

DERIVA GENÉTICA

Seja uma população de tamanho finito N , constante ao longo das gerações; sejam ainda p_0 e q_0 as frequências dos alelos $\underline{A}$ e $\underline{a}$ de um loco autossômico na geração 0; como o tamanho da população é constante, a geração 1 é formada da união de $2N$ gametas ao acaso dentre os indivíduos da geração 0:

$$(p_0 + q_0)^{2N};$$

q_1 pode tomar, portanto, qualquer um dos $(2N+1)$ valores seguintes:

$$\frac{0}{2N}; \frac{1}{2N}; \frac{2}{2N}; \dots; \frac{2N-2}{2N}; \frac{2N-1}{2N}; \frac{2N}{2N}$$

A probabilidade de que q tome o valor particular $q_j = j/2N$ é

$$\binom{2N}{j} \cdot p^{2N-j} \cdot q^j = \binom{2N}{j} \cdot (1-q)^{2N-j} \cdot q^j$$

$$\text{Onde } \binom{2N}{j} = \frac{(2N)!}{j!(2N-j)!} \quad (\text{combinação de } 2N \text{ elementos } j \text{ a } j)$$

Seja o seguinte exemplo: $N = 2$, $2N = 4$ genes; $p_0 = q_0 = 1/2$

$f(A) = p = 1$	$3/4$	$1/2$	$1/4$	0
$f(a) = q = 0$	$1/4$	$1/2$	$3/4$	1
estado $j = 0$	1	2	3	4

As probabilidades de que a população 1 esteja nos estados $j = 0, 1, 2, 3$ ou 4 são, respectivamente,

$$\begin{aligned} 0: & (1/2)^4 = 1/16 \\ 1: & 4(1/2)^3(1/2) = 1/4 \\ 2: & 6(1/2)^2(1/2)^2 = 3/8 \\ 3: & 4(1/2)(1/2)^3 = 1/4 \\ 4: & (1/2)^4 = 1/16 \end{aligned}$$

o que define o vetor da linha

$$Q(1) = (1/16 \ 1/4 \ 3/8 \ 1/4 \ 1/16).$$

Se a população 1 estiver no estado $j = 0$ ($p_1 = 1$, $q_1 = 0$), o que ocorre com uma probabilidade de $1/16$, as probabilidades de que a população 2 esteja nos estados $j = 0, 1, 2, 3, 4$ são, respectivamente,

$$\begin{aligned} 0: & 1 \\ 1: & 0 \\ 2: & 0 \\ 3: & 0 \\ 4: & 0. \end{aligned}$$

Se a população 1 estiver no estado $j=1$ ($p_1 = 3/4$, $q_1 = 1/4$), o que ocorre com uma probabilidade de $1/4$, as probabilidades de que a população 2 esteja nos estados $j = 0, 1, 2, 3, 4$ são, respectivamente,

$$\begin{aligned}0: & (3/4)^4 = 81/256 \\1: & 4(3/4)^3(1/4) = 27/64 \\2: & 6(3/4)^2(1/4)^2 = 27/128 \\3: & 4(3/4)(1/4)^3 = 3/64 \\4: & (1/4)^4 = 1/256.\end{aligned}$$

Se a população 1 estiver no estado $j=2$ ($p_1 = q_1 = 1/2$), o que ocorre com uma probabilidade de $3/8$, as probabilidades de que a população 2 esteja nos estados $j=0, 1, 2, 3, 4$ são respectivamente,

$$\begin{aligned}0: & (1/2)^4 = 1/16 \\1: & 4(1/2)^3(1/2) = 1/4 \\2: & 6(1/2)^2(1/2)^2 = 3/8 \\3: & 4(1/2)(1/2)^3 = 1/4 \\4: & (1/2)^4 = 1/16.\end{aligned}$$

Se a população 1 estiver no estado $j=3$ ($p_1 = 1/4$, $q_1 = 3/4$), o que ocorre com uma probabilidade de $1/4$, as probabilidades de que a população 2 esteja nos estados $j = 0, 1, 2, 3, 4$ são, respectivamente,

$$\begin{aligned}0: & (1/4)^4 = 1/256 \\1: & 4(1/4)^3(3/4) = 3/64 \\2: & 6(1/4)^2(3/4)^2 = 27/128 \\3: & 4(1/4)(3/4)^3 = 27/64 \\4: & (3/4)^4 = 81/256.\end{aligned}$$

Se a população 1 estiver no estado $j=4$ ($p_1 = 0$, $q_1 = 1$), o que ocorre com uma probabilidade de $1/16$, as probabilidades de que a população 2 esteja nos estados $j=0, 1, 2, 3, 4$ são, respectivamente,

$$\begin{aligned}0: & 0 \\1: & 0 \\2: & 0 \\3: & 0 \\4: & 1.\end{aligned}$$

Logo, as probabilidades de que a população 2 esteja nos estados $j = 0, 1, 2, 3, 4$ são, respectivamente,

$$0: \frac{1}{16} \times 1 + \frac{1}{4} \times \frac{81}{256} + \frac{3}{8} \times \frac{1}{16} + \frac{1}{4} \times \frac{1}{256} + \frac{1}{16} \\ \times 0 = \frac{85}{512} = 0,166016$$

$$1: \frac{1}{16} \times 0 + \frac{1}{4} \times \frac{27}{64} + \frac{3}{8} \times \frac{1}{4} + \frac{1}{4} \times \frac{3}{64} + \frac{1}{16} \\ \times 0 = \frac{27}{128} = 0,210938$$

$$2: \frac{1}{16} \times 0 + \frac{1}{4} \times \frac{27}{128} + \frac{3}{8} \times \frac{3}{8} + \frac{1}{4} \times \frac{27}{128} + \frac{1}{16} \\ \times 0 = \frac{63}{256} = 0,246094$$

$$3: \frac{1}{16} \times 0 + \frac{1}{4} \times \frac{3}{64} + \frac{3}{8} \times \frac{1}{4} + \frac{1}{4} \times \frac{27}{64} + \frac{1}{16} \\ \times 0 = \frac{27}{128} = 0,210938$$

$$4: \frac{1}{16} \times 0 + \frac{1}{4} \times \frac{1}{256} + \frac{3}{8} \times \frac{1}{16} + \frac{1}{4} \times \frac{81}{256} + \frac{1}{16} \\ \times 1 = \frac{85}{512} = 0,166016$$

o que define o vetor de linha

$$Q(2) = \left(\frac{85}{512} \quad \frac{27}{128} \quad \frac{63}{256} \quad \frac{27}{128} \quad \frac{85}{512} \right).$$

Sob forma matricial, as operações podem ser reescritas como

$$\begin{bmatrix} \frac{1}{16} & \frac{1}{4} & \frac{3}{8} & \frac{1}{4} & \frac{1}{16} \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{81}{256} & \frac{27}{64} & \frac{27}{128} & \frac{3}{64} & \frac{1}{256} \\ \frac{1}{16} & \frac{1}{4} & \frac{3}{8} & \frac{1}{4} & \frac{1}{16} \\ \frac{1}{256} & \frac{3}{64} & \frac{27}{128} & \frac{27}{64} & \frac{81}{256} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\ = \begin{bmatrix} \frac{85}{512} & \frac{27}{128} & \frac{63}{256} & \frac{27}{128} & \frac{85}{512} \end{bmatrix}$$

ou, abreviadamente, $Q(1) \cdot \underline{T} = Q(2)$, em que $\underline{T}$ é uma matriz transicional de probabilidades condicionais (matrizes desse tipo caracterizam-se por suas linhas somarem 1).

Generalizando, $Q(n) \cdot \underline{T} = Q(n+1)$.

O que foi visto foi a análise de um processo em que, dadas as condições de uma determinada população (tamanho e frequência), podemos determinar as probabilidades da população estar na mesma condição (frequências gênicas iguais) ou em condições diferentes. Como a deriva genética é um processo de amostragem casual, não podemos prever o que pode acontecer com a frequência gênica de uma determinada população pequena.

O que pode ser feito, no entanto, é estudar o comportamento de um número muito grande de populações com mesmo tamanho, em que podemos esperar que algumas aumentem as frequências gênicas, outras diminuam e outras permaneçam com frequências gênicas iguais.

A teoria da Estatística nos fornece meios de prever a dispersão das frequências gênicas em muitas populações. Para isso, usamos a medida da variância, na n -ésima geração, em um grupo de populações que na geração 0 têm as mesmas frequências gênicas. A previsão da variância no decorrer das gerações exige conhecimentos avançados de Estatística, mas está representada abaixo apenas para ilustração:

Na geração 0, não há variação, todas as frequências são idênticas, portanto:

$$\sigma_0^2 = 0$$

Na primeira geração (da distribuição binomial):

$$\sigma_1^2 = \frac{q_0(1-q_0)}{2N}$$

A média das frequências é igual a esperança:

$$E(q_1) = q_0 = q$$

e a variância:

$$\begin{aligned}\sigma_1^2 &= E(q_1^2) - q^2 = E(q_1^2) - q_0^2 \\ \therefore E(q_1^2) &= \sigma_1^2 + q_0^2 = q_0^2 + \frac{q_0(1-q_0)}{2N}\end{aligned}$$

A proporção de heterozigotos esperada na geração 1 é:

$$\begin{aligned}h_1 &= E(2p_1q_1) = 2E(q_1) - 2E(q_1^2) = 2q_0 - 2q_0^2 - \frac{2q_0(1-q_0)}{2N} = \\ 2q_0(1-q_0) - \frac{2q_0(1-q_0)}{2N} &= 2q_0(1-q_0)\left(1 - \frac{1}{2N}\right) = \\ h_0\left(1 - \frac{1}{2N}\right) &\therefore\end{aligned}$$

como a proporção de heterozigotos de uma geração é a multiplicação da quantidade de heterozigotos da geração seguinte por uma constante, na n-ésima geração:

$$h_n = h_0\left(1 - \frac{1}{2N}\right)^n$$

A variância na n-ésima geração:

$$\sigma_n^2 = E(q_n^2) - q_0^2$$

Colocando o termo $E(q_n^2)$ em termos de h_n (que já conhecemos) e q_0 :

$$\begin{aligned}h_n &= E(2p_nq_n) = E[2q_n(1-q_n)] = 2E(q_n) - 2E(q_n^2) = 2q_0 - 2E(q_n^2) \\ E(q_n^2) &= \frac{2q_0 - h_n}{2}\end{aligned}$$

$$\begin{aligned}
\sigma_n^2 &= q_0 - \frac{h_n}{2} - q_0^2 = q_0 - q_0^2 - q_0(1 - q_0) \left(1 - \frac{1}{2N}\right)^n = \\
&= q_0(1 - q_0) - q_0(1 - q_0) \left(1 - \frac{1}{2n}\right)^n = \\
&= q_0(1 - q_0) \left[1 - \left(1 - \frac{1}{2N}\right)^n\right]
\end{aligned}$$

$$\boxed{\sigma_n^2 = p_0 q_0 \left[1 - \left(1 - \frac{1}{2N}\right)^n\right]}$$

O limite de σ_n^2 quando n tende a infinito é $q_0(1 - q_0)$.

A Tabela abaixo mostra, para um número infinito de populações compostas, cada uma, por $N = 2$ indivíduos com $p_0 = q_0 = 1/2$ na geração inicial, os valores das probabilidades dos estados $j = 0, 1, 2, 3, 4$ e da variância σ_n^2 , calculados segundo os métodos mostrados anteriormente.

j	0	1	2	3	4	σ_n^2
geração						
0	0,000000	0,000000	1,000000	0,000000	0,000000	0,000000
1	0,062500	0,250000	0,375000	0,250000	0,062500	0,062500
2	0,166016	0,210938	0,246094	0,210938	0,166016	0,109375
3	0,248962	0,160400	0,181274	0,160400	0,248962	0,144531
4	0,311670	0,120506	0,135647	0,120506	0,311670	0,170898
5	0,358748	0,090399	0,101706	0,090399	0,358748	0,190674
10	0,466480	0,021453	0,024124	0,021453	0,466480	0,235922
15	0,492046	0,005091	0,005727	0,005091	0,492046	0,246659
20	0,498112	0,001208	0,001359	0,001208	0,498112	0,249207
25	0,499552	0,000287	0,000323	0,000287	0,499552	0,249812
∞	0,500000	0,000000	0,000000	0,000000	0,500000	0,250000