Agronomy 615
Statistical Genetics
W. E. Nyquist
March 2004

SOLUTIONS TO EXERCISES FOR CHAPTER 5

Exercise 5.1.

From equation (5.28) one can calculate the allelic frequencies:

$$p_1 = p_{11F} + \sum_{\substack{j=1 \\ j\neq 1}}^{m} p_{1jF} = 0.828 + 0.072 = 0.900$$

$$p_2 = p_{22F} + \sum_{\substack{j=1 \\ j\neq 2}}^{m} p_{2jF} = 0.028 + 0.072 = 0.100 = 1 - p_1 = 1 - .900$$

Then from either equation (5.29) or (5.33), we can obtain the inbreeding value. From (5.29)

$$F = \frac{1}{m-1}\left(\sum_{i=1}^{m} \frac{p_{iiF}}{p_i} - 1\right) = \tfrac{1}{2-1}\left(\frac{0.828}{0.900} + \frac{0.028}{0.100} - 1\right) = 0.92 + 0.28 - 1 = 1.2 - 1 = 0.20$$

Or, alternatively from (5.33)

$$F = \frac{H_0 - H_F}{H_0} \quad \text{where } H_0 = 2p_1p_2$$

$$= \frac{2(0.90)(0.10) - 0.144}{2(0.90)(01.0)} = \frac{0.18 - 0.144}{0.18} = 0.20$$

For another alternative way one can solve for F in equation (5.19) for either $i = 1$ or 2. For $i = 1$

$$p_{iiF} = p_i^2 + Fp_i\left(1 - p_i\right)$$

$$0.828 = \left(0.9\right)^2 + F(.9)(.10)$$

$$F = \frac{0.828 - \left(0.90\right)^2}{(.90)(.10)} = \frac{0.828 - 0.81}{0.09} = \frac{0.018}{0.09}$$

$$= 0.20$$

Exercise 5.2.

	A_1	A_2	
A_1	92	28	120
A_2	28	32	60
	120	60	180

$$\hat{p}_1 = \tfrac{120}{180} = \frac{92 + \frac{1}{2}56}{180} = \tfrac{2}{3}$$

$$\hat{p}_2 = \tfrac{60}{180} = \tfrac{1}{3}$$

Substituting in equation (5.29), we have

$$F = \frac{1}{m-1}\left[\sum_{i=1}^{m} \frac{p_{iiF}}{p_i} - 1\right] = \tfrac{1}{2-1}\left[\tfrac{92/180}{2/3} + \tfrac{32/180}{1/3} - 1\right] = \left[\frac{3(92/180)}{2} + \frac{3(32/180)}{1} - 1\right]$$

$$= \left[\tfrac{276}{360} + \tfrac{96}{180} - 1\right] = \frac{276 + 2(96)}{360} - \tfrac{360}{360} = \tfrac{108}{360} = 0.3$$

To plot the point in an equilateral triangle, we must first calculate the frequencies of the three genotypes:
$p_{A_1A_1} = \frac{92}{180} = 0.5111, p_{A_1A_2} = \frac{56}{180} = 0.3111, p_{A_2A_2} = \frac{32}{180} = 0.1777$
From Fig. 3.2, p. 3.34, we observe that the frequency of A_1A_1 is on the line ZX, ranging from zero on the right to 1 on the left, that the frequency of A_1A_2 is on the line XY, ranging from zero at the base to 1 at the top, and that the frequency of A_2A_2 is on the line YZ, ranging from zero at the top to 1 at the base. The point (0.5111, 0.3111, 0.1777) lies on a parabola for $F = 0.3$ similar to the lower parabola shown in Fig. 5.2, p. 5.13, for $F = 0.25$. The ratio of the line segments P_1P_2 to P_1Q equals F and the ratio of the line segments P_2Q to P_1Q equals the panmictic index $P = 1 - F$. I believe that equation (3.49) is also true for any value of F > 0, i.e., the perpendicular projection of the genotypic frequency point for the population onto the base divides the base into p and $1 - p$ (see p. 3.34).

Exercise 5.3.

The frequency of the mating type between heterozygotes was equal to twice the total frequency of mating between the two corresponding homozygotes [equation (5.45)] in the mating scheme during the following generations: i) in the initial random-mating population and in all subsequent generations until brother-sister mating commenced, and ii) in the population involving a certain proportion of brother-sister mating after a one-locus genotypic equilibrium distribution had been achieved and in all subsequent generations.

Exercise 5.4.

a. First, we calculate the allelic frequencies, using equation (5.28)

$$p_{A_1} = 0.68 + \tfrac{1}{2}(0.24) = 0.8$$

$$p_{A_2} = 0.08 + \tfrac{1}{2}(0.24) = 0.2$$

Then knowing the genotypic frequencies and the allelic frequencies, one may calculate F, using a number of different formulas:

Way 1: [equation (5.29)]

$$F = \frac{1}{m-1}\left[\sum_{i=1}^{m} \frac{p_{iiF}}{p_i} - 1\right] = \frac{1}{2-1}\left[\frac{0.68}{0.8} + \frac{0.08}{0.2} - 1\right] = \left[0.85 + 0.4 - 1\right] = 0.25$$

(this formula is possibly the best in an experimental situation since it may "average" the data in some sense)

Way 2: (from an expression in Table 5.1 (p. 5.10), e.g., $A_iA_i : (1-F)p_i^2 + Fp_i$)

$$p_{ii} = (1-F)p_i^2 + Fp_i$$

$$0.68 = (1-F)(0.8)^2 + F(0.8)$$

$$= (1-F)(0.64) \ + F(0.8)$$

$$= 0.64 - F(0.64) + F(0.8)$$

$$0.16F = 0.04$$

$$F = \frac{0.04}{0.16} = 0.25$$

Way 3: [equation (5.11)]

$$H_F = (1-F)H_0 \text{ where } H_0 = 2pq = 2(0.8)(0.02) = 0.32$$

$$0.24 = (1-F)0.32$$

$$= 0.32 - 0.32F$$

$$0.32F = 0.08$$

$$F = \frac{0.08}{0.32} = 0.25$$

b. The proportion of selfing in the population can be obtained from equation (5.52) by solving for *s:*

$$F_\infty = \frac{s}{2-s} \qquad \text{equation (5.52)}$$

$$s = \frac{2F_\infty}{1+F_\infty} = \frac{2(0.25)}{1+(0.25)} = \frac{0.50}{1.25} = 0.40 \qquad \text{equation (5.56)}$$

Exercise 5.5.

From equation (5.52) we have

$$F_\infty = \frac{s}{2-s} = \frac{0.99}{2-0.99} = \frac{0.99}{1.01} = 0.980198$$

The genotypic proportions from equations (5.19) and (5.20) are

$$p(A_1A_1) = p_1^2 + Fp_1(1-p_1) = \left(\tfrac{1}{2}\right)^2 + 0.980198\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = 0.4950495$$

$$2p(A_1A_2) = 2(1-F)p_1(1-p_1) = 2(1-0.980198)\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = 0.009901$$

$$p(A_2A_2) = (1-p_1)^2 + Fp_1(1-p_1) = \left(\tfrac{1}{2}\right)^2 + 0.980198\left(\tfrac{1}{2}\right)\left(\tfrac{1}{2}\right) = 0.4950495$$

Check: $0.4950495 + 0.009901 + 0.4950495 = 1.0000000$

Exercise 5.6.

a.i. Recalling that the frequency of the heterozygotes in the population in generation *t* is $H_t = (1-F_t)H_0$ [equation (6.7)], and that the inbreeding coefficient in generation *t* is $F_t = 1-\left(\tfrac{1}{2}\right)^t P_0$ or $F_t = 1-\left(\tfrac{1}{2}\right)^t$ [equation (6.23)], we substitute the latter expression for F_t in the expression for H_t to obtain $H_t = \left\{1-\left[1-\left(\tfrac{1}{2}\right)^t\right]\right\}H_0$. Then substituting the appropriate values in that expression, we solve for *t*.

$$0.001 = \left(\tfrac{1}{2}\right)^t 0.50$$

$$\left(\tfrac{1}{2}\right)^t = \frac{0.001}{0.50}$$

$$t\log\tfrac{1}{2} = \log 0.001 - \log 0.50$$

$$t = \frac{\log 0.001 - \log 0.50}{\log\frac{1}{2}}$$

$$= \frac{-3.00000-(9.69897-10.00000)}{9.69897-10.00000} = \frac{-3.00000-(-0.30103)}{-0.30103} = \frac{-2.69897}{-0.30103} = 8.96578 \text{ or 9 generations}$$

Therefore

$$F_t = 1-\left(\tfrac{1}{2}\right)^9 = 1-\tfrac{1}{512} = \tfrac{511}{512} = 0.9980469, \qquad \text{and}$$

$$H_t = (1-F_t)H_0 = \tfrac{1}{512}0.50 = \tfrac{1}{1024} = 0.0009766$$

a.ii. For calculating the frequency of the homozygotes at $t = 9$, no equation is actually given in the notes, but it is clear that the loss in frequency of heterozygotes is distributed equally to the two homozygotes for any system of mating. Thus, [see equation (6.11)]

$$p_{AAt} = p_{AA0} + \tfrac{1}{2}F_tH_0 = p_{AA0} + \tfrac{1}{2}(H_0 - H_t)$$
$$= 0.40 + \tfrac{1}{2}\left(0.50 - \tfrac{1}{1024}\right) = 0.40 + \tfrac{1}{2}(0.4990234) = 0.40 + 0.2495117 = 0.6495117$$
$$2p_{Aat} = (1 - F_t)H_0 = 0.0009766$$
$$p_{aat} = p_{aa0} + \tfrac{1}{2}F_tH_0 = p_{aa0} + \tfrac{1}{2}(H_0 - H_t)$$
$$= 0.10 + \tfrac{1}{2}\left(0.50 - \tfrac{1}{1024}\right) = 0.10 + \tfrac{1}{2}(0.4990234) = 0.10 + 0.2495117 = 0.3495117$$

b.i. The same number of generations required is the same as above: t = 8.97 or 9 generations.

b.ii.
$$p_{AAt} = 0.25 + \tfrac{1}{2}\left(0.50 - \tfrac{1}{1024}\right) = 0.25 + \tfrac{1}{2}(0.4990234) = 0.25 + 0.2495117 = 0.4995117$$
$$2p_{Aat} = 0.0009766$$
$$p_{aat} = 0.25 + \tfrac{1}{2}\left(0.50 - \tfrac{1}{1024}\right) = 0.25 + \tfrac{1}{2}(0.4990234) = 0.25 + 0.2495117 = 0.4995117$$

c. The number of generations is the same because the proportion of heterozygotes is the same in the two initial populations. The number of generations to reach a given level of heterozygotes is dependent only on the initial frequency of heterozygotes and is not dependent upon allelic frequencies.

The final equilibrium genotypic structures are different because the initial proportions of homozygotes or allelic frequencies are different. The frequencies of homozygotes approach the allelic frequencies as $t \to \infty$, and does not depend upon Hardy-Weinberg proportions initially.

Since the loss in heterozygotes is distributed equally to the two homozygotes (two-allele case) the difference between the frequency of the two homozygotes remains the same for $t = 0, 1, 2 \ldots \infty$. For example, for the first population

$p_{AAt} - p_{aat} = 0.40 - 0.10 = 0.30$ for $t = 0$, and

$p_{AAt} - p_{aat} = 0.6495117 - 0.3495117 = 0.30$ for $t = 9$

and for the second population

$p_{AAt} - p_{aat} = 0.25 - 0.25 = 0.00$ for $t = 0$, and

$p_{AAt} - p_{aat} = 0.4995117 - 0.4995117 = 0.00$ for $t = 9$

Exercise 5.7.

a. The single condition assumed is that the population size is very large (and has always been so). The reason is that F_{2t} is equal to the coefficient of coancestry between two different random individuals in the t – 1 generation. Thus, for that coefficient to be zero, the population size must be very large, so the chance of two related individuals mating is zero.

b. If we do not have a large population size and can not assume $F_{2t} = 0$, the recurrence relation to be used is [equation (6.133)]

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

It is commonly called a monoecious finite population with random self-fertilization.

Exercise 5.8.

a. In that portion of the population resulting from mating of related parents a certain proportion of the individuals will be nonidentical ($1 - F_2$) and will be mating at random, and the remaining proportion will be identical (F_2). Thus, the proportion of A_iA_i in the population from related parents is

$$\begin{pmatrix}\text{proportion of } A_iA_i \\ \text{homozygotes} \\ \text{in the population} \\ \text{that have related} \\ \text{parents}\end{pmatrix} = \frac{\begin{pmatrix}\text{proportion of} \\ A_iA_i \text{ homozygotes} \\ \text{in the population} \\ \text{from related parents}\end{pmatrix}}{\begin{pmatrix}\text{proportion of} \\ A_iA_i \text{ homozygotes} \\ \text{in the population} \\ \text{from random mating}\end{pmatrix} + \begin{pmatrix}\text{proportion of} \\ A_iA_i \text{ homozygotes} \\ \text{in the population} \\ \text{from related parents}\end{pmatrix}}$$

$$= \frac{a_2\left[p_i^2 + F_2 p_i(1-p_i)\right]}{a_1 p_i^2 + a_2\left[p_i^2 + F_2 p_i(1-p_i)\right]} \qquad \begin{bmatrix}\text{for the numerator, see equation (5.19)} \\ (1-F)p_i^2 + Fp_i = p_i^2 - Fp_i^2 + Fp_i \\ \qquad\qquad = p_i^2 + Fp_i(1-p_i)\end{bmatrix}$$

$$= \frac{a_2 p_i\left[p_i + F_2(1-p_i)\right]}{p_i\left\{a_1 p_i + a_2\left[p_i + F_2(1-p_i)\right]\right\}}$$

$$= \frac{a_2\left[p_i + F_2(1-p_i)\right]}{a_1 p_i + a_2 p_i + a_2 F_2(1-p_i)}$$

$$= \frac{a_2\left[p_i + F_2(1-p_i)\right]}{p_i + a_2 F_2(1-p_i)}$$

b. The inbreeding coefficient of the offspring from the mating of first cousins is 1/16. [This value 1/16 can be deduced from the pedigree for first cousins. It is also given in the title for Table 5.2 (p. 5.19)]. Thus,

$$\begin{pmatrix}\text{proportion of } A_iA_i \\ \text{homozygotes} \\ \text{in the population} \\ \text{that have related} \\ \text{parents}\end{pmatrix} = \frac{a_2\left[p_i + F_2(1-p_i)\right]}{p_i + a_2 F_2(1-p_i)}$$

$$= \frac{0.015\left[0.02 + \frac{1}{16}(1-0.02)\right]}{0.02 + 0.015\left(\frac{1}{16}\right)(1-0.02)} = \frac{0.0012187}{0.0209187} = 0.05826$$

That is, 5.8% of the affected individuals or homozygotes will have parents who are first cousins. If no first cousin marriages were permitted, one would reduce the frequency of affected individuals by 5.8%.