

SOLUTIONS TO EXERCISES FOR CHAPTER 6

**Exercise 6.1.**

$F$  is a useful parameter in population and quantitative genetics in the following ways:

- 1) Useful in determining the genotypic frequencies in a random-mating population which is subjected to some inbreeding, and also useful in an infinite set of closed, inbred subpopulations.
- 2) Frequency of heterozygotes decreases linearly as  $F$  increases. The loss in relative frequency of heterozygotes per locus is equal to  $F$ .  $F$  measures the relative decrease in heterozygosity in an inbreeding population at a single locus or in all loci.  $F$  gives the proportion of loci in a random individual which have become homozygous upon inbreeding of the population.
- 3) Useful in determining three and four-gene states of identity which in turn are useful in covariances of relatives.
- 4) In a population involving a mixture of two mating systems, random mating and self-fertilization,  $F$  can be used to measure the amount of self-fertilization.
- 5) The rate of inbreeding or relative change in  $F$  for different mating systems permits one to compare alternative mating systems.
- 6) Used in determining the expected amount of genetic variance in mating designs.
- 7)  $F$  value can be used to monitor a breeding population to avoid excessive inbreeding.
- 8)  $F$  can be used to predict the fitness of a particular individual if the organism suffers from inbreeding depression.  $F$  gives a quantitative measure of the unfavorable effect to be expected in a species where inbreeding is unfavorable.
- 9)  $F$  can be used to predict the increase in odds of a rare recessive trait being expressed in offspring of a mating of relatives compared to the mating of nonrelatives.
- 10) In developing homozygous lines we need to know how many generations of selfing, i.e., how much inbreeding is required to reach a certain proportion of homozygotes.
- 11)  $F$  equals the proportion of the population consisting of identical homozygotes.
- 12)  $F$  can presumably be used to determine how inbreeding affects the rate of loss in linkage disequilibrium.
- 13)  $F$  measures the amount of the dispersive process that has taken place since the base population (Falconer and Mackay, 1996, p. 58).
- 14) Measures the cumulative effect of random drift (Falconer and Mackay, 1996, p. 61)
- 15) Gives an idea of the distribution of gene frequency among subpopulations, when dispersion is balanced by mutation and migration (Falconer and Mackay, 1996, pp. 74-75).
- 16)  $F$  may help the breeder determine when outcrossing may become necessary.
- 17)  $F$  is independent of allelic frequencies.
- 18) Population mean declines linearly as  $F$  increases if there is dominance but no epistasis.
- 19) If there is both dominance and epistasis, the effect of inbreeding on the population mean may be a quadratic function in  $F$ .
- 20) Our own social relationships are partially determined by  $F$  in that our laws forbid the marriage of closely related individuals to prevent too much inbreeding.

**Exercise 6.2.**

a. The probability of heterozygosity after five generations of selfing (S) is

$$\begin{aligned} H_5 &= P_{S,5}H_0 \quad [\text{see equations (6.22) and (6.24)}] \\ &= \left(\frac{1}{2}\right)^5 H_0 \\ &= \frac{1}{32}(0.4) \\ &= \frac{1}{32} \frac{4}{10} = \frac{1}{80} \end{aligned}$$

The probability of homozygosity after five generations of selfing (S) is

$$1 - H_5 = \frac{79}{80} = 0.9875$$

b. The probability of heterozygosity after five generations of double-first-cousin mating (DFC) is

$$\begin{aligned} H_5 &= P_{DFC,5}H_0 \\ &= \frac{24}{32}(0.4) \\ &= \frac{24}{32} \frac{4}{10} = \frac{3}{10} \end{aligned}$$

The probability of homozygosity after five generations of double-first-cousin mating (DFC) is

$$1 - H_0 = \frac{7}{10} = 0.7$$

c. The proportion of independent loci originally segregating but are now nonsegregating loci after five generations of inbreeding is the same as the inbreeding coefficient, namely,

$$S: F = \frac{31}{32}$$

$$DFC: F = \frac{8}{32} = \frac{1}{4}$$

**Exercise 6.3.**

a. First we express the inbreeding coefficient for individuals A and E in generation  $t$  in terms of coefficients of coancestries in the  $t-1$  and  $t-2$  generations as was done in (6.54), namely,

$$\begin{aligned} F_t = F_{At} = F_{Et} = \theta_{ABt-1} &= \frac{1}{4}(\theta_{ACt-2} + \theta_{ADt-2} + \theta_{BCt-2} + \theta_{BDt-2}) \\ &= \frac{1}{4}(\theta_{DFCt-2} + \theta_{QSCt-2} + \theta_{QSCt-2} + \theta_{DFCt-2}) \\ &= \frac{1}{4}(F_{DFCt-2} + F_{QSCt-2} + F_{QSCt-2} + F_{DFCt-2}) \end{aligned}$$

Similar expressions exist for the other three remaining pairs of individuals in generation  $t$ , i.e., pair B and F, C and G, and D and H. The coefficients of coancestries between individuals A and C and between B and D in the above expression are double first cousins in generation  $t-2$  and we already know what that coancestry is equal to in terms of the inbreeding coefficients. Likewise the coefficients between A and D and between B and C are for quadruple second cousins in the  $t-2$  generation. That coefficient of coancestry is equal to the inbreeding coefficient for quadruple second cousins in the  $t-1$  generation. Substituting the recurrence relation for double first cousins (6.40) and the fact that  $\theta_{ADt-2} = \theta_{BCt-2} = F_{t-1}$  [equation (6.56)], or more basically equation (4.13), into the above expression gives the F recurrence relation, namely,

$$\begin{aligned} F_t &= \frac{1}{4} \left[ 2 \left( \frac{1}{8} + \frac{1}{2} F_{t-2} + \frac{1}{4} F_{t-3} + \frac{1}{8} F_{t-4} \right) + 2 F_{t-1} \right] \\ &= \frac{1}{2} \left( \frac{1}{8} + F_{t-1} + \frac{1}{2} F_{t-2} + \frac{1}{4} F_{t-3} + \frac{1}{8} F_{t-4} \right) \\ &= \frac{1}{16} + \frac{1}{2} F_{t-1} + \frac{1}{4} F_{t-2} + \frac{1}{8} F_{t-3} + \frac{1}{16} F_{t-4} \end{aligned}$$

Note that when we substitute the recurrence relation for double first cousins in the above expression, we change  $t-1$ ,  $t-2$ , and  $t-3$  in the recurrence relation to  $t-2$ ,  $t-3$ , and  $t-4$  because A and C, and B and D are double first cousins in generation  $t-2$ . If A and C, and B and D in generation  $t-2$  were mated, the inbreeding coefficient would be that for generation  $t-1$ . Hence, the recurrence relation is expressed in terms of the immediately preceding generations.

One can also expand the above expression for  $F_t$  into its coefficients of coancestry for  $t-3$  generation, substitute the recurrence relation for the corresponding coancestries, and obtain another recurrence relation as a function of different generations. The recurrence relation may not start with the immediate preceding generation, but the new recurrence relation should still give the correct numerical inbreeding values. When the above expression is expanded into coancestries in the  $t-3$  generation, one obtains

$$F_t = \frac{1}{4}(\theta_{AEt-3} + \theta_{AGt-3} + 2\theta_{AFt-3})$$

where the relations are those for full sibs, double first cousins, and quadruple second cousins. Substituting the recurrence relation for full sibs, that for double first cousins, and the inbreeding coefficient for the following generation for quadruple second cousins, one obtains

$$\begin{aligned} F_t &= \frac{1}{4} \left[ \left( \frac{1}{4} + \frac{1}{2}F_{t-3} + \frac{1}{4}F_{t-4} \right) + \left( \frac{1}{8} + \frac{1}{2}F_{t-3} + \frac{1}{4}F_{t-4} + \frac{1}{8}F_{t-5} \right) + 2F_{t-2} \right] \\ &= \frac{1}{4} \left( \frac{3}{8} + 2F_{t-2} + F_{t-3} + \frac{1}{2}F_{t-4} + \frac{1}{8}F_{t-5} \right) \\ &= \frac{3}{32} + \frac{1}{2}F_{t-2} + \frac{1}{4}F_{t-3} + \frac{1}{8}F_{t-4} + \frac{1}{32}F_{t-5} \end{aligned}$$

That this recurrence relation is an alternative expression for the one derived above can be demonstrated as follows: We add and subtract  $\frac{1}{2}F_{t-1}$  and break each coefficient into two parts.

$$\begin{aligned} F_t &= \frac{3}{32} + \frac{1}{2}F_{t-2} + \frac{1}{4}F_{t-3} + \frac{1}{8}F_{t-4} + \frac{1}{32}F_{t-5} \\ &= \frac{2}{32} + \frac{1}{32} + \frac{1}{2}F_{t-1} + \left( \frac{1}{4} + \frac{1}{4} \right)F_{t-2} + \left( \frac{1}{8} + \frac{1}{8} \right)F_{t-3} + \left( \frac{1}{16} + \frac{1}{16} \right)F_{t-4} + \frac{1}{32}F_{t-5} - \frac{1}{2}F_{t-1} \\ &= \frac{2}{32} + \frac{1}{2}F_{t-1} + \frac{1}{4}F_{t-2} + \frac{1}{8}F_{t-3} + \frac{1}{16}F_{t-4} \\ &\quad + \frac{1}{32} + \frac{1}{4}F_{t-2} + \frac{1}{8}F_{t-3} + \frac{1}{16}F_{t-4} + \frac{1}{32}F_{t-5} - \frac{1}{2}F_{t-1} \\ &= \frac{1}{16} + \frac{1}{2}F_{t-1} + \frac{1}{4}F_{t-2} + \frac{1}{8}F_{t-3} + \frac{1}{16}F_{t-4} \\ &\quad + \frac{1}{32} + \frac{1}{4}F_{t-2} + \frac{1}{8}F_{t-3} + \frac{1}{16}F_{t-4} + \frac{1}{32}F_{t-5} - \frac{1}{2}F_{t-1} \\ &= \frac{1}{16} + \frac{1}{2}F_{t-1} + \frac{1}{4}F_{t-2} + \frac{1}{8}F_{t-3} + \frac{1}{16}F_{t-4} \\ &\quad + \frac{1}{2} \left( \frac{1}{16} + \frac{1}{2}F_{t-2} + \frac{1}{4}F_{t-3} + \frac{1}{8}F_{t-4} + \frac{1}{16}F_{t-5} \right) - \frac{1}{2}F_{t-1} \\ &= \frac{1}{16} + \frac{1}{2}F_{t-1} + \frac{1}{4}F_{t-2} + \frac{1}{8}F_{t-3} + \frac{1}{16}F_{t-4} \\ &\quad + \frac{1}{2}F_{t-1} - \frac{1}{2}F_{t-1} \\ &= \frac{1}{16} + \frac{1}{2}F_{t-1} + \frac{1}{4}F_{t-2} + \frac{1}{8}F_{t-3} + \frac{1}{16}F_{t-4} \end{aligned}$$

which is the recurrence relation for quadruple second cousins derived above.

b. The mean coefficient of coancestry  $\bar{\theta}_t$  for generation  $t$  is derived on pp. 6.31-6.32. It is the mean coefficient of coancestry of all possible pairs  $\left[ \binom{8}{2} = 28 \right]$  of individuals in the population of size eight, namely ( $\bar{\theta}_t$  in (6.65) should have had a bar),

$$\bar{\theta}_t = \frac{4}{28}\theta_{FS_t} + \frac{8}{28}\theta_{DFC_t} + \frac{16}{28}\theta_{QSC_t}$$

Substituting the corresponding recurrence relation for full sibs and double first cousins, and  $F_{t+1}$  for quadruple second cousins, one obtains the recurrence relation

$$\begin{aligned}\bar{\theta}_t &= \frac{4}{28} \left( \frac{1}{4} + \frac{1}{2} F_t + \frac{1}{4} F_{t-1} \right) + \frac{8}{28} \left( \frac{1}{8} + \frac{1}{2} F_t + \frac{1}{4} F_{t-1} + \frac{1}{8} F_{t-2} \right) + \frac{16}{28} F_{t+1} \\ &= \frac{1}{28} (1 + 2F_t + F_{t-1} + 1 + 4F_t + 2F_{t-1} + F_{t-2} + 16F_{t+1}) \\ &= \frac{1}{28} (2 + 16F_{t+1} + 6F_t + 3F_{t-1} + F_{t-2}) \\ &= \left( \frac{1}{14} + \frac{4}{7} F_{t+1} + \frac{3}{14} F_t + \frac{3}{28} F_{t-1} + \frac{1}{28} F_{t-2} \right)\end{aligned}$$

#### Exercise 6.4.

The inbreeding coefficient for generation  $t$  for a half-sib mating system is

$$F_{At} = \theta_{ABt-1} = \theta_{(ABt-2)(ACt-2)} = \frac{1}{4} (\theta_{AAt-2} + \theta_{ABt-2} + \theta_{ACt-2} + \theta_{BCt-2}) \quad (1)$$

$$F_{Bt} = \theta_{ACt-1} = \theta_{(ABt-2)(ADt-2)} = \frac{1}{4} (\theta_{AAt-2} + \theta_{ABt-2} + \theta_{ADt-2} + \theta_{BDt-2}) \quad (2)$$

We observe that the last expression on the right-hand side of (1) is the same as that of (2), except that in (1) we have the coefficients of coancestry of A with C and B with C, whereas in (2) we have both A and B with D instead of with C. However, C and D are equivalent in that they are any two half-sib individuals. The coefficients of coancestry between any two pairs of random half sibs are equal, i.e.,

$$\theta_{ABt-2} = \theta_{ACt-2} = \theta_{ADt-2} = \theta_{BCt-2} = \theta_{BDt-2} = \theta_{CDt-2}$$

Therefore

$$F_t = F_{At} = F_{Bt} = \theta_{ABt-1} = \theta_{ACt-1}$$

We examine only  $F_{At}$  to obtain the recurrence relation

$$\begin{aligned}F_{At} = \theta_{ABt-1} &= \theta_{(ABt-2)(ACt-2)} = \frac{1}{4} (\theta_{AAt-2} + \theta_{ABt-2} + \theta_{ACt-2} + \theta_{BCt-2}) \\ &= \frac{1}{4} \left( \frac{1 + F_{At-2}}{2} + F_{t-1} + F_{t-1} + F_{t-1} \right) \\ &= \frac{1}{8} (1 + 6F_{t-1} + F_{t-2}) \\ &= \frac{1}{8} + \frac{3}{4} F_{t-1} + \frac{1}{8} F_{t-2}\end{aligned}$$

The recurrence series for generations 0 to 7 is

Generation	F
0	0
1	0
2	$\frac{1}{8} + \frac{3}{4}(0) + \frac{1}{8}(0) = \frac{1}{8}$
3	$\frac{1}{8} + \frac{3}{4}\left(\frac{1}{8}\right) + \frac{1}{8}(0) = \frac{7}{32}$
4	$\frac{1}{8} + \frac{3}{4}\left(\frac{7}{32}\right) + \frac{1}{8}\left(\frac{1}{8}\right) = \frac{39}{128}$
5	$\frac{1}{8} + \frac{3}{4}\left(\frac{39}{128}\right) + \frac{1}{8}\left(\frac{7}{32}\right) = \frac{195}{512}$
6	$\frac{1}{8} + \frac{3}{4}\left(\frac{195}{512}\right) + \frac{1}{8}\left(\frac{39}{128}\right) = \frac{919}{2048}$
7	$\frac{1}{8} + \frac{3}{4}\left(\frac{919}{2048}\right) + \frac{1}{8}\left(\frac{195}{512}\right) = \frac{4171}{8192}$

**Exercise 6.5.**

Backcrossing of the cross between two identical twins to one of the twins gives a constant inbreeding coefficient of  $1/2$  every successive generation of backcrossing. Referring to Fig. 6.7, p. 6.48, this backcrossing system involving two identical twins is discussed on p. 6.51 where it is stated that  $F_A = 0$  and the coefficient of coancestry between two bisexual (each twin can function as male and female as in plants with perfect flowers) identical twins,  $A$  and  $B$ , equals one half, i.e.,  $\theta_{AB} = \frac{1}{2}$ . Substituting those quantities in equation (6.100) gives  $1/2$  for  $t = 0, 1, \dots$

The bisexual identical twins have a coefficient of coancestry of  $1/2$ , because the coefficient is equivalent to a noninbred individual with itself (see pp. 4.13–4.17).

**Exercise 6.6.**

The formula for the effective population number for full-sib mating is not

$$\frac{1}{\frac{1}{4N_m} + \frac{1}{4N_f}} = \frac{1}{\frac{1}{4(1)} + \frac{1}{4(1)}} = \frac{1}{\frac{1}{2}} = 2.$$

It is the pseudo-effective population number  $N_s$  [see equations (6.176) and (6.182)]. It is too small by  $1/2$ . The true approximate effective population number for full sibs is  $N_e \cong 2\frac{1}{2}$ .

**Exercise 6.7.**

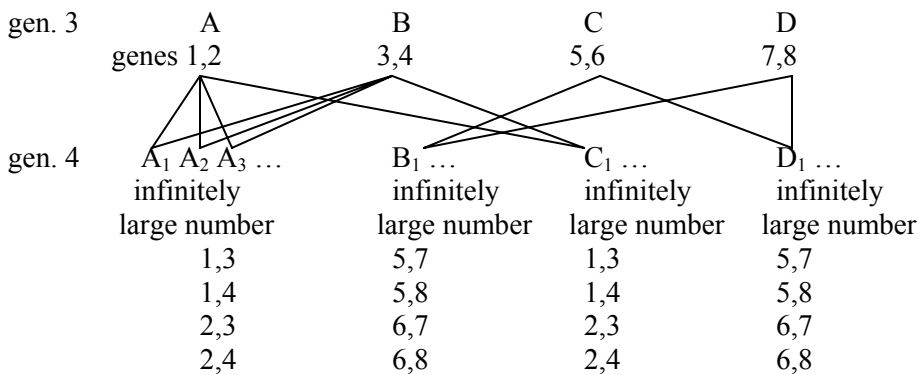
a. First, we note the values for the inbreeding coefficient for the double-first-cousin mating system from equation (6.42):

$$F_t: \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 & \dots \\ 0 & 0 & 0 & \frac{1}{8} & \frac{3}{16} & \frac{8}{32} & \dots \end{array}$$

and that for the mean coefficient of coancestry from equation (6.52)

$$\bar{\theta}_t: \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 & \dots \\ 0 & \frac{1}{12} & \frac{4}{24} & \frac{11}{48} & \frac{28}{96} & \frac{67}{192} & \dots \end{array}$$

From Fig. 6.4, we have the following mating pattern:



There are two ways to deduce the coefficient of inbreeding in generation 5 from random mating an infinitely large population of individuals in generation 4 of the double-first-cousin mating system:

Way 1. Calculate the coefficient of coancestry of the group of individuals in generation 3 with itself. This coefficient is equal to the mean coefficient of coancestry between the infinite number of individuals in generation 4, which in turn equals the coefficient of inbreeding in generation 5.

Way 2. Calculate the mean coefficient of coancestry itself between the infinite number of individuals in generation 4, which is equal to the coefficient of inbreeding in generation 5.

Way 1. Calculate the coefficient of coancestry of the group of four individuals in generation 3 with itself (see pp. 6.148 to 6.151).

gen. 3 individ		A		B		C		D	
genes		1	2	3	4	5	6	7	8
		1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8
A	1	IBD = 1	F <sub>3</sub> = 1/8	DFC		FS		DFC	
	2	(see F series)		$\theta_{AB_3} = F_4$ = 3/16		$\theta_{AC_3} = 5/16$ (see below)		$\theta_{AD_3} = F_4$ = 3/16	
B	3	DFC		IBD = 1	F <sub>3</sub> = 1/8	DFC		FS	
	4	$\theta_{AB_3} = F_4$ = 3/16		F <sub>3</sub> = 1/8		IBD = 1		$\theta_{BC_3} = F_4$ = 3/16	
C	5	FS		DFC		IBD = 1	F <sub>3</sub> = 1/8	DFC	
	6	$\theta_{AC_3} = 5/16$		$\theta_{BC_3} = F_4$ = 3/16		F <sub>3</sub> = 1/8		IBD = 1	
D	7	DFC		FS		DFC		IBD = 1	F <sub>3</sub> = 1/8
	8	$\theta_{AD_3} = F_4$ = 3/16		$\theta_{BD_3} = 5/16$		$\theta_{CD_3} = F_4$ = 3/16		F <sub>3</sub> = 1/8	
								IBD = 1	IBD = 1

$$\theta_{AC_3} = \theta_{BD_3} = \frac{1}{4} + \frac{1}{2}F_3 + \frac{1}{4}F_2 = \frac{1}{4} + \frac{1}{2}\left(\frac{1}{8}\right) + \frac{1}{4}(0) = \frac{5}{16} \quad [\text{see equation (6.38)}]$$

Therefore

$$\begin{aligned} \theta_{l, RM, 3} &= E(\theta_{ij4})^* = F_{5, RM} = \frac{1}{8}(1) + \frac{1}{8}F_3 + \frac{4}{16}FS_3 + \frac{8}{16}DFC_3 \\ &= \frac{1}{8} + \frac{1}{8}\left(\frac{1}{8}\right) + \frac{4}{16}\left(\frac{5}{16}\right) + \frac{8}{16}\left(\frac{3}{16}\right) \\ &= \frac{1}{8} + \frac{1}{64} + \frac{5}{64} + \frac{6}{64} = \frac{20}{64} = \frac{5}{16} \end{aligned}$$

\*This expectation is the mean of the infinite number of all possible coefficients of coancestry among an infinite number of individuals in generation 4.

Thus, the inbreeding coefficient in generation 5 is  $F_{5, RM} = \frac{5}{16}$ .

Inbreeding in generation  $\infty$  is equal to  $F_{5, RM}$  because the population is infinite in size, i.e.,  $F_{\infty} = F_{5, RM} = \frac{5}{16}$ .





Or, one may apply equation (6.280) and calculate  $\theta_{l,3}$  to obtain  $F_{5,RM}$  or  $F_\infty$  directly

$$\theta_{l,DFC,3} = F_{5,RM} = F_\infty = \frac{1 + F_3 - 2\bar{\theta}_3}{2N} + \bar{\theta}_3 = \frac{1 + \frac{1}{8} - 2\left(\frac{11}{48}\right)}{2(4)} + \frac{11}{48} = \frac{5}{16}$$

NOTE:  $\bar{\theta}_3$  can be calculated directly:  $\bar{\theta}_3 = \frac{1}{3}FS + \frac{2}{3}DFC = \frac{1}{3}\left(\frac{5}{16}\right) + \frac{2}{3}\left(\frac{3}{16}\right) = \frac{11}{48}$

Way 2.

Individuals in generation 4

	A <sub>1</sub> A <sub>2</sub> ... A <sub>∞</sub>	C <sub>1</sub> C <sub>2</sub> ... C <sub>∞</sub>	B <sub>1</sub> B <sub>2</sub> ... B <sub>∞</sub>	D <sub>1</sub> D <sub>2</sub> ... D <sub>∞</sub>
	1/4	1/4	1/4	1/4
A <sub>1</sub> A <sub>2</sub> freq = 1/4	 FS in gen 4	FS	DFC	DFC
A <sub>∞</sub> C <sub>1</sub> C <sub>2</sub> freq = 1/4	FS	 FS	DFC	DFC
C <sub>∞</sub> B <sub>1</sub> B <sub>2</sub> freq = 1/4	DFC	DFC	 FS	FS
B <sub>∞</sub> D <sub>1</sub> D <sub>2</sub> freq = 1/4	DFC	DFC	FS	 FS
D <sub>∞</sub>				

One can collapse the above diagram to

	A <sub>1</sub> ... C <sub>1</sub> ...	B <sub>1</sub> ... D <sub>1</sub> ...
	1/2	1/2
A <sub>1</sub> ... C <sub>1</sub> ... 1/2	FS in gen 4	DFC
B <sub>1</sub> ... D <sub>1</sub> ... 1/2	DFC	FS

In the above diagram we can ignore an individual with itself because its frequency goes to zero as the population becomes infinite.

$$F_{5, RM} = E(\theta_{ij4}) = \frac{\sum_{i=1}^4 \sum_{j=1}^4 \theta_{ij}}{16} = \frac{\sum_{i=1}^2 \sum_{j=1}^2 \theta_{ij}}{4} = 2\left(\frac{1}{4}\right)\theta_{FS_4} + 2\left(\frac{1}{4}\right)\theta_{DFC_4}$$

$$\text{where } \theta_{FS_4} = \frac{1}{4} + \frac{1}{2}F_4 + \frac{1}{4}F_3 = \frac{1}{4} + \frac{1}{2}\left(\frac{3}{16}\right) + \frac{1}{4}\left(\frac{1}{8}\right) = \frac{8+3+1}{32} = \frac{12}{32} = \frac{3}{8}$$

$$\theta_{DFC_4} = F_5 = \frac{8}{32} = \frac{1}{4}$$

$$\text{So } F_{5, RM} = \frac{1}{2}\theta_{FS_4} + \frac{1}{2}\theta_{DFC_4} = \frac{1}{2}\left(\frac{3}{8}\right) + \frac{1}{2}\left(\frac{1}{4}\right) = \frac{3+2}{16} = \frac{5}{16}$$

which is the same as that above in Way 1 and is the inbreeding coefficient in generation 5.

b. We have exactly the diagram shown in Way 1. Hence, the inbreeding in generation 5 and  $\infty$  is

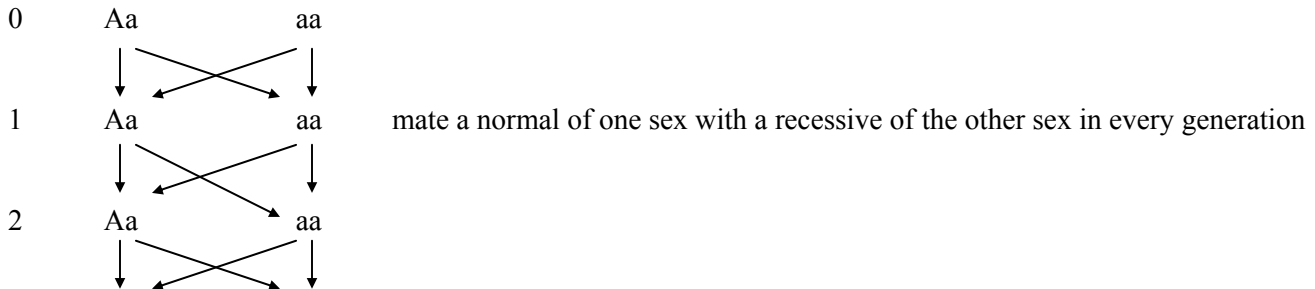
$$F_{4, RM} = F_{5, RM} = F_{\infty} = \frac{5}{16}$$

Inbreeding in generation 4 is larger (5/16) in part (b) than that ( $F_4 = 3/16$  for DFC) in part (a), but in generation 5 the coefficients of inbreeding are the same in both parts (a) and (b).

### Exercise 6.8.

1) Full-sib mating

In a full-sib mating system one desires to mate a heterozygous individual  $Aa$  by a homozygous recessive individual  $aa$  in every generation. We can initially identify a heterozygous individual by crossing several normal individuals to  $aa$ . Then after identifying an  $Aa$  individual, one can follow a full-sib mating system as follows:



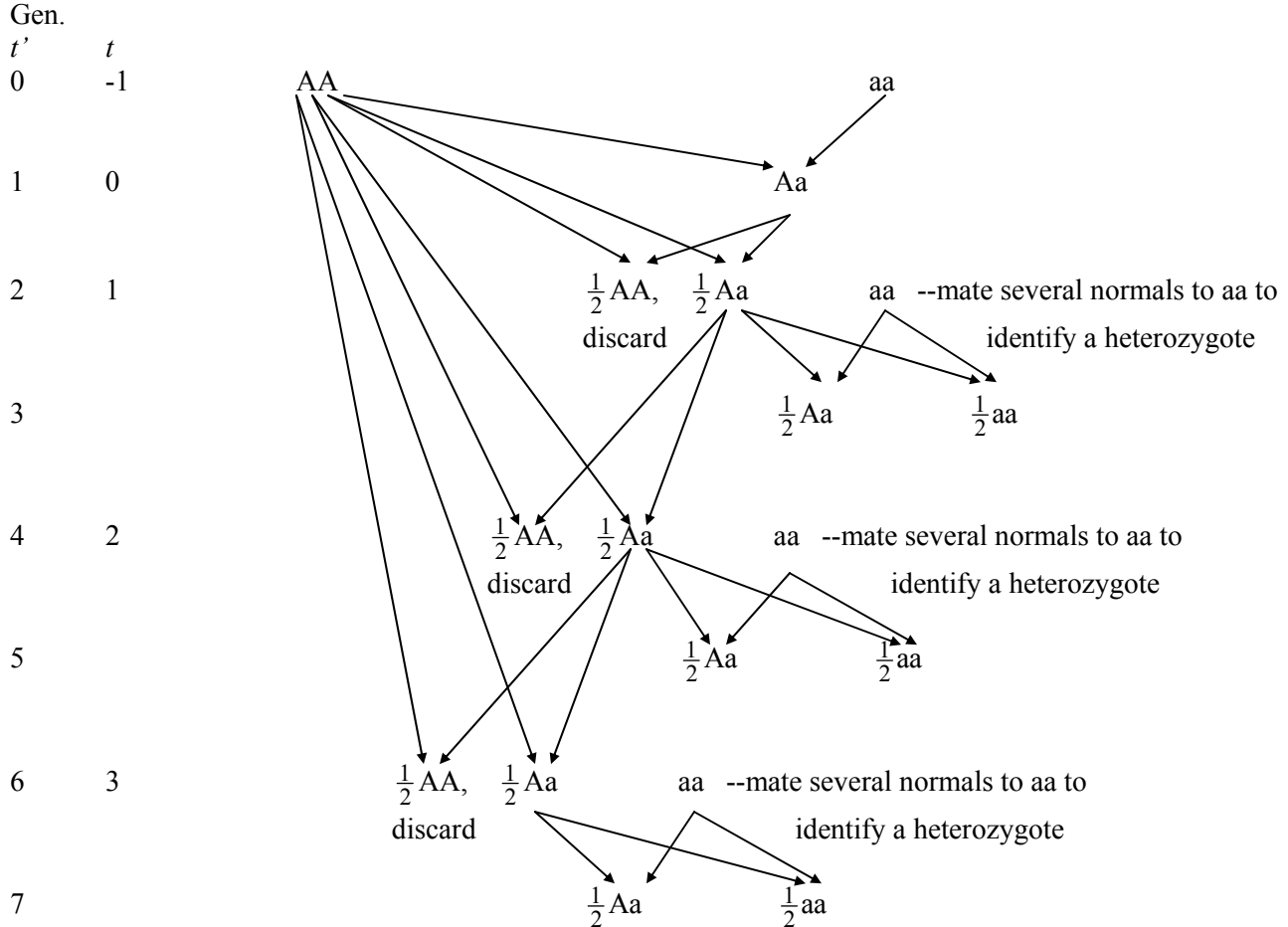
One simply pairs a normal, which will always be  $Aa$ , with a recessive individual generation after generation. It is a simple procedure, but it takes many generations to reach a high  $F$  value (14 generations to reach 0.95, 20 to reach 0.99, Doolittle, D.P., 1987, Population Genetics: Basic Principles, Springer-Verlag, p. 134), i.e., the recurrent series for  $F$  for generations 0 to 7 is

$t$ :	0	1	2	3	4	5	6	7
$F$ :	0	0	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{19}{32}$	$\frac{43}{64}$	$\frac{94}{128}$

To measure the three genotypes, one could measure a large number of  $Aa$  and  $aa$  from the last mating as well as  $AA$  individuals from the inbred line, but it is better to intermate the normal heterozygotes and produce all three genotypes in the same background. The normals would have to be progeny tested, i.e., crossed to  $aa$ , to identify heterozygotes from homozygous normal.

## 2) Backcrossing

A procedure which is repeated generation after generation is the following:



The  $F$  series corresponding to generations  $t'$  and  $t$  are:

$t'$ :	0	1	2	3	4	5	6	7	8	9
$t$ :	-1	0		1		2		3		4
$F$ :		0		$\frac{1}{2}$		$\frac{3}{4}$		$\frac{7}{8}$		$\frac{15}{16}$

vs. that for full sibs

	0	0	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{19}{32}$	$\frac{43}{64}$	$\frac{94}{128}$	...
--	---	---	---------------	---------------	---------------	-----------------	-----------------	------------------	-----

The backcrossing procedure would probably be preferred compared to full-sib mating, because the inbreeding coefficient increases more rapidly. However, with the backcross procedure one is able to compare the genotypic values of the three genotypes in only one background. With full-sib mating one has the possibility of simultaneously developing several full-sib lines and can compare the differences between the three genotypes in different genetic backgrounds, i.e., one can test for the presence of an interaction between genotypes and genetic backgrounds. We have not really compared the two procedures with respect to number of individuals required. Doolittle (D.P. Doolittle, 1987, Population Genetics: Basic Principles, Springer-Verlag, Lecture 29) briefly discussed the topic.

**Exercise 6.9.**

a. Inbreeding occurs from mating related individuals. In the case of a monoecious species this may occur from selfing or from mating two different individuals. The probability of selfing is determined by the number of parental individuals in generation  $t-1$ ,  $N_{t-1}$ , and is equal to  $\frac{1}{N_{t-1}}$ . The contribution to  $F_t$  from selfing is  $\frac{1+F_{t-1}}{2}$ . For mating between two different individuals the probability is  $1 - \frac{1}{N_{t-1}}$ , and its contribution to  $F_t$  is the average coefficient of coancestry between any two individuals,  $\theta_{t-1}$ . With no restrictions on random mating, two different genes may lie in the same individual or in two different individuals, so  $\theta_{t-1} = F_{t-1}$ . (Section 6.4.1.1, pp. 6.68 to 6.76, considered the amount of inbreeding in a monoecious population with random self-fertilization with constant  $N$  from generation to generation. See last paragraph, p. 6.76, which introduces variable population size and refers to Section 6.4.3.2.1 (1), equation (6.218) which shows that the effective population size in generation  $t$  is equal to the census number in generation  $t-1$ .)

b. Again inbreeding occurs only from mating of two different random parental individuals that are related. The parental individuals must be different in that selfing is avoided. Two parental individuals are related to the extent of the probability that a random gene from each is identical by descent or have a common ancestor which is a function of the number of individuals in generation  $N_{t-2}$ . This is also true for a dioecious species with either polygamy or monogamy. (Section 6.4.1.2, pp. 6.77 to 6.87, considered the amount of inbreeding in a monoecious population with avoidance of self-fertilization with constant  $N$  from generation to generation. See last paragraph, p. 6.87, which introduces variable population size and refers to Section 6.4.3.2.2 (1), equation (6.243) which shows that the pseudo-effective population size [see equations (6.239) and (6.242)] in generation  $t$  is equal to the census number in generation  $t-2$ .)

**Exercise 6.10.**

a. The probability of two genes in an individual in a monoecious population with random selfing in generation  $t$  coming from the same parental individual in generation  $t-1$  is  $\frac{1}{N}$ , assuming a constant population size. The conditional probability of those two random genes being IBD is the coefficient of coancestry of a random individual in generation  $t-1$  with itself, namely,  $\frac{1+F_{t-1}}{2}$ . Hence, the contribution to the inbreeding coefficient of that portion of the population resulting from selfing is equal to  $\frac{1}{N} \frac{1+F_{t-1}}{2}$ .

The probability of two gametes which form an individual in generation  $t$  coming from different parental individuals in generation  $t-1$  is  $1 - \frac{1}{N}$ . The conditional probability that the two random genes are identical by descent is simply the inbreeding coefficient in that parental generation  $t-1$ . The reason for this is that with random mating or with no restriction on union of gametes, the probability of two random genes from different individuals in  $t-1$  is the same as the probability of the random genes received by any individual in that  $t-1$  generation, namely  $F_{t-1}$ . Hence we have for the nonselving portion  $\left(1 - \frac{1}{N}\right)F_{t-1}$ . Putting the selfing and the nonselving portions together, we have

$$F_t = \frac{1}{N} \frac{1+F_{t-1}}{2} + \left(1 - \frac{1}{N}\right)F_{t-1} = \frac{1}{2N} + \left(1 - \frac{1}{2N}\right)F_{t-1} \text{ which is equation (6.133).}$$

The detailed derivation is presented in Section 6.4.1.1, pp.6.68 to 6.74.

b. Two genes in an individual in a monoecious population with avoidance of selfing in generation  $t$  can not come from the same individual in generation  $t - 1$ , because selfing is avoided. One must always go back to the generation of the common ancestor because it determines the IBD. Thus, two genes can come from the same individual in generation  $t - 2$  with probability  $\frac{1}{N}$  and the conditional probability of being IBD is  $\frac{1+F_{t-2}}{2}$ , so the combined probability is  $\frac{1}{N} \frac{1+F_{t-2}}{2}$ .

The probability that two genes in an individual in generation  $t$  come from different individuals is  $1 - \frac{1}{N}$  and the conditional probability is  $F_{t-1}$ . It is  $F_{t-1}$ , because the inbreeding in generation  $t - 1$  is the random union of genes from two different individuals in generation  $t - 2$ . Hence, we have

$$F_t = \frac{1}{N} \frac{1+F_{t-2}}{2} + \left(1 - \frac{1}{N}\right) F_{t-1} = \frac{1}{2N} + \left(1 - \frac{1}{N}\right) F_{t-1} + \frac{1}{2N} F_{t-2} \text{ which is equation (6.149).}$$

The detailed derivation is presented in Section 6.4.1.2, pp. 6.77 to 6.82.

c. With separate sexes, two genes in an individual in generation  $t$  cannot come from the same individual in generation  $t - 1$ , but they can come from the same individual in generation  $t - 2$ . They may come from the same male individual in generation  $t - 2$  (male grandparent) with probability  $\frac{1}{4N_m}$ . The  $1/4$  represents the probability that two genes (in a random individual in generation  $t$ ) came from male individuals (or male grandparents) in generation  $t - 2$ . (One-fourth is the probability that the two genes came from female individuals in generation  $t - 2$ , and  $1/2$  is the probability that they came from separate sexes in the  $t - 2$  generation.)  $1/N_m$  is the probability that the two genes came from the same male grandparent and the conditional probability that they are identical by descent is  $\frac{1-F_{t-2}}{2}$  which is the coefficient of coancestry of an individual with itself. Similarly the probability

that they come from the same female individual in generation  $t - 2$  (female grandparent) is  $\frac{1}{4N_f}$  and the

conditional probability is  $\frac{1+F_{t-2}}{2}$ . Thus, we have so far  $\left(\frac{1}{4N_m} + \frac{1}{4N_f}\right) \left(\frac{1+F_{t-2}}{2}\right)$ .

The probability that the genes come from separate sexes—one from a male grandparent and one from a female grandparent is  $\frac{1}{2}$ . The probability that they are IBD is the inbreeding in generation  $t - 1$ ,  $F_{t-1}$ , because it is the union of one random gene from the male parent and one from the female parent that determine the inbreeding any generation.

Now we must return to the case when two genes come from the male grandparents. They come from different male grandparents with probability  $1 - \frac{1}{N_m}$ . The conditional probability that two random genes—one

from one male grandparent and the other from another male grandparent is IBD is  $F_{t-1}$ , because the sex as determined by X and Y chromosomes is independent of autosome genes being IBD. So we have

$$\begin{aligned} & \frac{1}{4} \left(1 - \frac{1}{N_m}\right) F_{t-1} + \frac{1}{4} \left(1 - \frac{1}{N_m}\right) F_{t-1} + \frac{1}{2} F_{t-1} \\ &= \left(\frac{1}{4} - \frac{1}{4N_m} + \frac{1}{4} - \frac{1}{4N_f} + \frac{1}{2}\right) F_{t-1} = \left(1 - \frac{1}{4N_m} - \frac{1}{4N_f}\right) F_{t-1} \end{aligned}$$

Thus, in total we have

$$F_t = \left( \frac{1}{4N_m} + \frac{1}{4N_f} \right) \left( \frac{1+F_{t-2}}{2} \right) + \left( 1 - \frac{1}{4N_m} - \frac{1}{4N_f} \right) F_{t-1} \text{ which is equation (6.174).}$$

The detailed derivation is presented in Section 6.4.2.1, pp. 6.88 to 6.95.

d. Considering the full-sib structure, the probability that the two random genes in random individual in generation  $t$  come from two individuals, a male and a female parent, which are related to each other as full sibs is  $\frac{1}{N} = \frac{2}{2N}$ . Given a full-sib mating, the probability is  $\frac{1}{4}$  that the two genes come from the male grandparent and

have a conditional probability of  $\frac{1+F_{t-2}}{2}$ . Similarly the probability is  $\frac{1}{4}$  that the two genes come from the female grandparent and have a conditional probability of  $\frac{1+F_{t-2}}{2}$ . The probability of two genes coming from different grandparents—one from a male and one from a female is  $\frac{1}{2}$  and given that the two random genes have those origins, the conditional probability is  $F_{t-1}$  that they are IBD. So we have

$$\left( \frac{1}{N} \right) \left[ \frac{1}{4} \left( \frac{1+F_{t-2}}{2} \right) + \frac{1}{4} \left( \frac{1+F_{t-2}}{2} \right) + \frac{1}{2} F_{t-1} \right]$$

If they are non-full-sibs in generation  $t-1$ , that probability is  $1 - \frac{2}{N}$ . The probability is  $\frac{1}{4}$  that the two genes come from different male individuals in generation  $t-2$  with conditional probability of  $F_{t-1}$ . The probability is  $\frac{1}{4}$  that the two random genes in a random individual in generation  $t$  come from different female individuals in generation  $t-2$  with conditional probability of  $F_{t-1}$ . The probability is  $\frac{1}{2}$  that they come from different sexes in generation  $t-2$  and the probability of IBD is  $F_{t-1}$ . Thus, we have

$$\left( 1 - \frac{2}{N} \right) \left[ \frac{1}{4} F_{t-1} + \frac{1}{4} F_{t-1} + \frac{1}{2} F_{t-1} \right] = \left( 1 - \frac{2}{N} \right) F_{t-1}$$

In total, we have

$$\begin{aligned} & \left( \frac{1}{N} \right) \left[ \frac{1}{4} \left( \frac{1+F_{t-2}}{2} \right) + \frac{1}{4} \left( \frac{1+F_{t-2}}{2} \right) + \frac{1}{2} F_{t-1} \right] + \left( 1 - \frac{2}{N} \right) F_{t-1} \\ &= \frac{2}{N} \left( \frac{1+F_{t-2}}{4} \right) + \frac{2}{2N} F_{t-1} + \frac{N-2}{N} F_{t-1} \\ &= \left( \frac{1+F_{t-2}}{2N} \right) + \frac{2+2N-4}{2N} F_{t-1} \\ &= \frac{1}{2N} + \left( \frac{2N-2}{2N} \right) F_{t-1} + \frac{1}{2N} F_{t-2} \\ &= \frac{1}{2N} + \left( 1 - \frac{1}{N} \right) F_{t-1} + \frac{1}{2N} F_{t-2} \text{ which is equation (6.190).} \end{aligned}$$

The detailed derivation is presented in Section 6.4.2.2, pp. 6.98 to 6.102.

**Exercise 6.11.**

Bottlenecks or restrictions in population size in one or two generations have a great or dominant effect upon the effective population number. The effective population number of varying population numbers is the harmonic mean, and the harmonic mean is greatly influenced or dominated by the one or two small numbers. See Section 6.4.3.1, pp. 6.102 to 6.104.

**Exercise 6.12.**

a. The effective population number for a monoecious population with random self-fertilization from equation (6.216) is

$$N_{et} = \frac{\bar{k}(N_{t-1}\bar{k} - 1)}{\sigma_k^2 + \bar{k}(\bar{k} - 1)}$$

where  $\bar{k} = 2$

$$N_{t-1} = N_t = 10$$

$$\sigma_k^2 = 2N_t \left( \frac{1}{N_{t-1}} \right) \left( 1 - \frac{1}{N_{t-1}} \right) = 2 \cdot 10 \left( \frac{1}{10} \right) \left( 1 - \frac{1}{10} \right) = \frac{9}{5} \quad \text{equation (6.205)}$$

Substituting these numbers in the above formula for the effective population size, we have

$$N_{et} = \frac{\bar{k}(N_{t-1}\bar{k} - 1)}{\sigma_k^2 + \bar{k}(\bar{k} - 1)} = \frac{2(10 \cdot 2 - 1)}{\frac{9}{5} + 2} = \frac{38}{\frac{19}{5}} = 10$$

The rate of inbreeding from equation (6.143) is

$$\frac{1}{2N} = \frac{1}{2 \cdot 10} = \frac{1}{20} = 0.05$$

b. In this case, the effective population number for a monoecious population with random self-fertilization from equation (6.222) is

$$N_{et} = N_{t-1} \left( 1 + \frac{N_{t-1} - 1}{4N_t - N_{t-1} - 1} \right)$$

where  $N_{t-1} = N_t = 10$

Substituting these numbers in the above formula for the effective population size, we have

$$N_{et} = 10 \left( 1 + \frac{10 - 1}{4 \cdot 10 - 10 - 1} \right) = 10 \left( 1 + \frac{9}{29} \right) = 10 \left( \frac{29 + 9}{29} \right) = 10 \left( \frac{38}{29} \right) = 13.1034$$

Substituting the numerical value for effective population size for N in equation (6.143), we have the rate of inbreeding

$$\frac{1}{2N_{et}} = \frac{1}{2 \cdot 13.1034} = \frac{1}{26.2068} = 0.038158$$

c. NOTE: I am uncertain about some aspects of the following solution for part (c), given my current deficient state of understanding of this topic. Thus, the whole development for part (c) ought to be regarded with caution. I was going to delete part (c) from the exercise itself and not include any solution at all, but in the final stages I decided to include what I have developed for whatever merit it may have. I welcome your comments.

The number of male breeding individuals in every generation is 20 and that of female breeding individuals is 10. Let us symbolize the number of male breeding individuals as  $N_{m,t-1} = N_{m,t} = 20$  and that of females as  $N_{f,t-1} = N_{f,t} = 10$ .

Males:

The distribution of number of male gametes from every male parental individual is the binomial distribution

$$\binom{N_{m,t}}{k_m} \left( \frac{1}{N_{m,t-1}} \right)^{k_m} \left( 1 - \frac{1}{N_{m,t-1}} \right)^{N_{m,t}-k_m} \quad k_m = 0, 1, \dots, N_{m,t}$$

where  $N_{m,t-1} = N_{m,t} = 20$

Mean number of gametes per male individual is

$$\bar{k}_m = N_{m,t} \frac{1}{N_{m,t-1}} = \frac{N_{m,t}}{N_{m,t-1}} = \frac{20}{20} = 1$$

and the variance of the number of gametes per male individual is

$$\sigma_{km}^2 = N_{m,t} \left( \frac{1}{N_{m,t-1}} \right) \left( 1 - \frac{1}{N_{m,t-1}} \right) = 20 \left( \frac{1}{20} \right) \left( \frac{19}{20} \right) = \frac{19}{20}$$

Females:

The distribution of number of female gametes from every female parental individual is the binomial distribution

$$\binom{N_{f,t}}{k_f} \left( \frac{1}{N_{f,t-1}} \right)^{k_f} \left( 1 - \frac{1}{N_{f,t-1}} \right)^{N_{f,t}-k_f} \quad k_f = 0, 1, \dots, N_{f,t}$$

where  $N_{f,t-1} = N_{f,t} = 10$

Mean number of gametes per female individual is

$$\bar{k}_f = N_{f,t} \frac{1}{N_{f,t-1}} = \frac{N_{f,t}}{N_{f,t-1}} = \frac{10}{10} = 1$$

and the variance of the number of gametes per female individual is

$$\sigma_{kf}^2 = N_{f,t} \left( \frac{1}{N_{f,t-1}} \right) \left( 1 - \frac{1}{N_{f,t-1}} \right) = 10 \left( \frac{1}{10} \right) \left( \frac{9}{10} \right) = \frac{9}{10}$$

To better understand this situation I created the following hypothetical situation. The numbers in the individual cells are the number of male gametes from the particular male parent uniting with the random female plant selected for harvesting; only one mating has two male gametes uniting with a single female, namely, male parent 1 uniting with female 3—all other cells have only one.

		t - 1																				
		Female parents (S denotes a random plant selected for harvesting)																				
		1	2	S	4	S	6	7	S	S	10	11	S	S	S	16	S	S	19	S		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
t - 1 Male parents	$k_m$																					
	1			2																		2
	2															1						1
	3			1					1													2
	4																					0
	5															1						1
	6									1										1		2
	7																					0
	8																					0
	9																					0
	10													1						1		2
	11																					0
	12						1															1
	13									1				1						1		3
	14															1						1
	15			1															1	1		3
	16																		1			1
	17																					0
	18																					0
	19																			1		1
20																					0	
20																						
	$k_f$	-	-	4	-	1	-	-	3	0	-	-	2	-	1	2	-	2	5	-	0	
20	$k_m$	2	1	2	0	1	2	0	0	0	2	0	1	3	1	3	1	0	0	1	0	
20	$k_m + k_f$	2	1	6	0	2	2	0	3	0	2	0	3	3	2	5	1	2	5	1	0	
	40																					

The total number of gametes is (not clear)

$$\sum_{i=1}^{N_{m,t-1}} k_{mi} + \sum_{i=1}^{N_{f,t-1}} k_{fi} = N_{mt} + N_{ft} = 30$$

The total number of gametes by the  $i$ th parent is

$$k_i = k_{mi} + k_{fi} \quad \text{for } i = 1, \dots, N_{f,t-1}$$

The total number of ways in which two gametes from the same parent can unite is

$$\begin{aligned}
2 \sum_{i=1}^{N_{f,t-1}} \binom{k_i}{2} &= \sum_{i=1}^{N_{f,t-1}} k_i(k_i - 1) \\
&= \sum_{i=1}^{N_{f,t-1}} k_i^2 - \sum_{i=1}^{N_{f,t-1}} k_i \\
&= \sum_{i=1}^{N_{f,t-1}} k_i^2 - \sum_{i=1}^{N_{f,t-1}} \bar{k}
\end{aligned}$$

I guessed the answer to be equation (6.216).

$$N_{et} = \frac{\bar{k}(N_{t-1}\bar{k} - 1)}{\sigma_{\bar{k}}^2 + \bar{k}(\bar{k} - 1)}$$

$$\text{where } \bar{k} = 3 \text{ but } \bar{k} = \frac{\sum_{i=1}^{N_{f,t-1}} k_i}{N_{f,t-1}} = \frac{6+2+3+0+3+2+5+2+5+0}{10} = \frac{28}{10} = 2.8$$

$$N_{t-1} = N_{f,t-1} = 10$$

$$\begin{aligned}
\sigma_{\bar{k}}^2 &= \sigma_{km}^2 + \sigma_{kf}^2 = N_{m,t} \left( \frac{1}{N_{m,t-1}} \right) \left( 1 - \frac{1}{N_{m,t-1}} \right) + N_{f,t} \left( \frac{1}{N_{f,t-1}} \right) \left( 1 - \frac{1}{N_{f,t-1}} \right) \\
&= 20 \left( \frac{1}{20} \right) \left( \frac{19}{20} \right) + 10 \left( \frac{1}{10} \right) \left( \frac{9}{10} \right) \\
&= \frac{19}{20} + \frac{9}{10} \\
&= \frac{37}{20}
\end{aligned}$$

The effective population size from equation (6.216) is

$$N_{et} = \frac{\bar{k}(N_{t-1}\bar{k} - 1)}{\sigma_{\bar{k}}^2 + \bar{k}(\bar{k} - 1)} = \frac{3(10 \cdot 3 - 1)}{\frac{37}{20} + 3(3 - 1)} = \frac{3(29)}{\frac{37}{20} + 3(2)} = \frac{3(29)}{\frac{37+120}{20}} = \frac{60(29)}{37+120} = 11.0828$$

### Exercise 6.13.

All definitions of the group coefficients are straight-forward extensions of the same ideas associated with the coefficients of individuals.

The coefficient of inbreeding of an individual is defined to be the probability that the two random genes received by that individual at any locus are identical by descent. The coefficient of group inbreeding is the mean of the probabilities of all possible pairs among the  $2N$  homologous genes, drawn without replacement within the group, being identical by descent. When the group size is reduced to one ( $N = 1$ ), the group probability is equal to that for an individual.

The coefficient of coancestry between two individuals is defined to be the probability that one gene drawn at random from one individual is identical by descent to another gene drawn at random from another individual at the same locus. The coefficient of coancestry between two groups is the probability of a random gene from one group being identical by descent to a random homologous gene from another group. The similarities are: 1) drawing a gene randomly from one individual or one group, 2) homologous genes being identical by descent, and 3) when  $N_1 = 1$  and  $N_2 = 1$ , each group has only one individual and the definitions are the same.

The coefficient of coancestry of an individual with itself is defined to be the probability of a random gene from an individual being identical by descent to a random homologous gene from the same individual. Sampling is with replacement. The coefficient of coancestry of a group of individuals with itself is the probability that two homologous genes drawn randomly with replacement from the same group are identical by descent. The similarities are: 1) drawing a gene randomly with replacement, 2) homologous genes being identical by descent, and 3) when the group has only one individual, the definitions are the same.