

Biomatemática: que mistura é essa?

Biomathematics: what does that strange mix mean?

Armando G. M. Neves

Resumo

Explicamos de maneira bastante introdutória alguns modelos matemáticos para fenômenos da Biologia, mostrando a importância que a Matemática pode ter para o entendimento de algumas questões biológicas. Falamos em particular de modelos para epidemias, crescimento populacional, interações ecológicas e a evolução da espécie humana. Na maior parte do artigo os modelos ilustrados são determinísticos, mas também falamos um pouco sobre modelos estocásticos.

Palavras-chave: Modelos matemáticos. Equações diferenciais. Modelo de Wright-Fisher. Modelo SIR. Crescimento logístico.

Abstract

We explain in a rather introductory fashion some mathematical models for phenomena in Biology, highlighting the importance that Mathematics may have for understanding some biological questions. We introduce, in particular, models for epidemics, population growth, ecological interactions and the evolution of the human species. The models illustrated are mostly deterministic, but we also talk about stochastic models.

Keywords: Mathematical models. Differential equations. Wright-Fisher model. SIR model. Logistic growth.

¹ Universidade Federal de Minas Gerais, MG, Brasil - <http://orcid.org/0000-0003-1602-6881> - armandogmneves@gmail.com

1 Introdução

Este artigo é uma adaptação de uma palestra que apresentei na UFMG em 23/9/2019.

Vários fenômenos biológicos podem ser representados por modelos matemáticos. Queremos ilustrar aqui vários destes fenômenos e modelos, explicando a utilidade da Matemática em cada um dos casos.

Além de divulgar em nível básico uma interessante área da pesquisa em Matemática, um segundo objetivo é o de incentivar estudantes de cursos das áreas de Ciências Biológicas, Saúde e Ciências Humanas a estudarem mais matemática do que seus currículos padrão pressupõem.

Na mesma linha, acho importante mostrar aos estudantes de Ciências Exatas e Engenharia aplicações interessantes e menos conhecidas da matemática que estudam em seus cursos e também motivá-los a procurar mais conhecimentos matemáticos que os de seus currículos padrão.

2 As disciplinas básicas de Matemática na UFMG

Na UFMG, algumas disciplinas de Matemática, aqui brevemente elencadas, são oferecidas aos primeiros períodos dos cursos das áreas de Ciências Exatas e Engenharias. Uma seleção parecida é oferecida por outras instituições de ensino brasileiras e internacionais. Coloco aqui a lista para que os leitores que não cursam um curso dessas áreas, ou que estejam no início de um desses cursos, tenham alguma ideia de quanta Matemática devem estudar para entender com mais propriedade os exemplos que seguem neste artigo.

No primeiro semestre, em geral são cursados Cálculo Diferencial e Integral I (Cálculo I, abreviando) e Geometria Analítica e Álgebra Linear. No segundo semestre, Cálculo II. No terceiro semestre, Cálculo III e Equações Diferenciais A (EDA). No quarto semestre, Equações Diferenciais B (EDB).

Alguns cursos fazem, em lugar dos Cálculos II e III, a disciplina Cálculo de Várias Variáveis. Alguns cursos não fazem nenhuma disciplina de Equações Diferenciais. Alguns outros, fazem Equações Diferenciais C, que é uma mistura das EDA e EDB. Outros cursam somente EDA.

Obviamente, há disciplinas mais avançadas de Matemática, tanto para os cursos de graduação em Matemática e Matemática Computacional, quanto na pós-graduação em Matemática.

3 A derivada

A derivada de uma função f é uma nova função f' que a cada ponto x do domínio de f associa o coeficiente angular (inclinação) da reta tangente ao gráfico de f no ponto de abscissa x . A derivada $f'(x)$ representa no ponto de abscissa x a taxa de variação de f . Em particular, $f'(x) \geq 0$ se f é crescente em x e $f'(x) \leq 0$ nos pontos x em que f decresce.

A Fig. 1 ilustra para uma mesma função a derivada em vários pontos do gráfico.

Na UFMG a derivada e o conceito correlato de integral são introduzidos na disciplina Cálculo I. O cálculo do valor da derivada para uma função, por exemplo a ilustrada na Fig. 1, é aprendido nessa disciplina. Um maior desenvolvimento dos conceitos aparecem nas disciplinas subsequentes de Cálculo II e III. Uma referência de boa qualidade para essas disciplinas são os dois volumes de Stewart (2017).

Uma equação diferencial é uma equação em que se procura determinar uma função a partir de informações sobre sua derivada. Veremos neste artigo alguns exemplos de equações diferenciais e sistemas de equações diferenciais com aplicações biológicas. Equações diferenciais são fundamentais em várias outras áreas também, notadamente na Física e nas Engenharias. Na UFMG, as disciplinas EDA, EDB e EDC são disciplinas básicas sobre equações diferenciais. Uma boa referência básica é Boyce e diPrima (2015).

4 Dinâmica de doenças infecciosas

Atualmente o Brasil vive um grande surto de sarampo. O sarampo é causado por um agente patógeno, um vírus, que se transmite diretamente de pessoa a pessoa. Outras doenças desse tipo, em que o patógeno é transmitido diretamente de um indivíduo a outro, são os resfriados, a gripe (influenza), catapora, rubéola, caxumba, HIV, hepatites, SARS (síndrome respiratória aguda grave), Ebola, etc.

Uma outra característica do sarampo e também de várias outras das doenças acima nomeadas é que um indivíduo infectado adquire, após a cura, imunidade permanente à doença. Em outras doenças não há imunidade alguma, ou a imunidade é temporária.

Outra coisa a se levar em conta é por quanto tempo queremos estudar uma população. Do ponto de vista de um único surto ou epidemia, o tempo do estudo pode ser da ordem de poucas semanas. Neste caso, é razoável desprezar nascimentos que ocorram no período, ou mortes por motivos que não sejam a doença estudada.

Figura 1: A reta tangente e a derivada $f'(x)$ para alguns valores de x . Nestes exemplos usamos $f(x) = \frac{1}{2}x^3 - 4x + 1$.

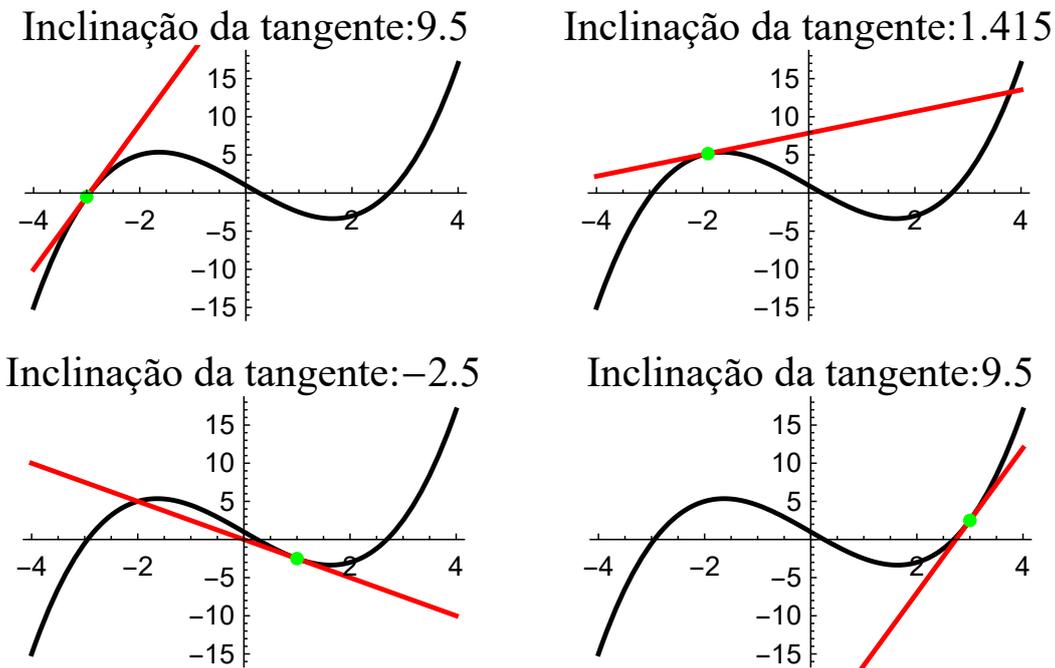


Figura 2: Esquema das transições entre os compartimentos no modelo SIR.



Do ponto de vista da modelagem que faremos, podemos dividir uma população em três classes (compartimentos) de indivíduos:

- Suscetíveis (S): os indivíduos que nunca tiveram a doença e podem adquiri-la.
- Infeciosos (I): os indivíduos que estão no momento com a doença e que podem transmiti-la aos suscetíveis.
- Recuperados (R): os indivíduos que já tiveram a doença e, ou se curaram, ou morreram. No caso de doenças que conferem imunidade permanente, como o sarampo, os indivíduos que se recuperam permanecem nesse compartimento por todo o tempo no futuro. Em outros casos, depois de algum tempo os recuperados podem voltar ao compartimento dos suscetíveis.

O modelo mais simples para a dinâmica de uma doença como o sarampo foi criado por Kermack e McKendrick em 1927 e é mais conhecido como o modelo SIR. O esquema da Fig. 2 mostra as possíveis transições entre os compartimentos do modelo. A figura mostra que estamos considerando imunidade permanente, pois não há saída de indivíduos do compartimento R. Mostra também que estamos desprezando nascimentos e mortes de indivíduos, o que significa que estamos interessados em um pequeno período de tempo.

Há muitas variações para o modelo básico. Além das já mencionadas – incluir nascimentos e mortes e imunidade temporária – mencionamos também que modelos mais complexos são necessários para descrever doenças em que o contágio envolve vetores. Por exemplo doenças importantíssimas como dengue, zika, chikungunya, febre amarela, malária, leishmaniose, doença de Chagas, etc. Modelos mais complexos são necessários também para doenças sexualmente transmissíveis, em que a população é dividida por classes de preferências para relações sexuais.

Descrevemos a seguir a versão mais básica do modelo SIR, aquela ilustrada na Fig. 3. Boas referências são Edelstein-Keshet (2005) ou Murray (2002).

Supomos então que uma população, incluindo indivíduos eventualmente mortos em consequência da doença, é de tamanho constante igual a N indivíduos, desprezando nascimentos na população e mortes por motivos alheios à doença. Se $S(t)$, $I(t)$ e

$R(t)$ representam os números de susceptíveis, infecciosos e recuperados no tempo t , então

$$S(t) + I(t) + R(t) = N .$$

A variação temporal das funções acima é prescrita pelo seguinte sistema de equações diferenciais:

$$\begin{cases} S'(t) &= -\frac{\beta}{N}S(t)I(t) \\ I'(t) &= \frac{\beta}{N}S(t)I(t) - \gamma I(t) \\ R'(t) &= \gamma I(t) \end{cases}$$

onde $\beta > 0$ e $\gamma > 0$ são parâmetros que explicaremos a seguir.

O termo $\frac{\beta}{N}S(t)I(t)$ que aparece na primeira e na segunda equações representa o contágio de susceptíveis por infecciosos. Para que ele não seja nulo, é necessário que haja tanto susceptíveis ($S(t) > 0$), quanto infecciosos ($I(t) > 0$). O termo aparece na primeira equação com um sinal menos, indicando que o contágio contribui negativamente para a variação do número de susceptíveis: o contágio retira indivíduos do compartimento S e os leva para o compartimento I. Por isto, o mesmo número de indivíduos aparece, agora com sinal mais, na segunda das equações.

O termo de contágio aparece multiplicado por $\frac{\beta}{N}$. Daí pode-se inferir que quanto maior β , mais contagiosa é a doença de que estamos tratando. Mais precisamente, pode-se interpretar β como o número de contatos por unidade de tempo de um indivíduo infeccioso com os demais indivíduos multiplicado pela probabilidade de que um tal contato leve à transmissão da infecção. O parâmetro β , portanto, depende tanto da doença modelada, quanto dos hábitos dos indivíduos.

A segunda equação no sistema representa tanto a entrada no compartimento I, por contágio, de susceptíveis, como já visto, quanto a saída para o compartimento R dos indivíduos que se curaram ou morreram. Quanto maior γ , mais indivíduos são recuperados por unidade de tempo. Pode-se mostrar que o parâmetro γ é o inverso do tempo médio de duração da infecção em um indivíduo. Portanto, quanto maior γ , tanto mais rápido diminui o número de infecciosos e aumenta o número de recuperados.

A imensa maioria das equações diferenciais, mesmo algumas muito simples, não pode ser resolvida exatamente e o sistema acima não é uma exceção. Apesar disto, temos métodos para obter soluções aproximadas com tanta exatidão quanto se deseje e também métodos que nos permitem entender de forma qualitativa as soluções. A Fig. 3 mostra, em diversas situações interessantes, os gráficos em função do tempo para as porcentagens na população de indivíduos susceptíveis e infecciosos. Para simplificar os gráficos, optamos por não mostrar a porcentagem de recuperados, que pode ser obtida subtraindo de 100% as porcentagens de susceptíveis e infecciosos. Os valores mostrados nos gráficos foram obtidos através de métodos de aproximação para a solução do sistema de equações diferenciais.

Podemos também explicar o comportamento das soluções pelos seguintes argumentos qualitativos.

Uma primeira possibilidade, simples demais e sem maior interesse, é que não haja nenhum indivíduo infeccioso ou nenhum susceptível. Enquanto não aparecer algum infeccioso e algum susceptível, o número de susceptíveis fica constante e os infecciosos se tornam recuperados. Vamos supor então que haja pelo menos um infeccioso e um susceptível iniciais, $I(0) > 0$ e $S(0) > 0$. Da primeira equação, $S'(t) = -\frac{\beta}{N}S(t)I(t)$, vê-se que $S'(t) < 0$. Isto significa que o número de susceptíveis sempre decresce, o que é o esperado, a não ser quando não houver mais susceptíveis ou infecciosos.

A segunda equação do sistema pode ser reescrita como, $I'(t) = (\frac{\beta}{N}S(t) - \gamma)I(t)$. Aqui temos duas situações possíveis: ou o número inicial $S(0)$ de indivíduos susceptíveis é menor que $\frac{\gamma N}{\beta}$, ou é maior que esse número.

No primeiro dos casos, ilustrado no gráfico abaixo e à direita na Fig. 3, o número de infecciosos irá sempre decrescer. Em muitas situações de interesse, porém, ocorre o contrário e o número de infecciosos crescerá enquanto se tiver $\frac{\beta}{N}S(t) > \gamma$. Mas como $S(t)$ é decrescente, então, em algum momento passará a valer $\frac{\beta}{N}S(t) < \gamma$ e o número de infecciosos começará a diminuir. Portanto, no segundo caso, o número de infecciosos cresce inicialmente e passa a decrescer somente quando $S(t)$ atingir o valor $\frac{\gamma N}{\beta}$. Isto é o que se conhece como uma epidemia. Os demais três gráficos na Fig. 3 ilustram epidemias com graus variados de severidade.

O gráfico acima e à esquerda ilustra uma epidemia bastante severa: sarampo em uma população inicialmente totalmente susceptível. Observe que em cerca de 10 dias praticamente toda a população foi atingida. No pior momento, cerca de 80% da população estava doente. O gráfico acima e à direita ilustra uma epidemia de influenza também em uma população totalmente susceptível no início. Observe que, do ponto de vista de contágio, essa epidemia é bem menos severa que a do sarampo. Ainda assim, no pior momento, por volta de 25 dias após os primeiros casos, cerca de 30% da população estava doente, o que é algo muito grave do ponto de vista da saúde pública. Em ambos os casos, usamos $\gamma = 1/7$, que significa que o tempo médio para recuperação de um indivíduo infeccioso é de 7 dias. Mas no caso do sarampo usamos $\beta = 2$ e no caso da influenza, usamos $\beta = 0.36$. Ambos os valores estão dentro das faixas indicadas na Tabela 1, que explicamos adiante. O gráfico abaixo e à esquerda ilustra ainda uma epidemia de influenza, com a diferença de que, nesse caso, a população não é inicialmente totalmente susceptível. Isto pode acontecer se uma parte da população for vacinada, ou se a mesma população já passou por uma epidemia recente que deixou parte da população na classe dos recuperados.

Caso uma população seja inicialmente totalmente susceptível, $S(0) \approx N$, a condição para crescimento do número de

Figura 3: Os gráficos abaixo mostram soluções aproximadas das equações do modelo SIR em quatro situações de interesse. Acima e à esquerda temos uma epidemia de sarampo ($\beta = 2, \gamma = 1/7$) em uma população totalmente susceptível e com 1% de infecciosos no instante inicial. Acima e à direita, temos uma epidemia de influenza (gripe). Os valores iniciais de susceptíveis e infecciosos são os mesmos do primeiro gráfico, assim como o valor $\gamma = 1/7$, mas para a influenza usamos $\beta = 0.36$. Abaixo e à esquerda mostramos uma epidemia de influenza em uma população em que 40% dos indivíduos foram vacinados e 10% eram infecciosos no instante inicial. Finalmente, abaixo e à direita temos ainda a influenza, mas em uma população em que 65% foram vacinados. Note que no último caso o número de infecciosos não cresce inicialmente. Não temos aí uma epidemia.

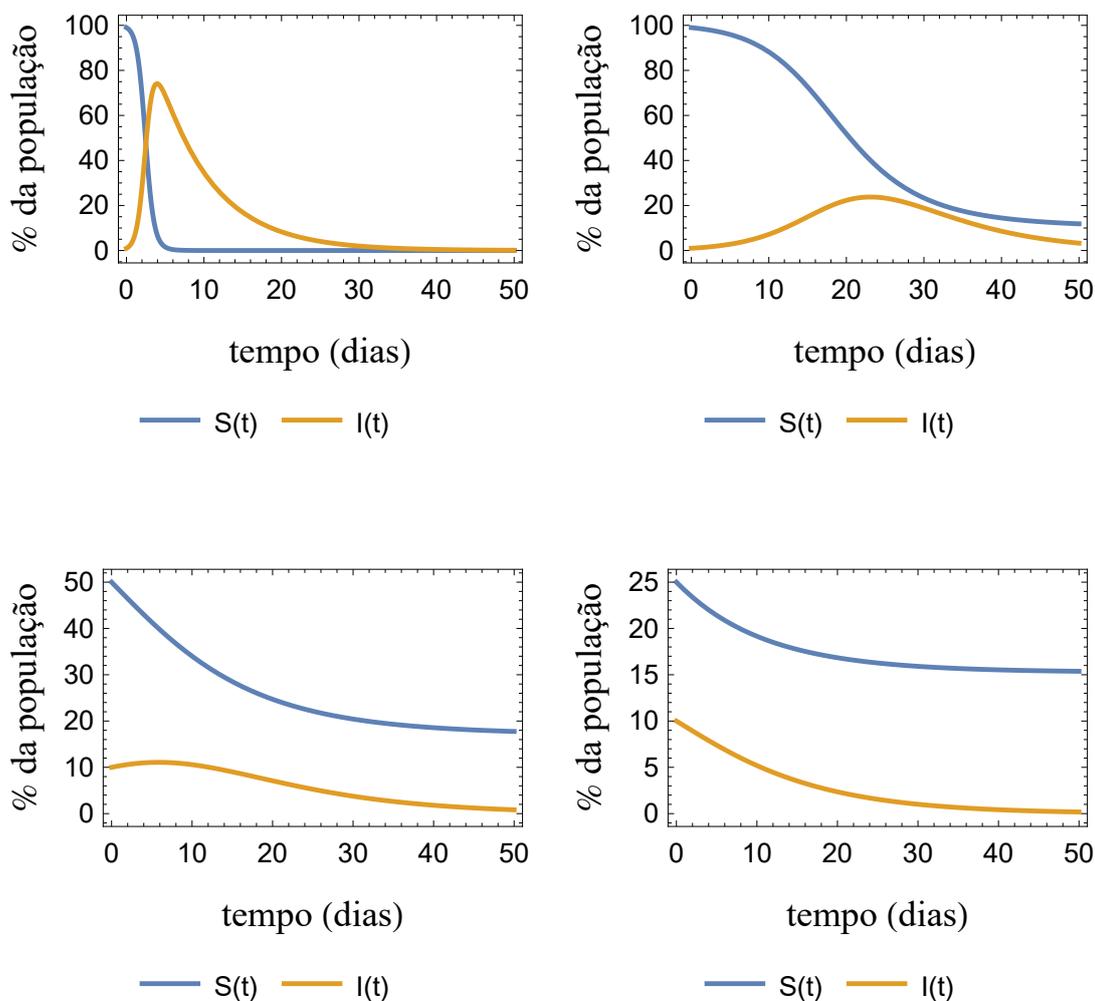


Tabela 1: Valores de R_0 para algumas doenças, adaptada de artigo em Wikipedia (Acessado em 23/10/2019)

Doença	R_0
Sarampo	12–18
Difteria	6–7
Varíola	5–7
Poliomielite	5–7
Rubéola	5–7
Caxumba	4–7
HIV/AIDS	2–5
Coqueluche	5.5
SARS	2–5
Influenza (pandemia de 1918)	2–3
Ebola (surto de 2014)	1.5–2.5

infecciosos é $\frac{\beta}{\gamma} > 1$. A razão

$$R_0 = \frac{\beta}{\gamma}$$

pode ser interpretada como o número médio de infecções secundárias causadas por um único indivíduo infeccioso em uma população inteiramente susceptível, antes que o indivíduo seja curado. Ela é chamada o número reprodutivo básico da infecção.

Uma doença que tenha uma duração muito curta (γ grande), ou que seja pouco contagiosa, ou que apareça em uma população com poucos contatos entre os indivíduos (β pequeno em ambos os casos) terá $R_0 < 1$ e não dará lugar a epidemias. Por outro lado, sempre que tivermos $R_0 > 1$ e uma população inteiramente susceptível, teremos ocorrência de epidemia e, quanto maior R_0 , mais severa será a epidemia.

A Tabela 1 mostra os valores de R_0 para algumas doenças. Conforme já argumentamos, R_0 depende tanto do agente patógeno, quanto do número de contatos entre os indivíduos da população, quanto do tempo médio para recuperação. Por este motivo, além de incertezas experimentais, a tabela fornece faixas de valores para R_0 . Por exemplo, o mesmo patógeno deve levar a um valor maior de R_0 em uma população mais concentrada (mais contatos entre os indivíduos) e a um valor menor em uma população mais dispersa. Uma boa estrutura hospitalar, medicamentos melhores ou a adoção de medidas de isolamento diminuem o tempo de recuperação, contribuindo também para a diminuição de R_0 . Observe ainda que o valor de R_0 não indica a gravidade de uma doença, mas o quanto essa é contagiosa. Observe ainda que de todas as doenças mostradas na tabela, o sarampo, uma grande preocupação de saúde pública no Brasil no momento em que este artigo é escrito, é exatamente a mais contagiosa.

Um dos ganhos consequências importantes do modelo simples que acabamos de descrever é entender o papel da vacinação na prevenção e enfrentamento de uma epidemia e também o de estimar a porcentagem da população que deve ser atingida por uma campanha de vacinação.

Do ponto de vista biológico, uma vacina é um preparado que induz a produção de anticorpos contra um patógeno no organismo do indivíduo que o recebe. Do ponto de vista matemático, segundo o modelo SIR, uma vacina é uma maneira de retirar da classe dos susceptíveis, transformando-os em recuperados, indivíduos que nunca tenham tido a doença.

Lembramos que a condição para que o número de infecciosos não cresça é $\frac{\beta}{N}S(t) < \gamma$. Usando a definição acima de R_0 , essa condição torna-se $R_0S(t)/N < 1$. Caso uma fração suficiente da população seja vacinada, de modo a ter $S(0)/N < 1/R_0$, o número de infecciosos não irá crescer. Este é o objetivo de uma campanha de vacinação. No caso do sarampo, com $R_0 \approx 15$, a fração da população a ser vacinada de modo a evitar uma epidemia, é portanto de $1 - \frac{1}{15}$, que dá aproximadamente 93%. Para a influenza, com $R_0 \approx 2.5$, a proporção é de 60%.

5 Modelos para crescimento populacional

Considere uma população cujo tamanho no tempo t seja $p(t)$. A equação diferencial que permite prever $p(t)$ é

$$p'(t) = R(t, p(t))p(t),$$

onde $R(t, p)$ é o número médio de descendentes por unidade de tempo de um indivíduo da população. É de se esperar que este número médio dependa tanto do tempo, quanto do tamanho da população.

A escolha mais simples possível é pensar que $R(t, p)$ seja uma constante r independente do tempo e do tamanho da população. Esta escolha nos leva ao chamado modelo de Malthus para o crescimento populacional, proposto em 1798.

Neste caso, é possível resolver facilmente a equação diferencial e obter $p(t) = p_0 e^{rt}$. O economista Malthus, baseado na sua hipótese simplista de que a população humana crescerá a uma taxa constante, previu que a população crescerá exponencialmente. Como os recursos para manutenção da população dificilmente crescem assim tão rápido, ele próprio percebeu que nos estudos econômicos é necessário levar em conta o tamanho da população.

Uma modificação simples e de grande alcance do modelo de Malthus é supor que, em vez de constante, $R(t, p)$ seja ainda independente de t , mas decrescente com p . O decréscimo de R com p representa a competição por recursos entre os indivíduos na população. Uma maneira simples de realizar essa proposta é tomar $R(p)$ um polinômio de 1º grau decrescente:

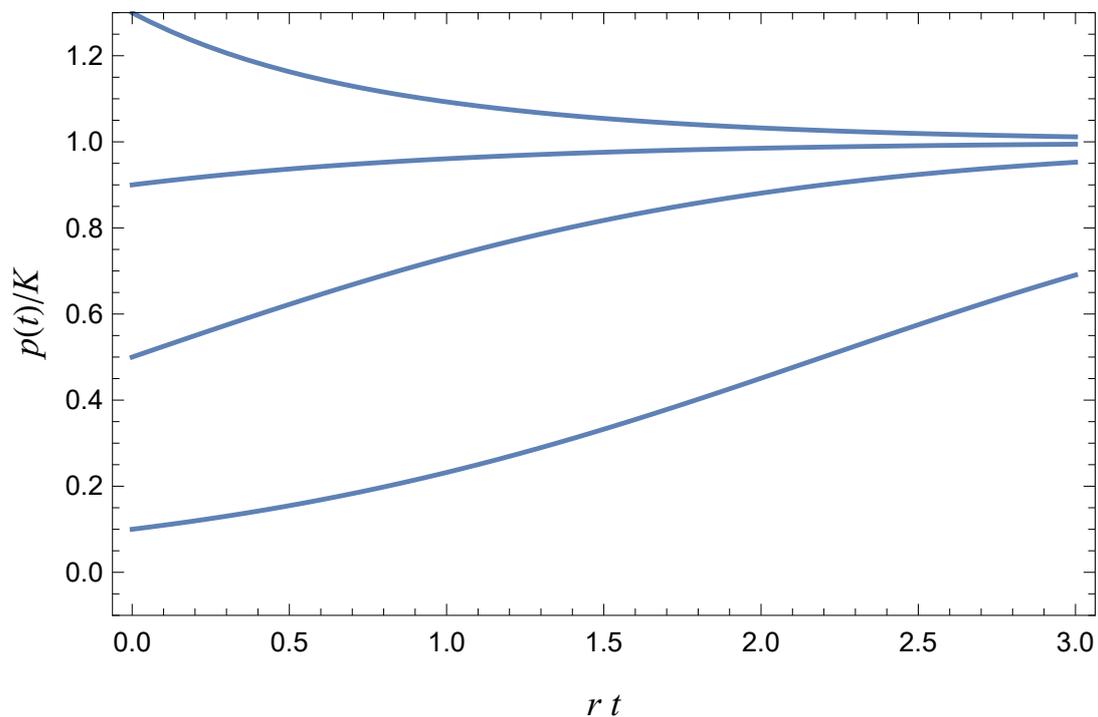
$$R(p) = r(1 - p/K).$$

Isto nos leva ao chamado modelo de Verhulst, ou logístico.

O parâmetro K na equação acima é chamado capacidade de suporte do meio. Enquanto p é muito menor que K , $R(p) \approx r$ e a população cresce de forma praticamente exponencial, como em Malthus. Mas isto dura pouco e à medida que a população cresce, o crescimento é desacelerado. Quando a população fica igual a K a taxa de crescimento torna-se nula. Pode-se resolver a equação diferencial e provar que K é a população limite para grandes valores de t .

A Fig. 4 mostra alguns gráficos de soluções da equação de crescimento logístico, mostrando que quer a população inicial seja menor que K , quer seja maior, depois de muito tempo a população atinge o limite K .

Figura 4: Os gráficos abaixo mostram soluções da equação de crescimento logístico para alguns valores da população inicial. Observe que em todos os casos $p(t)/K$ se aproxima de 1 quando t fica grande.



6 Modelos em Ecologia: populações em interação

Há diversos modelos para as várias possibilidades de interação entre indivíduos em uma população, por exemplo predação, competição, mutualismo, parasitismo, etc. Por ser uma extensão simples do modelo logístico que acabamos de mostrar, vamos ilustrar o tipo de modelos que aparece em Ecologia com um modelo para duas populações que competem por um mesmo recurso.

Suponha que duas populações de tamanhos $x(t)$ e $y(t)$ vivem em um mesmo habitat e competem por algum recurso ali presente. As equações que descrevem essa interação são

$$\begin{cases} x'(t) = r_1 \left(1 - \frac{x(t) + s_2 y(t)}{K_1} \right) x(t) \\ y'(t) = r_2 \left(1 - \frac{y(t) + s_1 x(t)}{K_2} \right) y(t) \end{cases} .$$

Os parâmetros nestas equações, todos positivos, têm os seguintes significados:

- K_1 e K_2 são as capacidades de suporte do ambiente para cada uma das duas populações.
- r_1 e r_2 são as taxas de crescimento das populações quando o ambiente é pouco povoado.
- s_2 representa o número de espaços do ambiente da população x ocupados por um único indivíduo da população y . Vice-versa, s_1 representa o número de espaços do ambiente da população y ocupados por um único indivíduo da população x . Podemos pensar s_2 como a agressividade dos indivíduos y para os x e, analogamente, s_1 é a agressividade dos indivíduos x para os y .

As equações diferenciais acima também não podem ser resolvidas exatamente. Podemos porém, assim como no caso do modelo SIR, estudar de forma qualitativa o comportamento de suas soluções. Uma pequena simplificação que ajuda nisto é reduzir o número de parâmetros. Um modelo simplificado, mantendo todas as características do inicial, mas com menos parâmetros, é

$$\begin{cases} x'(t) = [1 - (x(t) + ay(t))] x(t) \\ y'(t) = \rho [1 - (y(t) + bx(t))] y(t) \end{cases} ,$$

onde agora as capacidades de suporte são ambas iguais a 1, a agressividade da espécie y é a e a da espécie x é b . O parâmetro ρ é a taxa de crescimento da população y quando o ambiente está pouco povoado e, conforme se verá, não terá um papel importante no modelo.

Papel fundamental na análise qualitativa desse modelo é desempenhado pelas retas, ditas isóclinas, com equações $x + ay = 1$ e $bx + y = 1$. A primeira passa pelos pontos $(0, 1/a)$ e $(1, 0)$. A segunda passa pelos pontos $(1/b, 0)$ e $(0, 1)$. O leitor pode verificar que as isóclinas irão se interceptar em algum ponto do primeiro quadrante nas situações $a > 1$ e $b > 1$, caso em que ambas as agressividades são grandes, e $0 < a < 1$ e $0 < b < 1$, caso em que nenhuma das agressividades é grande.

A importância das isóclinas é que elas dividem o primeiro quadrante em regiões de crescimento e decréscimo das populações x e y . Nos pontos à esquerda da isóclina $x + ay = 1$ tem-se $x + ay < 1$ e, portanto, pela primeira equação do sistema, $x'(t) > 0$: a população x cresce. À direita da mesma isóclina, tem-se o contrário: a população x decresce. Analogamente, pela segunda equação do sistema, à esquerda da isóclina $bx + y = 1$ a população y cresce e à direita, decresce.

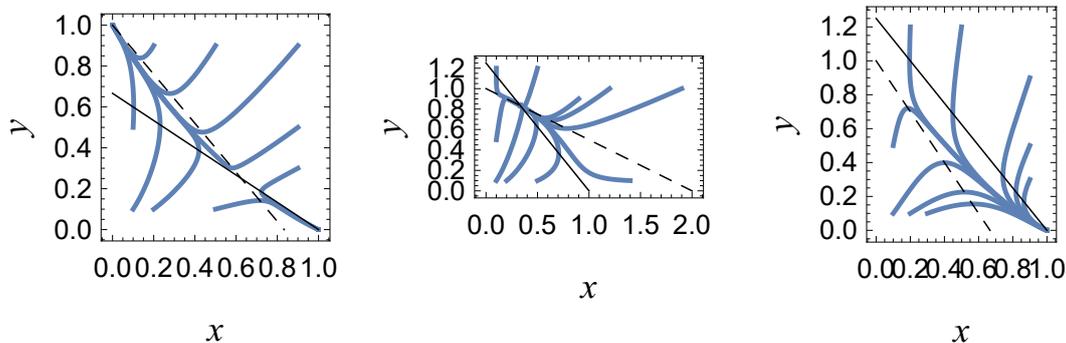
Os pontos $(1, 0)$ e $(0, 1)$ representam as possibilidades de ter uma das populações extinta, enquanto a outra está com valor igual à capacidade de suporte. São pontos de equilíbrio da dinâmica populacional. Quando as isóclinas se interceptam no primeiro quadrante, esse ponto de interseção é também um equilíbrio da dinâmica em que ambas as populações estão presentes. Finalmente, o ponto $(0, 0)$ é o equilíbrio em que ambas as populações foram extintas.

A Fig. 5 ilustra três situações desse modelo de competição. Nessa figura, a reta inclinada em traço contínuo é a isóclina $x + ay = 1$ e a reta tracejada é a isóclina $bx + y = 1$. As curvas são as trajetórias de diversas soluções $(x(t), y(t))$ das equações diferenciais do modelo.

Na situação mais à esquerda na Fig. 5, temos $a = 1.5$ e $b = 1.2$. As espécies são ambas muito agressivas. Como consequência, como mostra o quadro mais à esquerda da figura, as trajetórias parecem evitar o ponto de equilíbrio na interseção das isóclinas e vão terminar, ou no ponto $(1, 0)$, sobrevivência somente da espécie x , ou no ponto $(0, 1)$, sobrevivência somente da y . Uma análise qualitativa pode ser feita que mostra esse resultado: quando ambas as espécies são agressivas uma à outra e competem pelos recursos do ambiente, somente uma delas sobrevive. Este fato é conhecido em Ecologia como o *princípio da exclusão competitiva*.

No quadro central da Fig. 5 temos a situação em que $a = 0.8$ e $b = 0.5$. Nesse caso, as espécies são mutuamente pouco agressivas. Vê-se então que as trajetórias convergem para o ponto de equilíbrio da interseção das isóclinas. A conclusão é que quando as espécies não se agrirem, elas podem conviver em um mesmo ambiente e essa situação de convivência é estável.

Figura 5: Os gráficos exibem três situações distintas do modelo de competição entre espécies. Em cada um dos quadros, a reta inclinada em traço contínuo é a isóclina $x + ay = 1$ e a reta tracejada é a isóclina $bx + y = 1$. O quadro mais à esquerda ilustra as soluções do modelo com $a = 1.5$ e $b = 1.2$. O quadro central ilustra as soluções com $a = 0.8$ e $b = 0.5$. O quadro a direita é a ilustração de $a = 0.8$, $b = 1.5$.



No quadro mais à direita na Fig. 5, temos a situação $a = 0.8$ e $b = 1.5$. A espécie y é pouco agressiva e a espécie x é muito agressiva. Não existe aqui o ponto de equilíbrio na interseção das isóclinas no primeiro quadrante. Vê-se que todas as trajetórias levam à extinção da espécie y , pouco agressiva, e à sobrevivência da espécie x .

7 Modelos estocásticos

Os modelos ilustrados até agora são todos baseados em equações diferenciais e são, portanto, determinísticos. Isto quer dizer que o resultado final do modelo depende somente das condições iniciais, sem interveniência de nenhum fator fortuito.

Sabemos porém que processos biológicos podem depender de variáveis aleatórias. Em modelos, por exemplo, para evolução genética de populações, isto é comum. Em geral, mutações genéticas aparecem – em tempos aleatórios – em um único indivíduo em uma população. Caso a mutação seja benéfica, os mutantes terão mais descendentes que os não-mutantes. Se este é o caso, é de se esperar que em algum tempo os mutantes, inicialmente uma minoria, irão dominar a população. Mas isto pode não acontecer se os mutantes iniciais não tiverem sorte. Eles podem morrer antes de terem descendentes e a mutação, mesmo benéfica, pode ser perdida. Da mesma forma, mutações deletérias podem também, por sorte, se espalhar por uma população e dominá-la.

Portanto, a sobrevivência ou não de uma mutação em uma população é um fenômeno que deve ser estudado levando em conta fatores aleatórios. Uma boa referência para modelos para evolução genética de uma população é o livro de Ewens (1979).

Um exemplo simples para modelar a evolução genética de uma população é o processo estocástico de Wright-Fisher. Suponha uma população com tamanho fixo igual a N indivíduos. Cada indivíduo é, para uma determinada característica genética, de tipo A ou de tipo B. Consideramos que os indivíduos de tipo B possuem aptidão igual a 1 e que os indivíduos de tipo A possuem aptidão $r > 0$. Isto significa que os indivíduos A possuem em média r vezes mais descendentes que os de tipo B. Se $r > 1$, o tipo A é favorecido pela seleção natural com relação ao tipo B e deve aumentar sua proporção na população ao longo do tempo. Se $0 < r < 1$, a característica A é deletéria com relação a B e tende a ser extinta com a passagem do tempo. Caso $r = 1$, os indivíduos de tipo A e B possuem em média o mesmo número de descendentes e dizemos que a característica é *neutra*.

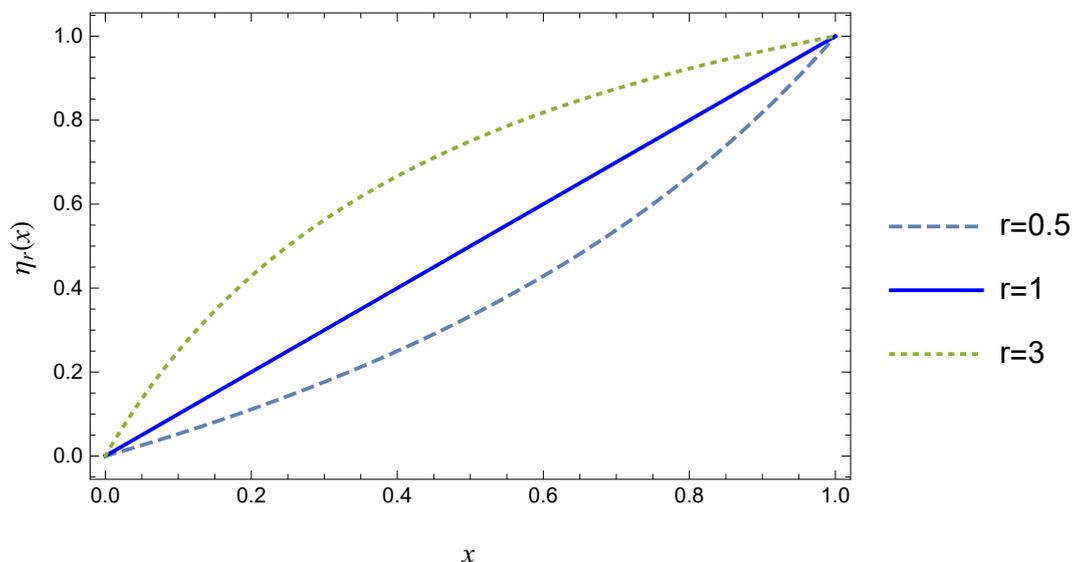
A versão mais simples do processo de Wright-Fisher, em que supomos que a reprodução é assexuada, funciona da seguinte maneira. Ao final de uma geração sorteia-se a composição da geração seguinte. O processo é que cada um dos N indivíduos da nova geração irá sortear de forma independente seu ancestral na geração anterior e herdará desse ancestral o tipo A ou B. O sorteio não é com iguais probabilidades: queremos que os indivíduos do tipo mais apto tenham maior probabilidade de serem sorteados como ancestrais. A maneira exata de atribuir pesos aos tipos é a seguinte: se x é a proporção de indivíduos de tipo A, a probabilidade de que um indivíduo sorteie um ancestral de tipo A será

$$\eta_r(x) = \frac{rx}{rx + 1 - x}.$$

Caso não sorteie um ancestral de tipo A, o indivíduo terá um ancestral de tipo B.

O número esperado de indivíduos A na próxima geração é $N\eta_r(x)$, mas o resultado do sorteio de uma geração é aleatório. A Fig. 6 ilustra o gráfico da função η_r para alguns valores de r . Caso $r > 1$, $\eta_r(x) > x$ para cada x entre 0 e 1, significando que os indivíduos de tipo A são selecionados com mais peso que os de tipo B no sorteio que irá formar a nova geração e provavelmente a proporção x de indivíduos A crescerá a cada nova geração. Observe-se, porém, que mesmo que r seja maior que 1, se x for pequeno $\eta_r(x)$ será pequeno também, indicando que o número de indivíduos de tipo A na geração seguinte será provavelmente pequeno. Ou seja, mesmo que sejam mais aptos que os de tipo B, o processo de Wright-Fisher corretamente prescreve que o número de indivíduos de tipo A muito provavelmente não irá aumentar de maneira súbita entre uma geração e outra.

Figura 6: Gráficos da função $\eta_r(x)$ para três valores de r . Note que o gráfico retilíneo é o do caso $r = 1$, em que $\eta_1(x) = x$.



A Fig. 7 mostra um exemplo de simulação do processo de Wright-Fisher com $r = 1.1$ em uma população com 100 indivíduos. O número inicial de indivíduos de tipo A é igual a 10. Observe que, apesar de os A serem mais aptos que os B e os B acabarem extintos após cerca de 60 gerações, às vezes o número de indivíduos A decresce de uma geração para outra.

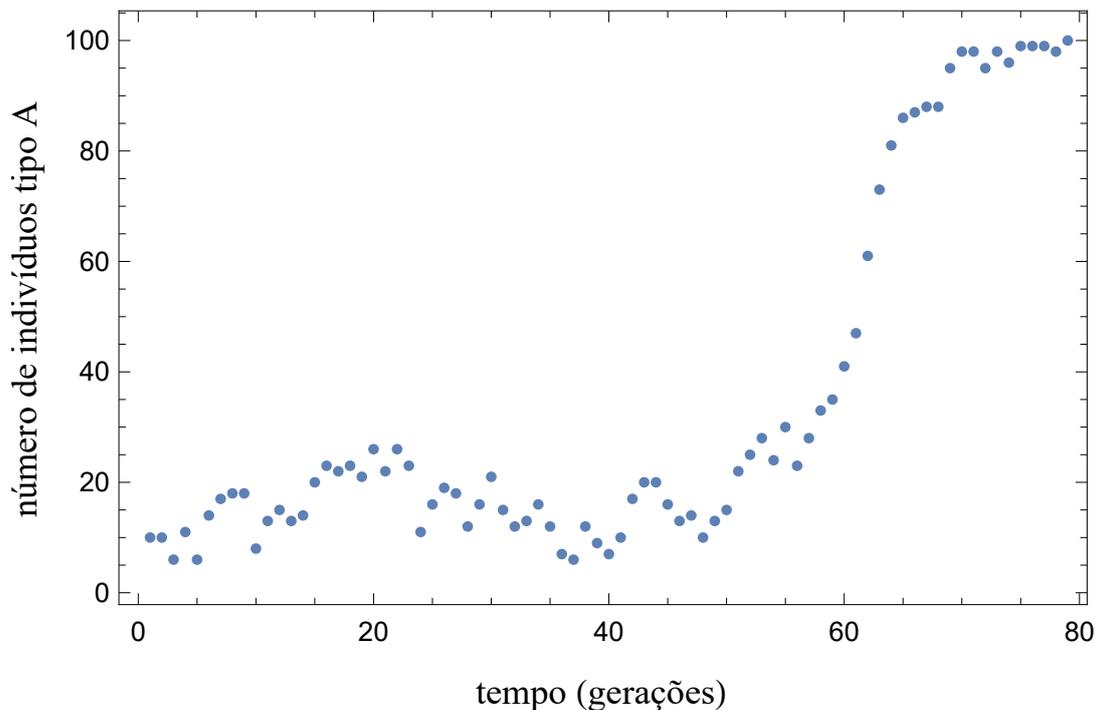
No caso neutro $r = 1$, quando a seleção natural não prefere nem A, nem B, pode-se ver facilmente que $\eta_r(x) = x$, significando que a proporção provável de indivíduos A na próxima geração é a mesma da geração atual.

O processo de Wright-Fisher faz parte de uma classe maior de processos estocásticos chamados cadeias de Markov. Da teoria das cadeias de Markov pode-se provar que com probabilidade 1, depois de um tempo aleatório, a população será toda do tipo A ou toda do tipo B. Diremos que haverá *fixação* do tipo A ou do tipo B.

Mesmo no caso neutro a fixação irá ocorrer com probabilidade 1. Este fenômeno é conhecido como *deriva genética*. Pode-se mostrar que para populações grandes, o tempo médio de fixação no caso neutro será substancialmente maior que no caso não-neutro.

A ocorrência de deriva genética é um fenômeno pouco intuitivo: se nem o tipo A é privilegiado sobre o tipo B, nem o contrário, por que um deles deverá ser extinto? No entanto, o modelo matemático descreve esta situação e ela é apontada como a explicação natural de grandes variações genéticas observadas em populações pequenas que ficam isoladas de uma população principal, por exemplo em ilhas.

Figura 7: Gráfico de uma simulação do processo de Wright-Fisher em uma população de 100 indivíduos de tipo A. Supomos que no tempo inicial há 10 indivíduos A e que $r = 1.1$, ou seja, em média os A têm 10% mais descendentes que os B.



A Fig. 8 ilustra duas simulações do processo de Wright-Fisher neutro em populações com 100 indivíduos. Em ambos os casos há inicialmente 50 indivíduos de tipo A e 50 de tipo B. No quadro da esquerda o tipo A foi fixado, enquanto no da direita, o tipo fixado foi o B.

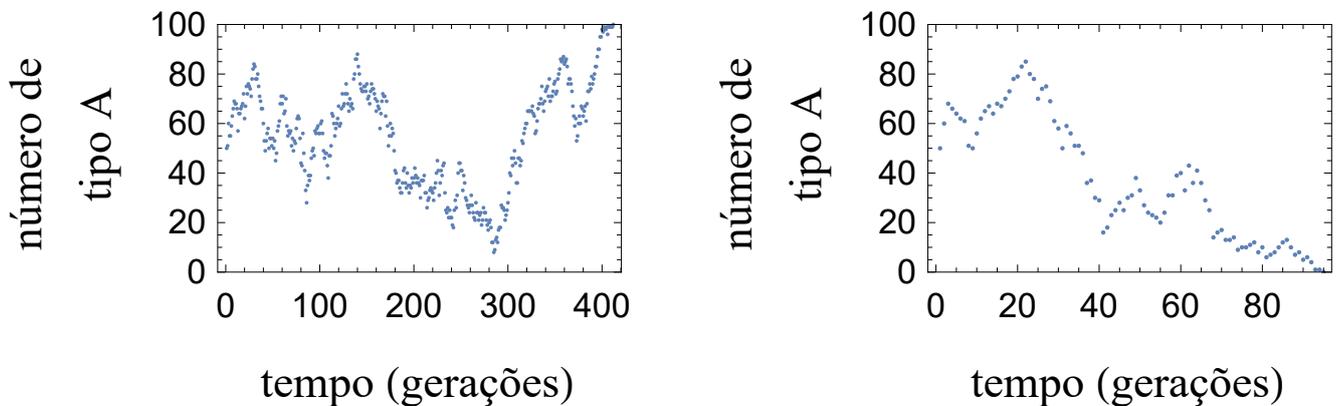
Uma referência razoavelmente elementar sobre cadeias de Markov e suas aplicações em Biologia é meu livro Neves (2014). Também nesse livro encontram-se mais detalhes que os que daremos aqui sobre uma aplicação recente do processo de Wright-Fisher à evolução da espécie humana. Nessa aplicação, o que se quer entender é o papel dos Neandertais na composição moderna da humanidade e explicar a sua extinção.

Até 2010 a imensa maioria dos pesquisadores acreditava na chamada *Teoria da Origem Única Africana*: a humanidade atual descendia unicamente de um pequeno grupo africano. A teoria, baseada principalmente em dados experimentais relativos ao DNA mitocondrial, dizia que não teria havido mistura genética desse grupo africano com outros grupos, por exemplo os Neandertais. Em Neves e Moreira (2006), eu e Carlos H. C. Moreira, da UFMG, afirmávamos que os dados de DNA mitocondrial não eram suficientes para se afirmar que não teria havido mistura genética com os Neandertais. Outro pesquisador que trabalhava na mesma linha era Maurizio Serva da Università dell'Aquila, Itália.

Em Green et al (2010) foi publicado um primeiro rascunho do DNA nuclear de um fóssil Neandertal e concluiu-se que cerca de 1 a 4% dos genes dos humanos atuais não-africanos são de origem Neandertal. Tal artigo jogou por terra a proposta radical da Teoria da Origem Única Africana e mudou o papel atribuído aos Neandertais na composição da humanidade atual. Além dos Neandertais, descobriu-se mais adiante que um outro grupo menos conhecido, os Denisovanos, também participou na composição genética dos humanos atuais.

Em Neves e Serva (2012), eu e Serva descrevemos um modelo para a mistura genética entre uma população de africanos anatomicamente modernos ancestrais dos humanos atuais e uma população de Neandertais. Usamos o processo de Wright-Fisher neutro como um modelo para a extinção dos Neandertais nesse encontro. Com este modelo, partindo do pressuposto de neutralidade entre africanos e Neandertais e do dado experimental de que os humanos atuais não-africanos possuem entre 1 e 4% de genes de origem Neandertal, pudemos quantificar que poucos eventos de reprodução entre modernos e Neandertais teriam sido suficientes para produzir os genes Neandertais encontrados hoje em grande parte dos humanos. Neste modelo, a extinção dos Neandertais teria sido puro azar, uma necessidade da deriva genética.

Figura 8: Duas simulações do processo de Wright-Fisher neutro em populações com 100 indivíduos, sendo inicialmente 50 de tipo A e 50 de tipo B.



8 Conclusões

A Biomatemática é hoje uma área bastante extensa. Muitos fenômenos biológicos que não descrevi aqui já foram modelados. E também, para os fenômenos aqui descritos, modelos mais completos podem ser formulados e análises mais avançadas podem ser feitas. Espero ter interessado algum leitor para que possa estudar algo mais sobre Biomatemática.

Referências

- Green et al, R. E. (2010). A draft sequence of the Neandertal genome. *Science*, 328, 710–722.
- Boyce, W. E., diPrima, R. C. (2015). *Equações Diferenciais Elementares e Problemas de Valores de Contorno*, 10º edn. LTC.
- Edelstein-Keshet, L. (2005). *Mathematical Models in Biology*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.
- Ewens, W. (1979). *Mathematical population genetics*. Biomathematics (Berlin), Springer-Verlag.
- Murray, J. D. (2002). *Mathematical Biology. An Introduction*, vol 1. Springer.
- Neves, A. G. M. (2014). *Aplicações biológicas de cadeias de Markov*. III Colóquio de Matemática da Região Norte.
- Neves, A. G. M., Moreira, C. H. C. (2006). Applications of the galton-watson process to human dna evolution and demography. *Physica A*, 368, 132.
- Neves, A. G. M., Serva, M. (2012). Extremely rare interbreeding events can explain Neanderthal DNA in modern humans. *PLoS ONE*, 7(10), e47,076.
- Stewart, J. (2017). *Cálculo*, 8º edn. Cengage Learning.
- Wikipedia (Acessado em 23/10/2019). Basic reproduction number. , URL https://en.wikipedia.org/wiki/Basic_reproduction_number.