CHAPTER 5

INBREEDING WITHIN A LARGE SINGLE POPULATION

Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>5.1</td>
</tr>
<tr>
<td>5.2 Genotypic structure of a large, random-mating population subject to inbreeding</td>
<td>5.3</td>
</tr>
<tr>
<td>5.3 Relation between frequency of heterozygotes or homozygotes in a large inbred population and the coefficient of inbreeding</td>
<td>5.14</td>
</tr>
<tr>
<td>5.4 Equilibrium in a large inbred population</td>
<td>5.19</td>
</tr>
<tr>
<td>5.5 Mixture of self-fertilization and random mating</td>
<td>5.23</td>
</tr>
</tbody>
</table>
CHAPTER 5
INBREEDING WITHIN A LARGE SINGLE POPULATION

5.1. Introduction. Assumptions:

The development of the coefficients of coancestry and inbreeding in Chapter 4, Sections 4.1 to 4.5, was independent of the genotypic constitution of the ancestors or founder individuals of a population and has been strictly in terms of pedigree relations. These pedigree relations would apply even to individuals in a homozygous line.

We now desire to relate those ideas of inbreeding to a population and deduce the consequences of inbreeding upon the genotypic structure of a large population as defined in Section 3.1. As Crow and Kimura (1970, p. 62) have pointed out, inbreeding may follow either of two patterns. The first pattern is when there is a certain amount of mating between relatives within a single large population which is otherwise random mating. This pattern of inbreeding involves a mixture of mating systems -- one or more inbreeding systems along with the random-mating system. An important characteristic of this pattern is that the inbred individuals in any generation are independent of the inbred individuals of the previous generation. A consequence of this is that the average coefficient of inbreeding for the population approaches a small equilibrium value, as the number of generations increases. The population never becomes completely composed of homozygotes. Furthermore, as soon as mating between relatives is discontinued in any generation and random mating occurs with respect to the whole population, the inbreeding coefficient returns to zero and the genotypic frequencies to Hardy-
Weinberg proportions. The inbreeding coefficient returns to zero, because the population remained large in size in all generations.

Two examples will serve to illustrate this pattern: (1) In the human species we might visualize, in some geographical area, a population or an isolate which is mating at random except for a certain proportion of matings which are first cousin matings. These first cousin matings do bring about a certain amount of inbreeding within the whole population. (2) In certain plant species such as cotton, random mating predominates but yet a certain proportion of selfing normally occurs.

Li (1976) has devoted a whole chapter to this kind of inbreeding, namely, Chapter 13, Equilibrium Populations with Inbreeding.

The second pattern is when inbreeding breaks the whole population into a very large number of subpopulations, each of which is finite in size and is, within itself, mating at random (see Section 6.1). Furthermore, inbreeding increases in each of these subpopulations from generation to generation. An extreme example is the subdividing of the whole population into subpopulations consisting of only one individual ($N = 1$) each of which mates randomly with itself, i.e., self-mating or fertilization. In subsequent generations one individual from the selfed progenies of each parental individual is self-fertilized again, and so on. A less extreme example would be a finite subpopulation of only two individuals ($N = 2$), somewhat analogous to the brother-sister or full-sib mating system to be discussed later (Section 6.3.2.1.). However, it might be pointed out here that the brother-sister mating system is one in which self-fertilization is excluded; matings of brother-by-brother and sister-by-sister do not occur. Thus, random mating within each subpopulation does not occur.

Falconer (1989) confines his attention entirely to this second pattern of inbreeding.
In this chapter, we consider the first pattern of inbreeding and the properties of a large inbred population. First, we consider the consequences of inbreeding upon the genotypic structures of the population. As before, we assume no migration, mutation, or selection. This whole chapter assumes only one locus with multiple alleles.

5.2. Genotypic structure of a large, random-mating population subject to inbreeding. Assumptions:

Let us consider a very large, random-mating population with one locus possessing multiple alleles as the initial base population. Its genotypic structure is given in (3.32) and the sum of all terms may be written as

\[(p_1 + \ldots + p_m)^2 = \sum_{i=1}^{m} p_i^2 + 2 \sum_{i=1}^{m} \sum_{j=1}^{m} p_i p_j = \bar{H}_0 + H_0 = 1 \tag{5.1}\]

where \(\bar{H}_0 = \sum_{i=1}^{m} p_i^2\) = frequency of homozygotes in the random-mating population (3.50),

\[H_0 = 1 - \sum_{i=1}^{m} p_i^2 = 2 \sum_{i=1}^{m} \sum_{j=1}^{m} p_i p_j = \text{frequency of heterozygotes in the random-mating population (3.51).}\]

Further, we suppose that in subsequent generations random mating of the whole population does not occur, but some inbreeding occurs by the mating of various kinds of relatives, e.g., selfing, full-sib by full-sib (brother × sister), half-sib by half-sib, first-cousin by first-cousin, etc. In addition, we suppose that in every successive generation, the individuals involved in each kind of inbreeding system are a random sample from the population. If the proportions of offspring from various mutually exclusive kinds of mating systems, including random mating, are constants, \(a_1, \ldots, a_k\), \(\sum_{i=1}^{k} a_i = 1\), from generation to generation with corresponding inbreeding coefficients \(F_{1t}, \ldots, F_{kt}, \ldots, F_{kt}\) in
generation $t$, where one $F_{it} = 0$ for random mating, then the mean or expected inbreeding coefficient of the population in generation $t$ is

$$F = F_t = \frac{F}{F_t} = \sum_{i=1}^{k} a_i F_{it}$$

(5.2)

where $F = \text{inbreeding coefficient in generation } t$ (note that the subscript $t$ is implied in this chapter).

This is the probability of two genes being identical by descent in a random individual (4.12) drawn from the population. It also means that in the whole population, the proportion $F$ of the population does possess two genes identical by descent or are identical homozygotes, and the proportion $(1 - F)$ of the population does not possess two genes identical by descent.

We desire to determine the genotypic structure of such a random-mating population subsequently inbred to the level $F$ (or $F_t$) in any given generation $t$. We will express the genotypic structure in three alternative ways as follows:

1. Genotypic structure defined as panmictic and fixed portions. The genotypic structure of such a population may be deduced as follows. Consider a random individual from the population. The probability that an individual is an identical homozygote is $F$, and the probability of the allele $A_i$ in the population is $p_i$. Since the allelic composition of the genotype is independent of the state of identity, the probability of an individual being an identical homozygote of genotype $A_iA_i$ is the product of the probabilities of the event of an identical homozygote and that of the allele $A_i$ (really these two probabilities and a conditional probability of one for the occurrence of the second $A_i$ allele given that an identical homozygote and one $A_i$ allele have occurred), namely,

$$p(\text{identical } A_iA_i) = Fp_i \quad \text{for } i = 1, \ldots, m$$

(5.3)

Similarly, the probability that an individual is not identical by descent is $(1 - F)$. Then, the probability of a nonidentical homozygote $A_iA_i$ is the product
of the three probabilities corresponding to the events of a nonidentical homozygote and the drawing of two $A_i$ alleles, since every event is independent. Thus,

$$p(\text{nonidentical } A_iA_i) = (1 - F)p_i^2$$

for $i = 1, \ldots, m$ \hspace{1cm} (5.4)

Finally, the probability of the homozygote $A_iA_i$ in the whole population is the sum of (5.3) and (5.4), namely,

$$A_iA_i: \quad p_{iiF} = Fp_i + (1 - F)p_i^2$$

for $i = 1, \ldots, m$ \hspace{1cm} (5.5)

where $p_{iiF}$ = probability of an individual of the homozygous genotype $A_iA_i$ in any generation of the population inbred to a coefficient of inbreeding $F$ (i.e., $F_t$).

In a similar manner, the probability that an individual is an unordered heterozygote $A_iA_j$ is the sum of the probabilities of the two ordered heterozygotes $A_iA_j$ and $A_jA_i$. The probabilities for each of the two ordered ways are the products of the probabilities of the three independent events, namely,

$$p(A_iA_j) = (1 - F)p_ip_j$$
$$p(A_jA_i) = (1 - F)p_jp_i$$

for $i, j = 1, \ldots, m; \ i < j$ \hspace{1cm} (5.6)

which upon summation gives the frequency of the heterozygote $A_iA_j$, namely,

$$A_iA_j: \quad 2p_{ijF} = (1 - F)p_ip_j + (1 - F)p_jp_i$$

$$= 2(1 - F)p_ip_j$$

for $i, j = 1, \ldots, m; \ i < j$ \hspace{1cm} (5.7)

where $2p_{ijF}$ = probability of an individual of the heterozygous genotype $A_iA_j$ in any generation of the population inbred to a coefficient of inbreeding $F$ (i.e., $F_t$).

The genotypic structure from (5.5) and (5.7) is

<table>
<thead>
<tr>
<th></th>
<th>Panmictic</th>
<th>Fixed</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_iA_i$</td>
<td>$p_{iiF} = (1 - F)p_i^2 + Fp_i$ for $i = 1, \ldots, m$</td>
<td></td>
</tr>
<tr>
<td>$A_iA_j$</td>
<td>$2p_{ijF} = 2(1 - F)p_ip_j$ for $i, j = 1, \ldots, m; \ i &lt; j$ \hspace{1cm} (5.8)</td>
<td></td>
</tr>
</tbody>
</table>

which may be written as a summation
\[(1 - F) \left[ \sum_{i=1}^{m} p_i^2 (A_iA_i) + 2 \sum_{i=1}^{m} \sum_{j=i+1}^{m} p_i p_j (A_iA_j) \right] + F \sum_{i=1}^{m} p_i^2 (A_iA_i) = 1 \quad (5.9)\]

Note that when \( F = 0 \), (5.9) reduces to the usual Hardy-Weinberg proportions (5.1).

The frequency of homozygotes in the inbred population is the sum of the frequencies of all homozygotes (5.5), namely,

\[
\bar{H}_F = \sum_{i=1}^{m} [(1 - F)p_i^2 + Fp_i] \]

\[
= (1 - F) \sum_{i=1}^{m} p_i^2 + F \sum_{i=1}^{m} p_i \quad \text{(see (5.1) and sub (2.4))} \]

\[
= (1 - F)\bar{H}_0 + F \quad (5.10)
\]

where \( \bar{H}_F \) = frequency of homozygotes in inbred population with inbreeding coefficient \( F \) (i.e., \( F_t \)),

\( \bar{H}_0 \) = frequency of homozygotes in original random-mating population (5.1),

and the frequency of heterozygotes is the sum of the frequencies of all heterozygotes (5.7), namely,

\[
H_F = 2(1 - F) \sum_{i=1}^{m} \sum_{j=i+1}^{m} p_i p_j = (1 - F)\bar{H}_0 \quad \text{(see (5.1))} \quad (5.11)
\]

where \( H_F \) = frequency of heterozygotes in inbred population with inbreeding coefficient \( F \) (i.e., \( F_t \)).

With only two alleles \( [p = p_1, (1 - p) = p_2] \) the frequency of the homozygotes in the noninbred random-mating population from (3.33)(3.34) is

\[
\bar{H}_0 = p^2 + (1 - p)^2 = p^2 + 1 - 2p + p^2 = 1 - 2p + 2p^2 = 1 - 2p(1 - p) \quad (5.12)
\]

and the frequency of the heterozygote from (3.33)(3.34) is

\[
H_0 = 2p(1 - p) \quad (5.13)
\]

With inbreeding, as always, the frequency of homozygotes is one minus the frequency of the heterozygotes in the same generation (5.11), namely,
\[ H_F = 1 - 2(1 - F)p(1 - p) \]  
and the frequency of the heterozygote from (5.11) is  
\[ H_F = 2(1 - F)p(1 - p) \]  
(5.15)

When the two alleles are equally frequent \([p = 1/2, (1 - p) = 1/2]\), the frequency of the homozygotes in the inbred population from (5.14) is  
\[ H_F = 1 - 2(1 - F)p(1 - p) = 1 - 2(1 - F)(1/2)(1/2) = 1 - \frac{1 - F}{2} = \frac{2 - 1 + F}{2} \]  
(5.16)

and the frequency of the heterozygote from (5.15) is  
\[ H_F = 2(1 - F)p(1 - p) = 2(1 - F)(1/2)(1/2) \]  
\[ = \frac{1 - F}{2} \]  
(5.17)

This means that regardless of the system(s) of inbreeding, individuals in a population inbred to \(F = 1/2\), say, from a noninbred random-mating population with allelic frequencies of 1/2 will have their genes alike or be homozygotes 3/4 of the time (5.16) and not alike or heterozygous 1/4 of the time (5.17). (An \(F\) value of 1/2 would normally never be reached with the pattern of inbreeding discussed in this chapter; it is the maximum possible value obtained from only complete self-fertilization and no random mating whatsoever. See Section 5.5, specifically (5.52) for \(s = 1\).)

The quantity \((1 - F)\) is often referred to as the panmictic index \(P\) because the genotypic frequencies in that portion of the population are those for a panmictic or random-mating population in Hardy-Weinberg equilibrium. Also, from (5.11) the quantity is equal to the ratio of the frequency of heterozygotes in an inbred population to the frequency of heterozygotes in a random-mating population. Thus,
\[ P = P_t - 1 - F_t - 1 - F = \frac{H_F}{H_0} \]  

where \( P = P_t \) = panmictic index in generation \( t \).

(2) **Genotypic structure defined as deviations from panmixia.** Equation (5.8) was written to reflect that the population is composed of a panmictic portion and a fixed portion. Alternatively, we may rewrite (5.8) in the form of deviations from panmixia. First, from (5.5) we have

\[ A_iA_i: \quad p_{iiF} = (1 - F)p_i^2 + Fp_i^2 = p_i^2 - Fp_i^2 + Fp_i = p_i^2 + Fp_i(1 - p_i) \quad \text{for } i = 1, \ldots, m \]  

and from (5.7)

\[ A_iA_j: \quad 2p_{ijF} = 2(1 - F)p_ip_j - 2Fp_ip_j \quad \text{for } i, j = 1, \ldots, m; \ i < j \]  

This expression clearly shows that the frequency of heterozygotes always declines with inbreeding.

The genotypic structure from (5.19) and (5.20) is

\[
\begin{array}{ll}
\text{Panmixia} & \text{Deviations from panmixia} \\
A_iA_i: & p_{iiF} = p_i^2 + Fp_i(1 - p_i) \quad \text{for } i = 1, \ldots, m \\
A_iA_j: & 2p_{ijF} = 2p_ip_j - 2Fp_ip_j \quad \text{for } i, j = 1, \ldots, m; \ i < j \\
\end{array}
\]  

The frequency of homozygotes is the sum of all homozygotes from (5.19), namely,

\[ H_F = \sum_{i=1}^{m} p_i^2 + F \sum_{i=1}^{m} p_i(1 - p_i) \]  

and the frequency of heterozygotes is the sum of all heterozygotes from (5.20), namely,
\[ H_F = 2 \sum_{i=1}^{m} \sum_{j=1 \atop i < j}^{m} p_{ij}F = 2 \sum_{i=1}^{m} \sum_{j=1 \atop i < j}^{m} p_i p_j - 2F \sum_{i=1}^{m} \sum_{j=1 \atop i < j}^{m} p_i p_j \quad \text{(sub. (5.1))} \]

\[ = H_0 - FH_0 \quad (5.23) \]

(3) **Genotypic structure defined as deviations from fixation.** Equation (5.8) may be rewritten again in the form of deviations from fixation. First, from (5.5) we have

\[ A_i A_i: \quad p_{ii}F = (1 - F)p_i^2 + Fp_i \]

\[ = p_i^2 - Fp_i^2 + Fp_i \]

\[ = p_i(p_i - Fp_i + F) \]

\[ = p_i(1 - 1 + p_i - Fp_i + F) \quad \text{(add and subtract 1)} \]

\[ = p_i[1 - (1 - Fp_i)] \]

\[ = p_i[1 - (1 - p_i)(1 - F)] \]

\[ = p_i - (1 - F)p_i(1 - p_i) \quad \text{for } i = 1, \ldots, m \quad (5.24) \]

and from (5.7)

\[ A_i A_j: \quad 2p_{ij}F = 2(1 - F)p_i p_j \]

\[ = 0 + 2(1 - F)p_i p_j \quad \text{for } i, j = 1, \ldots, m; \ i < j \quad (5.25) \]

Expression (5.25) clearly shows that if a population were completely inbred \((F = 1)\), it would be composed only of homozygotes whose frequencies equal their allelic frequencies. (However, an \(F\) value of one is never attained with this first pattern of inbreeding discussed in this chapter.)

The genotypic structure from (5.24) and (5.25) is

\[
\begin{array}{c|c|c}
\text{Fixation} & \text{Deviation} & \text{Deviation from fixation} \\
\hline
A_i A_i: & p_{ii}F = p_i - (1 - F)p_i(1 - p_i) & \text{for } i = 1, \ldots, m \\
A_i A_j: & 2p_{ij}F = 0 + 2(1 - F)p_i p_j & \text{for } i, j = 1, \ldots, m; \ i < j
\end{array}
\]

(5.26)
The frequency of homozygotes from (5.24) is

\[ H_F = \sum_{i=1}^{m} p_i - (1 - F) \sum_{i=1}^{m} p_i(1 - p_i) \]  

(5.27)

and the frequency of heterozygotes is as before (5.11).

Equations (5.8), (5.21), and (5.26) for the multiple-allelic case are presented in Table 5.1 along with the case of only two alleles \([p = p_1, q = (1 - p) = p_2]\). For this pattern of inbreeding within a population discussed herein, \(F\) is defined as in (5.2).

Table 5.1. Genotypic structure of an inbred population in any generation expressed in three equivalent ways for the multiple allelic case and for two alleles only, \(F = F_e\).  (After Li, 1976, p. 237.)

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Panmictic and fixed</th>
<th>Deviation from panmixia</th>
<th>Deviations from fixation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1A_1)</td>
<td>((1 - F)p_1^2 + Fp_1)</td>
<td>(p_1^2 + Fp_1(1 - p_1))</td>
<td>(p_1 - (1 - F)p_1(1 - p_1))</td>
</tr>
<tr>
<td>(A_1A_j)</td>
<td>(2(1 - F)p_1p_j)</td>
<td>(2p_1p_j - 2Fp_1p_j)</td>
<td>(0 + 2(1 - F)p_1p_j)</td>
</tr>
</tbody>
</table>

**Multiple alleles**

Two alleles

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Panmictic and fixed</th>
<th>Deviation from panmixia</th>
<th>Deviations from fixation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_1A_1)</td>
<td>((1 - F)p^2 + Fp)</td>
<td>(p^2 + Fpq)</td>
<td>(p - (1 - F)pq)</td>
</tr>
<tr>
<td>(A_1A_2)</td>
<td>((1 - F)pq)</td>
<td>(2pq - 2Fpq)</td>
<td>(0 + 2(1 - F)pq)</td>
</tr>
<tr>
<td>(A_2A_2)</td>
<td>((1 - F)q^2 + Fq)</td>
<td>(q^2 + Fpq)</td>
<td>(q - (1 - F)pq)</td>
</tr>
</tbody>
</table>

Regardless of the number of alleles, the frequency of any genotype is intermediate between that of a completely random-mating population and that of a completely inbred population. For only two alleles, this is clearly shown in Figure 5.1.
Figure 5.1. Genotypic proportions in populations with various degrees of inbreeding for given values of allelic frequencies. (Black = $A_1A_1$; white = $A_1A_2$; grey = $A_2A_2$.) (After Li, 1976, p. 236.)

If one is given the complete genotypic structure in any generation, one may calculate the allelic frequencies in a manner analogous to (2.81), namely,

$$P_i = P_{iiF} + \sum_{j=1, j\neq i}^{m} (2P_{ijF})(1/2) = P_{iiF} + \sum_{j=1}^{m} P_{ijF} \quad \text{for } i = 1, \ldots, m \quad (5.28)$$

Then using (5.28) one may calculate the coefficient of inbreeding in any generation as a function of the homozygous genotypic frequencies and allelic frequencies, namely, (see Box 5.1),

$$F = \frac{1}{m-1} \left[ \sum_{i=1}^{m} P_{iiF} - 1 \right] \quad (5.29)$$

In addition the coefficient of inbreeding in any generation can be calculated from the total frequency of homozygotes under random mating and that with inbreeding. From (5.10), we have

$$H_F = (1 - F)H_0 + F$$
$$= H_0 - FH_0 + F$$

$$F - FH_0 = H_F - H_0$$

$$F = \frac{H_F - H_0}{1 - H_0} \quad (5.29A)$$

Similarly, in terms of total frequency of heterozygotes from (5.11), we have
\[
H_F = (1 - F)H_0 \\
1 - F = \frac{H_F}{H_0} \\
F = 1 - \frac{H_F}{H_0}
\] (5.29B)

Box 5.1

Derivation of (5.29)

We start with an expression embodying the genotypic structure and the value \( F \) (5.22) for which we desire to solve eventually. Dividing both sides by \( p_i \), we write

\[
\frac{\sum_{i=1}^{m} p_{ii}F}{\sum_{i=1}^{m} p_i} = \frac{\sum_{i=1}^{m} p_i^2 + Fp_i(1 - p_i)}{\sum_{i=1}^{m} p_i} = \frac{\sum_{i=1}^{m} p_i[F + (1 - p_i)]}{\sum_{i=1}^{m} p_i} \quad \text{(sub. (5.19))}
\]

\[
= \sum_{i}[p_i + F(1 - p_i)] = \sum_{i} p_i + F\sum_{i}(1 - p_i) \quad \text{(sub. (2.4))}
\]

\[
= 1 + F(m - \sum_{i} p_i) = 1 + F(m - 1) \quad \text{(sub. (2.4))}
\]

\[
F = \frac{1}{m - 1} \left[ \frac{\sum_{i=1}^{m} p_{ii}F}{\sum_{i=1}^{m} p_i} - 1 \right]
\]

which is (5.29).

In a manner similar to that discussed in Section 3.3.4, populations inbred to a given value of \( F \) lie on a parabola below the parabola for random mating as shown in Figure 5.2 (see Figure 3.2 for a discussion of triangular representation).
Figure 5.2. The upper parabola represents all random-mating populations, while the lower parabola represents those populations in which $F = 1/4$. The two populations whose genotypic frequencies are represented by the points $p_1$ and $p_2$ have the same allelic frequencies ($p_A$ and $p_{\bar{A}}$).

If multiple alleles exist, one must consider the frequency $p_A$ of any particular allele versus the total frequency $p_A$ of all other alleles. (See Li, 1976, p. 239, for additional details.)

Example 5.1. Suppose that in a large population reproducing by a mixture of random mating and different inbreeding systems, the genotypic distribution was $0.08 A_1A_1$, $0.24 A_1A_2$, $0.68 A_2A_2$. Calculate the coefficient of inbreeding.

One way to obtain the inbreeding coefficient for this non Hardy-Weinberg distribution is to use (5.29). To use it, the allelic frequencies must be calculated from (5.28), namely,

$$p_1 = 0.08 + \frac{1}{2} (0.24) = 0.2$$
$$p_2 = 0.68 + \frac{1}{2} (0.24) = 0.8$$

Then substituting in (5.29), we obtain

$$F = \frac{1}{2} - \frac{1}{1} \left( \frac{0.08}{0.2} + \frac{0.68}{0.8} - 1 \right) = (0.4 + 0.85 - 1) = 0.25$$
Alternatively, the inbreeding coefficient can be obtained from (5.29A), i.e.,

\[ F = \frac{\overline{H}_F - \overline{H}_0}{1 - \overline{H}_0} = \frac{(0.08 + 0.68) - [(0.2)^2 + (0.8)^2]}{1 - [(0.2)^2 + (0.8)^2]} = 0.25 \]

or from (5.29B), i.e.,

\[ F = 1 - \frac{\overline{H}_F}{\overline{H}_0} = 1 - \frac{0.24}{2(0.2)(1 - 0.2)} = 0.25 \]

5.3. Relation between frequency of heterozygotes or homozygotes in a large inbred population and the coefficient of inbreeding. Assumptions:

From (5.23) we observe that the frequency of heterozygotes is a linear function of \( F \), namely,

\[ H_F = 2 \sum_{i=1}^{m} \sum_{j=1}^{m} p_i p_j - \left[ 2 \sum_{i=1}^{m} \sum_{j=1}^{m} p_i p_j \right] F \]

\[ = H_0 - H_0 F \]

where \( H_0 = 2 \sum_{i<j} p_i p_j \) = intercept of vertical axis,

\( -H_0 = -2 \sum_{i<j} p_i p_j \) = slope of straight line.

(Note that from the general form \( y = a + bx \), (5.30) is \( y = a - ax \) where \( b = -a \).)

This is shown graphically in Figure 5.3.
\[ H_0 = 2 \sum_{i<j} p_i p_j \]

Frequency of heterozygotes, \( H_F \) or \( 2p_{ij}F \)

\[ H_F = H_0 - H_0F \]

\[ 2p_{ij}F = 2p_i p_j - (2p_i p_j)F \]

Figure 5.3. The linear relation between the frequency of all heterozygotes or any particular heterozygote and the coefficient of inbreeding.

The change in frequency of heterozygotes in \( t \) generations may be defined in a manner similar to that for linkage disequilibrium in Section 3.4.3. as \( H_F - H_0 \), which from (5.30) is

\[
\begin{align*}
\text{change in frequency} & = H_F - H_0 = H_0 - H_0F - H_0 \\
\text{of heterozygotes in } t \text{ generations} & = -FH_0 
\end{align*}
\]

This change \( H_F - H_0 \) is always negative as (5.31) shows, because the frequency of heterozygotes always decreases with inbreeding. Hence, we consider the loss in frequency of heterozygotes \( (H_0 - H_F) \) which is the negative of the change in frequency of the heterozygotes (5.31),

\[
\begin{align*}
\text{loss in frequency} & = H_0 - H_F = - (H_F - H_0) = -(-FH_0) \\
\text{of heterozygotes in } t \text{ generations} & = FH_0 
\end{align*}
\]

Thus, the loss in frequency of heterozygotes is equal to the proportion \( F \) of the original frequency. This fact by itself makes the coefficient of inbreeding \( F \) a very useful measure in both theoretical and applied work.
By definition the **relative loss** in frequency of heterozygotes in \( t \) generations is equal to

\[
\frac{H_0 - H_F}{H_0} = \frac{FH_0}{H_0} = F \quad \text{(sub. (5.32))}
\]  

(5.33)

Although our attention has been confined to a single locus in this whole chapter, this property (5.33) is extendable to an arbitrary number of loci, because \( F \) is independent of any locus parameters such as allelic frequencies. That is, since \( F \) is the fraction of the frequency of heterozygotes lost at each and every locus, then we may say that \( F \) is the fractional loss of heterozygosity in the population considering any number of loci or all loci in the organism. Heterozygosity is the state of being heterozygous at one or more loci in an individual.

The above discussion considers the frequency of heterozygotes as a group. Turning our attention to the frequency of an individual heterozygote, the same relations exist as above for the heterozygotes as a group. From (5.20) the frequency of a single heterozygote is a linear function of \( F \) similar to that for the heterozygous group (5.30), namely,

\[
2p_{ij}F - 2p_{ij} = (2p_{ij})F \quad \text{for } i, j = 1, \ldots, m; \ i < j
\]  

(5.34)

where the slope of the line is \(-2p_{ij}\) or the negative of the initial frequency of the heterozygote \( A_iA_j \) in the random-mating population. This is also shown graphically in Figure 5.3. The **change** in frequency of the \( A_iA_j \) heterozygote in \( t \) generations is (compare (5.31))

\[
2p_{ij}F - 2p_{ij^0} = 2(1 - F)p_{ij} - 2p_{ij} = 2p_{ij}(1 - F - 1) \quad \text{(sub (5.34))}
\]

\[
= -F(2p_{ij}) \quad \text{for } i, j = 1, \ldots, m; \ i < j
\]  

(5.35)

where \( 2p_{ij^0} = 2p_{ij} \) = probability of an individual of the heterozygous genotype \( A_iA_j \) in the random mating population (3.31),
and the loss in frequency of the $A_iA_j$ heterozygote in $t$ generations is the negative of (5.35) (compare (5.32))

$$2p_{ij0} - 2p_{ijF} = F(2p_ip_j) \quad \text{for } i, j = 1, \ldots, m; \ i < j$$ (5.36)

The relative loss in $t$ generations is (compare (5.33))

$$\frac{2p_{ij0} - 2p_{ijF}}{2p_{ij0}} = \frac{F(2p_ip_j)}{2p_ip_j} - F$$ (5.37)

Thus, the relative loss in $t$ generations for any heterozygote is the same as that for all heterozygotes together (5.33).

With respect to homozygotes their frequency from (5.22) is again a straight line function of $F$ as expected, since it is $1 - H_F$ and $H_F$ is a linear function of $F$ (5.30). Thus,

$$H_F = \frac{m}{m} p_i^2 + \frac{m}{m} p_i(1 - p_i)F$$ (5.38)

This is shown graphically in Figure 5.4.

The relative change in frequency of homozygotes in $t$ generations is (5.22)

$$\frac{H_F - H_0}{H_0} = \frac{m}{m} p_i^2 + \frac{m}{m} p_i(1 - p_i) - \frac{m}{m} p_i^2$$ (5.39)
Figure 5.4. The linear relation between the frequency of all homozygotes or any particular homozygote and the coefficient of inbreeding.

The relative change is a function of $F$ as well as allelic frequency in contrast to that for the relative change or loss of frequency of heterozygotes (5.33).

For the frequency of an individual homozygote $A_i A_i$ we have

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

(5.40)

This is shown graphically in Figure 5.4.

The relative change in frequency of the homozygote $A_i A_i$ in $t$ generations is (5.19)

$$\frac{p_{iit} - p_{i0}}{p_{i0}} = \frac{p_i^2 + Fp_i(1 - p_i) - p_i^2}{p_i^2}$$

$$= F \frac{(1 - p_i)}{p_i}$$

(5.41)
The quantity \((1 - p_i)/p_i\) tends to be very large for alleles with low frequencies, as shown in Table 5.2. It is nearly inversely related to the frequency of the allele.

Table 5.2. The proportions of homozygous recessive individuals in the population with no inbreeding, \(F = 0\), and with inbreeding, \(F = 1/16\) (as though all matings in the whole population were between first cousins), and the relative change or increase of affected individuals. (After Crow and Kimura, 1970, p. 74.)

<table>
<thead>
<tr>
<th>Allelic frequency</th>
<th>(1 - p_i)</th>
<th>(1 - p_i)</th>
<th>Frequency of affected individuals</th>
<th>Change in frequency of affected individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p_i)</td>
<td>(p_i)</td>
<td>(p_{i0})</td>
<td>(p_{iF})</td>
<td>(p_{iF} - p_{i0})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(F = 0)</td>
<td>(F = 1/16)</td>
<td></td>
</tr>
<tr>
<td>.1</td>
<td>9</td>
<td>0.01</td>
<td>0.016</td>
<td>0.006</td>
</tr>
<tr>
<td>.01</td>
<td>99</td>
<td>0.0001</td>
<td>0.00072</td>
<td>0.00062</td>
</tr>
<tr>
<td>.005</td>
<td>199</td>
<td>0.000025</td>
<td>0.000336</td>
<td>0.000311</td>
</tr>
<tr>
<td>.001</td>
<td>999</td>
<td>0.000001</td>
<td>0.000063</td>
<td>0.000062</td>
</tr>
</tbody>
</table>

Thus, as the frequency of a rare recessive allele decreases, this quantity increases manyfold, so that with even a small amount of inbreeding, the relative change or relative increase (5.41) of affected individuals increases manyfold (Table 5.2). See Crow and Kimura, 1970, pp. 73 to 77, for additional considerations of this topic.

5.4. Equilibrium in a large inbred population. Assumptions:

When inbreeding occurs within a large single population and the number of generations tends to infinity, the coefficient of inbreeding approaches a small equilibrium value. Then the genotypic proportions, as given in Table 5.1, represents a stable genotypic structure for that equilibrium value of \(F\). When the
population has reached this stable genotypic composition, the genotypic proportions of the parental and offspring generations must be the same. This produces a certain balance between the frequencies of parental mating kinds such that for two alleles \((A_1 \text{ and } A_2)\), the frequency of the mating between two heterozygotes \((A_1A_2 \times A_1A_2)\) is two times the total frequency of the mating between two different homozygotes \((A_1A_1 \times A_2A_2 \text{ and } A_2A_2 \times A_1A_1)\) as discussed by Li (1976, p. 241).

This property will be proved first for only two alleles and then extended to multiple alleles. Let \(u_{k\ell}\) denote the frequency of the mating in an equilibrium population between parental genotypes \(k\) and \(\ell\) where \(k, \ell = 2, 1, 0\) refer to the number of \(A_1\) genes in the genotypes \(A_1A_1, A_1A_2, A_2A_2\), respectively. Thus, \(u_{22}\) is the frequency of the \(A_1A_1 \times A_1A_1\) mating in an equilibrium population. This is summarized below:

\[
\begin{array}{ccc}
A_1A_1 & A_1A_2 & A_2A_2 \\
2 & 1 & 0 \\
\hline
A_1A_1 (2) & u_{22} & u_{21} & u_{20} & P_{11} \\
A_1A_2 (1) & u_{12} & u_{11} & u_{10} & 2P_{12} \\
A_2A_2 (0) & u_{02} & u_{01} & u_{00} & P_{22} \\
\hline
P_{11} & 2P_{12} & P_{22} & 1
\end{array}
\]

We assume that all reciprocal matings are equal, namely,

\[u_{21} = u_{12}; \quad u_{20} = u_{02}; \quad u_{10} = u_{01}\]

(5.43)

The marginal values are the frequencies of the genotypes in the equilibrium population. In such an equilibrium population, the mating frequencies also equal or bear a simple relation to the frequencies of the corresponding offspring in the
next generation. In a manner similar to Table 3.1, the frequencies of the offspring can be calculated from the frequencies of the parents, namely,

<table>
<thead>
<tr>
<th>Kind of mating</th>
<th>Probability of mating</th>
<th>Probability of offspring</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_1A_1 \times A_1A_1$</td>
<td>$u_{22}$</td>
<td>$u_{22}$</td>
</tr>
<tr>
<td>$A_1A_1 \times A_1A_2$</td>
<td>$2u_{21}$</td>
<td>$u_{21}$</td>
</tr>
<tr>
<td>$A_1A_1 \times A_2A_2$</td>
<td>$2u_{20}$</td>
<td>$2u_{20}$</td>
</tr>
<tr>
<td>$A_1A_2 \times A_1A_2$</td>
<td>$u_{11}$</td>
<td>$1/4u_{11}$ $1/2u_{11}$ $1/4u_{11}$</td>
</tr>
<tr>
<td>$A_1A_2 \times A_2A_2$</td>
<td>$2u_{10}$</td>
<td>$u_{10}$</td>
</tr>
<tr>
<td>$A_2A_2 \times A_2A_2$</td>
<td>$u_{00}$</td>
<td>$u_{00}$</td>
</tr>
</tbody>
</table>

$p_{11} \quad 2p_{12} \quad p_{22}$

(5.44)

From (5.44) one observes that of the six mating kinds, four ($A_1A_1 \times A_1A_1$, $A_1A_1 \times A_1A_2$, $A_1A_2 \times A_2A_2$, $A_2A_2 \times A_2A_2$) produce offspring of the same genotypic frequencies as their parents. That is, mating kind $A_1A_1 \times A_1A_1$ occurs in frequency $u_{22}$ and yields only $A_1A_1$ offspring in the frequency $u_{22}$ in the offspring generation. Similarly, mating kind $A_1A_1 \times A_1A_2$ occurs in frequency $2u_{21}$, half of which consists of parents $A_1A_1$ with frequency $u_{21}$ and the other half of parents $A_1A_2$ with frequency $u_{21}$. Thus, since this mating produces offspring $A_1A_1$ and $A_1A_2$ in those same frequencies, they balance the parental frequencies, i.e., their sum equals the total frequency of that mating kind, and their sum is made up of the two genotypes in equal frequency required for the mating kind. Similarly, balancing exists for the mating kinds $A_1A_2 \times A_2A_2$ and $A_2A_2 \times A_2A_2$. Thus, we need to consider only the two remaining kinds ($A_1A_1 \times A_2A_2$ and $A_1A_2 \times A_1A_2$). In the $A_1A_1 \times A_2A_2$ mating kind with frequency $2u_{20}$, only the heterozygote is produced in the next generation. This kind produces no offspring of the parental genotypes.

Thus, all parental genotypes for this mating kind must be supplied by the last mating kind $A_1A_2 \times A_1A_2$ which is between heterozygotes. From this mating kind with frequency $u_{11}$, one-half of the offspring are homozygotes $A_1A_1$ and $A_2A_2$ in
equal frequency. Hence, this frequency \( (1/2)u_{11} \) must be equal to the frequency \( 2u_{20} \) of the other mating kind \( A_1A_1 \times A_2A_2 \) to supply the required proportion of homozygotes for that \( A_1A_1 \times A_2A_2 \) mating kind in the next generation. Thus,

\[
\begin{align*}
1/2 \ u_{11} &= 2u_{20} \\
u_{11} &= 2(2u_{20})
\end{align*}
\] (5.45)

\[
p(A_1A_2 \times A_1A_2) = 2[p(A_1A_1 \times A_2A_2) + p(A_2A_2 \times A_1A_1)]
\]
so the frequency of the heterozygous by heterozygous mating kind is two times the total frequency of the mating kind between two different homozygotes. This relation is independent of allelic frequency or amount of inbreeding. It is true for any population with a stable genotypic structure regardless of the mating system. It is clearly true for a random-mating population, i.e., \((2pq)(2pq) = 2(2p^2q^2)\) -- see Example 3.1.

This general theorem (5.45) is also true for multiple alleles in which case it may be stated as follows: The frequency of the heterozygote \( A_iA_j \) by heterozygote \( A_iA_j \) mating kind is two times the total frequency of the mating kind between the two corresponding different homozygotes \( A_iA_i \times A_jA_j \). Thus, we have

\[
p(A_iA_j \times A_iA_j) = 2[p(A_iA_i \times A_jA_j) + p(A_jA_j \times A_iA_i)] \quad \text{for } i, j = 1, \ldots, m; \ i \neq j
\] (5.46)

Note that this is true for every combination of \( i \) and \( j \).

The derivation of this general theorem is identical to that above for two alleles. We consider any two alleles or combination of \( i \) and \( j \), \( i < j \), in a multiple allelic series. Then, in the mating frequency \( u_{k\ell} \), let \( k \) and \( \ell \) denote the number of \( i \) alleles 2, 1, 0 in the genotypes \( A_iA_i \), \( A_iA_j \), and \( A_jA_j \), respectively. Then, the argument given above for only two alleles applies equally well to the case for any combination of \( i \) and \( j \).
5.5. **Mixture of self-mating and random mating.** Assumptions:

In many economically important plant species a mixture of mating systems occurs. The most important and simplest mixture is that when a certain proportion of offspring in each generation is derived from self-fertilization and the remaining proportion from random mating. The most noted economic species with the least disparity between the two proportions is cotton. In that species the proportion of selfing averages about 70%, but varies widely from 40 to 95%, increasing as the prevalence of insect vectors decreases. In most other species, such as maize or corn, the proportion of self-mating is much lower -- normally less than 1%. (See Bijlsma, Allard, and Kahler, 1986, for nonrandom mating in maize.) In normally self-fertilized species, such as barley and wheat, the reverse proportions normally occur, i.e., about 99% self-mating and 1% random mating.

We desire to derive the inbreeding coefficient $F_t$ in generation $t$. To do so we let $s$ be the proportion of the population that is produced by self-mating and $1 - s$ the proportion produced by random mating in every generation. In generation 1, $s$ is the proportion of the population obtained from selfing and that proportion would have an inbreeding coefficient $F_1 = 1/2$ (see Sections 4.3 and 4.4, and Example 4.3). In generation 2, $s^2$ would be the proportion of the population with an immediate history of two generations of selfing and that proportion would have an inbreeding coefficient of $F_2 = 3/4$, $(1 - s)s$ the proportion of the population with an immediate history of one generation of selfing and an inbreeding coefficient of $F_2 = 1/2$. In generation 3, $s^3$, $(1 - s)s^2$, $(1 - s)s$ are the proportions with an immediate history of 3, 2, and 1 generations of selfing, respectively, and those proportions would have inbreeding coefficient of $7/8$, $3/4$ and $1/2$, respectively. In general, in generation $t$ there is a proportion $s^t$ of the population with a history of continuous selfing for $t$ generations and that
proportion has an inbreeding coefficient \( F_t = 1 - \left( \frac{1}{2} \right)^t \) (see (6.23)). In addition, there are proportions \((1 - s)s^n\) with an immediate history of \(n = 1, 2, \ldots, t - 1\) generations of selfing, each with an inbreeding coefficient \(F_t = 1 - \left( \frac{1}{2} \right)^n\). Thus, the inbreeding coefficient is (see Box 5.1A).

\[
F_t = s^t \left[ 1 - \left( \frac{1}{2} \right)^t \right] + \sum_{n=1}^{t-1} (1-s)s^n \left[ 1 - \left( \frac{1}{2} \right)^n \right]
\]  

(5.46A)

### Box 5.1A

**Derivation of (5.46A)**

By definition

\[
F_t = s^t \left[ 1 - \left( \frac{1}{2} \right)^t \right] + \sum_{n=1}^{t-1} (1-s)s^n \left[ 1 - \left( \frac{1}{2} \right)^n \right]
\]

\[
= s^t \left[ 1 - \left( \frac{1}{2} \right)^t \right] + \sum_{n=1}^{t-1} s^n \left[ 1 - \left( \frac{1}{2} \right)^n \right] - s \sum_{n=1}^{t-1} s^n \left[ 1 - \left( \frac{1}{2} \right)^n \right]
\]

\[
= \sum_{n=1}^{t} s^n - \sum_{n=1}^{t} \left( \frac{s}{2} \right)^n - s \sum_{n=1}^{t-1} s^n \left[ 1 - \left( \frac{1}{2} \right)^n \right]
\]

\[
= \sum_{n=1}^{t} s^n - \sum_{n=2}^{t} \left( \frac{s}{2} \right)^n - s \sum_{n=1}^{t} \left[ \frac{s}{2} \right]^n
\]

\[
= s + 2 \sum_{n=2}^{t} \left( \frac{s}{2} \right)^n - s \sum_{n=2}^{t} \left[ \frac{s}{2} \right]^n
\]

\[
= s + 2 \sum_{n=2}^{t} \left( \frac{s}{2} \right)^n - \frac{s}{2} - s \sum_{n=2}^{t} \left[ \frac{s}{2} \right]^n
\]

\[
= \frac{s}{2} - \sum_{n=2}^{t} \left[ \frac{s}{2} \right]^n
\]

\[
= \sum_{n=2}^{t} \left[ \frac{s}{2} \right]^n
\]

which is (5.46A).
Alternatively, from (5.2) the inbreeding coefficient in generation \( t \) is

\[
F_t = sF_{1t} + (1 - s)F_{2t} \\
= sF_{1t} + (1 - 2)0 \\
= sF_{1t}
\]

(assume \( F_{2t} = 0 \))

where \( F_{1t} \) = mean inbreeding coefficient in generation \( t \) resulting from

inbreeding system 1, self-mating,

\( F_{2t} = 0 \) = inbreeding coefficient in generation \( t \), resulting from mating

system 2, random mating, in a large population.

The mean inbreeding due to selfing from one generation to the next can be

expressed as a recurrence relation \( F_{1t} = (1/2)(1 + F_{1t-1}) \). (6.17), so that upon

substitution in (5.46B), we have

\[
F_t = \frac{s}{2} (1 + F_{1t-1})
\]

(5.47)

(5.47)

By successive substitution of (5.47) we obtain

\[
F_t = \frac{s}{2} \left[ 1 + \frac{s}{2} \left( 1 + F_{1t-2} \right) \right] \\
= \frac{s}{2} \left[ 1 + \frac{s}{2} \left[ 1 + \frac{s}{2} (1 + F_{1t-3}) \right] \right] \\
= \frac{s}{2} + \left( \frac{s}{2} \right)^2 + \left( \frac{s}{2} \right)^3 + \left( \frac{s}{2} \right)^3 F_{1t-3} \\
= \frac{s}{2} + \left( \frac{s}{2} \right)^2 + \left( \frac{s}{2} \right)^3 + \ldots + \left( \frac{s}{2} \right)^t + \left( \frac{s}{2} \right)^t F_0 \quad (\text{assume } F_0 = 0)
\]

(5.48)

where \( s \) = proportion of population produced by self-fertilization,

\( t \) = number of generation, 0, 1, \ldots,

\( F_0 \) = coefficient of inbreeding of generation 0.

This derivation assumes that the individuals to be self-fertilized in every

generation are a random sample from the population, i.e., there is no tendency for

the offspring from self-fertilizing parents to self-fertilize again. The sum of
the first \( t \) terms is the sum of a geometric series which is given as (see any algebra book or mathematical handbook)

\[
S = \frac{a(1 - r^n)}{1 - r}
\]  
\[(5.49)\]

where \( S = \) sum of \( n \) terms in a geometric series,

\( a = \) value of first term in geometric series,

\( r = \) common ratio in geometric series,

\( n = \) number of terms in geometric series.

Thus, the sum of the first \( t \) terms in (5.48) is

\[
S = \frac{\frac{s}{2} \left[ 1 - \left( \frac{s}{2} \right)^t \right]}{1 - \frac{s}{2}} = \frac{2}{2 - s} \left( \frac{s}{2} \right)^t \left[ 1 - \left( \frac{s}{2} \right)^t \right] = \frac{s}{2 - s} \left[ 1 - \left( \frac{s}{2} \right)^t \right]
\]  
\[(5.50)\]

Substituting (5.50) in (5.48), we obtain

\[
F_t = \frac{s}{2 - s} \left[ 1 - \left( \frac{s}{2} \right)^t \right] + \left( \frac{s}{2} \right)^t F_0
\]  
\[(5.51)\]

As \( t \) increases, \( F_t \) approaches \( s/(2 - s) \), i.e.,

\[
F_\infty = \lim_{t \to \infty} F_t = \frac{s}{2 - s}
\]  
\[(5.52)\]

where \( F_\infty = \) inbreeding coefficient in the inbred equilibrium population as \( t \) approaches infinity.

As an example, suppose the amount of self-fertilization is 10%, \( s = 0.1 \), then the inbreeding coefficient from (5.51) in successive generations, \( t = 0, 1, 2, \ldots \), \( t \to \infty \) is equal to

\[
F_t: \quad \frac{1}{20} = 0.05; \quad \frac{399}{7600} = 0.0525; \quad \ldots \quad \frac{1}{19} = 0.0526315
\]  
\[(5.53)\]

Hence, from generation 1 to \( \infty \), \( F \) increases very little -- only 0.0026315 (= 0.0526315 - 0.05) or 0.26%.

The genotypic structure in any generation \( t \) is given by substituting \( F_t \) for \( F \) in any of the expressions given in Table 5.1.
The frequency of heterozygotes is given by (5.23), namely,

$$H_F = H_{F_t} = H_0 - F_t H_0$$  \hspace{1cm} (5.54)

Alternatively, by substituting (5.51) in (5.54) one may obtain $H_F$ in terms of $s$, namely (see Box 5.2),

$$H_F = \left[ \frac{2(1 - s)}{2 - s} + \frac{s}{2 - s} \left( \frac{s}{2} \right)^t \right] H_0$$  \hspace{1cm} (5.55)

**Box 5.2**

**Derivation of (5.55)**

Starting with (5.54) and substituting (5.51), we obtain (note $F_0 = 0$)

$$H_F = H_0 - F_t H_0 = (1 - F_t) H_0 = (1 - \left\{ \frac{s}{2 - s} \left( \frac{s}{2} \right)^t \right\} \left( \frac{s}{2} \right)^t F_0) H_0$$

$$= \left[ 1 - \frac{s}{2 - s} \left( \frac{s}{2} \right)^t \right] H_0 = \left[ \frac{2 - s}{2 - s} - \frac{s}{2 - s} \left( \frac{s}{2} \right)^t \right] H_0$$

$$= \frac{2 - s}{2 - s} \left( \frac{s}{2} \right)^t H_0 = \frac{2 - s - s \left( \frac{s}{2} \right)^t}{2 - s} H_0$$

$$= \left[ \frac{2 - 2s + s \left( \frac{s}{2} \right)^t}{2 - s} \right] H_0 = \left[ \frac{2(1 - s)}{2 - s} + \frac{s}{2 - s} \left( \frac{s}{2} \right)^t \right] H_0$$

which is (5.55).

Li (1976, p. 244) has a practical suggestion to plant breeders for a way to measure $s$, the amount of self-fertilization in any species. To do so one needs a population with a marker gene which is 'neutral' (not influenced by selection) and has incomplete dominance so that the frequency of the heterozygote as well as the corresponding homozygotes can be identified. The population is grown on an area of adequate size and harvested in bulk for enough generations so that the population approaches equilibrium with respect to genotypic proportions. Then
with complete genotypic classification the inbreeding coefficient $F_\infty$ at equilibrium can be calculated (5.29) (5.29A) (5.29B). Then by substituting $F_\infty$ in (5.52) we solve for the proportion of self-mating in the system, namely (see Box 5.3),

$$s = \frac{2F_\infty}{1 + F_\infty}$$  \hspace{1cm} (5.56)

**Box 5.3**

**Derivation of (5.56)**

Starting with (5.52), we solve for $s$,

$$F = \frac{s}{2 - s}$$

$$(2 - s)F = s$$

$$2F - sF = s$$

$$s + sF = 2F$$

$$s(1 + F) = 2F$$

$$s = \frac{2F}{1 + F}$$

which is (5.56).

Jain (1979) has reviewed many proposed methods for the estimation of the rate of selfing or outcrossing.

REFERENCES

Bijlsma, R., R. W. Allard, and A. L. Khaler

Cockerham, C. Clark.

Crow, J. F., and M. Kimura

Elandt-Johnson, R. C.

Falconer, D. S.

Jain, S. K.

Kempthorne, O.

Li, C. C.