uma classe de equivalência para cada possível combinação dos l alelos ⁶ e que qualquer uma delas que contenha dois ou mais alelos estará sujeita ao mesmo efeito que foi justificado pelas eqs. (2.2.5) e (2.2.7), que é a impossibilidade da cadeia permanecer indefinidamente na classe. Com probabilidade 1, eventualmente a cadeia deixará a classe de equivalência onde estava para nunca mais retornar, porque sua saída equivale à perda de um tipo de alelo, que nunca mais será visto na população. A cadeia de Markov terá ingressado, portanto, em uma outra classe de equivalência, com pelo menos um tipo de alelo a menos do que a classe em que se encontrava até então. Abandonando classe após classe, eventualmente a população atingirá um estado absorvente, correspondente à fixação de um alelo e ao fim de toda e qualquer transição entre estados.

Muito bem, agora iremos nos basear neste importante resultado para tentarmos calcular a probabilidade de fixação em um estado absorvente. Vamos considerar uma população de tamanho N que contém inicialmente N tipos de alelos, um exemplar de cada. Qual é a probabilidade de um particular alelo alcançar a fixação? Por simetria, é evidente que é $1/N$. Mas é importante que os leitores percebam que, como há apenas um exemplar de

⁶O número total de classes de equivalência é $\sum_{k=1}^l \binom{l}{k} = 2^l - 1$.

cada tipo de alelo inicialmente, a fixação de qualquer um deles equivale a, em algum momento, toda a população ser constituída por descendentes daquele particular alelo que integrava a população original, como está ilustrado na figura 2.2. Todas as linhagens dos demais $N - 1$ integrantes da população inicial se extinguíram. Imaginemos, agora, que o estado inicial da população contenha l variedades alélicas, com $l < N$, o que necessariamente faz com que pelo menos uma variedade alélica tenha mais de um representante. Imaginando que pudéssemos rastrear a genealogia de um particular alelo ao longo da passagem das gerações (explicitamente, ignorando eventuais alelos do mesmo tipo mas oriundos de outras linhagens), a probabilidade de que seus descendentes eventualmente constituíssem a totalidade da população em algum instante é $1/N$. Portanto, se inicialmente houver i_k alelos do tipo k , a probabilidade de fixação dessa classe é simplesmente a soma das probabilidades de fixação de cada um dos seus membros, que é $i_k(1/N) = i_k/N$.

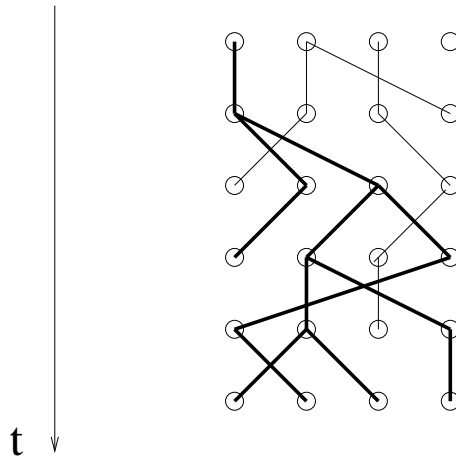


Figura 2.2: Possível genealogia de um alelo que atingiu fixação em uma população de Wright-Fisher de tamanho $N = 4$.

Apesar de toda a sua simplicidade, essa perspectiva genealógica é surpreendentemente útil na modelagem em Genética de Populações, tanto teórica quanto computacionalmente, pois essa idéia levou alguns pesquisadores a investigarem certas populações simulando sua evolução de trás para frente. Essa estratégia é muito eficiente porque, ao longo das gerações, só são representados os indivíduos que deixaram descendentes, tipicamente muito menos numerosos do que os membros das linhagens que se extinguíram.

Deve-se ressaltar que, no modelo de Wright-Fisher, um dos tipos de alelos atinge fixação mesmo em um contexto de **seleção neutra**, ou seja, uma classe sobrepuja a outra mesmo sem ter qualquer vantagem seletiva. Na verdade, a eliminação de um dos alelos é uma consequência direta da finitude da população. Dessa forma, apesar de ser uma representação bastante idealizada de um fenômeno biológico potencialmente muito complexo, que é a transmissão de material genético entre duas gerações de uma população, o modelo de Wright-Fisher consegue evidenciar a relevância da deriva genética, como descrita na subseção 1.2.3.

Mas e quanto às demais forças evolucionárias? Poderiam mutações e seleção natural serem incorporadas ao esquema de Wright-Fisher? Sim, outros efeitos biológicos podem ser considerados mediante alterações simples do modelo, embora ele seja particularmente ilustrativo em relação aos efeitos da deriva genética.

2.2.4 Variações no tema: mutações e seleção

Mutações podem ser vistas genericamente como alterações na natureza dos descendentes. Portanto, podemos incorporar esse fenômeno ao modelo de Wright-Fisher simplesmente considerando que um alelo sorteado para reproduzir-se não necessariamente dá origem a um descendente idêntico a si mesmo. Especificamente, em um modelo de Wright-Fisher com dois tipos de alelos em que cada indivíduo pode dar origem a um alelo do tipo oposto com probabilidade u , a probabilidade de transição de um estado com i alelos do tipo 1 para um estado com j alelos do mesmo tipo é

$$p_{i,j} = \binom{N}{j} p^j (1-p)^{N-j}, \quad (2.2.9)$$

em que

$$p = \left(\frac{i}{N}\right) (1-u) + \left(1 - \frac{i}{N}\right) u \quad (2.2.10)$$

é a probabilidade de um alelo do tipo 1 ingressar na população, seja como cópia direta, seja como mutação.

Exercício: O modelo de Wright-Fisher descrito acima apresenta quantos estados absorventes?

É claro que apenas as particularidades de um contexto biológico podem determinar se a hipótese da ocorrência de mutações como interconversões entre duas variedades alélicas é pertinente ou não. De fato, há um modelo famoso em Genética de Populações denominado modelo de infinitos alelos

em que, com base na enorme variedade de configurações observada no material genético de qualquer organismo, estipula-se que cada mutação leve ao surgimento de um novo alelo [2].

Exercício: Construa um modelo de Wright-Fisher com mutação que apresente um único estado absorvente. Escreva explicitamente a probabilidade de transição pertinente, como na eqs. (2.2.9) e (2.2.10).

É igualmente simples descrever a dinâmica de uma população de Wright-Fisher sujeita à seleção natural. Como vimos no capítulo 1, um organismo é mais apto do que os demais se ele se reproduz de forma eficiente que seus concorrentes e transmite essa capacidade a seus descendentes. Podemos incorporar essa característica ao modelo de Wright-Fisher promovendo sorteios tendenciosos para escolher os alelos que devem se reproduzir. Especificamente, podemos atribuir um valor adaptativo (que equivaleria a um peso probabilístico ⁷) a cada variedade alélica e escolher quem irá se reproduzir com probabilidade proporcional ao seu valor adaptativo.

Digamos que os alelos tenham valores adaptativos dados por w_1 ou w_2 , com índices auto-evidentes. Como sempre, imaginemos que em uma certa geração haja i alelos do tipo 1 e $N - i$ do tipo 2. A probabilidade de que um alelo do tipo 1 seja escolhido para deixar um descendente para a próxima geração é

$$p = \frac{iw_1}{iw_1 + (N - i)w_2} = \frac{i}{i + (N - i)s}, \quad (2.2.11)$$

onde $s = w_2/w_1$ é um coeficiente seletivo que indica se um alelo do tipo 2 é mais (caso $s > 1$) ou menos ($0 < s < 1$) apto do que um do tipo 1. O leitor deve perceber que, no esquema proposto, o valor adaptativo de uma das variedades alélicas pode ser arbitrariamente fixado em 1 e a intensidade da seleção pode ser “calibrada” por um único parâmetro (s , neste caso).

A matriz de transição da cadeia de Markov para este modelo de Wright-Fisher com seleção continua sendo dado pela eq. (2.2.9), mas agora com p dado pela eq. (2.2.11). Mesmo o modelo original, representado pela eq. (2.2.4), era um caso especial da eq. (2.2.9). Neste momento, já deve estar bem claro para os leitores que o modelo de Wright-Fisher geral é simplesmente um esquema de amostragem com reposição descrito pela distribuição multinomial da eq. (2.2.8), em que a probabilidade p_k depende única e exclusivamente das diversas forças evolucionárias que se queira estudar.

Exercício: O leitor deve notar que os estados 0 e N da cadeia de Wright-Fisher com seleção são absorventes, como no caso neutro, sem mutação. A

⁷Reveja o conceito biológico na subseção 1.2.3.

perspectiva genealógica do caso neutro poderia ser empregada para o cálculo das probabilidades de fixação no modelo com seleção? Por quê?

Exercício: Construa um modelo de Wright-Fisher que incorpore tanto a mutação quanto a seleção. Determine explicitamente a probabilidade de transição do seu modelo e verifique se ele tem estados absorventes.

2.3 Simulações

Na seção anterior, apresentamos cuidadosamente o modelo de Wright-Fisher, enfatizando sua capacidade de representar alguns aspectos qualitativos de certas populações idealizadas. Porém, deliberadamente evitamos calcular algumas grandezas que seriam claramente relevantes do ponto de vista de aplicações, como o tempo médio até a fixação (quando possível, como no caso neutro), por exemplo. O fato é que mesmo um modelo simples como o de Wright-Fisher pode impor grandes dificuldades a estudos analíticos exatos, e é bastante comum que seja necessário considerar aproximações. Na verdade, as probabilidades de mutação e seleção em contextos biológicos reais são tipicamente muito baixas, o que acaba sendo conveniente para a utilização de métodos perturbativos (como expansões em séries).

Muitas vezes, contudo, é conveniente simular computacionalmente o comportamento de um modelo, seja para ganhar intuição em relação a um particular comportamento de uma população (o que pode ajudar na elaboração de conjecturas que posteriormente levem a resultados rigorosos), seja para efetivamente calcular propriedades com uma base em uma metodologia estatística de simulação. Neste texto, não é possível discutir detalhadamente os conceitos da simulação estocástica, mesmo em nível básico. A intenção do autor, principalmente, é estimular os leitores a buscarem referências como [4] ou [18], caso se interessem pelo tema. Mesmo assim, durante o minicurso, um programa (que, à época, estará disponível no sítio mencionado no Prefácio) será descrito com detalhes à audiência.

No entanto, vamos comentar as idéias essenciais que subjazem a simulação de uma cadeia de Markov de tempo discreto. Antes de mais nada, é preciso obtermos um gerador de números aleatórios. Porém, esse termo, por si só, já deve deixar curioso quem não conhece o assunto. O que seria um número aleatório? Bem, antes de tudo, é preciso ter em mente que sempre desejamos, em uma dada aplicação, gerar números aleatórios segundo uma certa distribuição de probabilidade (discreta ou contínua) bem definida. Isto posto, podemos dizer que um número aleatório é simplesmente uma realização qualquer da variável aleatória cuja distribuição queremos amostrar. Como podemos, então, obter tal número em um computador?

Bem, o fato é que não precisamos obter números realmente aleatórios, mas apenas alguns que pareçam sê-los. Há diversas técnicas para obter esses **números pseudo-aleatórios** e as mais comuns são simples recorrências, perfeitamente determinísticas. Essas técnicas geram amostras de números que parecem ter sido originadas de distribuições específicas, mas transformações adequadas permitem que se obtenha (com maior ou menor esforço, tanto conceitual quanto computacional) virtualmente qualquer distribuição de probabilidade que se queira. Imaginemos que temos à nossa disposição um gerador da variável aleatória contínua uniformemente distribuída no intervalo $(0, 1)$ (até programas usuais como as modernas planilhas eletrônicas dispõem dessa funcionalidade). Geometricamente, podemos dizer que, cada vez que é utilizado, esse gerador “escolhe” uniformemente um ponto em um segmento de reta de comprimento unitário, independentemente dos resultados que possa ter gerado anteriormente.

No momento, nosso interesse é simular distribuições discretas de probabilidade. Portanto, podemos visualizar um segmento unitário como a união de vários outros segmentos, cada um representando um dos possíveis valores que a variável aleatória discreta em questão pode realizar e com comprimento igual à sua probabilidade de ocorrência. A probabilidade do nosso gerador de números aleatórios escolher um ponto que pertença a um específico segmento elementar é igual ao comprimento desse segmento, que é a probabilidade daquele particular evento ocorrer. Por exemplo, se considerarmos uma população de Wright-Fisher usual de tamanho $N = 10$, que em alguma particular geração tenha 3 alelos do tipo 1 e 7 do tipo 2, a probabilidade de escolhermos um organismo do tipo 2 para reproduzir-se é $7/10 = 0,7$. Vemos, na figura 2.3, que podemos simular essa escolha simplesmente sorteando uniformemente um ponto no segmento ilustrado naquela figura e que, com probabilidade 0,7, essa escolha recairá sobre o segmento correspondente à classe dos alelos do tipo 2.

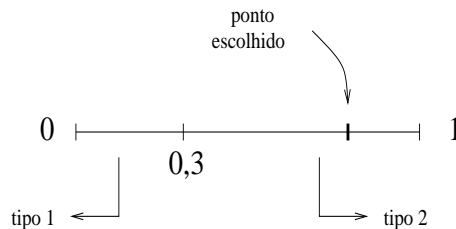


Figura 2.3: A probabilidade de um alelo do tipo 2 ser sorteado para reproduzir-se em uma população de Wright-Fisher com 3 alelos do tipo 1 e 7 do tipo 2 pode ser obtida pela escolha de um ponto no segmento $(0, 1)$.

Em princípio, a dinâmica de Wright-Fisher poderia ser implementada mediante a realização de tantos sorteios quantos fossem necessários para determinar toda a próxima geração. Entretanto, embora correto, tal algoritmo seria extremamente ineficiente do ponto de vista computacional, pois seriam necessários N sorteios para determinar uma nova geração. É possível realizar essa simulação de forma muito mais eficiente, com a geração de um único número aleatório. Afinal, o estado da cadeia de Markov pertinente é caracterizado por uma única grandeza, que é a quantidade de alelos do tipo 1. Embora o modelo de Wright-Fisher tenha sido descrito em termos de escolhas seqüenciais dos membros de cada geração (até por razões pedagógicas), a ordem com que os diferentes indivíduos ingressam na população é irrelevante, o que é evidente pelo termo $\binom{N}{j}$ na eq. (2.2.4).

Assim, dado que $X_t = i$, podemos imaginar o segmento unitário $(0, 1)$ dividido em $N + 1$ segmentos elementares, cada um correspondendo a um elemento j de \mathcal{S} . O j -ésimo segmento elementar tem comprimento $p_{i,j}$ dado pela eq. (2.2.4) e um único número r uniformemente distribuído no intervalo $(0, 1)$ determina X_{t+1} , pois, se r cair no j -ésimo segmento elementar, o que ocorre com probabilidade $p_{i,j}$, X_{t+1} realizar-se-á como j . A figura 2.4 ilustra algumas trajetórias dessa cadeia de Wright-Fisher com $N = 10$ e $X_0 = 5$, sem seleção ou mutação. Entretanto, embora a visualização de um processo estocástico possa ser interessante, é importante reafirmarmos que a verdadeira utilidade científica das simulações é a determinação de grandezas analiticamente desconhecidas mediante o emprego sistemático de metodologias estatísticas para análise dos dados.

2.4 O Modelo de Moran

Um dos fatores que torna difícil a obtenção de resultados analíticos exatos no modelo de Wright-Fisher é o fato dele permitir transições entre quaisquer elementos do espaço de estados em apenas um passo temporal. Além disso, esse clássico modelo só pode oferecer uma idéia aproximada do que seria a dinâmica detalhada de alguma população real, pois, por construção, ele não permite a convivência de organismos com seus descendentes, ou seja, não admite superposição de gerações.

Nesse sentido, o **modelo de Moran** [12] pode ser uma alternativa interessante ao modelo de Wright-Fisher em algumas situações, pois ele é caracterizado pela superposição de gerações, permite a obtenção de resultados analíticos em alguns casos e pode ser implementado tanto em tempo discreto quanto em tempo contínuo. A dinâmica é simples: um membro

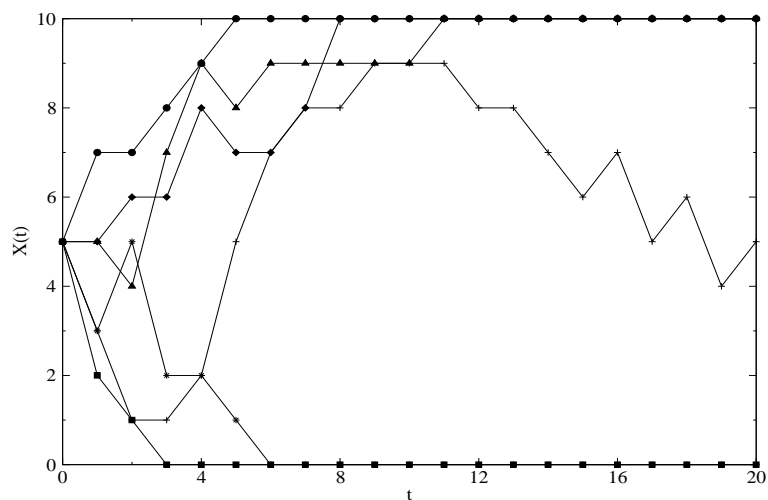


Figura 2.4: Algumas trajetórias de uma cadeia de Markov de uma população de Wright-Fisher com 2 alelos de tamanho $N = 10$ sem seleção ou mutação. As linhas cheias são apenas para guiar os olhares, o tempo é discreto.

da população é escolhido para deixar um descendente, que toma o lugar de um organismo já existente (pode ser o próprio ancestral). Esta eliminação é sempre uniforme, enquanto a escolha de quem se reproduzirá pode incorporar efeitos de seleção natural mediante a adoção de pesos probabilísticos, exatamente como no modelo de Wright-Fisher. Além disso, mutações podem ser facilmente incorporadas.

Como um exemplo concreto, determinemos a cadeia de Markov que representa uma população de tamanho constante N que evolui em tempo discreto segundo o modelo de Moran sem seleção ou mutação. Seus possíveis integrantes são alelos de dois tipos, 1 e 2. Como na apresentação do modelo de Wright-Fisher, vamos caracterizar a população pelo número de alelos do tipo 1. Se $X_t = i$, há i alelos do tipo 1 no instante t . Porém, ao invés de imaginarmos N sorteios para atualizar a população, consideraremos que a cadeia de Moran só pode

- i transicionar para o estado $i - 1$, caso um alelo do tipo 2 seja sorteado para deixar um descendente, que tomaria o lugar de um alelo do tipo 1, ou
- ii transicionar para o estado $i + 1$, caso um alelo do tipo 1 seja sorteado para deixar um descendente, que tomaria o lugar de um alelo do tipo

2, ou

- iii permanecer no estado i , caso um alelo de algum tipo seja sorteado para deixar um descendente, que tomaria o lugar de um alelo do mesmo tipo.

Portanto, estamos falando de uma cadeia de Markov de tempo discreto cuja matriz de transição é tal que $p_{i,j} = 0$ se $|i-j| \geq 2$ e, se $i = 1, \dots, N-1$,

$$p_{i,i+1} = \left(\frac{i}{N}\right) \left(\frac{N-i}{N}\right) = \left(\frac{N-i}{N}\right) \left(\frac{i}{N}\right) = p_{i,i-1} \quad (2.4.12)$$

e $p_{i,i} = 1 - p_{i,i+1} - p_{i,i-1}$. É fácil ver que 0 e N são estados absorventes, pois $p_{0,j} = \delta_{0,j}$ e $p_{N,j} = \delta_{N,j}$, onde $\delta_{i,j}$ vale 1, se $i = j$, ou 0, se $i \neq j$. Na eq. (2.4.12), em cada passagem, o primeiro termo entre parênteses é a probabilidade de um certo tipo de alelo ser escolhido para reproduzir-se, enquanto o segundo é a probabilidade de um certo tipo de alelo ser escolhido para ser eliminado.

Como a cadeia de Moran só pode permanecer onde se encontra ou mudar para estados adjacentes, ela é denominada um processo de nascimento e morte em tempo discreto. Essa característica faz com que o cálculo das probabilidades de fixação nos estados absorventes seja possível até mesmo em um contexto mais amplo. Seguindo [17], vamos considerar um caso geral que tem como casos particulares modelos envolvendo qualquer forma de seleção. Em outras palavras, não nos restringiremos às eqs. (2.4.12). Seja $\mathcal{S} = \{0, \dots, N\}$, como antes, e consideremos probabilidades de transição genéricas $p_{i,i+1}$ e $p_{i,i-1}$ para $i = 1, \dots, N-1$. Os estados 0 e N serão considerados absorventes. A probabilidade da cadeia permanecer no mesmo estado após um único passo é $1 - p_{i,i+1} - p_{i,i-1}$, pois vamos admitir que $p_{i,j} = 0$ se $|i-j| \geq 2$.

Seja x_i a probabilidade de uma realização dessa cadeia de Moran generalizada atingir o estado absorvente N dado que $X_0 = i$. É claro que toda a argumentação usada para justificar que o destino de toda população que obedece à dinâmica de Wright-Fisher sem mutação é eventualmente atingir um estado absorvente também se aplica a este modelo e, conseqüentemente, a probabilidade do estado absorvente 0 ser eventualmente atingido, dado que $X_0 = i$, é $1 - x_i$.

É evidente que $x_0 = 0$ e $x_N = 1$. Para $i \in C_{12} = \{1, \dots, N-1\}$, podemos obter uma recorrência para x_i , que é

$$x_i = p_{i,i-1}x_{i-1} + (1 - p_{i,i+1} - p_{i,i-1})x_i + p_{i,i+1}x_{i+1}. \quad (2.4.13)$$

Para demonstrar a validade da eq. (2.4.13), é preciso considerar todas as possibilidades para a primeira transição, usar as propriedades básicas da

probabilidade condicional e perceber que, pela homogeneidade temporal da cadeia de Markov considerada, a probabilidade de absorção em N depende do estado tomado como condição inicial, mas não depende do instante particular em que aquele estado inicial é tido como conhecido. Explicitamente,

$$\begin{aligned}
x_i &= \mathbb{P}(X_t = N \text{ para algum } t > 0 \mid X_0 = i) \\
&= \sum_{j \in \mathcal{S}} \mathbb{P}(X_t = N \text{ para algum } t > 0, X_1 = j \mid X_0 = i) \\
&= \sum_{j \in \mathcal{S}} \mathbb{P}(X_t = N \text{ para algum } t > 0 \mid X_1 = j, X_0 = i) \cdot \mathbb{P}(X_1 = j \mid X_0 = i) \\
&= \sum_{j \in \mathcal{S}} \mathbb{P}(X_t = N \text{ para algum } t > 0 \mid X_1 = j) \cdot p_{i,j} \\
&= \sum_{j=i-1}^{i+1} \mathbb{P}(X_t = N \text{ para algum } t > 0 \mid X_1 = j) \cdot p_{i,j} \\
&= \sum_{j=i-1}^{i+1} x_j \cdot p_{i,j} = p_{i,i-1}x_{i-1} + p_{i,i}x_i + p_{i,i+1}x_{i+1} \\
&= p_{i,i-1}x_{i-1} + (1 - p_{i,i+1} - p_{i,i-1})x_i + p_{i,i+1}x_{i+1},
\end{aligned}$$

onde o somatório se restringe aos termos correspondentes a probabilidades de transição não nulas.

A recorrência da eq. (2.4.13) pode ser reescrita como $p_{i,i+1}(x_{i+1} - x_i) = p_{i,i-1}(x_i - x_{i-1})$, que claramente sugere a introdução das variáveis

$$y_i = x_i - x_{i-1}, \quad (2.4.14)$$

se $i = 1, \dots, N$, e

$$\sigma_i = \frac{p_{i,i-1}}{p_{i,i+1}}, \quad (2.4.15)$$

que origina a equação

$$y_{i+1} = \sigma_i y_i, \quad (2.4.16)$$

e, conseqüentemente, como $y_1 = x_1$, a solução (para $i \geq 2$)

$$y_i = \left(\prod_{k=1}^{i-1} \sigma_k \right) y_1 = \left(\prod_{k=1}^{i-1} \sigma_k \right) x_1. \quad (2.4.17)$$

Como a eq. (2.4.14) mostra que

$$x_i = \sum_{j=1}^i y_j \quad (2.4.18)$$

e que

$$\sum_{j=1}^N y_j = x_N - x_0 = 1, \quad (2.4.19)$$

podemos combinar as eqs. (2.4.17) e (2.4.19) para concluir que

$$\begin{aligned} 1 = x_1 + \sum_{j=2}^N \left(\prod_{k=1}^{j-1} \sigma_k \right) x_1 &= x_1 \left(1 + \sum_{j=2}^N \prod_{k=1}^{j-1} \sigma_k \right) \Rightarrow \\ \Rightarrow x_1 &= \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^j \sigma_k} \end{aligned} \quad (2.4.20)$$

e combinar as eqs. (2.4.17) e (2.4.18) para concluir que

$$x_i = x_1 + \sum_{j=2}^i \left(\prod_{k=1}^{j-1} \sigma_k \right) x_1 = x_1 \left(1 + \sum_{j=1}^{i-1} \prod_{k=1}^j \sigma_k \right). \quad (2.4.21)$$

Finalmente, a combinação das eqs. (2.4.21) e (2.4.20) resulta na expressão explícita para a probabilidade de fixação no estado N condicionada ao fato de que $X_0 = i$,

$$x_i = \frac{1 + \sum_{j=1}^{i-1} \prod_{k=1}^j \sigma_k}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^j \sigma_k}. \quad (2.4.22)$$

Muitas vezes, é importante saber se uma população pura pode ser invadida por um organismo mutante ou migrante. No caso neutro (sem seleção), já vimos que a probabilidade de fixação de um particular indivíduo é $1/N$, onde N é o tamanho da população. Se a probabilidade um único alelo alienígena se fixar for maior do que $1/N$, ele certamente é mais apto do que a variedade pré-existente, o que, contudo, não garante sua fixação. De qualquer forma, é interessante estimarmos essa probabilidade, que denominaremos simplesmente como probabilidade de fixação (sem qualificação adicional em relação ao estado de partida). Se imaginarmos que, em uma população de N alelos do tipo 2 ocorre uma mutação e surge um alelo do tipo 1, a probabilidade de fixação do tipo 1 é

$$\rho = x_1 = \frac{1}{1 + \sum_{j=1}^{N-1} \prod_{k=1}^j \sigma_k}. \quad (2.4.23)$$

Esse resultado está relacionado a uma aplicação das cadeias de Markov de tempo discreto a um jogo em uma população finita que será discutido no capítulo 4.

Exercício: Refaça cuidadosamente todos os cálculos desta seção.

Exercício: Mostre que a probabilidade de fixação de um alelo do tipo 1 é

$$\rho = \frac{1 - 1/s}{1 - 1/s^N},$$

se ele tem valor adaptativo $s \neq 1$ e os alelos do tipo 2 tem valor adaptativo 1. *Dica:* determine a matriz de transição da cadeia de Markov e depois utilize a eq. (2.4.23).

Capítulo 3

Cadeias de Markov de Tempo Contínuo e os Processos de Nascimento e Morte

3.1 Introdução

Na seção 2.1, apresentamos os elementos essenciais da teoria das cadeias de Markov de tempo discreto em um contexto geral e, nas seções subsequentes, discutimos aplicações daqueles conceitos a modelos simples oriundos da Genética de Populações. Porém, o estudo da teoria geral dos processos em tempo contínuo requer o conhecimento de algumas ferramentas matemáticas (como a exponencial de uma matriz) que, embora simples, provavelmente são desconhecidas do público-alvo deste texto.

Na verdade, mesmo os textos que admitem que todos os requisitos sejam dominados pelos leitores adotam inicialmente um ponto de vista heurístico [1], [21], tentando fazer com que os estudantes adquiram alguma familiaridade com o tema antes de considerar a teoria geral. Neste texto, nós nos restringiremos à introdução heurística às cadeias de Markov de tempo contínuo, tentando ser rigorosos sempre que essa postura não comprometa a clareza que se espera de uma primeira exposição ao tema.

Para podermos usar uma linguagem precisa, precisamos introduzir um conceito que é recorrente em Matemática, embora provavelmente ainda desconhecido de parte do público-alvo deste texto.

3.1.1 Um pouco de Cálculo

Sempre que se deseja trabalhar com aproximações, é importante saber comparar precisamente os comportamentos de duas funções em certos limites. Dada uma função $g : \mathbb{R} \rightarrow \mathbb{R}$, podemos caracterizar toda uma família de funções $f : \mathbb{R} \rightarrow \mathbb{R}$ pela propriedade de que, para cada uma delas e para um certo $t_0 \in \mathbb{R}$, existe uma constante $K \in \mathbb{R}$ tal que $\lim_{t \rightarrow t_0} \frac{f(t)}{g(t)} = K$. Às vezes, esse conjunto de funções f é imprecisamente descrito como sendo “de ordem $g(t)$ ” quando $t \rightarrow t_0$, fato usualmente expresso como $f(t) = \mathcal{O}(g(t))$, ficando t_0 subentendido nesta notação de ‘ \mathcal{O} ’ grande. Trata-se de um abuso de notação, pois a idéia é que $\mathcal{O}(g(t))$ representa um conjunto e seria mais preciso escrevermos $f \in \mathcal{O}(g)$.

Por exemplo, quando $f : \mathbb{R} \rightarrow \mathbb{R}$ é uma função diferenciável (tantas vezes quanto se queira), ela pode ser expressa como um polinômio de Taylor de ordem arbitrária em torno de um ponto t_0 do seu domínio. Se nos restringirmos à primeira ordem,

$$f(t) = f(t_0) + f'(t_0)(t - t_0) + R_2(t),$$

em que $R_2(t)$ é uma função “resto” específica. Alternativamente, poderíamos escrever

$$f(t) = f(t_0) + f'(t_0)(t - t_0) + \mathcal{O}((t - t_0)^2) \quad (3.1.1)$$

para denotar que $f(t) - f(t_0) - f'(t_0)(t - t_0) \in \mathcal{O}((t - t_0)^2)$ e, sem nos preocuparmos com a forma exata do resto, caracterizar ¹ f nas vizinhanças de t_0 , pois $K = f''(t_0)/2$ é [6] tal que $\lim_{t \rightarrow t_0} \frac{R_2(t)}{(t - t_0)^2} = K$.

Se definirmos $\Delta t = t - t_0$, podemos reescrever a eq. (3.1.1) na forma mais concisa

$$f(t) = f(t_0) + f'(t_0)\Delta t + \mathcal{O}((\Delta t)^2), \quad (3.1.2)$$

em que

$$\lim_{\Delta t \rightarrow 0} \frac{\mathcal{O}((\Delta t)^2)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{\mathcal{O}((\Delta t)^2)}{(\Delta t)^2} \Delta t = K \cdot 0 = 0. \quad (3.1.3)$$

¹As propriedades da derivada de f em t_0 , por exemplo, dependem do comportamento do resto $R_2(t)$ em comparação com $t - t_0$, mas não da sua forma exata.

3.2 Taxas e o Processo de Poisson

3.2.1 Relembrando a dinâmica populacional determinística

Os leitores certamente estão familiarizados com as noções elementares de dinâmica populacional no contexto das equações diferenciais ordinárias. Especificamente, embora o tamanho de uma população seja uma grandeza discreta, imagina-se que ela possa ser descrita por uma variável contínua (e diferenciável!) caso o número de indivíduos na população seja suficientemente grande para que flutuações sejam desprezadas.

O modelo mais simples estudado nesse contexto é o de crescimento exponencial, em que uma população (como bactérias na ausência de competição, em ambientes não saturados) é caracterizada por uma taxa de crescimento λ que determina a relação de proporcionalidade entre o tamanho da população $N(t)$ em um instante t qualquer e sua taxa de variação $\frac{dN}{dt}(t)$. Explicitamente, essa dinâmica populacional é regida pela equação diferencial

$$\frac{dN}{dt}(t) = \lambda N(t), \quad (3.2.4)$$

que pode, no entanto, ser vista como decorrente de um raciocínio um pouco mais detalhado, envolvendo o conceito de aproximação discutido na seção anterior. Esperamos que esse raciocínio torne o conceito de taxa de variação ainda mais claro para os leitores e, principalmente, torne mais fácil a aprendizagem do conceito de taxa de probabilidade de um processo estocástico em tempo contínuo que estudaremos em breve.

A idéia que subjaz o modelo de crescimento exponencial é a de que cada membro da população reproduz-se continuamente no tempo e independentemente dos demais indivíduos com taxa λ , que deve ser interpretada como o número de descendentes a que cada indivíduo da população dá origem por unidade de tempo. O que isso quer dizer exatamente? A única interpretação precisa é que o número de descendentes de um organismo em um intervalo de tempo Δt é

$$\lambda \Delta t + \mathcal{O}((\Delta t)^2),$$

lembrando que a notação de ‘ \mathcal{O} ’ grande só faz sentido no contexto de um limite, que, no caso, ocorre quando $\Delta t \rightarrow 0$. Portanto, a expressão acima só faz sentido para Δt “pequeno”, no sentido da subseção 3.1.1.

Assim, no instante t , vamos concentrar nossas atenções em um único elemento da população que indexaremos por i , embora todos partilhem das

mesmas propriedades. Se os descendentes desse indivíduo não se reproduziram durante um pequeno intervalo de tempo Δt , o tamanho daquela linhagem ao final deste período seria

$$N_i^{(1)}(t + \Delta t) = 1 + \lambda\Delta t + \mathcal{O}((\Delta t)^2), \quad (3.2.5)$$

onde o fator 1 na soma refere-se ao próprio indivíduo i , já presente em t . O problema (aparente, como veremos) com a eq. (3.2.5) é que ela não é condizente com a noção de reprodução contínua, pois não contempla a possibilidade dos descendentes de i também poderem se reproduzir durante o intervalo Δt . Por exemplo, poderíamos dizer que, ao final do intervalo $\Delta t/2$, já há $1 + \lambda\Delta t/2 + \mathcal{O}((\Delta t)^2)$ organismos na linhagem iniciada por i , e que cada um deles dá origem a mais $1 + \lambda\Delta t/2 + \mathcal{O}((\Delta t)^2)$ indivíduos, resultando em

$$N_i^{(2)}(t + \Delta t) = \left[1 + \frac{\lambda\Delta t}{2} + \mathcal{O}((\Delta t)^2) \right]^2 = 1 + \lambda\Delta t + \mathcal{O}((\Delta t)^2) \quad (3.2.6)$$

membros naquela linhagem em $t + \Delta t$, após a expansão do binômio. Evidentemente, esse raciocínio pode ser efetuado indefinidamente e os leitores já devem ter percebido que $N_i(t + \Delta t)$, o tamanho em $t + \Delta t$ da linhagem iniciada pelo solitário indivíduo i no instante t e continuamente expandida, está relacionado ao tamanho $N_i^{(m)}(t + \Delta t)$ que a linhagem iniciada por i em t alcança em $t + \Delta t$ após m passos por

$$\begin{aligned} N_i(t + \Delta t) &= \lim_{m \rightarrow \infty} N_i^{(m)}(t + \Delta t) \\ &= \lim_{m \rightarrow \infty} \left[1 + \frac{\lambda\Delta t}{m} + \mathcal{O}((\Delta t)^2) \right]^m = 1 + \lambda\Delta t + \mathcal{O}((\Delta t)^2), \end{aligned} \quad (3.2.7)$$

pois toda a dependência em m reside nos termos de ordem maior ou igual à quadrática. A importantíssima mensagem que deve ser extraída da discussão acima é a seguinte: dado que o tempo é considerado contínuo, **não é preciso que nos preocupemos com possíveis subdivisões do intervalo Δt , pois elas só introduzem correções de ordem mais alta, que anular-se-ão quando $\Delta t \rightarrow 0$.**

Finalmente, como as linhagens originadas pelos $N(t)$ membros da população já presentes no instante t são independentes, o tamanho total $N(t + \Delta t)$

da população no instante $t + \Delta t$ é simplesmente a soma

$$\begin{aligned}
 N(t + \Delta t) &= \sum_{i=1}^{N(t)} N_i(t + \Delta t) \\
 &= \sum_{i=1}^{N(t)} [1 + \lambda \Delta t + \mathcal{O}((\Delta t)^2)] \\
 &= N(t) [1 + \lambda \Delta t + \mathcal{O}((\Delta t)^2)] \\
 &= N(t) + \lambda N(t) \Delta t + \mathcal{O}((\Delta t)^2), \tag{3.2.8}
 \end{aligned}$$

que leva à eq. (3.2.4) pois

$$\begin{aligned}
 \frac{dN}{dt}(t) &= \lim_{\Delta t \rightarrow 0} \frac{N(t + \Delta t) - N(t)}{\Delta t} \\
 &= \lim_{\Delta t \rightarrow 0} \left\{ \lambda N(t) + \frac{\mathcal{O}((\Delta t)^2)}{\Delta t} \right\} = \lambda N(t),
 \end{aligned}$$

em que a eq. (3.1.3) foi utilizada.

3.2.2 Taxas probabilísticas e o processo de Poisson

As aplicações da teoria elementar das cadeias de Markov em tempo contínuo à dinâmica populacional estão fundamentadas na hipótese de que todo e qualquer evento aleatório (que pode estar associado a um membro da população, como a geração de descendentes ou a morte daquele indivíduo, ou não, como em migrações) é caracterizado por uma **taxa de probabilidade** $\lambda > 0$, de modo que a probabilidade de um evento ocorrer em um intervalo de tempo Δt é

$$\lambda \Delta t + \mathcal{O}((\Delta t)^2), \tag{3.2.9}$$

enquanto a probabilidade do mesmo evento não ocorrer é

$$1 - \lambda \Delta t + \mathcal{O}((\Delta t)^2). \tag{3.2.10}$$

Muitas interpretações equivocadas deste conceito decorrem de estudantes perderem de vista o fato que essa expressão só faz sentido no contexto de um limite quando $\Delta t \rightarrow 0$. Pensar em algum valor fixo de Δt abre imediatamente a possibilidade de que a taxa seja suficientemente alta para constituir uma absurda “probabilidade” maior que 1.

Para refinarmos e aplicarmos esse conceito, consideremos inicialmente uma população cujos membros não se reproduzem nem morrem (!), constituída única e exclusivamente por um processo migratório em que a chegada

de cada novo elemento ocorre com taxa λ . O espaço de estados, diferentemente do capítulo anterior, é $\mathcal{S} = \{0, 1, \dots\}$, o que naturalmente suscita questões matemáticas que, embora importantes, não serão abordadas neste texto. Seja $p_n(t) \equiv \mathbb{P}[N(t) = n]$ a probabilidade de que a população tenha tamanho n no instante t . Por simplicidade, vamos adotar em um primeiro momento a condição inicial $p_n(0) = \delta_{n,0}$, ou seja, inicialmente a população tem um tamanho bem definido N_0 . Se $n \geq 1$, e considerando todos os possíveis estados no instante t ,

$$\begin{aligned}
p_n(t + \Delta t) &= \mathbb{P}[N(t + \Delta t) = n] \\
&= \sum_{i \in \mathcal{S}} \mathbb{P}[N(t + \Delta t) = n, N(t) = i] \\
&= \sum_{i=0}^n \mathbb{P}[N(t + \Delta t) = n \mid N(t) = i] \cdot \mathbb{P}[N(t) = i] \\
&= [1 - \lambda\Delta t + \mathcal{O}((\Delta t)^2)] \cdot p_n(t) + [\lambda\Delta t + \mathcal{O}((\Delta t)^2)] \cdot p_{n-1}(t) + \\
&\quad + \sum_{i=0}^{n-2} [\lambda\Delta t + \mathcal{O}((\Delta t)^2)]^{n-i} \cdot p_i(t) \\
&= (1 - \lambda\Delta t) \cdot p_n(t) + \lambda\Delta t \cdot p_{n-1}(t) + \mathcal{O}((\Delta t)^2)
\end{aligned} \tag{3.2.11}$$

Da equação acima, decorre que

$$\begin{aligned}
p_n(t + \Delta t) &= (1 - \lambda\Delta t) \cdot p_n(t) + \lambda\Delta t \cdot p_{n-1}(t) + \mathcal{O}((\Delta t)^2) \Rightarrow \\
&\Rightarrow \frac{dp_n}{dt}(t) = \lim_{\Delta t \rightarrow 0} \frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} \\
&= \lim_{\Delta t \rightarrow 0} \left\{ \lambda p_{n-1}(t) - \lambda p_n(t) + \frac{\mathcal{O}((\Delta t)^2)}{\Delta t} \right\} \\
&\quad \therefore \frac{dp_n}{dt}(t) = \lambda p_{n-1}(t) - \lambda p_n(t). \tag{3.2.12}
\end{aligned}$$

O caso $n = 0$ deve ser analisado separadamente. Não é difícil ver que

$$\begin{aligned}
p_0(t + \Delta t) &= \mathbb{P}[N(t + \Delta t) = 0] \\
&= \mathbb{P}[N(t + \Delta t) = 0, N(t) = 0] \\
&= \mathbb{P}[N(t + \Delta t) = 0 \mid N(t) = 0] \cdot \mathbb{P}[N(t) = 0] \\
&= [1 - \lambda\Delta t + \mathcal{O}((\Delta t)^2)] \cdot p_0(t) \\
&= (1 - \lambda\Delta t) \cdot p_0(t) + \mathcal{O}((\Delta t)^2)
\end{aligned} \tag{3.2.13}$$

que leva a

$$\begin{aligned}
 p_0(t + \Delta t) &= (1 - \lambda \Delta t) \cdot p_0(t) + \mathcal{O}((\Delta t)^2) \Rightarrow \\
 \Rightarrow \frac{dp_0}{dt}(t) &= \lim_{\Delta t \rightarrow 0} \frac{p_0(t + \Delta t) - p_0(t)}{\Delta t} \\
 &= \lim_{\Delta t \rightarrow 0} \left\{ -\lambda p_0(t) + \frac{\mathcal{O}((\Delta t)^2)}{\Delta t} \right\} \\
 \therefore \frac{dp_0}{dt}(t) &= -\lambda p_0(t). \tag{3.2.14}
 \end{aligned}$$

Esse processo estocástico é denominado **processo de Poisson** e provavelmente é a cadeia de Markov de tempo contínuo mais simples que se possa imaginar. O leitor atento já deve ter notado que a eq. (3.2.12) não é uma simples equação diferencial. Trata-se de uma **equação diferencial de diferenças**, por envolver não apenas derivadas mas também dois elementos distintos da seqüência $(p_n(t))$. Em geral, uma equação como essa requer o conhecimento de certas técnicas provavelmente desconhecidas da maioria dos leitores. No entanto, a eq. (3.2.14) é uma equação diferencial usual, análoga à de crescimento exponencial, cuja solução, considerando a condição inicial estipulada, $p_n(0) = \delta_{n,0}$, é $p_0(t) = \exp[-\lambda t]$. Esta função pode ser substituída na eq. (3.2.12) quando $n = 1$, o que a reduz a uma equação diferencial ordinária, linear e não homogênea, cuja solução pode ser determinada facilmente. Esse procedimento pode ser iterado e a solução geral é dada por uma distribuição de Poisson de parâmetro λt ,

$$p_n(t) = \frac{\exp[-\lambda t](\lambda t)^n}{n!}. \tag{3.2.15}$$

Exercício: Demonstre, por indução finita, a validade da eq. (3.2.15).

Às vezes, um processo como o de Poisson é denominado um processo de contagem, por caracterizar a distribuição de probabilidade do número de ocorrências de um certo evento até um certo instante. É evidente que, uma vez que a população tenha atingido um certo tamanho, ela não pode diminuir. Além disso, a probabilidade de ocorrência de duas ou mais migrações em um mesmo intervalo temporal tende a zero no limite de tempo contínuo.

É possível mostrar que não apenas esses fatos mas também a estrutura detalhada das eqs. (3.2.12) e (3.2.14) e, em geral, todas as propriedades dos processos de tempo contínuo caracterizados pelas taxas de probabilidade definidas nesta seção, podem ser completamente caracterizados por um diagrama de estados análogo àquele introduzido no capítulo anterior, em que

as pertinentes taxas de probabilidade de transição entre os estados (ao invés das probabilidades de transição, que só estariam definidas em - fixos e arbitrários - intervalos temporais) indexam os arcos, como pode ser visto na figura 3.1. É essencial que os leitores percebam que não há laços (um arco cuja origem e destino são o mesmo nó) em um diagrama de estados em tempo contínuo, em contraposição à figura 2.1.

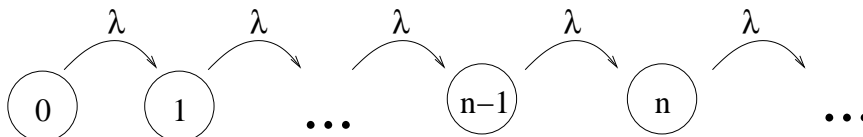


Figura 3.1: Diagrama de estados de um processo de Poisson de taxa λ .

Dessa forma, após termos apresentado rigorosamente os conceitos que subjazem os processos de tempo contínuo mediante a análise detalhada do processo de Poisson, adotaremos de agora em diante um ponto de vista heurístico, que nos permitirá obter por inspeção o sistema de equações diferenciais de diferenças que descreve um processo qualquer. Por razões pedagógicas, vamos fazer uma analogia com um problema bem conhecido do estudo de aplicações de equações diferenciais.

Em um reservatório que inicialmente contém V_0 litros de uma mistura salina homogênea, mas do qual vazam u_- litros por minuto, despeja-se uma solução salina de concentração c_+ gramas por litro com vazão de $u_+ > u_-$ litros por minuto. Ensina-se heurísticamente (a justificativa rigorosa é análoga à dada para o processo de Poisson) que a taxa de variação da massa de sal $m(t)$ nesse sistema é decorrente de um balanço entre uma taxa de entrada e uma taxa de saída de sal do sistema, de modo que

$$\frac{dm}{dt}(t) = c_+ u_+ - \frac{m(t)}{V_0 + (u_+ - u_-)t} u_-.$$

Tanto o termo de entrada quanto o de saída tem a estrutura de um produto de concentração por vazão, apenas a concentração do termo de saída não é constante. Intuitivamente, a concentração representa um “conteúdo de massa” e a vazão, uma “utilização” daquele conteúdo. O seu produto constitui uma taxa ou, alternativamente, um fluxo.

A equação diferencial de diferenças que rege o comportamento de $p_n(t)$ pode ser escrita diretamente, mediante a observação dos arcos ligados ao nó correspondente ao estado n . Arcos incidentes nesse estado contribuem com “vazões positivas de probabilidade”, enquanto os que dele emanam representam “vazões negativas de probabilidade”. Essas vazões são multiplicadas

pelas correspondentes probabilidades da cadeia encontrar-se em certo estado (as “concentrações”) para gerar os “fluxos de probabilidade” pertinentes. O leitor deve observar novamente as equações que descrevem o processo de Poisson e compará-las com o diagrama de estados para certificar-se do significado exato dessas palavras.

Exercício: Construa o diagrama de estados pertinente e determine explicitamente a solução do processo de Poisson de taxa λ com condição inicial $p_n(0) = \delta_{n, N_0}$.

Porém, é importante ressaltar que só podemos obter a solução explícita do processo de Poisson mediante as técnicas usuais de equações diferenciais ordinárias porque ele é um modelo bastante simples. Em geral, o estudo analítico das cadeias de Markov em tempo contínuo requer o uso de uma ferramenta denominada **função geradora**, que obedece uma equação diferencial parcial (EDP), objeto matemático com o qual não se admite familiaridade neste texto. Para ilustrar o conceito, demonstraremos sua utilidade no contexto simples do processo de Poisson, que levará a uma EDP suficientemente simples para não nos desviar demais dos nossos propósitos. Processos mais elaborados, que descreveremos na próxima seção, serão estudados apenas mediante simulações computacionais, como veremos no fim deste capítulo.

3.2.3 Funções geradoras

De forma análoga à transformada de Laplace, que é um objeto matemático associado biunivocamente a certas funções com domínios contínuos, uma função geradora é associada a uma seqüência (ou seja, uma função com domínio discreto). Explicitamente, a função geradora $g(z)$ associada à seqüência (a_n) , $n = 0, 1, \dots$, é $g(z) = \sum_{n=0}^{\infty} a_n z^n$ e questões de convergência não são tão importantes quanto podem parecer, pois a função geradora pode ser expandida como uma série formal de potências, em que cada coeficiente de potência corresponde a um termo da seqüência associada sem que se realizem somas que poderiam divergir. No presente contexto, sobretudo, a função geradora sempre converge se $|z| < 1$.

Como estamos trabalhando com seqüências cujos termos dependem do parâmetro t , o objeto com que devemos trabalhar é uma função geradora com dois parâmetros,

$$g(z, t) = \sum_{n=0}^{\infty} p_n(t) z^n. \quad (3.2.16)$$

Mediante manipulações análogas àquelas que surgem na solução de equações diferenciais por séries infinitas, os dois lados da eq. (3.2.12) podem ser

multiplicados por z^n e somados para todos os valores de n , levando a

$$\frac{\partial g}{\partial t}(z, t) = \lambda(z - 1)g(z, t), \quad (3.2.17)$$

admitindo-se que as operações de diferenciação em relação a t e de soma infinita possam ser invertidas. A condição inicial $p_n(0) = \delta_{n,0}$ faz com que $g(z, 0) = 1$ e a eq. (3.2.17) pode ser integrada diretamente para determinar a função geradora

$$g(z, t) = \exp[\lambda t(z - 1)], \quad (3.2.18)$$

cuja expansão em série de potências de z leva à eq. (3.2.15).

3.3 Nascimentos e Mortes

O processo de Poisson é um membro de uma classe geral de processos estocásticos denominados **processos de nascimento**. Estes processos caracterizam-se pelo fato das mudanças de estado só ocorrerem no sentido dos estados de índices mais altos (crescimento das populações, na interpretação biológica) e apenas entre estados adjacentes, $k \rightarrow k + 1$. Porém, cada estado n da cadeia de Markov pode dispor de sua própria taxa de nascimento, λ_n . O diagrama de estados de um processo de nascimento genérico está representado na figura 3.2 e obedece às equações

$$\begin{aligned} \frac{dp_n}{dt}(t) &= \lambda_{n-1}p_{n-1}(t) - \lambda_n p_n(t), \quad n \geq 1, \\ \frac{dp_0}{dt}(t) &= -\lambda_0 p_0(t). \end{aligned} \quad (3.3.19)$$

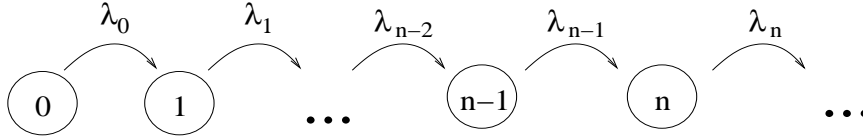


Figura 3.2: Diagrama de estados de um processo de nascimento genérico.

Um processo de nascimento, em particular, pode ser considerado o análogo estocástico do modelo determinístico de crescimento exponencial. Trata-se do processo de nascimento linear representado no diagrama da figura 3.3, onde é evidente que se considera que cada membro da população reproduz-se independentemente dos demais, cada um com taxa λ .

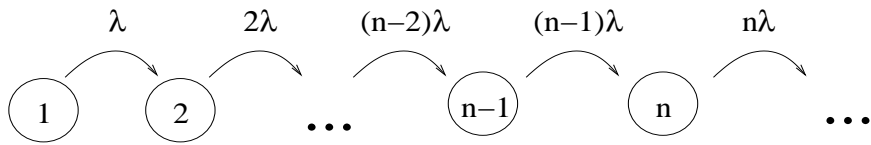


Figura 3.3: Diagrama de estados de um processo de nascimento linear com taxa λ .

É claro que um modelo populacional mais próximo da realidade de qualquer população real deve considerar a possibilidade de mortes e é natural que se considere toda uma família de **processos de nascimento e morte**, como ilustrado na figura 3.4. Os mesmos argumentos que nos levaram à conclusão que dois ou mais eventos da mesma natureza nunca serão observados em um mesmo intervalo de tempo, suficientemente pequeno, aplicam-se também a combinações de eventos distintos, de modo que apenas transições entre estados adjacentes são admissíveis. A hipótese de independência entre os elementos da população leva naturalmente a um processo linear como aquele representado na figura 3.5.

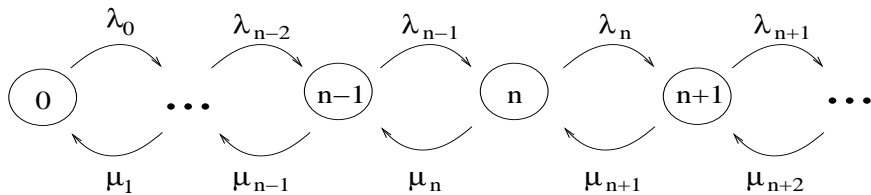


Figura 3.4: Diagrama de estados de um processo de nascimento e morte genérico.

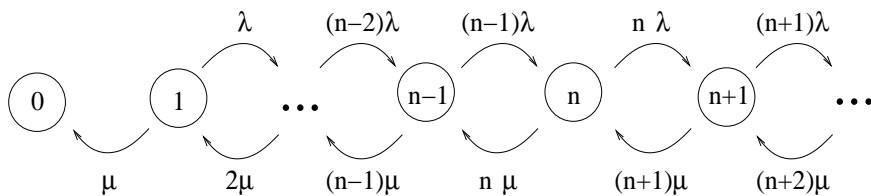


Figura 3.5: Diagrama de estados de um processo de nascimento e morte linear, com taxa de nascimento λ e taxa de morte μ .

Muitas vezes, é possível determinar analiticamente algumas características dos processos de nascimento e morte, como a média e a variância do

tamanho da população em cada instante t . Os processos lineares são particularmente tratáveis. Porém, embora importante, a determinação dessas propriedades demandaria muitos cálculos e não seria muito instrutiva neste momento. Pelo caráter introdutório deste minicurso, agora parece mais conveniente introduzir aos leitores uma técnica de simulação computacional de processos estocásticos de tempo contínuo que é amplamente utilizada em pesquisa, em diversas áreas.

3.4 Algoritmo de Gillespie

Pela definição de uma taxa de probabilidade, o algoritmo mais direto para simular um processo estocástico em tempo contínuo estaria baseado na definição de um pequeno intervalo temporal Δt , “suficientemente pequeno”, e no teste da ocorrência de cada tipo de evento, considerando suas particulares taxas, nesse período. Esses testes seriam efetuados pelo sorteio de números pseudo-aleatórios, no sentido descrito no capítulo anterior, que seriam comparados com o produto de Δt por cada taxa envolvida.

Entretanto, tal algoritmo seria impreciso e extremamente ineficiente. Impreciso porque a escolha de valores fixos para Δt poderia levar à indesejável ocorrência de múltiplos eventos nessa janela temporal. De fato, a definição de taxa de probabilidade só faz sentido como um limite, e a múltipla ocorrência sempre seria possível, embora sua probabilidade possa ser reduzida mediante a escolha de valores bem pequenos (em relação às taxas) de Δt . Mas isso tornaria ineficiente o algoritmo, pois seriam necessárias mais iterações para simular o processo em um dado período e, na maior parte dos intervalos testados, não ocorreria qualquer evento. Em termos práticos, vários números pseudo-aleatórios seriam desperdiçados, no sentido em que não resultariam em transições da cadeia de Markov.

Uma alternativa eficiente foi desenvolvida por D. T. Gillespie [5]. Esse algoritmo será descrito segundo o tratamento de [19], uma excelente introdução às aplicações de processos estocásticos e simulações à questão da Origem da Vida, além de outros temas em Biologia Teórica. A idéia básica do algoritmo de Gillespie consiste em usar apenas 2 números pseudo-aleatórios para descrever cada evento. O primeiro determina quando deve ocorrer o próximo evento, enquanto o segundo indica qual o tipo de evento que será realizado, escolhido entre M processos componentes, com taxas λ_i , $i = 1, \dots, M$. É conveniente definir $\lambda_0 \equiv \sum_{i=1}^M \lambda_i$.

Inicialmente, determina-se a probabilidade de não ocorrer qualquer evento

ao longo de um intervalo temporal t , $P_0(t)$, mediante a equação

$$\begin{aligned} P_0(t + \Delta t) &= P_0(t) \cdot P_0(\Delta t) = P_0(t) \prod_{i=1}^M [1 - \lambda_i \Delta t + \mathcal{O}((\Delta t)^2)] \\ &= P_0(t) (1 - \lambda_0 \Delta t) + \mathcal{O}((\Delta t)^2) \Rightarrow \frac{dP_0(t)}{dt} = -\lambda_0 P_0(t) \\ &\quad \therefore P_0(t) = e^{-\lambda_0 t}. \end{aligned} \quad (3.4.20)$$

Alguns fatos implícitos na última passagem do cálculo acima são a independência temporal de λ_0 entre eventos (pois, por exemplo, algum dos λ_i poderia ser uma taxa total de nascimentos, dependente do tamanho da população) e a condição inicial $P_0(0) = 1$.

Assim, se $P(t)$ for a densidade de probabilidade do intervalo entre 2 eventos consecutivos quaisquer ser t , então

$$\begin{aligned} P(t) \Delta t &= P_0(t) [1 - P_0(\Delta t)] = e^{-\lambda_0 t} [1 - e^{-\lambda_0 \Delta t}] \\ &= e^{-\lambda_0 t} [\lambda_0 \Delta t + \mathcal{O}((\Delta t)^2)] \\ &\quad \therefore P(t) = \lambda_0 e^{-\lambda_0 t}. \end{aligned} \quad (3.4.21)$$

Por outro lado, a densidade de probabilidade conjunta $P(i, t)$ de que um evento demore um tempo t após o último evento para ocorrer e seja do tipo i é tal que

$$\begin{aligned} P(i, t) \Delta t &= P_0(t) \lambda_i \Delta t \\ \therefore P(i, t) &= \lambda_i e^{-\lambda_0 t} \end{aligned} \quad (3.4.22)$$

de modo que a probabilidade condicional de que um evento seja do tipo i , dado que ele ocorre no instante t , $P(i|t)$, é

$$P(i|t) = \frac{P(i, t)}{P(t)} = \frac{\lambda_i}{\lambda_0}. \quad (3.4.23)$$

O primeiro passo do algoritmo de Gillespie consiste simplesmente em gerar um número aleatório com a distribuição exponencial (3.4.21), que indica o intervalo entre o último evento já realizado e o próximo evento a ocorrer. O segundo passo determina a natureza do evento mediante o sorteio de um outro número aleatório, r , uniformemente distribuído no intervalo $(0, \lambda_0)$. Geometricamente, esse intervalo é constituído pelos N segmentos λ_i , devidamente ordenados. Se r cair no segmento i , o que ocorre com probabilidade dada por (3.4.23), o evento é do tipo i .

Veremos agora os resultados de algumas simulações realizadas segundo o algoritmo de Gillespie de um processo linear de imigração e morte (taxas

α e μ , respectivamente), descrito pelas equações

$$\begin{aligned} \frac{dp_n}{dt}(t) &= \alpha p_{n-1}(t) + (n+1)\mu p_{n+1}(t) - (\alpha + n\mu) p_n(t), \quad n \geq 1, \\ \frac{dp_0}{dt}(t) &= \mu p_1(t) - \alpha p_0(t) \end{aligned} \quad (3.4.24)$$

e de um processo linear de nascimento, imigração e morte (taxas λ , α e μ , respectivamente), regido pelas equações

$$\begin{aligned} \frac{dp_n}{dt}(t) &= [\alpha + (n-1)\lambda] p_{n-1}(t) + (n+1)\mu p_{n+1}(t) \\ &\quad - [\alpha + n(\lambda + \mu)] p_n(t), \quad n \geq 1, \\ \frac{dp_0}{dt}(t) &= \mu p_1(t) - \alpha p_0(t). \end{aligned} \quad (3.4.25)$$

Exe
imigi

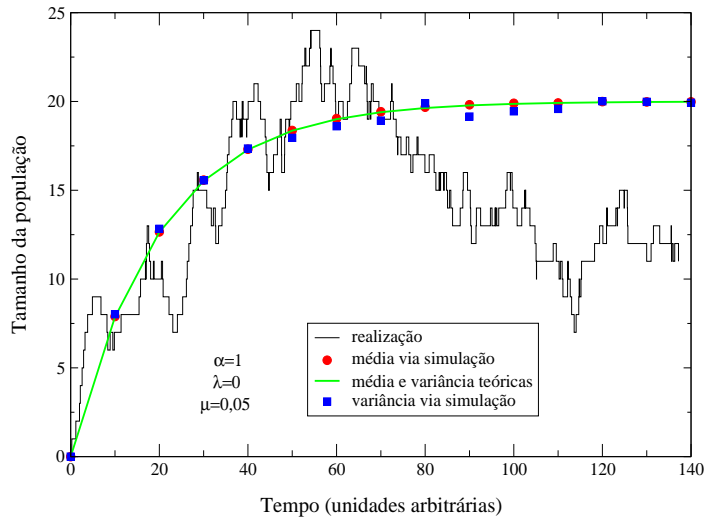


Figura 3.6: Aplicação do algoritmo de Gillespie para simular um processo de imigração e morte. A população inicial é nula e foram geradas 10^4 realizações do processo (uma das quais é exibida) para a determinação da média e da variância.

As figuras abaixo mostram curvas obtidas a partir de resultados analíticos que não foram apresentados neste texto, mas que podem ser encontrados

em [1], por exemplo. A figura 3.6 mostra uma realização do processo de imigração e morte oscilando em torno do valor médio previsto analiticamente. Curiosamente, a média e a variância deste processo são idênticas em qualquer instante t . Mas esse resultado não é válido, em geral, para o processo de imigração e morte. Ele decorre do fato da população inicial ser nula, o que faz com que o tamanho da população sempre seja descrito por uma distribuição de Poisson. A figura 3.7 ilustra algumas propriedades de um processo de imigração, nascimento e morte, e é digna de nota a concordância entre

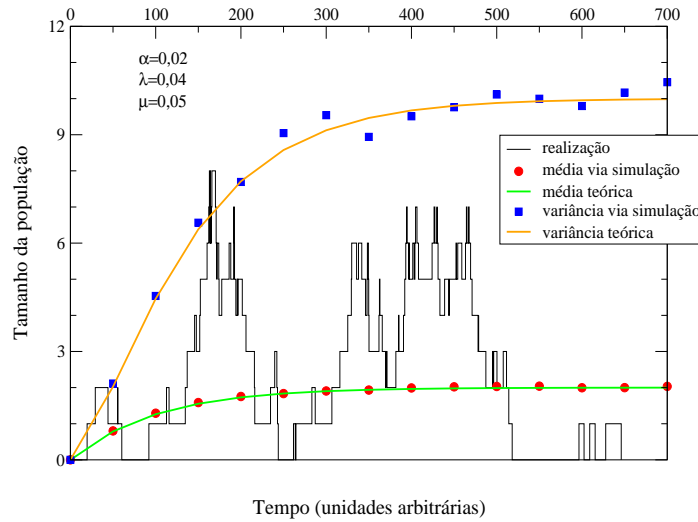


Figura 3.7: Aplicação do algoritmo de Gillespie a um processo de imigração, nascimento e morte. A população inicial é nula e foram geradas 10^4 realizações do processo (uma das quais é exibida) para a determinação da média e da variância.

Capítulo 4

Aplicações: Evolução do Câncer e Jogos em Populações Finitas

Agora, veremos como as técnicas que já aprendemos, apesar de simples, podem ser utilizados na construção de modelos de sistemas bem complexos. Em particular, estudaremos a construção de um modelo matemático para reproduzir os aspectos básicos do desenvolvimento de um câncer e vamos descrever o comportamento de uma população finita composta por dois tipos de indivíduos (os “jogadores”) cuja adaptabilidade depende da composição da própria população. Este capítulo segue fielmente alguns trechos do livro [17], onde os leitores podem encontrar também maiores detalhes sobre os modelos aqui discutidos.

4.1 Um Modelo Matemático para o Câncer

4.1.1 Mais um pouco de Biologia

Cada célula de um ser vivo contém um aparato completo de estruturas biológicas que lhe confere a capacidade de auto-reprodução. No decurso da história da Terra, certamente as primeiras células que surgiram precisavam reproduzir-se eficientemente para sobreviverem em um ambiente já dominado por organismos não-celulares. Portanto, essas protocélulas teriam maiores chances de sobrevivência se apresentassem taxas de replicação relativamente altas. No entanto, em algum outro momento da nossa

história evolutiva, surgiram organismos multicelulares, em que diferentes células precisaram cooperar no sentido em que uma competição reprodutiva desenfreada entre as próprias unidades constituintes desses organismos inviabilizaria sua sobrevivência ¹.

Atualmente, sabe-se que existe um complexo sistema de regulação gênica que controla toda a atividade celular. Podemos dizer que a identidade de uma célula é determinada pela expressão dos genes nela contidos ², que depende do ambiente em que ela se encontra. Em particular, os genes relacionados à divisão celular devem estar sob estrito controle em células funcionais, de modo que a replicação ocorra somente quando isso for condizente com o processo normal de desenvolvimento do tecido em questão.

De fato, cada célula recebe continuamente sinais químicos das demais, que indicam se os processos bioquímicos estão ocorrendo de forma normal no tecido. Em geral, quando há sinalização de algum erro sério, desenvolve-se um processo de morte celular programada denominado apoptose. Em outras palavras, trata-se de um suicídio “pelo bem do grupo”, para impedir a proliferação de células defeituosas. Estas podem surgir mediante a ocorrência de mutações que comprometam a estabilidade do mecanismo de cooperação celular, que podem levar ao surgimento de variedades celulares que se reproduzem de forma excessivamente rápida. Muitas vezes, esses “egoístas livres” são exterminados pela “inanição” decorrente da apoptose dos seus vizinhos, mas alguns deles escapam desse mecanismo e dão origem à proliferação desenfreada de novas células tão inconvenientes para o tecido quanto elas próprias. Esse aglomerado de células mutantes pode transformar-se em um tumor se atingir um tamanho suficientemente grande e o processo patológico que se segue é o que se denomina câncer.

¹ Não se sabe como ocorreu o surgimento das células primitivas, muito menos a origem dos organismos multicelulares. Porém, essa lacuna em nosso conhecimento não é apenas de natureza biológica. É possível avaliar a viabilidade do surgimento de um organismo inédito em um ambiente mediante abstrações que desconsideram completamente os detalhes bioquímicos potencialmente envolvidos nesses processos. Mesmo assim, não se conhecem propostas matemáticas para a dinâmica dessas populações de organismos que sustentem a hipótese do seu surgimento. Dessa forma, a Modelagem Matemática pode desempenhar um papel importante na resolução dessas importantes questões científicas ainda em aberto.

² As notórias células-tronco podem ser vistas como células cujos genes estão em um estado “mais facilmente programável”, prontas para desempenhar qualquer função que lhes seja “atribuída” pelo tecido em que se encontram, inclusive substituir células “danificadas” - daí as especulações sobre seu possível uso terapêutico. Enquanto as células-tronco embrionárias seriam naturalmente boas candidatas para essas aplicações, muitos pesquisadores afirmam que as células-tronco adultas, presentes em alguns tecidos, e as células-tronco obtidas por uma “reprogramação” de células já diferenciadas teriam certas limitações, embora ainda haja muita controvérsia científica (além da religiosa...) sobre esse tema.

Empiricamente, a idéia de que o câncer seria alguma forma de doença genética começou a tomar forma no fim do século XIX, quando David van Hansemann observou que células cancerosas apresentavam evidências de processos anormais de divisão celular. Posteriormente, descobriu-se que a maioria das células cancerosas apresenta um número anormal de cromossomos (aneuploidia). Herman Muller, em 1927, descobriu que radiações ionizantes, já reconhecidamente carcinogênicas, eram também mutagênicas, o que lhe rendeu o Prêmio Nobel de Medicina de 1946. Muitos anos mais tarde, em 1951, Muller associou diretamente esses fenômenos quando propôs que uma célula deveria tornar-se cancerosa após sofrer múltiplas mutações.

Hoje, sabemos que existem pelo menos 3 mecanismos básicos que podem levar ao surgimento do câncer. O primeiro, complicadíssimo, corresponde na verdade a todo um conjunto de possíveis mutações em alguns das centenas de genes que regulam o processo de divisão celular e é denominado **instabilidade cromossômica**. O segundo é a ativação de certos genes que aumentam a taxa de proliferação celular se um alelo sofre mutação ou é expresso incorretamente. Esse conceito de **oncogenes** foi proposto em 1976 por J. Michael Bishop e Harold E. Varmus e os levou ao Prêmio Nobel de Medicina de 1989. Finalmente, em 1971, Alfred Knudson propôs a existência dos posteriormente descobertos **genes supressores de tumor** (GST's), ao estudar os padrões de incidência etária do retinoblastoma, um tipo de câncer infantil que leva ao surgimento de tumores nos olhos das crianças.

Knudson descobriu que o número de casos aumenta de forma quadrática com a idade entre as crianças que apresentam um único tumor, mas linearmente (e mais rapidamente) entre aquelas com múltipla incidência de tumores em ambos os olhos. Ele propôs que os portadores de múltiplos tumores teriam herdado um alelo inativo de um dos seus pais, mas que uma célula só se tornaria cancerosa após a ocorrência de uma mutação que desativasse o segundo alelo. Por outro lado, o crescimento mais lento do número de casos na outra subpopulação seria explicado pela necessidade da ocorrência de duas mutações. De fato, já se conhecem diversos GST's e eles apresentam a propriedade de que a primeira mutação é neutra (ou quase), enquanto a segunda realmente altera substancialmente o fenótipo celular, aumentando sua taxa reprodutiva.

Neste texto, vamos apenas construir um processo Markoviano para descrever a evolução de um conjunto de células portadoras de GST's. Porém, seu estudo analítico não é simples, e não nos preocuparemos em estudar suas conseqüências.

4.1.2 Um modelo estocástico simples

Vamos considerar uma população de N células, cada uma das quais pertencente a uma entre 3 categorias. As células do tipo 0 são aquelas que apresentam os dois alelos de um gene supressor de tumor ativos (não cancerosas, portanto). Aquelas do tipo 1 possuem exatamente um alelo ativo. Elas não são cancerosas, mas estão mais próximas de atingir essa condição do que as células do tipo 0. Finalmente, as células do tipo 2 são as cancerosas, com ambos os alelos inativos.

Embora admitamos que as células do tipo 1 tenham a mesma taxa reprodutiva que as variedades do tipo 0, isso não quer dizer que as taxas de ocorrência de mutações nessas duas subpopulações sejam idênticas. De fato, há evidências experimentais de que a taxa u_1 com que células do tipo 0 adquirem sua primeira mutação seja menor do que a taxa u_2 com que células do tipo 1 tornam-se cancerosas, pois certos processos só são possíveis após a ocorrência da primeira inativação alélica.

Adotando a dinâmica de Moran de tempo discreto discutida no capítulo 2, vamos construir uma cadeia de Markov de tempo discreto cujo espaço de estados é $\mathcal{S} = \{0, 1, \dots, N + 1\}$. Se pelo menos uma das N células for cancerosa (tipo 2) no instante t , diremos que $X_t = N + 1$. Isso é suficiente para iniciar o surgimento de um tumor e será considerado um estado absorvente dessa cadeia de Markov. Por outro lado, por $X_t = i$ denotaremos o estado em que i células são do tipo 1 e $N - i$ são do tipo 0.

Dessa forma, a construção da matriz de transição da cadeia de Markov é imediata. Expondo-as de uma forma um pouco menos elegante do que seria possível, mas mais conveniente para a descrição da dinâmica, as probabilidades de transição não nulas são $p_{N+1, N+1} = 1$ e

$$p_{i, i-1} = \left(\frac{N-i}{N}\right) (1-u_1) \left(\frac{i}{N}\right), \quad i = 1, \dots, N-1, \quad (4.1.1)$$

$$p_{i, i} = \left(\frac{N-i}{N}\right) (1-u_1) \left(\frac{N-i}{N}\right) + \left(\frac{i}{N}\right) (1-u_2) \left(\frac{i}{N}\right) + \left(\frac{N-i}{N}\right) u_1 \left(\frac{i}{N}\right), \quad i = 0, \dots, N, \quad (4.1.2)$$

$$p_{i, i+1} = \left(\frac{N-i}{N}\right) u_1 \left(\frac{N-i}{N}\right) + \left(\frac{i}{N}\right) (1-u_2) \left(\frac{N-i}{N}\right), \quad i = 0, \dots, N-1 \quad (4.1.3)$$

e

$$p_{i,N+1} = \left(\frac{i}{N}\right) u_2, \quad i = 1, \dots, N. \quad (4.1.4)$$

Analisemos a eq. (4.1.2), por exemplo. O primeiro termo no lado direito dessa equação corresponde à probabilidade de, na dinâmica de Moran, uma célula do tipo 0 ser escolhida para reproduzir-se, multiplicada pela probabilidade dessa descendente não se transformar em uma célula do tipo 1 e pela probabilidade dessa nova célula substituir uma outra semelhante a si própria, deixando inalterado o estado da cadeia de Markov. Por outro lado, a cadeia também permaneceria no mesmo estado se uma célula do tipo 1 substituísse uma congênere, o que está representado no segundo termo. Finalmente, o terceiro termo corresponde à probabilidade de uma célula do tipo 0 se reproduzir e substituir uma do tipo 1, mas apenas após ter sofrido uma mutação que a torna idêntica à outra que ela substitui e mantém inalterado o estado do sistema. As demais expressões admitem interpretações análogas.

Exercício: Certifique-se de que você compreendeu a construção de todas as probabilidades de transição do modelo descrito logo acima.

4.2 Um Jogo Evolucionário

Não é possível descrever cuidadosamente os princípios da Teoria de Jogos nestas notas. Um leitor interessado nesse tema pode consultar [14] para conhecer uma abordagem rigorosa, mas introdutória e despida de aplicações, ou [7], para conhecer em detalhes a abordagem evolucionária. No próximo parágrafo, introduziremos alguns termos com significados que serão bem claros neste contexto e faremos algumas relações para sugerir aos leitores porque esse formalismo tornou-se relevante para o estudo de certos problemas biológicos.

A Teoria de Jogos é um formalismo matemático idealizado para descrever o comportamento de sistemas compostos por diferentes agentes, os assim chamados **jogadores**, que atuam em um determinado contexto em busca de uma **recompensa** (*payoff*) que depende também das **ações** dos demais jogadores, guiadas por **estratégias** previamente adotadas. Respetivamente, em sistemas biológicos, é possível vislumbrarmos diferentes **indivíduos** em um ambiente competitivo com suas **adaptabilidades** relativas dependendo do resto da população, que se **reproduzem** de acordo com suas **características intrínsecas** (=“genótipo”). Se os “indivíduos” fossem diferentes alelos, por exemplo, suas identidades estariam muito bem

definidas e seriam constantes na ausência de mutações, que fariam o papel de alterações de estratégias.

No modelo que vamos estudar, não admitiremos a possibilidade de mutações; a recompensa de um jogador ao interagir com um outro dependerá única e exclusivamente das suas identidades. Explicitamente, vamos considerar uma população de tamanho constante N cuja composição é completamente especificada pela quantidade i de jogadores do tipo 1, os restantes $N - i$ sendo classificados como do tipo 2. A matriz de recompensas \mathbf{M} é tal que $M_{\mu\nu}$ é a recompensa de um jogador do tipo μ quando ele se “confronta” com um jogador do tipo ν , $\mu, \nu = 1, 2$. No entanto, ao contrário da Teoria de Jogos usual, onde as freqüências dos diferentes tipos de jogadores são mantidas constantes, permitiremos que essas freqüências mudem no tempo, associando a adaptabilidade de um agente à matriz de recompensas. Trataremos, portanto, de seleção dependente das freqüências, o que insere nosso estudo na assim chamada Teoria Evolucionária de Jogos [7].

Como definir as adaptabilidades a partir de \mathbf{M} ? Seguiremos [17], introduzindo um coeficiente de seleção $w \in [0, 1]$ que regula a influência da matriz de recompensas (note o que ocorre quando $w = 0$ ou $w = 1$) e, para w fixo, definindo a adaptabilidade de um agente do tipo μ quando a população está no estado i como

$$f_i^\mu(w) = 1 - w + wr_i^\mu, \quad (4.2.5)$$

onde r_i^μ é a recompensa média obtida por um jogador do tipo μ quando há i jogadores do tipo 1. Embora a matriz de recompensas seja constante, r_i^μ depende do estado da população porque os 2 pesos usados no cálculo da recompensa média são as probabilidades de um agente μ encontrar-se com um outro, que pode ser como ele próprio ou do tipo oposto. Explicitamente,

$$r_i^1 = \left(\frac{i-1}{N-1} \right) M_{11} + \left(\frac{N-i}{N-1} \right) M_{12} \quad (4.2.6)$$

e

$$r_i^2 = \left(\frac{i}{N-1} \right) M_{21} + \left(\frac{N-i-1}{N-1} \right) M_{22}. \quad (4.2.7)$$

Dadas as adaptabilidades, a escolha da dinâmica populacional deve ser guiada pelo princípio básico de uma boa modelagem, que é balancear a tratabilidade do modelo com sua capacidade de representação dos fenômenos em análise. Novamente, vamos adotar uma dinâmica de Moran de tempo discreto agindo na população, como descrito na seção 2.4. A cada iteração, um jogador é sorteado para gerar um descendente idêntico a ele, que substituirá

um dos N membros já presentes (que pode ser até seu próprio ancestral) escolhido aleatória e uniformemente. Porém, o sorteio do indivíduo que se reproduzirá não é uniforme, mas sim ponderado pelas adaptabilidades dos diferentes jogadores. É importante ressaltar que a seleção atua somente na etapa de reprodução. Se, além disso, as adaptabilidades afetassem as “probabilidades de encontro” presentes nas eqs. (4.2.6-4.2.7) ou a escolha do agente a ser substituído, teríamos mais dificuldades em identificar os efeitos seletivos por “multiplicidade de atuação”.

Muito bem, agora só precisamos descrever o comportamento do sistema ao longo do tempo. Como a população é caracterizada apenas pelo número i e a dinâmica de Moran só permite a substituição de um dos N membros da comunidade a cada passo temporal, estamos trabalhando com uma cadeia de Markov de tempo discreto cuja matriz de transição é tal que $p_{i,j} = 0$ se $|i - j| \geq 2$ e, se $i = 1, \dots, N - 1$,

$$p_{i,i+1} = \frac{if_i^1}{if_i^1 + (N-i)f_i^2} \cdot \frac{N-i}{N}, \quad (4.2.8)$$

$$p_{i,i-1} = \frac{(N-i)f_i^2}{if_i^1 + (N-i)f_i^2} \cdot \frac{i}{N} \quad (4.2.9)$$

e $p_{i,i} = 1 - p_{i,i+1} - p_{i,i-1}$. Evidentemente, $p_{0,0} = 1 = p_{N,N}$ e $p_{0,k} = 0 = p_{N,l}$ para $k \neq 0$ e $l \neq N$.

Portanto, temos mais um modelo que é um caso particular do processo geral de nascimento e morte em tempo discreto que estudamos na seção 2.4. Relembrando aquela notação, as eqs. (4.2.8) e (4.2.9) fazem com que $\sigma_i = \frac{f_i^2}{f_i^1}$ e a eq. (2.4.23) mostra automaticamente que a probabilidade de fixação de um agente do tipo 1 quando todos os $N - 1$ indivíduos são do tipo 2, ou seja, a probabilidade de que uma população no estado $i = 1$ eventualmente atinja o estado $i = N$, é

$$\rho = \frac{1}{1 + \sum_{k=1}^{N-1} \prod_{j=1}^k \frac{f_j^2}{f_j^1}}. \quad (4.2.10)$$

Maiores progressos de agora em diante dependeriam de detalhes da matriz de recompensas \mathbf{M} ou de aproximações convenientes. Em [17], admite-se que w seja próximo de 0 e alguns resultados bem interessantes são obtidos com base em expansões assintóticas (no tamanho da população, mesmo finita). Infelizmente, não temos espaço para trilhar esse caminho. Além disso, essas análises subsequentes, embora relevantes, desviariam nossa atenção

do *processo de modelagem*, que foi o foco deste minicurso. De qualquer forma, os leitores interessados em maiores detalhes podem procurar aquela referência.

Apêndice A

Teoria Axiomática de Probabilidade

Como todos os leitores destas notas devem saber, um experimento que requer uma análise probabilística é aquele cujo resultado não é determinado univocamente, mas sim pode ser qualquer elemento ω (denominado **evento elementar**) de um **conjunto amostral** Ω . Dessa forma, é natural que se busquem formas de associar a cada **evento** possível (qualquer subconjunto de Ω) um número que indique ou meça, em algum sentido, o quão freqüentemente esse evento deveria ocorrer se o experimento fosse realizado diversas vezes. Além disso, espera-se que a noção intuitiva de uma **variável aleatória** como sendo uma função que associa um número (real) associado a cada possível evento também seja bem conhecida (apesar de ser denominada uma “variável”), bem como as mais comuns distribuições e densidades de probabilidades. Todos esses conceitos estão descritos de forma elementar em [11], embora todos devessem estudar algum texto do nível de [8] mais cedo ou mais tarde.

Isso já é suficiente para formar uma idéia intuitiva do que é um processo estocástico, como foi feito no capítulo 1. Porém, este apêndice deve ser apreciado pelos leitores mais ávidos por rigor.

Para enunciarmos precisamente a definição de um processo estocástico, precisamos abstrair a noção de um modelo probabilístico com que se iniciou o capítulo 1 para aquela de um **espaço de probabilidade**, que é um trio $(\Omega, \mathcal{A}, \mathbb{P})$ onde

1. Ω é o familiar espaço amostral,
2. \mathcal{A} é uma coleção de subconjuntos de Ω conhecida tecnicamente como

uma σ -álgebra (ou simplesmente “família de eventos”) e

3. \mathbb{P} é uma medida de probabilidade, uma função $\mathbb{P} : \mathcal{A} \rightarrow \mathbb{R}$.

Previsivelmente, \mathcal{A} e \mathbb{P} devem satisfazer certas propriedades.

Surpreendentemente para os neófitos, \mathcal{A} não precisa ser necessariamente o conjunto de todos os subconjuntos de Ω . Na verdade, em muitos casos isso não é sequer desejável, embora não possamos discutir isso nestas notas. Uma σ -álgebra \mathcal{A} deve apenas ser tal que

- i. $\emptyset \in \mathcal{A}$ e $\Omega \in \mathcal{A}$,
- ii. para todo $M \in \mathcal{A}$ também seja verdade que $M^c \in \mathcal{A}$, onde M^c é o complemento de M em Ω , $M^c = \{\omega \in \Omega : \omega \notin M\}$, e
- iii. para qualquer família de conjuntos $M_j \in \mathcal{A}$, com $j \in \mathbb{N}$, também seja verdade que $(\cup_{j=1}^{\infty} M_j) \in \mathcal{A}$.

Pelo menos as duas primeiras propriedades são bem intuitivas. Tipicamente, um evento pode ocorrer ou não. Mas a não ocorrência de um evento equivale à ocorrência do seu complemento, o que é expresso por (ii). O evento certo Ω e o evento impossível \emptyset poderiam suscitar dúvidas, mas, desde que a eles sejam atribuídas probabilidades “coerentes” (1 e 0, respectivamente), a propriedade (i) é evidente.

Vejamos, então, quais são os **axiomas de Kolmogorov** para \mathbb{P} . Essa medida deve ser tal que

- i. para todo $M \in \mathcal{A}$, $\mathbb{P}(M) \geq 0$,
- ii. $\mathbb{P}(\Omega) = 1$ e
- iii. para qualquer família de conjuntos $M_j \in \mathcal{A}$, dois a dois disjuntos (tais que $M_i \cap M_k = \emptyset$ se $i \neq k$), vale a propriedade $\mathbb{P}(\cup_{j=1}^{\infty} M_j) = \sum_{j=1}^{\infty} \mathbb{P}(M_j)$.

As propriedades (i) e (ii) são bem intuitivas, enquanto a (iii) equivale à famosa “regra do OU” que muitos estudantes são obrigados a decorar. Das (iii), segue que $\mathbb{P}(\emptyset) = 0$, por exemplo.

Como se vê, existe muita coisa por trás daquela vaga noção de “atribuir probabilidades a eventos”. Da mesma forma, o conceito de variável aleatória deve ser refinado. Os leitores certamente já calcularam algo como a “probabilidade de uma certa variável aleatória X ser menor que 2”, por exemplo. Entretanto, o que devemos obter são as probabilidades de eventos. Qual é o significado de um cálculo como aquele? Para respondermos essa questão, o próprio conceito de variável aleatória precisa ser definido precisamente.

Pois bem, uma variável aleatória (real) é uma função $X : \Omega \rightarrow \mathbb{R}$ tal que $\{\omega : X(\omega) \leq x\} \in \mathcal{A}$ para todo $x \in \mathbb{R}$. Portanto, quando dizemos, em um abuso de linguagem, que queremos calcular $\mathbb{P}(X < 2)$, o que na verdade queremos dizer é $\mathbb{P}(\{\omega : X(\omega) < 2\})$. Uma variável aleatória é discreta se a imagem da função X , $X(\Omega)$, for um conjunto enumerável.

Finalmente, temos condições de definir precisamente o que é um processo estocástico.

Definição A.1. *Dado qualquer conjunto \mathcal{T} , um **processo estocástico** com conjunto indicial \mathcal{T} é uma coleção $\{X_t : t \in \mathcal{T}\}$ de variáveis aleatórias indexadas por \mathcal{T} e definidas em um mesmo espaço de probabilidade.*

A classificação dos diferentes processos estocásticos e suas propriedades estão descritos no corpo do texto, no capítulo 1.

Bibliografia

- [1] N. T. J. Bailey, “The Elements of Stochastic Processes: With Applications to the Natural Sciences”, 1st edition, John Wiley & Sons, 1964.
- [2] R. Durrett, “Probability Models for DNA Sequence Evolution”, 1st edition, Springer Verlag, 2002.
- [3] W. J. Ewens, “Mathematical Population Genetics”, 1st edition, Springer Verlag, New York, 1979.
- [4] J. E. Gentle, “Random Number Generation and Monte Carlo Methods”, 2nd edition, Springer Verlag, 2004.
- [5] D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, *J. Comp. Phys.*, **22** (1976), 403-434.
- [6] H. L. Guidorizzi, “Um Curso de Cálculo - vol. 1”, 5a. edição, Editora LTC, 2001.
- [7] J. Hofbauer and K. Sigmund, “Evolutionary Games and Population Dynamics”, 1st edition, Cambridge University Press, 1998.
- [8] B. R. James, “Probabilidade: Um Curso em Nível Intermediário”, 2a. edição, IMPA, Rio de Janeiro, 2002.
- [9] S. Karlin and H. M. Taylor, “A First Course in Stochastic Processes”, 2nd edition, Academic Press, 1975.
- [10] L. P. Maia, “Dinâmica Adaptativa, Genealogias e Testes Estatísticos de Neutralidade em Evolução Molecular”, Tese de Doutorado, Instituto de Física de São Carlos da Universidade de São Paulo, São Carlos, 2004.

- [11] P. L. Meyer, “Probabilidade: Aplicações à Estatística”, 2a. edição, LTC Editora, Rio de Janeiro, 1983.
- [12] P.A.P. Moran, Random processes in genetics, *Proc. Camb. Phil. Soc.* **54** (1958), 60-71.
- [13] A. C. Morgado, J. B. P. de Carvalho, P. C. P. Carvalho e P. Fernandez, “Análise Combinatória e Probabilidade”, 9a. edição, Publicação SBM, 1991.
- [14] P. Morris, “Introduction to Game Theory”, 1st edition, Springer, 1994.
- [15] J. D. Murray, “Mathematical Biology: I. An Introduction”, 3rd edition, Springer, 2002.
- [16] J. D. Murray, “Mathematical Biology: II. Spatial Models and Biomedical Applications”, 3rd edition, Springer, 2003.
- [17] M. A. Nowak, “Evolutionary Dynamics: Exploring the Equations of Life”, 1st edition, Belknap Press, 2006.
- [18] S. M. Ross, “Simulation”, 4th edition, Academic Press, Oxford, 2006.
- [19] G. Rowe, “Theoretical Models in Biology”, Clarendon Press, Oxford, 1994.
- [20] J. P. O. Santos, M. P. Mello e I. T. C. Murari, “Introdução à Análise Combinatória”, 1a. edição, Editora Ciência Moderna, 2008.
- [21] H. C. Tuckwell, “Elementary Applications of Probability Theory”, 2nd edition, Chapman & Hall, 1995.

Índice

- fitness*, 16
- loci*, 14
- locus*, 14

- alelo, 14
- algoritmo de Gillespie, 52
- apoptose, 58

- câncer, 57, 58
- cadeia, 12
- cadeia de Markov, 13, 19, 41
- cadeia homogênea, 20
- classe de equivalência, 23, 25
- cromossomo, 14

- deriva genética, 17, 22
- diplóide, 15
- dominância, 15

- equação diferencial de diferenças, 47
- estado absorvente, 25

- fenótipo, 14
- fixação, 27, 38
- função geradora, 49

- Genética de Populações, 15
- genótipo, 15
- gene, 14
- gene supressor de tumor, 59
- genealogia, 29

- haplóide, 15

- heterozigoto, 15
- homozigoto, 15

- instabilidade cromossômica, 59

- Jogos, Teoria de, 61

- Markov, cadeia de, 13, 19, 41
- Markov, processo de, 13
- modelo de Moran, 34, 61
- modelo de Wright-Fisher, 22
- modelo sem superposição de gerações, 22
- mutação, 15, 30

- notação \mathcal{O} , 42

- oncogene, 59

- processo de Markov, 13
- processo de nascimento, 50
- processo de nascimento e morte, 51
- processo de Poisson, 47
- processo estocástico, 11, 12, 65, 67

- seleção natural, 16, 31
- simulação, 32

- taxa de probabilidade, 45

- valor adaptativo, 16
- variável aleatória, 65

- Wright-Fisher, modelo de, 22

NOTAS EM MATEMÁTICA APLICADA

1. Restauração de Imagens com Aplicações em Biologia e Engenharia
Geraldo Cidade, Antônio Silva Neto e Nilson Costa Roberty
2. Fundamentos, Potencialidades e Aplicações de Algoritmos Evolutivos
Leandro dos Santos Coelho
3. Modelos Matemáticos e Métodos Numéricos em Águas Subterrâneas
Edson Wendlander
4. Métodos Numéricos para Equações Diferenciais Parciais
Maria Cristina de Castro Cunha e Maria Amélia Novais Schleicher
5. Modelagem em Biomatemática
Joyce da Silva Bevilacqua, Marat Rafikov e Cláudia de Lello
Courtouke Guedes
6. Métodos de Otimização Randômica: algoritmos genéticos e “simulated annealing”
Sezimária F. Pereira Saramago
7. “Matemática Aplicada à Fisiologia e Epidemiologia”
H.M. Yang, R. Sampaio e A. Sri Ranga
8. Uma Introdução à Computação Quântica
Renato Portugal, Carlile Campos Lavor, Luiz Mariano Carvalho
e Nelson Maculan
9. Aplicações de Análise Fatorial de Correspondências para Análise de Dados
Homero Chaib Filho

10. Modelos Matemáticos baseados em autômatos celulares para Geoprocessamento
Marilton Sanchotene de Aguiar, Fábria Amorim da Costa,
Graçaliz Pereira Dimuro e Antônio Carlos da Rocha Costa
11. Computabilidade: os limites da Computação
Regivan H. N. Santiago e Benjamã R. C. Bedregal
12. Modelagem Multiescala em Materiais e Estruturas
Fernando Rochinha e Alexandre Madureira
13. Modelagem em Biomatemática (Coraci Malta ed.)
 - 1 - “Modelagem matemática do comportamento elétrico de neurônios e algumas aplicações”
Reynaldo D. Pinto
 - 2 - “Redes complexas e aplicações nas Ciências”
José Carlos M. Mombach
 - 3 - “Possíveis níveis de complexidade na modelagem de sistemas biológicos”
Henrique L. Lenzi, Waldemiro de Souza Romanha e Marcelo Pelajo- Machado
14. A lógica na construção dos argumentos
Angela Cruz e José Eduardo de Almeida Moura
15. Modelagem Matemática e Simulação Numérica em Dinâmica dos Fluidos
Valdemir G. Ferreira, Hélio A. Navarro, Magda K. Kaibara
16. Introdução ao Tratamento da Informação nos Ensinos Fundamental e Médio
Marcilia Andrade Campos, Paulo Figueiredo Lima
17. Teoria dos Conjuntos Fuzzy com Aplicações
Rosana Sueli da Motta Jafelice, Laércio Carvalho de Barros,
Rodney Carlos Bassanezi
18. Introdução à Construção de Modelos de Otimização Linear e Inteira
Socorro Rangel

19. Observar e Pensar, antes de Modelar
Flavio Shigeo Yamamoto, Sérgio Alves, Edson P. Marques Filho,
Amauri P. de Oliveira
20. Frações Contínuas: Propriedades e Aplicações
Eliana Xavier Linhares de Andrade, Cleonice Fátima Bracciali
21. Uma Introdução à Teoria de Códigos
Carlile Campos Lavor, Marcelo Muniz Silva Alves, Rogério
Monteiro de Siqueira, Sueli Irene Rodrigues Costa
22. Análise e Processamento de Sinais
Rubens Sampaio, Edson Cataldo, Alexandre de Souza Brandão
23. Introdução aos Métodos Discretos de Análise Numérica de EDO e
EDP
David Soares Pinto Júnior
24. Representações Computacionais de Grafos
Lílian Markenzon, Oswaldo Vernet
25. Ondas Oceânicas de Superfície
Leandro Farina
26. Técnicas de Modelagem de Processos Epidêmicos e Evolucionários
Domingos Alves, Henrique Fabrício Gagliardi
27. Introdução à teoria espectral de grafos com aplicações
Nair Maria Maia de Abreu, Renata Raposo Del-Vecchio, Cybele
Tavares Maia Vinagre e Dragan Stevanović
28. Modelagem e convexidade
Eduardo Cursi e Rubens Sampaio
29. Modelagem matemática em finanças quantitativas em tempo discreto
Max Oliveira de Souza e Jorge Zubelli
30. Programação não linear em dois níveis: aplicação em Engenharia
Mecânica
Ana Friedlander e Eduardo Fancello

31. Funções simétricas e aplicações em Combinatória

José Plínio de Oliveira Santos e Robson da Silva

32. Semigrupos aplicados a sistemas dissipativos em EDP

Carlos Raposo da Cunha