3rd term — Population genetics

Genetic drift and mutation

Problems

(1)

A p: 0'875

a q: 0'125

4 individuals → 8 alleles

No a → 8A → $0'875^8 = 0'344$ → probability of no a ($P(A) = P(\text{no a})$)

No A → 8a → $0'125^8 = 5'96 \cdot 10^{-8}$ → probability of no A ($P(a) = P(\text{no A})$)

~~$p^2 + 2pq + q^2 = 1 \rightarrow 2pq = 1 - p^2 + q^2 \rightarrow 2pq = 1 - (0'125^2 + 0'875^2) = 0'219$~~

$P(A \cap a)$ $P(A \cap a) = 1 - P(A) - P(a) \rightarrow P(A \cap a) = 1 - 0'344 - 5'96 \cdot 10^{-8} = 0'656$

(2)

Aa → 12 individuals → 24 alleles → all heterozygous

90% heterozygous → number of generations → $H_o = 1$, $H_t = 1 - 0'9 = 0'1$ → 10% heterozygous

$H_t = H_o \left(1 - \frac{1}{2N}\right)^t \rightarrow \ln\left(\frac{H_t}{H_o}\right) = t \cdot \ln\left(1 - \frac{1}{2N}\right) \rightarrow t = \frac{\ln(H_t / H_o)}{\ln(1 - (1/2N))} \rightarrow$

$t = \frac{\ln(0'1/1)}{\ln(1 - (1/24))} = 54'1$ → it will take 54 generations for the population to be heterozygous in 10%

$N = 240$

$t = \frac{\ln(0'1/1)}{\ln(1 - (1/2 \cdot 240))} = 1104'1$ → in 1104 generations of a population of 240 individuals you will have 10% heterozygosity

(3) $N = 95$

1 new allele

$P_o \rightarrow \frac{1}{2N} = P_o \rightarrow \frac{1}{2 \cdot 95} = 0'0053$ → there is a 0'53% probability that it will drift to fixation

$P_{lost} = 1 - P_o \rightarrow P_{lost} = 1 - 0'0053 = 0'995$ → There is a 99'5% probability of loss

$T_{fix} = -4N \frac{(1-p)\ln(1-p)}{p} = -4 \cdot 95 \frac{(1-0'0053)\ln(0'995)}{0'0053} = 378'9$ → Average fixation time 379 generations

$T_{loss} = -4N \frac{0'995 \ln(0'995)}{1 - 0'995} = 357'6$ → Average allele loss time is of 358 generations

④

$P_A = 0'75$

$N = 40$

$P_A' = 0'75$

$P(k \text{ alleles in } n \text{ alleles}) = \frac{n!}{k!(n-k)!} \cdot p^k \cdot q^{n-k}$

$k = 2N \cdot P_A' \rightarrow k = 2 \cdot 40 \cdot 0'75 = 60$

$q = 1 - p \rightarrow q = 1 - 0'75 = 0'25$

$n = 2N \Rightarrow n = 2 \cdot 40 = 80$

$P(60 \text{ in } 80) = \frac{80!}{60!(80-60)!} \cdot 0'75^{60} \cdot 0'25^{(80-60)} = 0'1025$

There is a 10'25% chance that in the next generation the allele A frequency is of 0'75

⑤

$N = 12 \rightarrow 6M$, $6F$

$A = 0'5 = P_0$

$a = 0'5 = q_0$

$P_1^1 = 0'458$ $q_1^1 = 1 - 0'458 = 0'542$

$P_2^1 = 0'583$ $q_2^1 = 1 - 0'583 = 0'417$

$k = 2N \cdot P_0 \rightarrow k = 12 \cdot 0'5 = 6$

$n = 2N \Rightarrow 2 \cdot 6 = 12$

$P_1 = \frac{12!}{6!(12-6)!} \cdot 0'458^6 \cdot 0'542^{(12-6)} = 0'2162 \rightarrow$ Probability of 21'62% for population 1

$P_2 = \frac{12!}{6!(12-6)!} \cdot 0'583^6 \cdot 0'417^{(12-6)} = 0'19077 \rightarrow$ Probability of 19'07% in population 2

$P_d = 0'375$

$q_d = 1 - 0'375 = 0'625$

$k_d = 12 \cdot 0'375 = 4'5 \simeq 5$

$P = \frac{12!}{5!(12-5)!} \cdot 0'375^5 \cdot 0'625^{(12-5)} = 0'22$

$P(0'375\% \text{ in } P_1 \text{ and } P_2) = 0'22 \cdot 0'22 = 0'04787 \rightarrow 4'787\%$ of that P_1 and P_2 have $p = 0'375$ in the second generation

(6) p1s Genetic drift theory

$p_0 = 0'5$ $q_0 = 0'5$ → 2A 2a

$t = 2$

$p_2 = 0'25$ $q_2 = 0'75$ → 1A 3a

$P(0'25|0'5) = P(1A|2A)$ * $= P($ 1allele A in gen 2$) =$ ~~$0'4219$~~

* → table

	0	1	2	3	4	nº gen
nº alleles in gen *	1	0	0'4219	0'0625 0'25 0'375 0'25 0'0625	0'0469	0

$P($1all in gen 2$) = (0 \cdot 0'0625) + (0'4219 \cdot 0'25) + (0'25 \cdot 0'375) + (0'0469 \cdot 0'25) + (0 \cdot 0'0625) = 0 + 0'1055 + 0'0938 + 0'01173 + 0 = 0'21103$

The probability that a population of 2 individuals (4 alleles) that start with $p_0 = 0'5$ have $p_2 = 0'25$ in 2 generations is of 0'21103

(7)

$N_m = \frac{55000}{50} = 1100$ males

$N_F = 55000 - 1100 = 53900$ females

} assuming ~~N constant~~ adult individuals all in groups of 50F to 1M

$N_e = \frac{4 N_m N_f}{N_m + N_f} \rightarrow N_e = \frac{4 \cdot 1100 \cdot 53900}{1100 + 53900} = 4312$

Assuming equal probability of M/F individual independency

27500 = M 27500 = F (there will be males without a group)

(8)

$N_0 = 104$ $N_1 = 62$ $N_2 = 10$ $N_3 = 110$ $t = 4$

$\frac{1}{N_e} = \frac{1}{t} \cdot \left(\frac{1}{N_0} + \frac{1}{N_1} + \frac{1}{N_2} + \frac{1}{N_3}\right) \Rightarrow \frac{1}{N_e} = \frac{1}{4} \cdot \left(\frac{1}{104} + \frac{1}{62} + \frac{1}{10} + \frac{1}{110}\right) = 0'03371 \rightarrow$

$\rightarrow N_e = \frac{1}{0'03371} = 29'6658$ → The effective size of the population is 30

Genetic drift

7

$N = 55000$

Assuming that all males controls a group of exactly 50 females, that all individuals are adults and that there are no lone individuals. Otherwise you could assume an even (50/50) split.

$N_m = \frac{55000}{50} = 1100$

all others must be females

$N_f = 55000 - 1100 = 53900$

a) $N_e = \frac{4 \cdot 1100 \cdot 53900}{1100 + 53900} = 4312 = N_e \rightarrow N_e = \frac{4 N_m \cdot N_f}{N_m + N_f}$

b) $N_m = \frac{20}{50} = 0'4 \rightarrow$ round up, for the species to continue reproducing you need 1 male at least

$N_m = 1$

$N_f = N - N_m \rightarrow N_f = 20 - 1 = 19$

$N_e = \frac{4 \cdot 1 \cdot 19}{1+19} = 3'8 \simeq 4$.

c) $H_0 = 0'028$

$t = 5$

Assuming that the bottleneck didn't change the heterozygosity proportion

Assuming that the population grows to contain 50 individuals in 5 generations

$H_t = H_0 \left(1 - \frac{1}{2N}\right)^t \rightarrow H_5 = 0'028 \left(1 - \frac{1}{250}\right)^5 = 0'0266$

Mutation

①

$A \rightarrow a$
$\mu = 0'00002$

$p_0 = 0'5$

p_{10} ? $\quad p_t = p_0 (1-\mu)^t$

$p_{10} = 0'5\,(1-0'00002)^{10} = 0'49990009$ → Frequency of A in 10 generations

$p_{100} = 0'5\,(1-0'00002)^{100} = 0'4990099$ → Frequency of A in 100 generations

$p_{1000} = 0'5\,(1-0'00002)^{1000} = 0'49009$ → Frequency of A in 1000 generations

$p_{10000} = 0'5\,(1-0'00002)^{10000} = 0'409365$ → Frequency of A in 10000 generations

Its a very slow change in allele frequency

②

$\mu = 0'005 \qquad v = 0'0015$

$p_0 = 0'8 \qquad q_0 = 0'2$

$$p_t = \frac{v}{\mu+v} + \left(p_0 - \frac{v}{\mu+v}\right)(1-\mu-v)^t$$

$$q_1 = \frac{0'0015}{0'005+0'0015} + \left(0'2 - \frac{0'0015}{0'005+0'0015}\right)(1-0'005-0'0015)^1 = 0'2002$$

In the first generation $q = 0'2002$

$$q_2 = \frac{0'0015}{0'005+0'0015} + \left(0'2 - \frac{0'0015}{0'005+0'0015}\right)(1-0'005-0'0015)^2 = 0'200398$$

In the second generation $q = 0'200398$

$$q_3 = \frac{0'0015}{0'005+0'0015} + \left(0'2 - \frac{0'0015}{0'005+0'0015}\right)(1-0'005-0'0015)^3 = 0'2005961$$

In the third generation $q = 0'2005961$

$$p = \frac{v}{\mu+v} \rightarrow p = \frac{0'0015}{0'005+0'0015} = 0'2308 \qquad q = 1-p \rightarrow q = 1-0'2308 = 0'7692$$

Allele frequencies in equilibrium are $q = 0'7692$ $\quad p = 0'2308$