

Modelos deterministas e estocásticos em Evolução

AULA

10

Meta da aula

Diferenciar determinismo e acaso na Evolução; relacionar o acaso com o destino evolutivo dos alelos em populações pequenas.

objetivos

Ao final desta aula, você deverá ser capaz de:

- Discriminar modelos deterministas e estocásticos em Ciência.
- Correlacionar tamanho amostral com variância.

INTRODUÇÃO

Nas próximas aulas, você aprenderá como a seleção natural pode determinar o destino evolutivo de um alelo. Você verá que se um alelo tiver vantagem adaptativa sobre outro (ou um menor coeficiente seletivo), sua frequência na população aumentará até que atinja a frequência na qual o nível adaptativo médio da população será maior. A evolução das frequências gênicas, nesses casos, será determinada pelas forças evolutivas agindo sobre os alelos (no caso, a seleção natural). Se o alelo *A* de um gene, por exemplo, tiver um valor adaptativo (w) de 5% a mais do que o outro (se *A* tiver $w = 1,05$, enquanto que *a* tenha $w = 1,00$), ele aumentará progressivamente até se fixar (**Figura 10.1**).

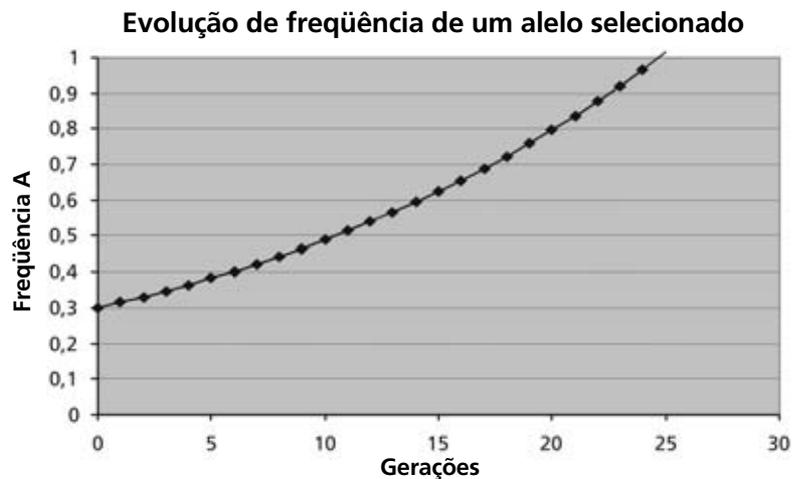


Figura 10.1

Modelos assim, em que o resultado da equação é calculado como média, e é mantido constante, são chamados *modelos deterministas*. Por que será que eles têm esse nome? Pense e responda:



Eles têm esse nome porque a evolução da variável (por exemplo, a frequência do alelo *A* no caso anterior) é *determinada* diretamente pela equação, em função de seus parâmetros (no nosso caso, o tempo e a vantagem adaptativa). A maior parte dos modelos que usamos em Ciência é determinista. Alguns exemplos bem conhecidos de modelos deterministas são aqueles de movimento dos corpos desenvolvidos por Newton, os modelos de cinética enzimática, em Bioquímica, e os modelos de crescimento populacional, em Ecologia.

Os modelos deterministas são extremamente úteis em Ciência, por representarem propriedades gerais dos fenômenos estudados.

Veja, por exemplo, a relação entre o comprimento total e o número de raios branquiais no bagre *Clarias anguillaris* (**Figura 10.2**). A partir dos dados apresentados, observe o gráfico e veja se pode responder: quantos rastos branquiais terão os bagres de 300mm? (dica: use uma régua).

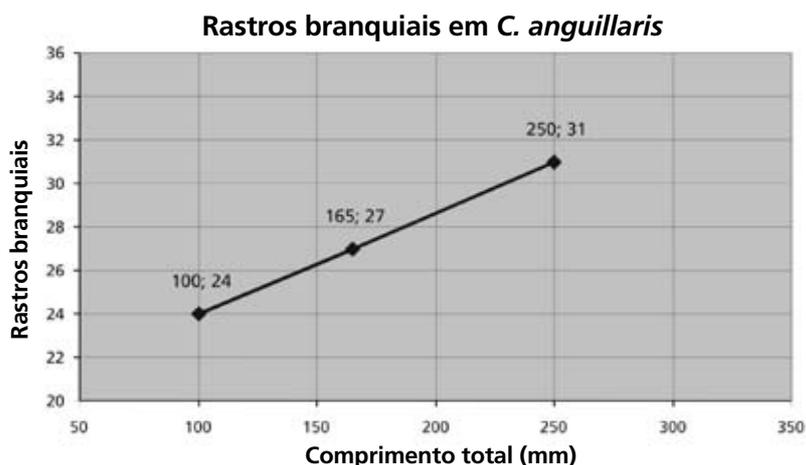


Figura 10.2



Você deve ter visto que, como a relação apresentada é linear, você pode *extrapolar* a partir da linha apresentada. Assim, o ponto em que a linha passa pelo valor projetado, no eixo das ordenadas, a 300mm de comprimento, vai corresponder, no eixo das abscissas, a 33 rastros branquiais.

Nos modelos deterministas, a probabilidade da mudança na variável é mantida constante e pode ser estimada, em um determinado ponto, a partir do comportamento daquela variável nos pontos anteriores. Ou seja, no caso da **Figura 10.1**, se tivéssemos somente os pontos entre a geração 0 e a geração 10 (quando a frequência de A aumentou de 0,30 para 0,49), poderíamos estimar, com uma boa confiança, os pontos futuros da geração 11 em diante. Modelos deterministas permitem facilmente que, a partir dos dados de uma curva, sejam feitas *extrapolações* (previsões de pontos exteriores a ela) e *interpolações* (previsões de pontos interiores a ela). Uma interpolação seria, por exemplo, prever que bagres de 200mm teriam cerca de 29 rastros branquiais (**Figura 10.2**).

Se por um lado os modelos deterministas são muito poderosos, por outro eles não levam em conta o acaso e consideram somente as tendências predominantes na evolução das variáveis. Considere, por exemplo, que a probabilidade de chuva para cada dia do mês de junho, na cidade do Rio de Janeiro, seja de 17% (www.inmet.gov.br). Considere também, para efeito de simplificação, que em cada dia de chuva caiam 16mm de água. Com Base nesses dados, qual seria sua previsão determinista para a quantidade de milímetro de água que cairia, em média, durante os meses de junho, no Rio de Janeiro?





ANDREI ANDREYEVICH MARKOV

Nasceu na Rússia em 1856 e ficou conhecido por seus trabalhos em Matemática e Estatística. Suas pesquisas principais foram sobre o desenvolvimento de modelos matemáticos, em que a evolução dos valores das variáveis é calculada de acordo com as probabilidades associadas a cada passo. A evolução dos valores dessas variáveis é conhecida como cadeia de Markov.

Como o mês de junho tem 30 dias e a probabilidade de chuva é de 17% diários, teríamos, deterministicamente, 5 dias de chuva ($30 \times 0,17$). Como dissemos que em cada dia de chuva caem 16mm de água, teríamos, em um mês médio, 80mm ($5 \text{ dias} \times 16\text{mm}$) de água de chuva caindo sobre o Rio de Janeiro. Poderíamos mesmo dizer que, em média, caem $80/30 = 2,7\text{mm}$ de água por dia durante o mês de junho no Rio de Janeiro. No entanto, apesar de esses valores serem médios e fundamentados em probabilidades, não podemos saber se em um dia específico (por exemplo, no dia de São João, 24 de junho) irá chover. O matemático russo **ANDREI MARKOV** decidiu estudar, no século XIX, como seria a evolução dos modelos matemáticos em que fossem consideradas as probabilidades de ocorrência dos eventos, não como médias do período total mas, sim, como o resultado da soma, passo a passo, desses eventos.

Os modelos em que a cada nova medida (cada dia do mês de junho, por exemplo) são estimadas as probabilidades dos eventos (como a chance de chover ou não) e o resultado líquido de tal probabilidade (milímetros de chuva) é computado e somado ao do evento anterior (no nosso caso, aumentando o total de milímetros de chuva acumulados no mês) são chamados de **modelos estocásticos**. Os modelos estocásticos são, mais complexos e mais próximos da realidade. Por outro lado, os gráficos resultantes da aplicação desses modelos não são bonitinhos, com linhas lisas retas ou curvas (como as **Figuras 10.1** e **10.2**). Ao incorporar o acaso a cada passo, os modelos estocásticos produzem linhas irregulares, mais difíceis de analisar. Quando o número de repetições de amostragem de um modelo estocástico tende ao infinito, transforma-se em modelo determinista, como veremos a seguir.

A cerveja e as cadeias de Markov

Entender modelos estocásticos é fundamental para compreender uma das forças evolutivas mais importantes que existem: a deriva gênica (que você verá na Aula 11). Então, para fixarmos bem as diferenças entre modelos deterministas e estocásticos, vamos falar de cerveja!

Imagine que um bêbado precise andar 200 metros do bar até sua casa. Vamos tentar modelar essa tarefa, em função da quantidade de cervejas ingeridas e suas conseqüências no sentido da orientação desse bêbado. Na verdade, como somos cientistas, vamos tentar generalizar esse modelo para o comportamento locomotor de todos os bêbados (assumindo que não existam diferenças individuais importantes entre eles, o que é um pressuposto grande, mas necessário para nossa modelagem). A unidade de progresso no nosso modelo será o número de metros que o bêbado conseguiu caminhar, em direção a sua casa, em 10 segundos.

Vamos considerar, também, que exista, a cada período de dez segundos, uma probabilidade X de que ele se dirija no sentido correto da casa, e que exista uma probabilidade Y (no caso, $Y = 1 - X$) de que ele se dirija no sentido oposto à sua casa, por estar desorientado (Figura 10.3).

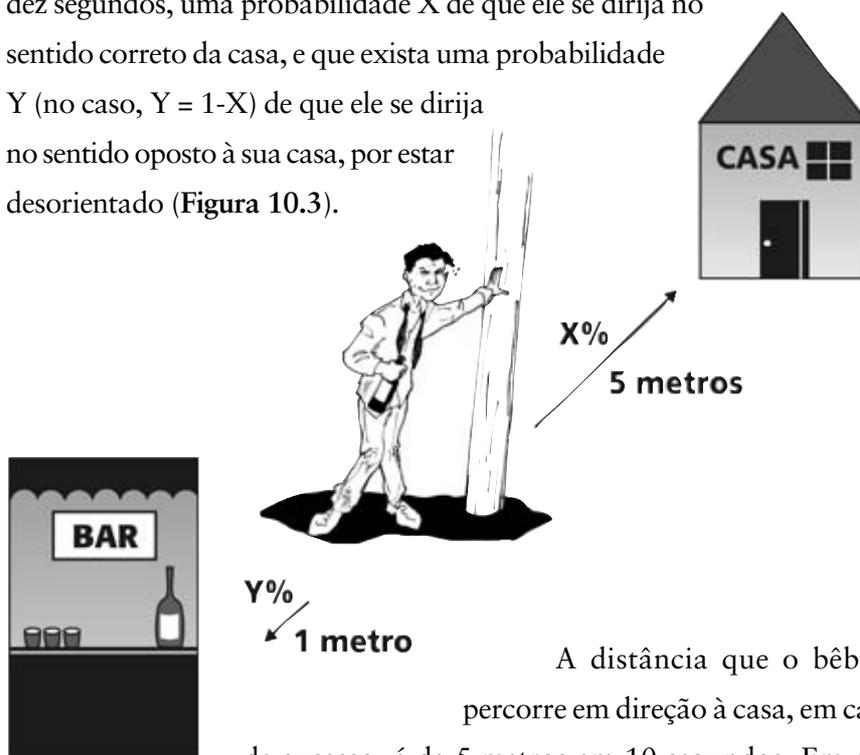


Figura 10.3: Os parâmetros de nosso modelo: A é a probabilidade de o bêbado deslocar-se em direção de sua casa; R é a probabilidade de ele desorientar-se e, conseqüentemente, deslocar-se para longe de casa.

A distância que o bêbado percorre em direção à casa, em casos de sucesso, é de 5 metros em 10 segundos. Em caso de desorientação, ele recua 1 metro naqueles 10 segundos. Finalmente, vamos dizer que X e Y dependem da quantidade de álcool ingerida, de modo que podemos criar uma tabela de probabilidades (Tabela 10.1):

Tabela 10.1: Probabilidades de avanços (A) ou retrocessos (R) em bêbados caminhando em direção às suas casas, em função do número C de garrafas de cerveja bebidas.

Cervejas (C)	A	R
2	0,95	0,05
4	0,80	0,20
6	0,60	0,40
8	0,30	0,70

Observe bem a tabela. Ela indica que a chance de caminhar na direção certa (A) de nossos bêbados padrão diminui drasticamente com o número de cervejas ingeridas. Tomar duas cervejas tem pouco efeito no deslocamento (apenas 5% de chance de se desviar do seu caminho), enquanto oito cervejas fazem com que os bêbados errem seu caminho 70% das vezes. Em Estatística, existe um termo para incorporar a probabilidade de um evento e seu resultado. Você lembra qual é esse termo?



Esse termo chama-se *Esperança*, e é calculado como o produto da probabilidade e do resultado. Assim, se a chance de alguém ganhar na loteria é de 0,02%, e o prêmio é de R\$ 1.000, a Esperança dessa pessoa ganhar é de 0,0002 (ou seja, 0,02% representado na escala de 0 a 1) vezes R\$ 1.000, ou $E_{(\text{ganhar})} = 0,0002 \times 1.000 = \text{R\$ } 0,20$. Ou seja, na média, se o bilhete da loteria custa R\$ 1, a pessoa terá um retorno de 20 centavos para cada real apostado (os outros 80 centavos são o lucro das pessoas que criaram o jogo, para cada real apostado).

No caso de pessoas que ingeriram quatro cervejas, qual a Esperança média de deslocamento? Nesse caso, temos duas probabilidades; veja a **Tabela 10.1** e lembre-se de que ir no sentido certo corresponde a andar cinco metros para a frente, e ir no sentido errado corresponde a 1 metro para trás. Calcule a Esperança para cada uma e some os dois resultados.



A probabilidade A de andar no sentido certo, após a ingestão de quatro cervejas, é de 0,80, e a de andar no sentido errado é de 0,20. Assim,

$$E_{(A)} = 0,80 \times (5 \text{ metros}) = 4 \text{ metros};$$

$$E_{(R)} = 0,20 \times (-1 \text{ metro}) = -0,20 \text{ metros.}$$

O resultado líquido dessas esperanças é que, a cada 10 segundos, uma pessoa que tomou quatro cervejas se deslocará $E_{(A)} + E_{(R)}$, ou seja, $4 - 0,20$ metros.

Assim, em seu cambalear, o bêbado se aproximaria da casa, em média, em 3,80 metros a cada 10 segundos, de modo que ele precisaria de $(200\text{m} / 3,80\text{m}) \times 10\text{segundos} = 526$ segundos (ou 8,8 minutos) para chegar a casa.

Então, vamos lá: calcule agora quantos minutos os bêbados levariam para chegar a casa, em função das quantidades diferentes de cerveja ingeridas. Primeiramente, vamos calcular as Esperanças para o número de metros percorridos a cada dez segundos para cada nível de embriaguez, ou seja, os valores de $E_{(2 \text{ cervejas})}$, $E_{(6 \text{ cervejas})}$ e $E_{(8 \text{ cervejas})}$. O caso da Esperança para quatro cervejas (3,80 metros a cada 10 segundos) nós já calculamos. Use a fórmula geral

$$E_{(C \text{ cervejas})} = (A_C \times 5) + (R_C \times -1) \text{ metros.}$$

O que é a mesma coisa que

$E_{(C \text{ cervejas})} = (A_C \times 5) - (R_C \times 1)$ metros, onde C é a linha correspondente, na **Tabela 10.1**, à quantidade de cervejas.

No nosso exemplo das quatro cervejas, $A_4 = 0,80$ e $R_4 = 0,20$. Assim, a conta seria:

$$E_{(4 \text{ cervejas})} = (0,80 \times 5) - (0,20 \times 1) = 4 - 0,2 = 3,8 \text{ metros}$$

Agora divide a distância do bar até a casa (200 metros) pela distância percorrida em 10 segundos para cada quantidade de cervejas (ou seja, as Esperanças). Como essas distâncias eram as esperadas para um deslocamento durante 10 segundos, você precisa multiplicar por 10, para ter o valor total em segundos (no caso das quatro cervejas, como vimos, seriam 526 segundos). Vamos lá, preencha a **Tabela 10.2** com seus resultados.

Tabela 10.2: A Esperança dos bêbados

Cervejas (C)	A	R	E(C)	Segundos para chegar a casa
2	0,95	0,05		
4	0,80	0,20	3,80	526
6	0,60	0,40		
8	0,30	0,70		



Tabela 10.3: Respostas da Esperança dos bêbados

Cervejas (C)	A	R	E(C)	Segundos	= Minutos
2	0,95	0,05	4,70	425	7,1
4	0,80	0,20	3,80	526	8,8
6	0,60	0,40	2,60	769	12,8
8	0,30	0,70	0,80	2.500	41,7

Esse modelo que fizemos é determinista ou estocástico?



Como estamos trabalhando com a Esperança média dos deslocamentos, o modelo é determinista. Assumimos que, a cada 10 segundos, um bêbado com 6 cervejas andará, em média, 2,6 metros em direção a casa. Não estamos levando em conta quanto ele andará nos primeiros 10 segundos ou entre os 50 e os 60 segundos desde que começou. No modelo determinista, o que interessa é a constância da regra ao longo do tempo. Veja, por exemplo, como ficariam esses deslocamentos entre bar e casa para os quatro níveis de consumo de

cerveja (Figura 10.4). As várias linhas são lisinhas, bem comportadas. Elas são baseadas em Esperanças, ou seja, são idealizações da relação entre consumo de cerveja e deslocamento, e representam o que seria o movimento de bêbados idealizados, que seguem perfeitamente o modelo que desenhamos.

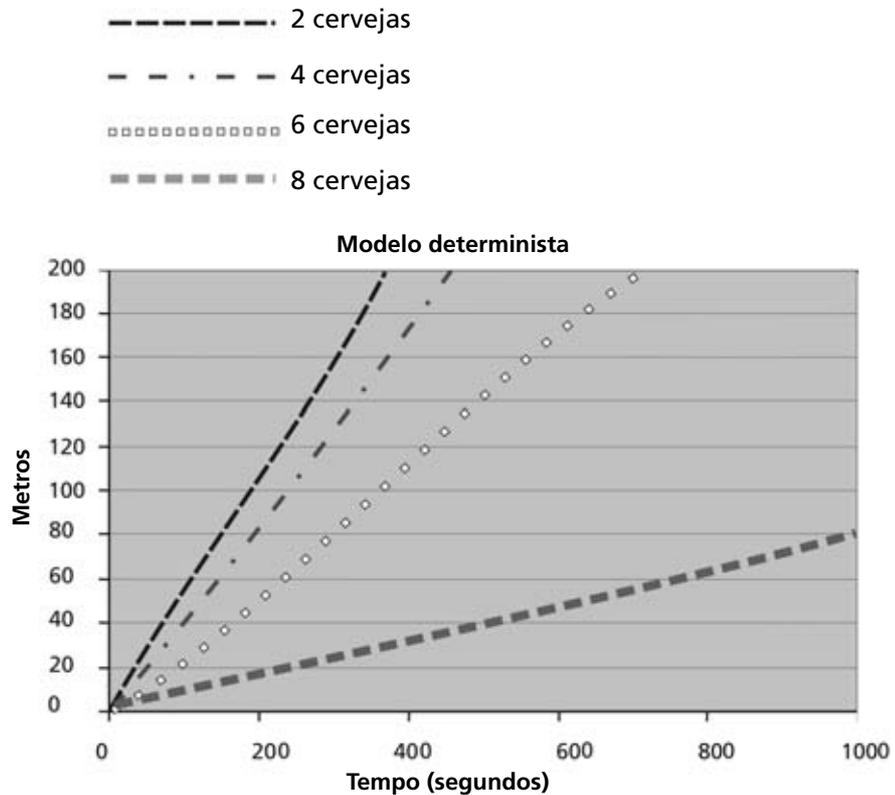


Figura 10.4: O bêbado e o determinista. Relações deterministas entre consumo de cerveja e o deslocamento entre o bar e casa.

Qual seria o equivalente estocástico a esse modelo?

Para um modelo estocástico, temos de considerar o resultado do movimento de nosso bêbado a cada período de tempo (no nosso caso, um período de 10 segundos). Como o resultado do movimento (ir para a frente 5 metros ou ir para trás 1 metro) tem duas probabilidades associadas (A e R no nosso modelo, que dependem da quantidade de cerveja ingerida), precisamos sortear, a cada período, um número aleatório entre 0 e 1, para decidir se, naquele período, o bêbado andou para a frente (aleatório < A) ou para trás (aleatório > A). Veja a Tabela 10.4. Nela calculamos, em períodos sucessivos de 10 segundos, os deslocamentos de uma dada pessoa que consumiu 6 garrafas de cerveja.

Tabela 10.4: Deslocamentos estocásticos de uma pessoa, de acordo com o modelo para ingestão de 6 garrafas de cerveja. Unidades de tempo = 10 segundos. Como, no modelo, a chance de andar na direção certa é de 60%, o sentido do deslocamento será considerado certo (5 metros para a frente) quando o número aleatório for menor ou igual a 0,60, e será considerado errado (1 metro para trás) quando for superior a 0,60.

Tempo	Número aleatório	Sentido	Metros	Percurso em direção a casa
1	0,601	errado	-1,0	-1
2	0,431	certo	5,0	4
3	0,438	certo	5,0	9
4	0,027	certo	5,0	14
5	0,567	certo	5,0	19
6	0,387	certo	5,0	24
7	0,631	errado	-1,0	23
8	0,530	certo	5,0	28
9	0,447	certo	5,0	33
10	0,992	errado	-1,0	32
11	0,131	certo	5,0	37
12	0,061	certo	5,0	42
13	0,259	certo	5,0	47
14	0,017	certo	5,0	52
15	0,686	errado	-1,0	51
16	0,661	errado	-1,0	50
17	0,411	certo	5,0	55
18	0,172	certo	5,0	60
19	0,675	errado	-1,0	59
20	0,014	certo	5,0	64
21	0,904	errado	-1,0	63
22	0,327	certo	5,0	68
23	0,266	certo	5,0	73
24	0,434	certo	5,0	78
25	0,078	certo	5,0	83
26	0,447	certo	5,0	88
27	0,929	errado	-1,0	87
28	0,438	certo	5,0	92
29	0,065	certo	5,0	97
30	0,800	errado	-1,0	96
31	0,226	certo	5,0	101
32	0,962	errado	-1,0	100
33	0,711	errado	-1,0	99
34	0,186	certo	5,0	104
35	0,430	certo	5,0	109
36	0,229	certo	5,0	114
37	0,521	certo	5,0	119
38	0,399	certo	5,0	124
39	0,220	certo	5,0	129
40	0,957	errado	-1,0	128
41	0,491	certo	5,0	133
42	0,684	errado	-1,0	132

43	0,434	certo	5,0	137
44	0,277	certo	5,0	142
45	0,335	certo	5,0	147
46	0,849	errado	-1,0	146
47	0,782	errado	-1,0	145
48	0,402	certo	5,0	150
49	0,621	errado	-1,0	149
50	0,677	errado	-1,0	148
51	0,51	certo	5,0	153
52	0,85	errado	-1,0	152
53	0,47	certo	5,0	157
54	0,11	certo	5,0	162
55	0,96	errado	-1,0	161
56	0,49	certo	5,0	166
57	0,12	certo	5,0	171
58	0,10	certo	5,0	176
59	0,39	certo	5,0	181
60	0,90	errado	-1,0	180
61	0,50	certo	5,0	185
62	0,26	certo	5,0	190
63	0,84	errado	-1,0	189
64	0,79	errado	-1,0	188
65	0,30	certo	5,0	193
66	0,06	certo	5,0	198
67	0,83	errado	-1,0	197
68	0,85	errado	-1,0	196
69	0,04	certo	5,0	201
70	0,24	certo	5,0	206
71	0,57	certo	5,0	211
72	0,21	certo	5,0	216
73	0,07	certo	5,0	221
74	0,25	certo	5,0	226
75	0,29	certo	5,0	231
76	0,82	errado	-1,0	230
77	0,76	errado	-1,0	229
78	0,52	certo	5,0	234
79	0,34	certo	5,0	239
80	0,47	certo	5,0	244
81	0,93	errado	-1,0	243
82	0,36	certo	5,0	248
83	0,41	certo	5,0	253
84	0,12	certo	5,0	258
85	0,84	errado	-1,0	257
86	0,90	errado	-1,0	256
87	0,30	certo	5,0	261
88	0,68	errado	-1,0	260
89	0,54	certo	5,0	265
90	0,94	errado	-1,0	264
91	0,80	errado	-1,0	263

92	0,84	errado	-1,0	262
93	0,20	certo	5,0	267
94	0,50	certo	5,0	272
95	0,38	certo	5,0	277
96	0,84	errado	-1,0	276
97	0,81	errado	-1,0	275
98	0,88	errado	-1,0	274
99	0,05	certo	5,0	279
100	0,82	errado	-1,0	278

Os números utilizados para as decisões das probabilidades foram obtidos a partir de um gerador de números aleatórios do programa Excel®, mas também poderiam ter sido retirados de tabelas de números aleatórios existentes em livros de Estatística. Lembre-se de que números aleatórios são números que variam ao acaso.

Repare que, se repetíssemos essa simulação, teríamos resultados diferentes. Essa é uma diferença importante entre os modelos estocásticos e os modelos deterministas: como os modelos deterministas são fundamentados em tendências constantes, sempre dão os mesmos resultados, desde que os parâmetros não sejam mudados. Os modelos estocásticos dão resultados diferentes para cada cálculo, mesmo quando os parâmetros são mantidos constantes.

Veamos como fica representada, graficamente, a evolução de nossa pessoa entre o bar e o lar, depois das 6 cervejas, segundo o modelo estocástico (Figura 10.5).

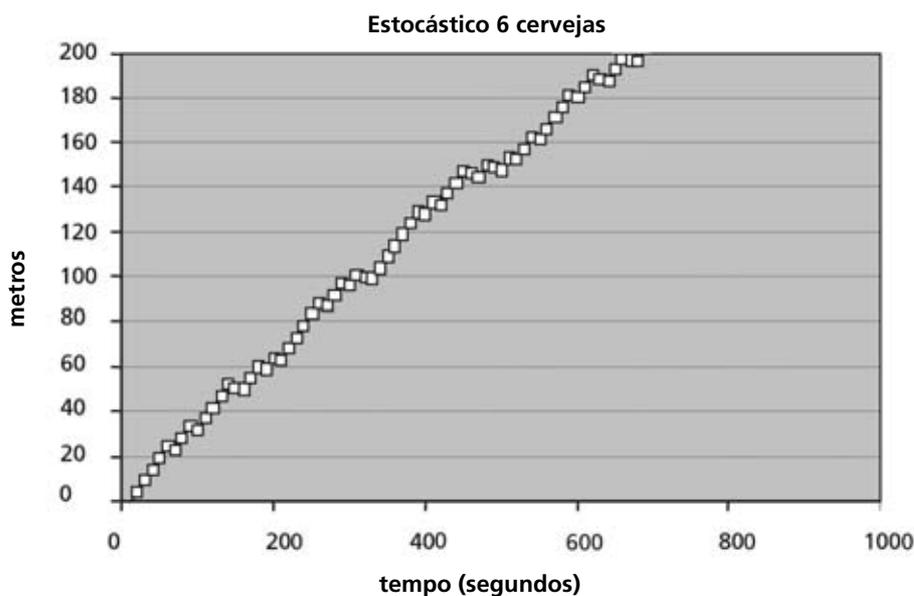


Figura 10.5: Deslocamentos estocásticos de uma pessoa entre o bar e sua casa, após seis garrafas de cerveja.

Repare que, ao contrário das linhas dos modelos deterministas, a linha da **Figura 10.6** não é reta ou regular, apesar de mostrar tendência ao deslocamento em direção à casa. Se refizéssemos dez vezes os cálculos de deslocamentos sob o efeito de 6 garrafas de cerveja, veríamos 10 trajetos diferentes, mas que oscilariam em torno de uma média, que seria uma linha bem mais regular do que as outras (**Figura 10.6**).

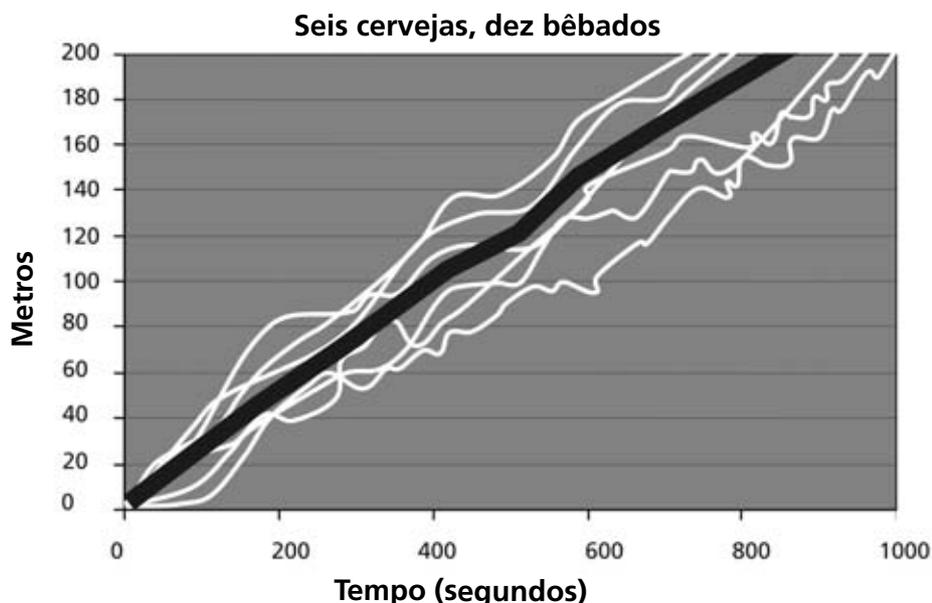


Figura 10.6: Deslocamento de 10 pessoas, conforme nosso modelo estocástico para movimentos de pessoas sob efeito de seis garrafas de cerveja. A linha grossa representa a média dos deslocamentos das 10 pessoas.

Entendeu, então, as diferenças entre os modelos deterministas e estocásticos?

Compare as **Figuras 10.4** e **10.5**. Se conhecêssemos apenas os deslocamentos nos primeiros 400 segundos do deslocamento da pessoa que tomou 6 cervejas, teríamos como determinar, a partir da **Figura 10.4**, quantos metros ele teria percorrido após um total de 500 segundos? E a partir da **Figura 10.5**?



Pois essa é a grande diferença entre os dois tipos de modelo! Como vimos antes, no modelo determinista, o conhecimento de uma parte da curva permite extrapolar ou interpolar outros valores de maneira bastante exata. Nos modelos estocásticos isso é impossível (apesar de podermos ter uma idéia geral do padrão).

Chega de cerveja!

Muito bem, usamos o nosso modelo dos deslocamentos sob o efeito do álcool para diferenciar, de maneira simples, as propriedades dos

dois tipos de modelo. Mas o que isso tem a ver com Evolução? Bom, a resposta é... tudo, é claro! Os cientistas dos vários campos da Ciência se preocupam com padrões gerais, que são bem representados por modelos deterministas. Mas, em Evolução, o acaso é uma coisa importante demais para ser desprezada. Mais do que um “ruído” nos gráficos, como seria visto por deterministas, a variação é a fonte da Evolução.

Mesmo que o comportamento dos genes, na média, acabe seguindo um modelo determinista, a evolução acontece a cada passo, como aquele bêbado que queria voltar para casa. Assim, se queremos entender como procede a evolução das espécies, devemos entender tanto dos modelos deterministas (como quando estudamos seleção natural) como dos estocásticos (quando estudamos deriva gênica, que será nossa próxima aula). A grande diferença entre o exemplo do bêbado e a evolução dos genes nas populações é que, no caso do bêbado, o que provocava a variação em seus passos era um efeito externo (a quantidade de álcool ingerido); no caso da evolução das populações, a maior fonte de variação aleatória é o número de indivíduos da população.

Por que será que os genes nas populações pequenas variam mais ao acaso do que nas populações grandes? Repare, de novo, na **Figura 10.6**. Cada bêbado segue um caminho diferente, e a média dos caminhos dos bêbados é bem mais regular que o caminho de cada um individualmente. Se tivéssemos 100 bêbados, as linhas continuariam bem espalhadas, mas a linha média seria ainda mais regular. Com um número infinito de bêbados, a linha média seria igual à linha esperada no modelo determinista da **Figura 10.4**. Por que a linha média de 10 bêbados é mais irregular que a linha média de 100 bêbados?



Porque as médias amostrais se aproximam cada vez mais das médias populacionais, quando se aumenta o tamanho da amostra. No nosso caso, a curva média populacional é a estimada pelo modelo determinista, e a curva média amostral é a calculada fazendo-se as médias dos deslocamentos dos nossos bêbados a cada passo (**Tabela 10.5**). No caso dos genes nas populações, a cada geração existe uma amostragem dos genes que irão fazer parte da geração seguinte; é como se fosse um sorteio, em que o que é sorteado são os alelos.

Imaginemos o caso de um loco com dois alelos. Vamos representar esses dois alelos por grãos de feijão de duas cores (feijão preto, que

chamaremos de "alelo P", e feijão carioquinha, "alelo C", por exemplo). Vamos imaginar uma população de 50 indivíduos (ou seja, 100 alelos – feijões – pois cada indivíduo é diplóide, ou seja, tem dois alelos para cada loco). Vamos dizer que as freqüências dos alelos sejam $f_p = 0,6$ e $f_c = 0,4$. Como seria a proporção de feijões P (pretos) e C (carioquinhas) nos meus 100 feijões?



AMOSTRAGEM COM REPOSIÇÃO

Em Estatística, temos dois tipos de amostragem. Quando cada amostra retirada não é colocada de volta, dizemos que a amostragem é sem reposição. Quando ela é colocada de volta, dizemos que é com reposição. Por que alguém iria se preocupar em colocar de volta alguma coisa que retirou para contar/medir/pesar etc. Na verdade, colocar de volta é uma maneira que temos de simular um tamanho infinito, se não, a cada vez que retiramos um objeto da amostragem, modificamos a probabilidade para a retirada do próximo objeto. Por exemplo, se temos três bolas brancas e sete bolas pretas em um saco, a probabilidade de retirar uma bola branca é de $3/10 = 30\%$. Digamos, então, que retiramos, ao acaso, uma bola do saco, e ela foi branca. Qual a chance de que a próxima bola seja branca? Se a amostragem fosse com reposição, a chance continuaria a ser de 30% . No entanto, se não for com reposição, a proporção de bolas brancas e pretas no saco vai ter se modificado, pois agora teremos 2 bolas brancas e 7 pretas, o que dá uma proporção de bolas brancas de $2/9 = 0.222$ ou $22,2\%$.

Se as freqüências são 0,6 e 0,4, isso significa que existem 60% de P e 40% de C na população. Como são 100 alelos, isso significa que teremos 60 alelos pretos e 40 alelos carioquinhas. Então faça isso: coloque 60 alelos P e 40 alelos C em um saco e misture bem. Essa é a nossa população. Vamos fazer um experimento de **AMOSTRAGEM COM REPOSIÇÃO**.

Se pegarmos 10 alelos, qual a Esperança para cada tipo de alelo que vamos retirar? Como a probabilidade de pegar cada tipo de alelo é a sua freqüência (ou seja, quanto mais freqüente é o tipo de alelo, mais provável é pegá-lo), as Esperanças são $E(P) = f_p \times N$; e $E(C) = f_c \times N$, onde N é o número de feijões que pegamos. As Esperanças, então, são de que peguemos 6 alelos pretos e 4 alelos carioquinhas, certo? Ou seja, em um modelo determinista, o esperado seria que, todas as vezes que eu retirasse 10 alelos, seis fossem P e quatro fossem C.

Vamos ver o que acontece quando incorporamos o acaso. Vamos amostrar 10 "alelos". Como a amostragem é com reposição, pegue os dez feijões, anote as cores e os devolva ao saco (sacudindo bem, para misturar). O que você encontrou? Faça isso de novo mais 9 vezes e preencha a **Tabela 10.5**. Faça também o total dos 10 experimentos de amostragem.

Tabela 10.5: Amostragem aleatória dos alelos P (preto) e C (carioquinha). São feitos 10 experimentos de amostragem de 10 feijões, com reposição.

	1	2	3	4	5	6	7	8	9	10	total
Alelo P											
Alelo C											

Qual a proporção média dos alelos P e C na sua amostragem?



A proporção média dos alelos é calculada pelo número total de cada alelo, dividido pelo total de alelos. No nosso caso, como foram amostrados 10 vezes 10 alelos, o total será de 100 alelos. Então $f_p = (\text{Total P}) / 100$.

Coloque no gráfico da **Figura 10.7** todos os pontos das suas 10 amostragens. Nas colunas “P” e “C” coloque pontos correspondentes às frequências dos alelos em cada repetição (se houver valores iguais, coloque um ao lado do outro). Você terá, no final, 10 pontos na coluna C e 10 pontos na coluna P. Repare que você deve registrar, no gráfico, as frequências relativas dos alelos P e C (ou seja, não coloque o número total de alelos, e sim a proporção deles em cada amostragem. No nosso caso, como eram 10 alelos, a proporção de cada alelo será a quantidade de alelos dividida por 10; ou seja, se você encontrou 3 feijões pretos, a frequência do alelo P é 3/10, ou seja, 0,3).

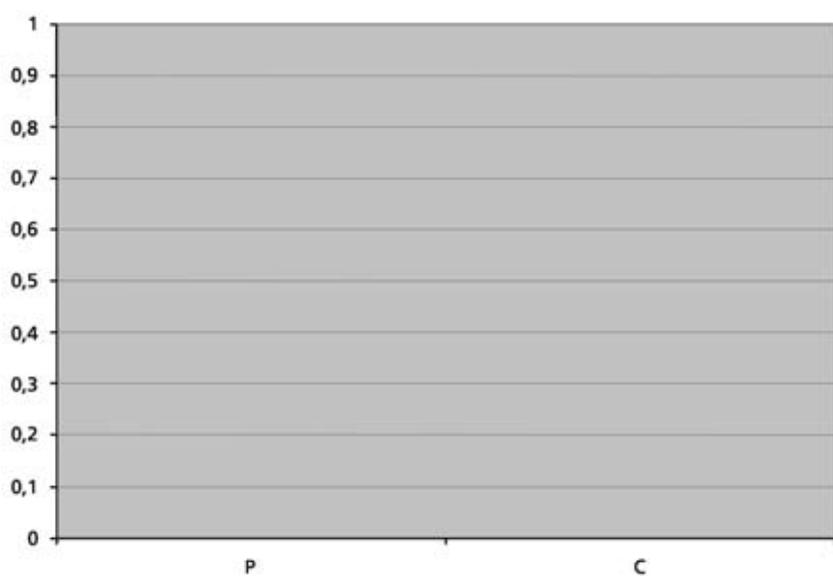


Figura 10.7: Frequências dos alelos P e C em 10 amostragens sucessivas, de 10 alelos, com reposição.

Nos dez casos, você deve ter tido desvios em relação ao esperado, que era de 6 alelos P e 4 alelos C. No entanto, é provável que a média dos alelos tenha se aproximado mais da Esperança do que cada uma das 10 amostragens. Isso é semelhante ao que você havia observado no caso dos bêbados da **Figura 10.6**. Coloque os valores das frequências médias de P e C na **Figura 10.7**, com uma outra cor para ressaltá-los.

Se você tivesse pego cinco feijões ao acaso, em vez de 10, o que você acha que teria acontecido com a variação das frequências de P e M em cada experimento? Vamos experimentar! (**Tabela 10.7**)

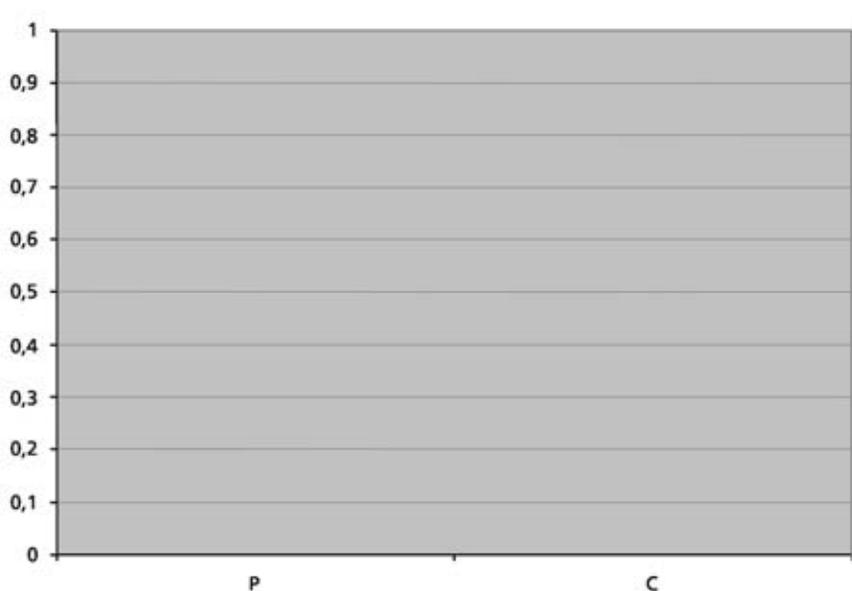
Tabela 10.7. Amostragem aleatória dos alelos P (preto) e C (carioquinha). São feitos 10 experimentos de amostragem de 5 feijões, com reposição.

	1	2	3	4	5	6	7	8	9	10	total
Alelo P											
Alelo C											

Não se esqueça de que a média, agora, é o total dos feijões dividido por 50.

Faça a mesma coisa que você fez com a **Figura 10.7** agora, na **Figura 10.8**. Não se esqueça de que os valores das frequências devem ser calculados como número de alelos dividido pelo total (agora 5 em vez de 10). Ou seja, se você encontrou 3 feijões-pretos, a frequência do alelo P é $3/5$, ou 0,6. Registre também, no gráfico, o valor correspondente às médias. Use a mesma cor que você usou para a média na **Figura 10.7**.

Figura10.8: Frequências dos alelos P e C em 10 amostragens sucessivas, cada uma de 5 alelos, com reposição.



Compare as Figuras 10.7 e 10.8. O que aconteceu com a variação dos pontos em torno da média em cada um? Provavelmente ela foi maior na Figura 10.8 do que na Figura 10.7. Essa variação representa a variância das frequências, que vai ser tanto menor quanto maior for o tamanho amostral (no nosso caso, o número de feijões retirados do saco). Então, não se esqueça: quanto maior for o número de alelos amostrados de cada vez, mais próximas serão as proporções dos alelos das proporções originais na população. Ou, quanto mais feijões a gente amostra, maior a chance de que a proporção de cada um seja parecida com as quantidades dentro do saco.

Na Aula 11, sobre deriva gênica, você verá como em populações pequenas os genes evoluem de maneira mais “bêbada”, enquanto em populações grandes eles evoluem de maneira mais regular.

RESUMO

Uma maneira eficiente de representar os fenômenos naturais é o uso de modelos. Os modelos em que o acaso não é considerado, de modo que o comportamento das variáveis resultantes é determinado por seus parâmetros, são chamados deterministas. Os modelos em que os resultados de cada ponto são estimados sem levar em conta os pontos anteriores são chamados *estocásticos*. Os modelos *estocásticos* levam em conta o acaso, e seu comportamento depende, fundamentalmente, do tamanho amostral a cada passo. Em outras palavras, eles dependem da variância amostral.

ATIVIDADES FINAIS

1. Qual a diferença entre modelos deterministas e estocásticos?

RESPOSTA COMENTADA

Modelos deterministas não levam em conta o acaso. Nos modelos deterministas, é possível fazer extrapolações e interpolações. Nos modelos estocásticos, o acaso é considerado a cada passo, e extrapolações e interpolações exatas não são possíveis.

2. A chance de alguém ganhar em dada loteria é de um em um milhão (esse valor parece pequeno; na verdade, é bem próximo das chances de ganhar na maioria das loterias federais). Se o prêmio for de trezentos mil reais e o custo da aposta for de R\$ 2, qual a Esperança de vitória, em reais, para cada real apostado?

RESPOSTA COMENTADA

A Esperança é calculada como Probabilidade vezes Resultado (no nosso caso, o prêmio em reais). Então, a Esperança, para cada aposta, é de um em um milhão ($1/1.000.000 = 0,000001$) vezes o prêmio (R\$ 300.000), ou seja, $300.000 \times 0,000001 = \text{R\$ } 0,30$ por aposta. Como a aposta custa R\$ 2, a Esperança, por real, é de R\$ 0,15 (os outros R\$ 0,85 são os lucros do governo, da loja etc.).

3. Você resolveu ver a proporção de torcedores do Flamengo e do Fluminense em um Fla x Flu. Para isso, ficou na roleta e contou quantas pessoas entravam no estádio com camisas rubro-negras e tricolores. Essa amostragem é com ou sem reposição?

RESPOSTA COMENTADA

É uma amostragem sem reposição: cada pessoa só passa uma vez pela roleta.

4. Vamos supor que, em uma população de índios brasileiros, tenham sido observadas as seguintes freqüências gênicas: $Rh+ = 0,45$ e $Rh- = 0,55$ nos anos 1960. Nos anos 1990, foram amostradas as freqüências dos mesmos genes em crianças desses índios, observando-se as freqüências: $Rh+ = 0,53$ e $Rh- = 0,47$. Um colega seu, ao ver esses resultados, falou que isso devia ser porque o alelo $Rh+$ era vantajoso na população, e aumentou por seleção natural. Que outras explicações você poderia dar para esse aumento?

RESPOSTA COMENTADA

O aumento da freqüência poderia ser devido à seleção, mas pode ter acontecido, também, por acaso, principalmente se se considerar que o tamanho populacional dos grupos indígenas é, em geral, bastante reduzido.

5. Em um estudo com populações de salmão, Hendry (2001) observou que, a $9^{\circ} C$, eram necessários 110 dias para que eclodissem 10% dos ovos. Depois de 120 dias, 80% dos ovos haviam eclodido. No intervalo entre 5% e 95% de eclosão, o autor observou que a relação entre tempo e porcentagem de eclosão era, na média, linear. Considerando-se um modelo determinista, quantos ovos eclodiram em 115 dias? E seguindo um modelo estocástico?

RESPOSTA COMENTADA

Em um modelo determinista, consideraríamos que os valores apresentados seriam dois pontos em uma reta. Assim, se em 10 dias (entre o dia 110 e o dia 120) a porcentagem de eclosão havia aumentado 70% (de 10% para 80%), então, em cinco dias ela iria aumentar a metade (35%). Ou seja, com 115 dias a porcentagem de eclosões seria de $10\% + 35\% = 45\%$. Em um modelo estocástico tal previsão não seria possível, pois seria necessário levar em conta a probabilidade, a cada dia, de haver eclosões, e essas probabilidades poderiam acontecer ou não, pois precisaríamos considerar o acaso.

AUTO-AVALIAÇÃO

Esta aula é fundamental para que você possa entender um dos processos mais importantes na evolução das populações naturais: a deriva gênica. Ela também é importante na questão filosófica que distingue os modelos deterministas e estocásticos. Se você entendeu bem essa diferença, poderá compreender mais facilmente vários modelos em Ecologia e outros ramos da Biologia. O mais importante, nesta aula, é verificar, com seus próprios cálculos, como a variância dos pontos em torno da média é maior quando o tamanho amostral é menor. Isso deve ter ficado claro na comparação entre as **Figuras 10.7 e 10.8**. Pode até ter acontecido, no seu caso, que essas variâncias não tenham sido muito diferentes (afinal, o acaso pode levar a muitas coisas estranhas!). Mas o mais provável é que a nuvem de pontos em volta do valor de 0,6 para o "alelo" P (os feijões pretos) tenha sido mais espalhada na **Figura 10.8** (tamanho amostral de 5) que na **Figura 10.7** (tamanho amostral de 10). Da mesma forma, a média na **Figura 10.7** deve ter ficado mais próxima do valor populacional original (0,6) do que a média na **Figura 10.8**. Observar os dados que você mesmo produziu é a melhor maneira de perceber esse fenômeno. Então, se você não preencheu as tabelas (talvez não haja feijões à mão neste momento), procure preenchê-las depois. Não é obrigatório que sejam feijões, podem ser pedrinhas, bolas-de-gude, botões, fichas etc. O importante é que sejam mais ou menos do mesmo tamanho, para que o sorteio seja ao acaso.

Evolução

Referências

ANDRADE, Carlos Drummond de. No meio do caminho. *Revista Antropofagia*, São Paulo, 1928.

AGNESE, J. F., TEUGELS G. G., GALBUSERA P., GUYOMARD, R., VOLCKAERT F. *Morphometric and genetic characterization of sympatric populations of Clarias gariepinus and C. anguillaris from Senegal. Journal of Fish Biology* 50, p. 1143-1157, 1997

AVISE, J. C., *Molecular markers, Natural history and Evolution*. London: Chapman & Hall, 1994.

CARLSON, E. A., The Drosophila group. *Genetics*, Maryland, v. 79, p. 15-27, 1975.

CHAMBERS, R. *Vestiges of the natural history of creation and other evolutionary writings*. London: The University of Chicago Press, 1845.

COYNE, J. A., Ernest Mayr and the origin of species. *Evolution*, v. 48, n. 1, p. 19-30. 1994.

CROW, J. F., Population genetics history: A personal view. *Annual Review of Genetics*, v. 21, p. 1-22, 1987.

DARWIN, C. R. *A origem das espécies*. Belo Horizonte: Editora Itatiaia, 2002.

DOBZHANSKY, T., *Genética do processo evolutivo*. São Paulo: EDUSP/Editora Polígono, 1973.

FREIRE-MAIA, N. *Teoria da evolução: de Darwin a teoria sintética*. Belo Horizonte: Editora Itatiaia, 1988.

FISHER, R.A. *The genetical theory of natural selection*. Oxford: Oxford University Press, 1930.

FUTUYAMA, D. J., *Biologia Evolutiva*. 2. ed. Ribeirão Preto: Sociedade Brasileira de Genética/CNPq, 1992.

GRIFFITHS, Anthony J.F., *et al. Introdução à genética*. 7. ed. Rio de Janeiro: Guanabara Koogan, 2002, 794p.

GOULD, S. J., Ernest Mayr and the centrality of species. *Evolution*, v. 48, n. 1, p. 31-35, 1994. *Ever since Darwin*. London: Penguin Books, 1977.

HALDANE, J. B. S., *The causes of evolution*. New Jersey: Princeton University Press, 1990.

HARTL, Daniel L.; CLARK, Andy G. *Principles of population genetics*. 3. ed. Sunderland: Sinauer Associates, 1997, 542 p.

HENDRY ANDREW P. Adaptive divergence and the evolution of reproductive isolation in the wild: an empirical demonstration using introduced sockeye salmon. *Genetica* 112, pp. 515–534. 2001

HILLIS, D. M., J. J. BULL, White, M.E., Badgett, M.R., Molineux, I.J. (1992) Experimental phylogenetics: Generation of a known phylogeny. *Science* 255: 589-592.

HILLIS, D.M.; MORITZ, C. & MABLE, B.K. *Molecular Systematics*. 2 ed. Sunderland, Massachusetts: Sinauer Associates, Inc, 1996

MACHUCA, R., JORGENSEN, L.B., THEILADE, P., NIELSEN, C. Molecular investigation of transmission of Human Immunodeficiency Virus Type 1 in a criminal Case. *Clinical and Diagnostic Laboratory Immunology* 8: 884-890, 2001

MATIOLI, S. R. *Biologia Molecular e Evolução*. Ribeirão Preto: Holos Editora, 2001.

MAYR, E. *Populações, espécies e evolução*. São Paulo: EDUSP, 1977. What was the Evolutionary Syntesis? *Trends in Ecology and Evolution*, v. 8, n. 1, p. 31-34, 1993.

SILVA, E.P. & RUSSO, C. Techniques and statistical data analysis in molecular population genetics. *Hydrobiologia*, 420:119-135, 2000.

SILVA, E.P. Genética Marinha. In: PEREIRA, R.C. & SOARES-GOMES, A. (ed.). *Biologia Marinha*. Rio de Janeiro: Editora Interciência. Pp. 333-351, 2002

SNUSTAD, D. Peter.; SIMMONS, Michael J. *Fundamentos de genética*. 2. ed. Rio de Janeiro: Guanabara Koogan, 2001. 756 p.

TANAKA, Y., ASHIKARI, T., SHIBANO, Y., AMACHI, T., YOSHIZUMI, H., & MATSUBARA, H. (Construction of a human cytochrome c gene and its functional expression in *Saccharomyces cerevisiae*. *J Biochem* 103: 954-61, 1988

WRIGHT, S. Evolution in Mendelian populations. *Genetics*, Maryland, v. 16, p. 97-159, 1931.

ZHU, T., B. KORBER, B. “An African HIV-1 sequence from 1959 and implications for the origin of the epidemic.” *Nature* 391: 594-597, 1998

ISBN 85-7648-065-4



9 788576 480655



UENF
Universidade Estadual
do Norte Fluminense



Universidade Federal Fluminense
uff



UNIRIO



Ministério
da Educação

BRASIL
UM PAÍS DE TODOS
GOVERNO FEDERAL