TABLE OF CONTENTS

LIST OF TABLES .......................................................... vi
LIST OF FIGURES .......................................................... viii
1. INTRODUCTION ......................................................... 1
2. GENETIC DISTANCE AND THE COANCESTRY COEFFICIENT .......... 3
   2.1 Background ...................................................... 3
   2.2 The Model ........................................................ 8
   2.3 Notation .......................................................... 9
3. ESTIMATION OF THE COANCESTRY COEFFICIENT USING MULTILOCUS GENOTYPIC DATA .............................................. 15
   3.1 Two Loci Each with Two Codominant Alleles ................. 15
      3.1.1 The Variance-Covariance Matrix of the Estimators of the Variance Components .......................... 24
      3.1.2 The Asymptotic Distribution of $g$ as $a \to \infty$ .... 34
      3.1.3 The Estimation of $\gamma$ .................................. 35
      3.1.4 Large Sample Variances of Some Estimators of $\theta$ .. 42
   3.2 Many Loci Each with Two Codominant Alleles ............... 44
   3.3 A Special Case--Random Mating ................................ 48
      3.3.1 One Locus .................................................. 62
      3.3.2 Two Loci .................................................. 64
4. A SIMULATION STUDY ................................................ 71
   4.1 Discussion of Programming ..................................... 71
   4.2 Results .......................................................... 75
5. DOMINANCE, MULTIPLE ALLELES AND OTHER EXTENSIONS .......... 82
   5.1 Dominance ....................................................... 82
   5.2 Multiple Alleles ................................................ 85
   5.3 Mutation and Migration ......................................... 89
6. DISCUSSION .......................................................... 92
7. SUMMARY .............................................................. 98
8. LIST OF REFERENCES ................................................ 100
TABLE OF CONTENTS (Continued)

9. APPENDICES ....................................................... 103

9.1 Derivation of the Variance–Covariance Matrix of the Vector $\bar{s}$ ............................................. 104
9.2 Asymptotic Multivariate Normality of $s$ as $a \rightarrow \infty$ .................. 116
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Descent measures for one locus</td>
<td>7</td>
</tr>
<tr>
<td>2.2</td>
<td>Nonidentity descent measures for two loci</td>
<td>12</td>
</tr>
<tr>
<td>2.3</td>
<td>Frequencies of gene combinations at two loci</td>
<td>13</td>
</tr>
<tr>
<td>3.1</td>
<td>A bivariate analysis of variance of the random vector X_{ijk}</td>
<td>18</td>
</tr>
<tr>
<td>3.2</td>
<td>Gametic counts for the i^{th} sampled population</td>
<td>20</td>
</tr>
<tr>
<td>3.3</td>
<td>The elements of the submatrices of Λ</td>
<td>26</td>
</tr>
<tr>
<td>3.4</td>
<td>The elements of the submatrices of Ξ</td>
<td>30</td>
</tr>
<tr>
<td>3.5</td>
<td>The submatrices of Λ and Ξ for the special case of random mating</td>
<td>51</td>
</tr>
<tr>
<td>3.6</td>
<td>Values of contrasts in two locus nonidentity measures for a monoecious population size N = 1000, with linkage parameter λ, at various times (t)</td>
<td>61</td>
</tr>
<tr>
<td>3.7</td>
<td>Asymptotic standard deviations (sd) of ṭ and ᴭ for subdivided monoecious populations of size N = 100 at various generation times (t)</td>
<td>65</td>
</tr>
<tr>
<td>3.8</td>
<td>Asymptotic standard deviations (sd) of 新房, ṭ, ᴥ and ᴭ for 30 subdivided monoecious populations of size N = 100 with linkage parameter λ = 0.0</td>
<td>65</td>
</tr>
<tr>
<td>3.9</td>
<td>Asymptotic standard deviations (sd) of 新房, ṭ, ᴥ and ᴭ for 30 subdivided monoecious populations of size N = 100 with linkage parameter λ = 1.0</td>
<td>67</td>
</tr>
<tr>
<td>4.1</td>
<td>Realized means, standard deviations, mean square errors and variances for four estimators of θ for the case of two monoecious populations diverging from an initial reference population with parameters α₁ = α₂ = .25 and λ = 0</td>
<td>76</td>
</tr>
<tr>
<td>4.2</td>
<td>Realized means, standard deviations, mean square errors and variances for four estimators of θ for the case of two monoecious populations diverging from an initial reference population with parameters α₁ = α₂ = .25 and λ = 0.9</td>
<td>77</td>
</tr>
<tr>
<td>4.3</td>
<td>Realized means, standard deviations, mean square errors and variances for four estimators of θ for the case of two monoecious populations diverging from an initial reference population with parameters α₁ = .24, α₂ = .09 and λ = 0</td>
<td>78</td>
</tr>
</tbody>
</table>
LIST OF TABLES (Continued)

4.4 Realized means, standard deviations, mean square errors and variances for four estimators of $\theta$ for the case of two monoecious populations diverging from an initial reference population with parameters $\alpha_1 = .24$, $\alpha_2 = .09$ and $\lambda = 0.9$. 79

9.1 Translations of redundant two-locus notation to one-locus notation .............................. 108
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The model</td>
<td>8</td>
</tr>
<tr>
<td>3.1</td>
<td>Asymptotic coefficients of variation (CV) of $\tilde{\theta}$, $\tilde{\theta}$, $\hat{\theta}$ and $\hat{\theta}_W$ for samples of size 50 from 20 monoecious populations of size $N = 1000$ with $\alpha_1 = 0.0196$, $\alpha_2 = 0.2496$ and $\lambda = 0.9$</td>
<td>70</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

There is some interest among geneticists in measurements, based on electrophoretic data, of the genetic diversity among populations and in the relationship of such measurements to particular models for the evolution of populations. The study of such measures of genetic diversity or distance has been described by Kirk (1977) as a "major growth industry" and that author has added that "new measures of distance come off the assembly line almost as frequently as new car models." This work is concerned with a preexisting measure of genetic distance and addresses the problem of estimating this distance when multilocus data are available.

The need for a genetic distance, from which time since divergence can be recovered, for instances of short-term evolution is described in Chapter 2. The model upon which such a genetic distance is based is explicated and the notation utilized in the rest of the thesis is also described in that chapter.

In Chapter 3 an analysis of variance with a structure reflecting the model presented in Chapter 2 is described. Such an analysis of variance was presented by Cockerham (1969) for the case of one locus and Cockerham and Weir (1977) for the case of two loci. Variances and covariances of the estimated variance components in this analysis of variance are presented. The expressions for these variances and covariances are completely general and apply to any regular system of mating. They are consequently rather tedious to derive and complete details of their derivations are put in an appendix. These expressions may be translated into functions of descent measures and initial gene frequencies
and this translation is simplified if linkage equilibria are assumed to exist in the initial reference population at all pairs of scored loci. Four estimators of the distance measure are introduced and some limited properties of these estimators are presented, the most notable being that for large numbers of replicate populations the estimators are asymptotically normal. The asymptotic variances of these estimators are compared for the special case of monoecious random mating populations with random selfing. The results of a small simulation study designed to illuminate the small sample properties of these estimators for the case of monoecy and two-scored loci are presented in Chapter 4.

The effects of some failures of assumptions on the methods presented in Chapters 3 and 4 are briefly described in Chapter 5 and a discussion of the relationship of this study to previous studies and recommendations for future research are presented in Chapter 6. A summary, Chapter 7, completes the thesis.
2. GENETIC DISTANCE AND THE COANCESTRY COEFFICIENT

A genetic distance is a scalar quantity which aims to measure how different a pair of populations, or possibly several populations, are in their genetic constitution. The literature concerning genetic distances is vast but has been thoroughly reviewed by, for example, Goodman (1972), Smith (1977) and Nei (1978a). In this chapter, a brief discussion of previous studies is presented, the use of gene identity measures as genetic distances for isolated populations which have evolved over a short time is described, and the model and notational tools upon which this thesis is founded are presented.

2.1 Background

Prior to 1970 most genetic distances (see, for example, Smith, 1977) were constructed to serve solely as scalar-valued discriminants between currently constituted populations. As such, they were closely related to geometric distances with populations being represented by points in Euclidean space. If two populations had similar "genetic constitutions," that is, similar gene frequencies, their genetic distance was by construction small. Often small genetic distance was construed to mean that the pair of populations were recently descended from a common ancestral population (the underlying model of the evolution of the populations being that of a tree rather than a web). The recovery of time since divergence from these distances was next to impossible as the distances were based on analyses which were local in time.

The need for a genetic distance from which divergence time could be recovered was initially met by Nei (1972, 1973), Latter (1973) and Morton (1973) although such distances were inherent in the work of
Malécot (1948, 1969), Wright (1965) and Cockerham (1967, 1969). The "standard genetic distance" proposed by Nei (1972) for the study of the long-term evolution of two populations X and Y is given by

\[ D = -\log_e I \]

where

\[ I = \frac{J_{XY}}{\sqrt{J_X J_Y}}, \]  

(2.1.1)

\( J_{XY} \) is the probability that a gene chosen at random from population X is identical in state to a gene chosen at random from population Y, and \( J_X \) is the probability that two randomly chosen genes from population X are identical in state, with a similar definition for \( J_Y \). If more than one locus is studied, the \( J \)'s are arithmetically averaged over loci before inclusion in (2.1.1). The model of evolution which Nei used to arrive at his distance measure appears to have the following features:

(i) Initially, at time zero, a random mating population in Hardy-Weinberg equilibrium, in linkage equilibrium at all pairs of loci, and in drift-mutation equilibrium "splits" into two identical populations which remain isolated.

(ii) At the unknown time \( t \) at which the genetic distance between the two populations is required, the populations are still in the three types of equilibrium mentioned in (i).

(iii) All new mutations are unique (an infinite-alleles model).

(iv) No selection.

(v) Equal population sizes or equal effective population sizes.

The effect of assumption (i) is to eliminate a source of sampling variation due to the fission process and to ensure that \( I \) at time zero is equal to one. The infinite-alleles model and this splitting process
also ensure that "identity in state" is equivalent to "identity by descent" so that Nei is able to use the recursion for identity given by Kimura and Crow (1964) to obtain at time $t$,

$$I \sim e^{-2vt}$$

where $v$ is the mutation rate. Thus,

$$D \sim 2vt$$

and this standard genetic distance is approximately a linear function of divergence time. In Nei's treatment of the variance of estimators of $I$ (Nei and Roychoudhury, 1974; Nei, 1978b) the total variance of an estimator of $I$ can be partitioned into components for loci and genes within loci, because loci are assumed to be independent and mating is assumed to be random. Cockerham, Weir and Reynolds (1981), however, show that when these assumptions do not hold, the partition of variation of identity measures generates labels for populations, loci, loci by populations, individuals within populations, and loci by individuals within populations. Moreover, in the partition of the total variance of actual inbreeding, the component of variance for loci is zero as each locus has the same expected inbreeding, but how these components relate to the components of variance for estimators of $I$ is not clear.

Latter (1973) suggested that the mean coefficient of kinship within populations or the average coancestry coefficient, where the averaging is essentially over all pairs of individuals within populations and all replicate populations, may be used as a measure of genetic distance between two populations for the case of drift alone and drift plus mutation to novel alleles. He presented an estimator of this distance which, apart from a small correction factor, was a weighted average over
all scored alleles and all scored loci of the estimator of the co-
ancestry coefficient proposed by Cockerham (1969) for the case of two
alleles at one locus.

Cockerham (1979) proposed that the coancestry coefficient, denoted
by $\bar{\theta}_1$ or $\theta$ in the sequel and defined as the average of the $N(N-1)/2$
coancestries $\theta_{XY}$ of pairs of individuals $X$ and $Y$ in a population of
size $N$ could be used as a genetic distance between a pair or among
several populations. The coancestry $\theta_{XY}$ of a pair of individuals $X$
and $Y$ within a population is defined in Table 2.1 which is adapted from
Cockerham (1971). One feature of the coancestry coefficient is that
for many mating systems involving finite populations it is a mono-
tonically increasing function of generation time $t$ so that given
information about the mating system and making some assumptions about
the fission process, time since fission or divergence can be recovered
from knowledge of $\bar{\theta}_1$. For example, for a monoecious population of
size $N$ mating at random including a random amount of selfing, $\bar{\theta}_1$ at
time $t$ is equal to

$$1 - \left(1 - \frac{1}{2N}\right)^t$$

so that

$$t = \frac{\log_e (1-\bar{\theta}_1)}{\log_e \left(1 - \frac{1}{2N}\right)}$$

$$= -2N \log_e (1-\bar{\theta}_1)$$

while for a monoecious population of size $N$ mating at random but
excluding selfing $\bar{\theta}_1$ at time $t$ is equal to

$$1 - \left(\frac{1+r}{2}\right) \left[\frac{1}{2} \left(\frac{N-1}{N} + r^{-1}\right)\right]^t - \left(\frac{1-r}{2}\right) \left[\frac{1}{2} \left(\frac{N-1}{N} - r^{-1}\right)\right]^t$$
Table 2.1 Descent measures for one locus.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Identity relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_X$</td>
<td>$x \equiv x'$</td>
</tr>
<tr>
<td>$\theta_{XY}$</td>
<td>$x \equiv y$</td>
</tr>
<tr>
<td>$\gamma_{XY}$</td>
<td>$x \equiv x' \equiv y$</td>
</tr>
<tr>
<td>$\gamma_{XYZ}$</td>
<td>$x \equiv y \equiv z$</td>
</tr>
<tr>
<td>$\delta_{XY}$</td>
<td>$x \equiv x' \equiv y \equiv y'$</td>
</tr>
<tr>
<td>$\delta_{XYZ}$</td>
<td>$x \equiv x' \equiv y \equiv z$</td>
</tr>
<tr>
<td>$\delta_{XYZW}$</td>
<td>$x \equiv y \equiv z \equiv w$</td>
</tr>
<tr>
<td>$\Delta_{X+i}$</td>
<td>$x \equiv y, x' \equiv y'$</td>
</tr>
<tr>
<td>$\Delta_{X,y}$</td>
<td>$x \equiv x', y \equiv y'$</td>
</tr>
<tr>
<td>$\Delta_{X+YZ}$</td>
<td>$x \equiv y, x' \equiv z$</td>
</tr>
<tr>
<td>$\Delta_{X,YZ}$</td>
<td>$x \equiv x', y \equiv z$</td>
</tr>
<tr>
<td>$\Delta_{XY,ZW}$</td>
<td>$x \equiv y, z \equiv w$</td>
</tr>
</tbody>
</table>

*a*Different upper case subscripts denote different random individuals.

*b*Lower case letters denote random genes from corresponding upper case individuals. Primes denote different genes. Nothing is implied about genes not shown to be identical by descent.
where \( r = \frac{N}{\sqrt{N^2 + 1}} \) and \( t \) may be recovered by tabular or graphical means.

Recognizing that the assumptions concerning equilibria in the sampled populations, which are made in the derivation of Nei's measure of genetic distance are inappropriate for populations which have only recently split from a common ancestral population, the remainder of this chapter, and the thesis, is devoted to the use of the coancestry coefficient and other functions of gene identity measures as genetic distances.

2.2 The Model

The model for the fission process and other sampling processes is presented diagrammatically in Figure 2.1. An initial reference population,

![Diagram showing the model](image)

**Figure 2.1** The model. Straight arrows represent sampling of offspring from progeny arrays. Jagged arrows represent the formation of the isolated populations.
noninbred, essentially infinite, in Hardy-Weinberg equilibrium at each locus, but not necessarily in linkage equilibrium at every pair of loci, is imagined. Replicate populations begin as independent random samples of size $N$ from the initial reference population at time zero. Generations are discrete, there is no selection, and replicate populations are assumed to remain isolated and constant in size. Each replicate population at a certain generation time is a random sample of size $N$ from the conceptually infinite offspring array generated according to the rules of the mating system by that replicate population in the previous generation. The samples at generation $t$ are from the conceptually infinite offspring arrays generated by the replicate populations in generation $t-1$ so that the sample size, $n$, may be greater than the size, $N$, of each replicate population.

This model, which has a long history in population genetics and essentially dates back to the work of Wahlund (see, for example, Cockerham, 1973) differs in several key respects to the model utilized by Nei (1972, 1978a). Firstly, the founder populations are not necessarily identical and as independent random samples of finite size $N$ may not be in Hardy-Weinberg equilibrium or linkage equilibrium. Secondly, this model is not restricted to binary fission, and thirdly, mating systems other than monoecy can be subsumed under the model.

2.3 Notation

In the sequel, several notations are used to denote various probabilities of states of identity between and among genes. While all these notations have appeared in the literature before, they are gathered here for ease of reference. The necessary one-locus descent
measures have been given in Table 2.1. These one-locus measures are expectations over all replicate populations, and over all pairs, triples or quadruples of individuals (whichever the case may be) in each replicate population.

Initially, in Chapter 3, the digenic descent measures $\frac{1}{1} F_1$, $\frac{1}{1} \bar{F}$ and $\frac{1}{1} \bar{\theta}$ are used to express equivalence relationships between genes at different loci. Definitions of these descent measures can be found in Cockerham and Weir (1977). Later in Chapter 3 and the rest of the thesis, particularly in the derivation of variances and covariances of estimated variance components, linkage disequilibria between pairs of loci in the initial reference population are assumed to be zero so that these descent measures and their precursors in Cockerham and Weir (1973) are not required in the translation of frequencies of gene combinations to functions of descent measures and gene frequencies.

The notation for the frequencies of various gene combinations is given in Weir and Hill (1980) and Weir, Cockerham and Reynolds (1981). For combinations of genes at one locus, horizontal and vertical bars separate genes from different individuals, so that, for example,

$$ p_A^A $$

is the probability that a random gene from a locus in a randomly chosen individual is allele A and that a randomly chosen gene from the same locus in another randomly chosen individual is also allele A. This probability is given by

$$ p_A^A = \frac{1}{1} p_A + (1 - \frac{1}{1}) p_A^2 $$
where $p_A$ is the frequency of allele A in the initial reference population. As another example, consider the probability that three genes chosen at random from the same locus in three different randomly chosen individuals are all allele A. This probability, denoted by

$$\frac{A|A}{A}.$$

is given by

$$\gamma_{XYZ}p_A^3 + 3(\theta-\gamma_{XYZ})p_A^2 + (1-3\theta+2\gamma_{XYZ})p_A^3.$$

For combinations of genes at two loci, the necessary notation is given in Tables 2.2 and 2.3 which are both adapted from Weir and Hill (1980). In Table 2.3, horizontal and vertical bars separate genes in different individuals while diagonal bars separate genes on different gametes within the same individual. For example, assuming that linkage disequilibrium in the initial population is zero, the frequency of the double homozygote AABB is given by

$$p_{AB}^{AB} = F_{11}p_Ap_B + F_{10}p_A^2p_B + F_{01}p_A^2p_B + F_{00}p_A^2p_B$$

(2.3.1)

$$= (1-2\theta+\theta_1)p_Ap_B + (\theta-\theta_1)(p_A^2p_B + p_A^2p_B) + \theta_1p_A^2p_B^2.$$

(2.3.2)

Equation (2.3.1) can be deduced from Table IV of Cockerham and Weir (1973) and involves probabilities of identity by descent. The equivalent expression in equation (2.3.2) involves the probabilities of nonidentity given in Tables 2.2 and 2.3 and reduces to the appropriate expression in Table 2.3.

As another example consider the probability that in two randomly chosen individuals, one is homozygous for allele A at the first locus and a randomly chosen gene from the second locus of this individual is
Table 2.2 Nonidentity descent measures for two loci.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Gene arrangements</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Theta_1 )</td>
<td>((ab</td>
</tr>
<tr>
<td>( \Theta_2 )</td>
<td>((ab)(a'b'))</td>
</tr>
<tr>
<td>( \Gamma_1 )</td>
<td>((ab)(a'</td>
</tr>
<tr>
<td>( \Gamma_2 )</td>
<td>((ab</td>
</tr>
<tr>
<td>( \Gamma_3 )</td>
<td>((ab)(a')(b'))</td>
</tr>
<tr>
<td>( \Delta_1^* )</td>
<td>((a</td>
</tr>
<tr>
<td>( \Delta_2^* )</td>
<td>((a</td>
</tr>
<tr>
<td>( \Delta_3^* )</td>
<td>((a</td>
</tr>
<tr>
<td>( \Delta_4^* )</td>
<td>((a</td>
</tr>
<tr>
<td>( \Delta_5^* )</td>
<td>((a)(a')(b)(b'))</td>
</tr>
</tbody>
</table>

\( a^* \) The probability that genes \( a, a' \) at one locus are not identical by descent, and genes \( b, b' \) at a second locus are also not identical by descent.

\( b \) Bars separate genes on separate gametes and parentheses separate genes in separate individuals.
### Table 2.3 Frequencies of gene combinations at two loci.

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Coefficient of</th>
<th>( P_{AB} )</th>
<th>( P_{AB}(2 - P_A - P_B) )</th>
<th>( P_A(1 - P_A)P_B(1 - P_B) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{AB} ) + ( P_{AB} )</td>
<td>2</td>
<td>- ( p )</td>
<td>( \theta_1 )</td>
<td></td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- ( p )</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- ( \Pi )</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- (( P + \Pi ))</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- ( \Pi )</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- (( P + \Pi ))</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- ( \Pi )</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- ( \Pi )</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- (( P + \Pi ))</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- ( \Pi )</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- (( P + \Pi ))</td>
</tr>
<tr>
<td>( P_{A</td>
<td>B} + P_{A</td>
<td>B} )</td>
<td>2</td>
<td>- ( \Pi )</td>
</tr>
</tbody>
</table>

\( a \quad P = 1 - F_X, \quad \Pi = 1 - \theta_{XY} \).
allele B, while a randomly chosen gene from the second locus of the second individual is allele B. This probability is given by

\[ P_{AB}^{A/B} = (1-P-\Pi+\tau_2)P_{A}P_{B} + (\Pi-\tau_2)P_{A}P_{B}^2 + (P-\tau_2)P_{A}^2P_{B} + \tau_2P_{A}^2P_{B} \]

\[ = P_{A}P_{B} - P_{A}P_{B}(1-P_{A}) - \Pi P_{A}P_{B}(1-P_{B}) + \tau_2P_{A}(1-P_{A})P_{B}(1-P_{B}). \]

The advantage of working with nonidentity rather than identity when studying gene combinations involving two loci is simply that the transition equations for the two-locus nonidentity measures are homogeneous (Weir, Avery and Hill, 1980).
3. ESTIMATION OF THE COANCESTRY COEFFICIENT USING
MULTILOCUS GENOTYPIC DATA

Estimation of the coancestry coefficient in the one-locus case,
when the unit of observation is the gene, has been treated extensively
by Cockerham (1969, 1973). As a paradigm for further work, the estimation
of the coancestry coefficient when two loci each with two codominant alleles are scored, is discussed in some detail.

3.1 Two Loci Each with Two Codominant Alleles

For the k\textsuperscript{th} gamete in the j\textsuperscript{th} individual in the i\textsuperscript{th} population, consider the random vector of indicator random variables

\[ x_{ijk} = \begin{bmatrix} x_{ijkl} \\ x_{ijk2} \end{bmatrix} \quad (3.1.1) \]

where

\[ x_{ijkl} = 1 \text{ if gamete } k \text{ in individual } j \text{ in population } i \text{ carries} \]

allele A at the first locus,

\[ = 0 \text{ otherwise; } \]

and where,

\[ x_{ijk2} = 1 \text{ if gamete } k \text{ in individual } j \text{ in population } i \text{ carries} \]

allele B at the second locus,

\[ = 0 \text{ otherwise. } \]

Such a random vector is defined in Cockerham and Weir (1977). The expectation of this random vector over the three stages of sampling (a population from an infinite array of replicate populations, n individuals from an infinite progeny array in each replication population and a complete sample or census of the two haplotypes or gametes within each
sampled individual) is simply the vector of frequencies of allele A and allele B in the initial population,

$$E(x_{ijk}) = \mu = \begin{bmatrix} p_A \\ p_B \end{bmatrix}.$$  

Expectations of various products of these random vectors are as follows:

$$E(x_{ijk}^T x_{ijk}) = \begin{bmatrix} p_A & p_{AB} \\ p_{AB} & p_B \end{bmatrix} = \begin{bmatrix} p_A & p_A p_B + \bar{F}_1 \Delta_{AB} \\ p_A p_B + \bar{F}_1 \Delta_{AB} & p_B \end{bmatrix}$$

$$= \Sigma + \mu \mu^T,$$

$$E(x_{ijk}^T x_{i'jk'}) = \begin{bmatrix} p_A & p_{AB} \\ p_{AB} & p_B \end{bmatrix} = \begin{bmatrix} p_A + \bar{F}_1 p_A (1-p_A) & p_A p_B + \bar{F}_1 \Delta_{AB} \\ p_A p_B + \bar{F}_1 \Delta_{AB} & p_B + \bar{F}_1 p_B (1-p_B) \end{bmatrix}$$

$$= c_{ab} + \mu \mu^T, \quad k \neq k'$$

$$E(x_{ijk}^T x_{ij'k'}) = \begin{bmatrix} p_A & p_{AB} \\ p_{AB} & p_B \end{bmatrix} = \begin{bmatrix} p_A^2 & p_A p_B + \bar{\theta} \Delta_{AB} \\ p_A p_B + \bar{\theta} \Delta_{AB} & p_B^2 \end{bmatrix}$$

$$= c_{a} + \mu \mu^T, \quad j \neq j'$$

$$E(x_{ijk}^T x_{i'j'k'}) = \begin{bmatrix} 2p_A & p_A p_B \\ p_A p_B & p_B^2 \end{bmatrix} = \mu \mu^T, \quad i \neq i'.$$
In Table 3.1 a bivariate analysis of variance for the random vector \( x_{ijk} \) is presented. Two decompositions of the bivariate analogs of expected mean squares are given. The first decomposition is standard (see, for example, Bock, 1975) and is in terms of variance-covariance component matrices. The second decomposition is a bivariate analog of the covariance representation given by Cockerham (1969). It should be noted that the first decomposition differs a little from that given by Cockerham (1969, 1973) and Cockerham and Weir (1977). Those authors write expected mean squares as linear combinations of components of variation which are equivalent to the "polykays" of Tukey (1956) and the "cap sigmas" of Zyskind (1962). The decomposition, in Table 3.1, which takes into account the complete sampling at the third stage (gametes within individuals) may be slightly more customary. Certainly the condition for a parametrically negative component of variation for individuals within populations is more stringent with the present decomposition. The condition is \( \vec{F}_1 < \bar{\theta}_1 \) with the previous authors' decomposition but \( \frac{1}{2}(1+\vec{F}_1) < \bar{\theta}_1 \) with the present decomposition.

Quadratic unbiased estimators of the variance-covariance component matrices \( E_c, E_b \) and \( E_a \) are, respectively:

\[
S_c = \frac{1}{an} \left[ \sum_{i,j,k} x_{ijk}^T x_{ijk} - \frac{1}{2} \sum_{i,j} \left( \sum_k x_{ijk}^T x_{ijk} \right) \right]
\]

\[
= \frac{1}{2an} \left[ \frac{N_{11}}{N_1} + \frac{N^2}{N_1} - \frac{1}{2} \left( N_{11} + N_{21} \right) \right]
\]

\[
\left[ \begin{array}{cc}
\frac{1}{N_2} + \frac{1}{N_{11}} & \frac{1}{N_2} + \frac{1}{N_{11}} \\
\frac{1}{N_2} + \frac{1}{N_{11}} & \frac{1}{N_2} + \frac{1}{N_{11}} \\
\frac{N_{12} + 2 N_{21} + N_{21}}{N_{12}} & \frac{N_{12} + 2 N_{21} + N_{21}}{N_{12}} \\
\frac{N_{12} + 2 N_{21} + N_{21}}{N_{12}} & \frac{N_{12} + 2 N_{21} + N_{21}}{N_{12}} \\
\end{array} \right]
\]
Table 3.1 A bivariate analysis of variance of the random vector $X_{ijk}$.

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Matrices of Sums of Squares and Cross Products (SSP)</th>
<th>$\delta$ (SSP/df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1</td>
<td>$\frac{1}{2an} \sum_{i,j,k} (X_{i} X_{j} X_{k} X_{ijkl})$</td>
<td>$2\lambda_{b} + 2n\lambda_{a} + 2an\gamma_{AB} = \gamma_{T} + C_{ab} + 2(n-1)C_{a} + 2(n-1)\gamma_{AB}$</td>
</tr>
<tr>
<td>Populations</td>
<td>a</td>
<td>$\frac{1}{2n} \sum_{i,j,k} (X_{i} X_{j} X_{k} X_{ijkl}) - \frac{1}{2an} \sum_{i,j,k} (X_{i} X_{j} X_{k} X_{ijkl})$</td>
<td>$2\lambda_{b} + 2n\lambda_{a} = \gamma_{T} + C_{ab} + 2(n-1)C_{a}$</td>
</tr>
<tr>
<td>Individuals within Populations</td>
<td>a(n-1)</td>
<td>$\frac{1}{2} \sum_{i,j,k} (X_{i} X_{j} X_{k} X_{ijkl}) - \frac{1}{2n} \sum_{i,j,k} (X_{i} X_{j} X_{k} X_{ijkl})$</td>
<td>$2\gamma_{b} = \gamma_{T} + C_{ab} - 2C_{a}$</td>
</tr>
<tr>
<td>Gametes within Individuals</td>
<td>an</td>
<td>$\sum_{i,j,k} (X_{i} X_{j} X_{k} X_{ijkl})$</td>
<td>$\gamma_{c} = \gamma_{T} - C_{ab}$</td>
</tr>
<tr>
<td>Total</td>
<td>2an</td>
<td>$\sum_{i,j,k} (X_{i} X_{j} X_{k} X_{ijkl})$</td>
<td>$\frac{1}{2} \gamma_{c} + \gamma_{b} + \gamma_{a} + \gamma_{AB} = \gamma_{T} + \gamma_{AB}$</td>
</tr>
</tbody>
</table>

\[
\Sigma_{a} = \begin{bmatrix}
\hat{\delta}_{1}p_{A}(1-p_{A}) & 1 & 0 & \delta_{AB} \\
\hat{\delta}_{1}p_{B}(1-p_{B}) & 0 & \delta_{AB} & 0
\end{bmatrix}, \quad \Sigma_{b} = \frac{1}{2} \begin{bmatrix}
(1 + \hat{\delta}_{1} - 2\hat{\delta}_{0})p_{A}(1-p_{A}) & (\hat{\delta}_{1} + 1 - 2\hat{\delta}_{1})p_{A} & (1 - \hat{\delta}_{1})p_{A} & (\hat{\delta}_{1} - 1)p_{A} \\
(\hat{\delta}_{1} + 1 - 2\hat{\delta}_{0})p_{B}(1-p_{B}) & (1 + \hat{\delta}_{1} - 2\hat{\delta}_{0})p_{B} & (1 - \hat{\delta}_{1})p_{B} & (\hat{\delta}_{1} - 1)p_{B}
\end{bmatrix}, \quad \Sigma_{c} = \begin{bmatrix}
(1 - \hat{\delta}_{1})p_{A}(1-p_{A}) & (\hat{\delta}_{1} - 1)p_{A} \\
(\hat{\delta}_{1} - 1)p_{B}(1-p_{B}) & (1 - \hat{\delta}_{1})p_{B}
\end{bmatrix}
\]
\[ S_b = \frac{1}{2a(n-1)} \left[ \frac{1}{2} \sum_{i,j,k} (\Sigma_{i,k} x_{i,j,k}^T)(\Sigma_{i,j,k}) - \frac{1}{2n} \sum_i (\Sigma_{i,k} x_{i,j,k}^T)(\Sigma_{i,k} x_{i,j,k}) \right] \]

\[ = \frac{1}{4a(n-1)} \begin{bmatrix}
4 N_{11}^1 + N_{11}^2 + N_{11}^3 \\
N_{11}^1 + N_{11}^2 + 2 N_{11}^3 \\
- \frac{1}{n} \sum_i (2_i^1 N_{12}^i + 2_i^1 N_{11}^i + N_{11}^2) \\
- \frac{1}{n} \sum_i (2_i^1 N_{12}^i + 2_i^1 N_{11}^i + N_{11}^2) \\
x (2_i^1 N_{12}^i + 2_i^1 N_{11}^i + N_{11}^2)
\end{bmatrix}

\[ N_{11}^1 + N_{11}^2 + 2 N_{11}^3 \\
N_{11}^1 + N_{11}^2 + 2 N_{11}^3 \\
4 N_{11}^1 + N_{11}^2 + N_{11}^3 \\
4 N_{11}^1 + N_{11}^2 + N_{11}^3 \\
x (2_i^1 N_{12}^i + 2_i^1 N_{11}^i + N_{11}^2)
\]
where the genotype counts (the N's) are defined in Table 3.2. The estimators of the variances are in fact minimum-variance-quadratic-unbiased as the third and fourth moments of the components of $x_{ijk}$ are finite (Graybill and Hultquist, 1961).

Table 3.2 Gametic counts for the $i^{th}$ sampled population.

<table>
<thead>
<tr>
<th>Gamete 1</th>
<th>Gamete 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>AB $''$</td>
</tr>
<tr>
<td>$i_{11}$</td>
<td>$i_{11}$</td>
</tr>
<tr>
<td>$i_{12}$</td>
<td>$i_{12}$</td>
</tr>
<tr>
<td>$i_{21}$</td>
<td>$i_{21}$</td>
</tr>
<tr>
<td>$i_{22}$</td>
<td>$i_{22}$</td>
</tr>
<tr>
<td>A'B</td>
<td>A'B $''$</td>
</tr>
<tr>
<td>$i_{11}$</td>
<td>$i_{11}$</td>
</tr>
<tr>
<td>$i_{12}$</td>
<td>$i_{12}$</td>
</tr>
<tr>
<td>$i_{21}$</td>
<td>$i_{21}$</td>
</tr>
<tr>
<td>$i_{22}$</td>
<td>$i_{22}$</td>
</tr>
</tbody>
</table>

It should be noted that while the off-diagonal elements of these observed variance-covariance component matrices require the detection of coupling and repulsion double heterozygotes, the diagonal elements do not share this requirement and merely require single locus genotypic counts.

Several approaches which may lead to the recovery of the coancestry coefficient from these observed variance-covariance matrices are possible. One approach, as $\bar{c}_1$ in the one-locus case is an intraclass correlation, is to search for bivariate analogs of the intraclass correlation and take
appropriate scalar functions of them. Two possibilities for bivariate analogs of the intraclass correlation are (where \( \div \) denotes elementwise division):

\[
R_1 = \Sigma_a + \Sigma_T = \begin{bmatrix}
\bar{\theta}_1 & 1 \bar{\theta}/\bar{F}^1 \\
1 \bar{\theta}/\bar{F}^1 & \bar{\theta}_1
\end{bmatrix}
\]

and

\[
R_2 = \Sigma_a \Sigma_T^{-1} = \frac{1}{\hat{\alpha}_1 \hat{\alpha}_2 - (\bar{F}^1 \Delta_{AB})^2} \begin{bmatrix}
\bar{\theta}_1 \hat{\alpha}_1 \hat{\alpha}_2 - \frac{1}{\bar{\theta}} \bar{F}^1 \Delta_{AB} & (1 - \bar{\theta}^1 \bar{F}^1) \Delta_{AB} \hat{\alpha}_1 \\
(1 - \bar{\theta}^1 \bar{F}^1) \Delta_{AB} \hat{\alpha}_2 & \hat{\alpha}_1 \hat{\alpha}_2 - \frac{1}{\bar{\theta} \bar{F}^1 \Delta_{AB}}
\end{bmatrix}
\]

where \( \hat{\alpha}_1 = p_A (1 - p_A) \) and \( \hat{\alpha}_2 = p_B (1 - p_B) \).

Three scalar functions which either extricate \( \bar{\theta}_1 \) or at least come close to extricating \( \bar{\theta}_1 \) are as follows:

\[
\frac{1}{2} \text{tr}(R_1) = \bar{\theta}_1 , \quad (3.1.2)
\]

\[
\frac{1}{2} \text{tr}(R_2) = \frac{\bar{\theta}_1 \hat{\alpha}_1 \hat{\alpha}_2 - \frac{1}{\bar{\theta}} \bar{F}^1 \Delta_{AB}}{\hat{\alpha}_1 \hat{\alpha}_2 - (\bar{F}^1 \Delta_{AB})^2} , \quad (3.1.3)
\]

and

\[
\lambda_{\text{max}}(R_2) = \frac{\bar{\theta}_1 \hat{\alpha}_1 \hat{\alpha}_2 - \frac{1}{\bar{\theta}} \bar{F}^1 \Delta_{AB} + (1 - \bar{\theta}^1 \bar{F}^1) \Delta_{AB} \sqrt{\hat{\alpha}_1 \hat{\alpha}_2}}{\hat{\alpha}_1 \hat{\alpha}_2 - (\bar{F}^1 \Delta_{AB})^2} , \quad (3.1.4)
\]

where tr denotes the trace of a matrix and \( \lambda_{\text{max}} \) denotes the largest eigenvalue. When there is no linkage disequilibrium in the initial population (\( \Delta_{AB} = 0 \)), all three scalar functions are equal to \( \bar{\theta}_1 \). Both equation (3.1.3) and equation (3.1.4) along with eight other scalar functions have been suggested by Ahrens (1976) as possible generalizations of the intraclass correlation coefficient. The procedure
suggested by Ahrens for estimating such quantities is simply to replace \( \Sigma_a \) and \( \Sigma_T \) in equations (3.1.3) and (3.1.4) by \( S_a \) and \( S_T \), respectively, where \( S_T = \frac{1}{2} S_c + S_b + S_a \). Thus, for example, the quantity in equation (3.1.2) and subsequently \( \bar{\theta}_1 \) is estimated by an unweighted average, namely,

\[
\bar{\theta} = \frac{1}{2} \frac{S_a (1,1)}{S_T (1,1)} + \frac{S_a (2,2)}{S_T (2,2)} \quad (3.1.5)
\]

Another suggestion, due to Cockerham (1979), is to use the weighted average,

\[
\tilde{\theta} = \frac{S_a (1,1) + S_a (2,2)}{S_T (1,1) + S_T (2,2)} \quad (3.1.6)
\]

as an estimator of the coancestry coefficient. The rationale behind this suggestion is that when the observed variance components in equation (3.1.6) are replaced by their expectations, the result is \( \text{tr}(\Sigma_a) / \text{tr}(\Sigma_T) = \bar{\theta}_1 \). As estimators of the coancestry coefficient, these quantities in equations (3.1.5) and (3.1.6) do not readily appear to have their genesis in some optimization problem.

Recognizing the nonlinear structure of \( \Sigma_a \), \( \Sigma_b \) and \( \Sigma_c \), another approach is to adapt the methodology of Jöreskog (1970) and Browne (1974), reviewed in Jöreskog (1978), for estimating the parameters in covariance matrices with nonlinear structure, to the problem of estimating the coancestry coefficient.

In order to simplify the discussion some changes in notation are made. Let,

\[
S_a = \begin{bmatrix} z_1 & y \\ y & z_2 \end{bmatrix}, \quad \Sigma_a = \begin{bmatrix} \theta_1 & D_a \\ D_a & \theta_2 \end{bmatrix}
\]
\[ S_b = \begin{bmatrix} w_1 & v \\ v & w_2 \end{bmatrix}, \quad \Sigma_b = \begin{bmatrix} (\frac{1+F}{2} - \theta)\alpha_1 & D_b \\ D_b & (\frac{1+F}{2} - \theta)\alpha_2 \end{bmatrix} \]

\[ S_c = \begin{bmatrix} r_1 & u \\ u & r_2 \end{bmatrix}, \quad \text{and} \quad \Sigma_c = \begin{bmatrix} (1-F)\alpha_1 & D_c \\ D_c & (1-F)\alpha_2 \end{bmatrix} \]

where \( \alpha_1 = p_A(1-p_A), \alpha_2 = p_B(1-p_B), D_a = 1/\Delta_{AB}, D_b = \frac{1}{2}(\bar{F}^1 + \bar{F}^2)\Delta_{AB}, \) and \( D_c = (\bar{F}^2 - \bar{F})\Delta_{AB} \) and where the subscripts and bars have been deleted from the inbreeding coefficient (\( \bar{F}^1 \)) and the coancestry coefficient (\( \bar{\theta}^1 \)).

Noting that the expectations of \( y, v \) and \( u \) do not depend on \( F, \theta, \alpha_1 \) or \( \alpha_2 \), only the diagonal elements of the variance-covariance component matrices need to be considered. Let,

\[ s^T = [z_1, z_2, w_1, w_2, r_1, r_2] \quad (3.1.7) \]

Then,

\[ E(s^T) = \sigma^T(\gamma) = [\theta\alpha_1, \theta\alpha_2, (\frac{1+F}{2} - \theta)\alpha_1, (\frac{1+F}{2} - \theta)\alpha_2, (1-F)\alpha_1, (1-F)\alpha_2] \]

where,

\[ \gamma^T = [\alpha_1, \alpha_2, \theta, F] \quad (3.1.8) \]

The estimation problem can now be viewed as the fitting of the vector \( \sigma(\gamma) \), which depends on four parameters, to the observed vector \( s \) containing six observations and bears a close resemblance to the type of problem considered by Browne (1974).

Before proceeding further, the variance-covariance matrix of the observed vector \( s \) is presented in the next section. Knowledge of this matrix allows asymptotic variances of estimators such as \( \tilde{\gamma} \) and \( \tilde{\delta} \) to be derived and, of course, influences the choice of objective function to be minimized in the fitting of \( \sigma(\gamma) \) to \( s \).
3.1.1 The Variance-Covariance Matrix of the Estimators of the Variance Components

The variance-covariance matrix of the vector $\mathbf{s}$ of the estimated variance components can be written in the form:

$$
\text{Var}(\mathbf{s}) = \frac{1}{a^2} + \frac{1}{an} + \frac{1}{an(n-1)} \text{T} + \frac{1}{a(a-1)} \phi + \frac{1}{a(a-1)n} \psi + \frac{1}{a(a-1)n^2} \Omega.
$$

(3.1.9)

The component matrices $\phi$, $\psi$ and $\Omega$ are as follows:

$$
\phi = 2 \begin{bmatrix}
(p_A^2 - p_A^2) & (p_A^2 - p_A p_B^2) \\
(p_A^2 - p_A p_B^2) & (p_B^2 - p_B^2)
\end{bmatrix}
$$

$$
\psi = 2 \begin{bmatrix}
(p_A + p_A^2 - 2p_A^2) & (p_A^2 + p_A^2 - 2p_A^2) \\
(p_A + p_A^2 - 2p_A^2) & (p_B + p_B^2 - 2p_B^2)
\end{bmatrix}
$$

$$
\Omega = \frac{1}{2} \begin{bmatrix}
(p_A + p_A^2 - 2p_A^2)^2 & (p_A + p_A^2 - 2p_A^2)^2 \\
(p_A + p_A^2 - 2p_A^2)^2 & (p_B + p_B^2 - 2p_B^2)^2
\end{bmatrix}
$$

The component matrix $T$ is as follows:

$$
T = \frac{1}{2} \begin{bmatrix}
2^A & -2^A \\
-2^A & 2^A \\
& & 4^0 \\
& & & 4^0
\end{bmatrix}
$$
where the submatrix $A$ is given by,

$$
A = \begin{bmatrix}
    p^A_A + 2p^A_A^2 & p^A_A & p^A_A + p^A_A \frac{p_{AB}}{p_{AB}} & p^A_A + 2p^A_A \frac{p_{AB}}{p_{AB}} + p^A_A \\
    -4(p^A_A^2 + p^A_A - p^A_A^3) & -4(p^A_A \frac{p_{AB}}{p_{AB}} + p^A_A \frac{p_{AB}}{p_{AB}} - p^A_A) \\
    p^A_A \frac{p_{AB}}{p_{AB}} + 2p^A_A \frac{p_{AB}}{p_{AB}} + p^A_A \frac{p_{AB}}{p_{AB}} & p^A_A \frac{p_{AB}}{p_{AB}} + 2p^A_A \frac{p_{AB}}{p_{AB}} + p^A_A \frac{p_{AB}}{p_{AB}} \\
    -4(p^A_A \frac{p_{AB}}{p_{AB}} + p^A_A \frac{p_{AB}}{p_{AB}} - p^A_A) & -4(p^A_A \frac{p_{AB}}{p_{AB}} + p^A_A \frac{p_{AB}}{p_{AB}} - p^A_A)
\end{bmatrix}
$$

The component matrices $\Lambda$ and $\Xi$ may be partitioned into $2 \times 2$ submatrices in the manner,

$$
\Lambda = \begin{bmatrix}
    K & Q & R \\
    Q^T & L & S \\
    R^T & S^T & M
\end{bmatrix}
$$

$$
\Xi = \begin{bmatrix}
    B & E & G \\
    E^T & C & H \\
    G^T & H^T & D
\end{bmatrix}
$$

The elements of the submatrices of $\Lambda$ and $\Xi$ are presented in Tables 3.3 and 3.4, respectively. Complete details of the derivation of these variances and covariances of the elements of the vector $\mathbf{s}$ are relegated to Appendix 9.1.

It is clear that the variance-covariance matrix of $\mathbf{s}$ depends not only on the parameter vector $\gamma^T = [\alpha_1, \alpha_2, \theta, F]$ but also on the three-gene, four-gene and two-gene-pair probability functions for one locus introduced by Cockerham (1971) and on the two locus descent measures discussed in Weir, Avery and Hill (1980) and Weir and Hill (1980).

Note, however, that the assumption of linkage equilibrium in the initial population has simplified the expression, in terms of descent measures, of those elements of the matrix which involve two loci. If this assumption is false, the more general formulation of Cockerham and Weir (1973) is required.
### Table 3.3: The elements of the submatrices of $A$.

<table>
<thead>
<tr>
<th>Elements</th>
<th>Expression in terms of frequencies of gene combinations</th>
<th>Expression in terms of descent measures $^\dagger$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K(1,1)$</td>
<td>$\frac{\lambda_{AA}}{A</td>
<td>A} - (\frac{\lambda_{AA}}{A})^2 - 4\rho_{AA}(\frac{\lambda_{AB} + \lambda_{BA}}{A</td>
</tr>
<tr>
<td>$K(1,2) = K(2,1)$</td>
<td>$\frac{\lambda_{AB}}{A</td>
<td>B} - (\frac{\lambda_{AB}}{A</td>
</tr>
<tr>
<td>$K(2,2)$</td>
<td>$\rho_{AB}^2 (\frac{\lambda_{AB}}{B</td>
<td>B} - 4\rho_{AB}(\frac{\lambda_{AB}}{B</td>
</tr>
<tr>
<td>$L(1,1)$</td>
<td>$\frac{1}{4} \left[ \frac{\lambda_{AA}}{A} + 2\rho_{AB} - 4 \left( \frac{\lambda_{AA}}{A</td>
<td>A} + \frac{\lambda_{AB}}{A</td>
</tr>
<tr>
<td>$L(1,2) = L(2,1)$</td>
<td>$\frac{1}{4} \left[ \frac{\lambda_{BB}}{B} + 2\rho_{AB} - 2 \left( \frac{\lambda_{AB}}{A</td>
<td>B} + \frac{\lambda_{BA}}{A</td>
</tr>
<tr>
<td>$L(2,2)$</td>
<td>$\frac{1}{4} \left[ \frac{\lambda_{BB}}{B} + 2\rho_{AB} - 4 \left( \frac{\lambda_{AB}}{A</td>
<td>B} + \frac{\lambda_{BA}}{A</td>
</tr>
<tr>
<td>Elements</td>
<td>Expression in terms of frequencies of gene combinations</td>
<td>Expression in terms of descent measures $^7$</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>M(1,1)</td>
<td>$p_A^2 - 2p_A^1 + 1 - p_A^A + p_A^A - (p_A - p_A)^2$</td>
<td>$(\theta - 2\gamma_{XY} + 6\gamma_{XY})p_A(1-p_A) + (-4\gamma + 8\gamma_{XY} + 2\gamma_{X,Y} - 6\gamma_{XY} - 2\gamma_{XYZ})^2 p_A(1-p_A)^2$</td>
</tr>
<tr>
<td>M(1,2) = M(2,1)</td>
<td>$p_A^B - (p_A^A + p_A^A)^B + p_A^A + (p_A - p_A^A)(p_B - p_B^A)$</td>
<td>$[\Delta^2 - (1-\gamma)^2]p_A(1-p_A)p_B(1-p_B)$</td>
</tr>
<tr>
<td>M(2,2)</td>
<td>$p_B^B - 2p_B^B + p_B^B - (p_B - p_B)^2$</td>
<td>$(\theta - 2\gamma_{XY} + 6\gamma_{XY})p_B(1-p_B) + (-4\gamma + 8\gamma_{XY} + 2\gamma_{X,Y} - 6\gamma_{XY} - 2\gamma_{XYZ})^2 p_B(1-p_B)^2$</td>
</tr>
<tr>
<td>Q(1,1)</td>
<td>$\left[ \frac{1}{2} \left[ (p_A^A + p_A^A) - 2p_A^2 + 2p_A^1 + (p_B^A - p_B^A)(p_B - p_B^2)</td>
<td></td>
</tr>
</tbody>
</table><p>ight] - 2p_A(p_A^A + p_A^A - 2p_A^1) \right]$ | $\left[ \frac{1}{2} \left[ (\gamma_{XYZ} + 6\gamma_{XYZ} - 2\gamma_{XYZ}^2)p_A(1-p_A) + (\gamma_{X,Y} - 2\gamma_{XYZ} - 6\gamma_{XYZ} - 6\gamma_{XY} - 2\gamma_{XYZ}^2) \right] + 12\gamma_{XYZ} - 2\gamma_{XYZ} + 2\gamma_{XYZ} + 2\gamma_{XYZ} - 2\gamma_{XYZ} - 2\gamma_{XYZ} - 2\gamma_{XYZ} \right] + \Delta^2 - (1-\gamma)^2 \right]$ |
| Q(1,2)   | $\left[ \frac{1}{2} \left[ (p_A^B + p_A^B) - 2p_A^1 + 2p_A^1 - 2p_A^2 \right] - 2p_A(p_A^B + p_A^B - 2p_A^1) \right]$ | $\left[ \frac{1}{2} \left[ (1+F - 2\gamma)(1-\gamma) + \gamma_{X,Y} - 2\gamma_{XYZ} + \gamma_{XYZ} - 6\gamma_{XYZ} - 6\gamma_{XY} - 2\gamma_{XYZ} \right] + 12\gamma_{XYZ} - 2\gamma_{XYZ} + 2\gamma_{XYZ} + 2\gamma_{XYZ} - 2\gamma_{XYZ} - 2\gamma_{XYZ} - 2\gamma_{XYZ} \right] + \Delta^2 - (1-\gamma)^2 \right]$ |
| Q(2,1)   | $\left[ \frac{1}{2} \left[ (p_A^B + p_A^B) - 2p_A^1 + 2p_A^1 - 2p_A^2 \right] - 2p_A(p_A^B + p_A^B - 2p_A^1) \right]$ | $\left[ \frac{1}{2} \left[ (1+F - 2\gamma)(1-\gamma) + \gamma_{X,Y} - 2\gamma_{XYZ} + \gamma_{XYZ} - 6\gamma_{XYZ} - 6\gamma_{XY} - 2\gamma_{XYZ} \right] + 12\gamma_{XYZ} - 2\gamma_{XYZ} + 2\gamma_{XYZ} + 2\gamma_{XYZ} - 2\gamma_{XYZ} - 2\gamma_{XYZ} - 2\gamma_{XYZ} \right] + \Delta^2 - (1-\gamma)^2 \right]$ |</p>
<table>
<thead>
<tr>
<th>Elements</th>
<th>Expression in terms of frequencies of gene combinations</th>
<th>Expression in terms of descent measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q(2,2)</td>
<td>( \frac{1}{2} \left[ \frac{B}{A} + \frac{B}{B} - 2 \frac{B}{A} \right] )</td>
<td>( \frac{1}{2} \left[ \frac{B}{A} + \frac{B}{A} - 2 \frac{B}{A} \right] )</td>
</tr>
<tr>
<td>R(1,1)</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
</tr>
<tr>
<td>R(1,2)</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
</tr>
<tr>
<td>R(2,1)</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
</tr>
<tr>
<td>R(2,2)</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
<td>( \left( \frac{M}{A} + \frac{M}{A} \right)^2 )</td>
</tr>
</tbody>
</table>
Table 3.3 (Continued)

<table>
<thead>
<tr>
<th>Elements</th>
<th>Expression in terms of frequencies of gene combinations</th>
<th>Expression in terms of descent measures†</th>
</tr>
</thead>
</table>
| S(1,1)   | $\frac{1}{2} \left[ p_A^A - p_A^A A - 2 \left( p_A^A A - p_A^A A \right) - \left( p_A - p_A^A \right) \left( p_A + p_A^A - 2 p_A^A \right) \right]$ | $\frac{1}{2} \left[ \left( B - 2 \delta_{XY} - 2 \gamma_{XYZ} - \delta_{XYZ} \right) p_A \left( 1 - p_A \right) \right.$  \[+ \left. \left( -4 \delta_{XY} + 6 \delta_{XY} + 2 \delta_{XYZ} + 8 \gamma_{XYZ} + 2 \left( \delta_{XY} + 2 \gamma_{XYZ} \right) \right. \right.$  \[- \left. 12 \delta_{XYZ} + p^2 - 2p \delta \right) p_A^A \left( 1 - p_A^A \right)^2 \right] |}
| S(1,2)   | $\frac{1}{2} \left[ p_A^B + p_A^B - 2 \left( p_A^B - p_A^B \right) - \left( p_A - p_A^B \right) \left( p_A + p^B - 2 p_B \right) \right]$ | $\frac{1}{2} \left[ \left( 2 \delta_{XY}^A - \delta_{XY}^A \right) + \left( -1 + \gamma \right) \left( 1 + p - 2 \theta \right) \right] p_A \left( 1 - p_A \right) p_B \left( 1 - p_B \right)$ |}
| S(2,1)   | $\frac{1}{2} \left[ p_A^B + p_A^B - 2 \left( p_A^B + p_A^B \right) - \left( p_A - p_A^B \right) \left( p_A + p^B - 2 p_B \right) \right]$ | $\frac{1}{2} \left[ \left( 2 \delta_{XY}^B - \delta_{XY}^B \right) + \left( -1 + \gamma \right) \left( 1 + p - 2 \theta \right) \right] p_A \left( 1 - p_A \right) p_B \left( 1 - p_B \right)$ |}
| S(2,2)   | $\frac{1}{2} \left[ p_B^B + p_B^B - 2 \left( p_B^B + p_B^B \right) - \left( p_B - p_B^B \right) \left( p_B + p^B - 2 p_B^B \right) \right]$ | $\frac{1}{2} \left[ \left( 2 \delta_{XY}^B - \delta_{XY}^B \right) + \left( -1 + \gamma \right) \left( 1 + p - 2 \theta \right) \right] p_A \left( 1 - p_A \right) p_B \left( 1 - p_B \right)$ |}

† No initial linkage disequilibrium has been assumed.
Table 3.4 The elements of the submatrices of $E$.

<table>
<thead>
<tr>
<th>Elements</th>
<th>Expression in terms of frequencies of gene combinations</th>
<th>Expression in terms of descent measures$^\dagger$</th>
</tr>
</thead>
</table>
| B(1,1)   | $2 \left[ \frac{p_A^2 + p_A B}{A} - \frac{p_A^2 + p_A B}{A} - 2 p_A \left( \frac{p_A + p_A}{A} - 2 \frac{p_A^2 + p_A}{A} \right) \right] p_A \left( p_A + p_A^2 \right)$ | $2 \left[ \left( \gamma_{xyz} + 6 \xi_{xyz} - 2 \delta_{xyzw} \right) p_A \left( 1 - p_A \right) \\
+ \left( \beta - 4 \gamma_{xyz} + \Delta_X YZ + 2 \Delta_X YZ - 6 \delta_{xyz} - 6 \delta_{xyzw} + 12 \xi_{xyzw} \right) p_A^2 \left( 1 - p_A \right)^2 \right]$ |
| B(1,2) = B(2,1) | $2 \left[ \frac{p_A + p_A B}{A} - 2 \left( \frac{p_A + p_A B}{A} - 2 \frac{p_A^2 + p_A}{B} \right) \right] p_A \left( p_A + p_A \right)$ | $2 \left( \frac{\Gamma_3 - 2 \Delta_X^2 + \Delta_X^3}{p_A} \right) p_A \left( 1 - p_A \right) p_B \left( 1 - p_B \right)$ |
| B(2,2)   | $2 \left[ \frac{p_A + p_A B}{A} - 2 \left( \frac{p_A + p_A B}{A} - 2 \frac{p_A^2 + p_A}{B} \right) \right] p_A \left( p_A + p_A \right)$ | $2 \left[ \left( \gamma_{xyz} + \Delta_X YZ - 2 \delta_{xyz} \right) p_B \left( 1 - p_B \right) \\
+ \left( \beta - 4 \gamma_{xyz} + \Delta_X YZ + 2 \Delta_X YZ - 6 \delta_{xyz} - 6 \Delta_X YZ + 12 \xi_{xyzw} \right) p_B^2 \left( 1 - p_B \right)^2 \right]$ |
| C(1,1)   | $\frac{1}{8} \left[ 7 p_A + 2 \left( 5 p_A + 14 p_A B \right) + p_A \right]$ | $\frac{1}{8} \left[ \left( 1 + 7 p - 2 \left( 50 + 14 \gamma_{xy} + 5 \xi_{xy} \right) + 32 \left( \gamma_{xyz} + 6 \xi_{xyz} + 2 \delta_{xyzw} \right) \right) p_A \left( 1 - p_A \right) \\
+ \left( -2 - 2 \gamma_{xy} \right) + \left( 5 + 14 \gamma_{xy} \right) + 2 \left( \Delta_X YZ - 2 \gamma_{xy} \right) + 2 \Delta_X YZ - 2 \gamma_{xy} \right) p_A \left( 1 - p_A \right) \right]$ |
| C(1,2) = C(2,1) | $\frac{1}{8} \left[ \left( \frac{p_A + p_A B}{A} + 2 \left( \frac{p_A + p_A B}{A} + 2 \left( \frac{p_A + p_A B}{A} \right) + \frac{p_A + p_A B}{A} \right) \right] p_A \left( p_A + p_A \right)$ | $\frac{1}{8} \left[ \left( 2 \Gamma_1 - 2 \Delta_X^2 \right) - 16 \left( \Gamma_2 - \Delta_X^2 \right) + 16 \left( \gamma_{xy} - 2 \Delta_X^3 \right) p_A \left( 1 - p_A \right) \right]$ |
Table 3.4 (Continued)

<table>
<thead>
<tr>
<th>Elements</th>
<th>Expression in terms of frequencies of gene combinations</th>
<th>Expression in terms of descent measures(^1)</th>
</tr>
</thead>
</table>
| C(2,2)   | \[
\frac{1}{8} \left[ p_B^2 + 7p_B^B - 2 \left( 5p_B^B + 14p_B^B + p_B^B \right) + 32 \left( p_B^B + p_B^B + p_B^B - p_B^B \right) \right]
\] | \[
\frac{1}{8} \left[ 1 + 7F - 2(50 + 14\gamma_{XY} + 8\alpha_{XY} + 32(\gamma_{XY} + 8\gamma_{XY} - 6\gamma_{XY})p_B^2(1-p_B) \\
+ [-2 - 28F + 8(50 + 14\gamma_{XY} + 6\delta_{XY}) - 2(\delta_{XY} + 2\delta_{XY} + 2\gamma_{XY}) \\
+ 32(-4\gamma_{XY} + \alpha_{XY} - 2\alpha_{XY} - 6\gamma_{XY}) \\
- 3\delta_{XY} - 2\delta_{XY} + 6\delta_{XY})p_B^2(1-p_B)^2 \right]
\] |
| D(1,1)   | \[
\frac{1}{2} \left[ p_A^A - p_A^A - 2 \left( p_A^A - 2p_A^A + p_A^A \right) \right]
\] | \[
\frac{1}{2} \left[ 1 - F - 28 + 4\gamma_{XY} - 2\delta_{XY} \right]p_B^2(1-p_B)
\] |
| D(1,2) = D(2,1) | \[
\frac{1}{2} \left[ p_A^A + p_A^A - 2 \left( p_A^A + p_A^A + 2p_A^A - 2 \left[ p_A^A + p_A^A + p_A^A \right] \right) \right]
\] | \[
\frac{1}{2} \left[ 1 - F - 28 + 4\gamma_{XY} - 2\delta_{XY} \right]p_B^2(1-p_B)
\] |
| D(2,2)   | \[
\frac{1}{2} \left[ p_B^B - p_B^B - 2 \left( p_B^B + p_B^B + p_B^B \right) \right]
\] | \[
\frac{1}{2} \left[ 1 - F - 28 + 4\gamma_{XY} - 2\delta_{XY} \right]p_B^2(1-p_B)
\] |
| E(1,1)   | \[
\frac{1}{2} \left[ p_A^A + 3p_A^A - 2 \left[ 3 \left( p_A^A + p_A^A \right) - 4p_A^A \right] \right] \\
- p_A^A \left[ p_A^A + 3p_A^A - 2 \left( 3p_A^A + p_A^A \right) - 4p_A^A \right]
\] | \[
\frac{1}{2} \left[ (\theta + 3\gamma_{XY} - 6\delta_{XY} + 6\delta_{XY} - 8\delta_{XY})p_A^2(1-p_A) \\
+ 6[-6 - 2\gamma_{XY} + 4\gamma_{XY} - (\delta_{XY} + 2\delta_{XY}) \\
+ 6\delta_{XY} + 4\delta_{XY} - 2\delta_{XY} - 8\delta_{XY})p_A^2(1-p_A)^2 \right]
\] |
### Table 3.4 (Continued)

<table>
<thead>
<tr>
<th>Elements</th>
<th>Expression in terms of frequencies of gene combinations</th>
<th>Expression in terms of descent measures $^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>E(1,2)</td>
<td>$\frac{1}{2} \left( p_{AB}^{A} + 2p_{A}^{A}B - p_{A}^{A}B \right) - 2 \left( p_{AB}^{A} + p_{A}^{A}B - 2p_{A}^{A}B \right)$</td>
<td>$\left[ G_2 - \delta_{xy}^{A} - 2(G_3 + \delta_{xy}^{A}) + 4\delta_{xy}^{A} \right] p_A (1-p_A) p_B (1-p_B)$</td>
</tr>
<tr>
<td>E(2,1)</td>
<td>$\frac{1}{2} \left( p_{AB}^{A} + 2p_{A}^{A}B - p_{A}^{A}B \right) - 2 \left( p_{AB}^{A} + p_{A}^{A}B - 2p_{A}^{A}B \right)$</td>
<td>$\left[ G_2 - \delta_{xy}^{A} - 2(G_3 + \delta_{xy}^{A}) + 4\delta_{xy}^{A} \right] p_A (1-p_A) p_B (1-p_B)$</td>
</tr>
<tr>
<td>E(2,2)</td>
<td>$\frac{1}{2} \left( p_{AB}^{A} + 3p_{A}^{A}B - 2 \left( 3(p_{AB}^{A} + p_{A}^{A}B - 4p_{A}^{A}B) \right) \right.$</td>
<td>$\left( G + 2\delta_{xy}^{A} - 6\delta_{xy}^{A} + 6\delta_{xy}^{A} \right) p_B (1-p_B)$</td>
</tr>
<tr>
<td>G(1,1)</td>
<td>$p_{A}^{A} - p_{A}^{A} \left( p_{A}^{A} + p_{A}^{A} \right) - p_{A}^{A} - p_{A}^{A} \left( p_{A}^{A} + p_{A}^{A} \right)$</td>
<td>$\left[ G - \delta_{xy}^{A} - 2(G + \delta_{xy}^{A}) \right] p_A (1-p_A)$</td>
</tr>
<tr>
<td>G(1,2)</td>
<td>$p_{AB}^{A} - 2p_{A}^{A}B - p_{A}^{A}B \left( p_{A}^{A} + p_{A}^{A}B - 2p_{A}^{A}B \right)$</td>
<td>$-2 \left( G_2 - \delta_{xy}^{A} \right) p_A (1-p_A) p_B (1-p_B)$</td>
</tr>
<tr>
<td>G(2,1)</td>
<td>$p_{AB}^{A} - 2p_{A}^{A}B - p_{A}^{A}B \left( p_{A}^{A} + p_{A}^{A}B - 2p_{A}^{A}B \right)$</td>
<td>$-2 \left( G_2 - \delta_{xy}^{A} \right) p_A (1-p_A) p_B (1-p_B)$</td>
</tr>
<tr>
<td>G(2,2)</td>
<td>$p_{B}^{B} - 2(p_{B}^{B}B - p_{B}^{B}B) - p_{B}^{B} \left( p_{B}^{B} - p_{B}^{B} \right)$</td>
<td>$\left[ G - \delta_{xy}^{A} - 2(G + \delta_{xy}^{A}) \right] p_B (1-p_B)$</td>
</tr>
</tbody>
</table>

Note: The expressions are based on the frequencies of gene combinations and the descent measures, which are used to calculate the probabilities in genetic studies.
Table 3.4 (Continued)

<table>
<thead>
<tr>
<th>Elements</th>
<th>Expression in terms of frequencies of gene combinations</th>
<th>Expression in terms of descent measures$^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(1,1)$</td>
<td>$\frac{1}{4} \left[ p_A - p_A^A - 6p_A^A + 4p_A^{AB} + 2p_A^{A</td>
<td>A} + 8 \left( p_A^{A</td>
</tr>
<tr>
<td>$H(1,2)$</td>
<td>$\frac{1}{4} \left[ (p^{AB} + p_{AB}) - 2 (p^{AB} + p_{AB}) - 2 \left[ p^{A</td>
<td>B} + p^{A</td>
</tr>
<tr>
<td>$H(2,1)$</td>
<td>$\frac{1}{4} \left[ (p^{AB} + p_{AB}) - 2 (p^{AB} + p_{AB}) - 2 \left[ p^{A</td>
<td>B} + p^{A</td>
</tr>
<tr>
<td>$H(2,2)$</td>
<td>$\frac{1}{4} \left[ p_B - p_B^B - 6p_B^B + 4p_B^{B</td>
<td>B} + 2p_B^{B</td>
</tr>
</tbody>
</table>

$^+$ No initial linkage disequilibrium has been assumed.
As an aside, it can be noted that an estimator of average heterozygosity for the two loci is

\[ \hat{H} = r_1 + r_2 \].

This estimator has variance,

\[
\text{Var}(\hat{H}) = \text{Var}(r_1) + \text{Var}(r_2) + 2 \text{Cov}(r_1, r_2)
\]

\[ = \frac{1}{a} \left[ M(1,1) + M(2,2) + \frac{1}{n} \left( D(1,1) + D(2,2) \right) \right] 
\]

\[ + \frac{2}{a} \left[ M(1,2) + \frac{1}{n} D(1,2) \right] . \]

This expression for the variance of \( \hat{H} \) can be recovered from the general formula given in Weir, Cockerham and Reynolds (1981).

3.1.2 The Asymptotic Distribution of \( \mathbf{s} \) as \( a \to \infty \)

The vector \( \mathbf{s} \) has expectation \( \mathbf{g}(\gamma) \) and a variance-covariance matrix given by equation (3.1.9). As the distribution of \( X_{ijk} \) does not appear to have a closed form for the simplest of mating structures, namely drift, an exact closed form for the distribution of \( \mathbf{s} \) for any mating system seems to be an unreachable goal. The discovery of an asymptotic distribution for \( \mathbf{s} \) for a fixed number \( a \) of sampled populations but for a large number of sampled individuals within each population \( n \to \infty \) would be a useful contribution as most analyses of population structure involve only a small number of populations but large samples within each population. Unfortunately, only gametes sampled from different populations are assumed to be independent. Gametes sampled from the same population are related. This results in a finite number \( a \) of sequences of \( (n-1) \)-dependent random variables, where \( n \to \infty \), but central limit
Theorems for such random variables do not appear to be available (Hoeffding and Robbins, 1948, page 775).

The only remaining case is that of an asymptotic distribution for \( \bar{s} \) as \( a \to \infty \) when \( n \) is fixed. It is standard to show that as \( a \to \infty \) the vector \( \bar{s} \) has an asymptotic distribution which is multivariate normal with mean vector \( g(\gamma) \) and variance-covariance matrix,

\[
V = \frac{1}{a^2} \Lambda + \frac{1}{an} \Xi + \frac{1}{an(n-1)T}
\]

(3.1.10)

where \( \Lambda, \Xi \) and \( T \) are the matrices defined in section 3.1.1. All the ingredients for the proof of this result are contained in Cramér (1946, Chapter 28) or Serfling (1980, section 2.2). Nevertheless a proof outline is included, for the sake of completeness, in Appendix 9.2.

### 3.1.3 The Estimation of \( \gamma \)

The fitting of \( g(\gamma) \) to \( \bar{s} \) may be accomplished in several ways (Jöreskog, 1978). Two methods are considered here. Ordinary least squares consists in minimizing with respect to \( \gamma \) the square of the Euclidean norm

\[
Q = || \bar{s} - g(\gamma) ||^2 = [\bar{s} - g(\gamma)]^T [\bar{s} - g(\gamma)].
\]

(3.1.11)

Weighted least squares consists in minimizing with respect to \( \gamma \) the norm,

\[
Q_W = [\bar{s} - g(\gamma)]^T W [\bar{s} - g(\gamma)],
\]

(3.1.12)

where \( W \) is a symmetric matrix of weights that is in some sense close to the inverse of the variance-covariance matrix of \( \bar{s} \). If \( W \) does equal the inverse of the variance-covariance matrix of \( \bar{s} \), then the minimization of \( Q_W \) is more correctly described as generalized least squares.
Some extra notation and assumptions are introduced by way of the verification of some regularity conditions:

(i) The elements of $\sigma(\gamma)$ and all partial derivatives of the first three orders with respect to the elements of $\gamma$ are continuous and bounded in a neighborhood of the true value of $\gamma$ designated $\gamma_0$. In fact,

$$
\frac{\partial \sigma(\gamma)}{\partial \gamma} = \begin{pmatrix}
\frac{\partial \sigma_i}{\partial \gamma_j}
\end{pmatrix}_{ij}
= \frac{\partial \sigma}{\partial \gamma}
= \begin{bmatrix}
\theta & 0 & \alpha_1 & 0 \\
0 & \theta & \alpha_2 & 0 \\
\frac{1+\epsilon}{2} - \theta & 0 & -\alpha_1 & \frac{1}{2} \alpha_1 \\
0 & \frac{1+\epsilon}{2} - \theta & -\alpha_2 & \frac{1}{2} \alpha_2 \\
1-\epsilon & 0 & 0 & -\alpha_1 \\
0 & 1-\epsilon & 0 & -\alpha_2
\end{bmatrix},
$$

is equal to a constant ($-1$, $0$, $\frac{1}{2}$ or $1$),

and

$$
\frac{\partial^2 \sigma_i}{\partial \gamma_j \partial \gamma_k} = 0 \quad \text{for all } i \text{ and any } j, k \text{ and } l.
$$

(ii) The matrix $\frac{\partial \gamma}{\partial \gamma^T}$ is of full column rank so that the matrix,

$$
G = \begin{pmatrix}
\frac{\partial \sigma}{\partial \gamma} \\
\frac{\partial \sigma}{\partial \gamma^T}
\end{pmatrix}
= \begin{pmatrix}
\frac{\partial \sigma}{\partial \gamma^T} \\
\frac{\partial \sigma}{\partial \gamma}
\end{pmatrix}
$$
\[
\begin{bmatrix}
\theta^2 + \left(\frac{1+F}{2} - \theta\right)^2 + (1-F)^2 & 0 & (2\theta - \frac{1+F}{2}) \alpha_1 & \left(\frac{3}{4} \theta - \frac{1}{2} \theta - \frac{3}{4}\right) \alpha_1 \\
0 & \theta^2 + \left(\frac{1+F}{2} - \theta\right)^2 + (1-F)^2 & (2\theta - \frac{1+F}{2}) \alpha_2 & \left(\frac{5}{4} \theta - \frac{1}{2} \theta - \frac{3}{4}\right) \alpha_2 \\
\left(\frac{2\theta - \frac{1+F}{2}}{2}\right) \alpha_1 & (2\theta - \frac{1+F}{2}) \alpha_2 & 2(\alpha_1^2 + \alpha_2^2) & -\frac{1}{2} (\alpha_1^2 + \alpha_2^2) \\
\left(\frac{5}{4} \theta - \frac{1}{2} \theta - \frac{3}{4}\right) \alpha_1 & \left(\frac{5}{4} \theta - \frac{1}{2} \theta - \frac{3}{4}\right) \alpha_2 & -\frac{1}{2} (\alpha_1^2 + \alpha_2^2) & \frac{3}{4} (\alpha_1^2 + \alpha_2^2)
\end{bmatrix}
\]

is positive definite and, of course, nonsingular. Also if $W$ is positive definite then

\[\frac{\partial \sigma}{\partial \gamma} W \frac{\partial \sigma}{\partial \gamma}^T\] is positive definite.

(iii) The parameter vector $\gamma$ is uniquely determined by $\sigma(\gamma)$. This can be verified by noting that the equations,

\[
\sigma(\gamma) = \begin{bmatrix}
\sigma_1 \\
\sigma_2 \\
\sigma_3 \\
\sigma_4 \\
\sigma_5 \\
\sigma_6
\end{bmatrix} = \begin{bmatrix}
\theta \alpha_1 \\
\theta \alpha_2 \\
(\frac{1+F}{2} - \theta) \alpha_1 \\
(\frac{1+F}{2} - \theta) \alpha_2 \\
(1-F) \alpha_1 \\
(1-F) \alpha_2
\end{bmatrix}
\]

with the overidentifying conditions $\sigma_1/\sigma_2 = \sigma_3/\sigma_4 = \sigma_5/\sigma_6$ have solutions

\[
\alpha_1 = \sigma_1 + \sigma_3 + \frac{1}{2} \sigma_5 ,
\]

\[
\alpha_2 = \sigma_2 + \sigma_4 + \frac{1}{2} \sigma_6 ,
\]

\[ \theta = \frac{\sigma_1}{\sigma_1 + \sigma_3 + \frac{1}{2} \sigma_5} = \frac{\sigma_2}{\sigma_2 + \sigma_4 + \frac{1}{2} \sigma_6} \]

\[ F = 1 - \frac{\sigma_5}{\sigma_1 + \sigma_3 + \frac{1}{2} \sigma_5} = 1 - \frac{\sigma_6}{\sigma_2 + \sigma_4 + \frac{1}{2} \sigma_6} . \]

Now,

\[ \frac{\partial Q}{\partial Y^T} = -2(s - \sigma(Y))^T \frac{\partial \sigma(Y)}{\partial Y^T} \]

and this equals the null vector if and only if

\[ a_1 = \frac{z_1 \theta + w_1 \left( \frac{1+F}{2} - \theta \right) + r_1 (1-F)}{\theta^2 + \left( \frac{1+F}{2} - \theta \right)^2 + (1-F)^2} \]

(3.1.13)

\[ a_2 = \frac{z_2 \theta + w_2 \left( \frac{1+F}{2} - \theta \right) + r_2 (1-F)}{\theta^2 + \left( \frac{1+F}{2} - \theta \right)^2 + (1-F)^2} \]

(3.1.14)

\[ \theta - \left( \frac{1+F}{2} - \theta \right) = \frac{a_1(z_1-w_1) + a_2(z_2-w_2)}{a_1^2 + a_2^2} \]

(3.1.15)

\[ \left( \frac{1+F}{2} - \theta \right) - 2(1-F) = \frac{a_1(w_1-2r_1) + a_2(w_2-2r_2)}{a_1^2 + a_2^2} \]

(3.1.16)

Substituting equations (3.1.13) and (3.1.14) into equations (3.1.15) and (3.1.16) gives two simultaneous cubic equations in \( \theta \) and \( (1+F)/2 - \theta \):

\[ \xi^T K_1 \xi = 0 \]

\[ \xi^T K_2 \xi = 0 \]

where
\[
\begin{align*}
\zeta^T &= \begin{bmatrix}
\theta^3, & (\frac{1+F}{2} - \theta)^3, & \theta^2(\frac{1+F}{2} - \theta), & \theta(\frac{1+F}{2} - \theta)^2, & \theta^2, & (\frac{1+F}{2} - \theta)^2, \\
\theta(\frac{1+F}{2} - \theta), & \theta, & \frac{1+F}{2} - \theta, & 1
\end{bmatrix}, \\
\zeta^T &= \begin{bmatrix}
z_1^2 + z_2^2, & w_1^2 + w_2^2, & r_1^2 + r_2^2, & z_1w_1 + z_2w_2, & z_1r_1 + z_2r_2, & w_1r_1 + w_2r_2
\end{bmatrix},
\end{align*}
\]

\[
K_1 = \begin{bmatrix}
-4 & 0 & 4 & 5 & 6 & -10 \\
0 & 4 & -4 & -5 & 10 & -6 \\
-9 & 5 & 4 & 5 & 26 & -30 \\
-5 & 9 & -4 & -5 & 30 & -26 \\
8 & 0 & -8 & -8 & -22 & 26 \\
0 & -8 & 8 & 8 & -26 & 22 \\
8 & -8 & 0 & 0 & -52 & 52 \\
-4 & 0 & 4 & 4 & 24 & -24 \\
0 & 4 & -4 & -4 & 24 & -24 \\
0 & 0 & 0 & 0 & -8 & 8
\end{bmatrix},
\]

and

\[
K_2 = \begin{bmatrix}
4 & 0 & -4 & -5 & -6 & 10 \\
0 & 0 & 0 & 0 & 0 & 0 \\
5 & -5 & 0 & 0 & -20 & 20 \\
0 & -4 & 4 & 5 & -10 & 6 \\
-4 & 0 & 4 & 8 & 16 & -26 \\
0 & 4 & -4 & 0 & 0 & -6 \\
0 & 8 & -8 & 0 & 20 & -32 \\
0 & 0 & 0 & -4 & -8 & 24 \\
0 & -4 & 4 & 0 & 0 & 16 \\
0 & 0 & 0 & 0 & 0 & -8
\end{bmatrix}.
\]
These equations may be solved for $\theta$ and $(1+F)/2 - \theta$ by iteration using as starting values

$$\tilde{\theta} = (z_1 + z_2)/[z_1 + z_2 + w_1 + w_2 + \frac{1}{2} (r_1 + r_2)]$$

and

$$\frac{1+F}{2} - \theta = (w_1 + w_2)/[z_1 + z_2 + w_1 + w_2 + \frac{1}{2} (r_1 + r_2)],$$

respectively. These solutions may then be substituted into equations (3.1.13) and (3.1.14) to obtain solutions for $\alpha_1$ and $\alpha_2$. The numerical solutions so obtained give rise to the ordinary least squares estimator of $\gamma$ denoted by $\hat{\gamma}$. Now the matrix,

$$\frac{\partial^2 Q}{\partial \gamma \gamma^T} = 2 \begin{bmatrix}
\theta^2 + \frac{(1+F)}{2} - \theta
& 0 & (\frac{5}{2} F - \theta - \frac{3}{2}) \alpha_1
& (\theta - \frac{1}{2} w_1 - r_1)

\frac{5}{2} F - \theta - \frac{3}{2}
& (\frac{5}{2} F - \theta - \frac{3}{2}) \alpha_2

\alpha_1 - (z_1 - w_1)
& \alpha_2 - (z_2 - w_2)

(\frac{5}{2} F - \theta - \frac{3}{2}) \alpha_1
& (\frac{5}{2} F - \theta - \frac{3}{2}) \alpha_2

\frac{1}{2} (\alpha_1^2 + \alpha_2^2)
& \frac{1}{2} (\alpha_1^2 + \alpha_2^2)

\frac{5}{4} (\alpha_1^2 + \alpha_2^2)
& \frac{5}{4} (\alpha_1^2 + \alpha_2^2)
\end{bmatrix}$$

is required to be positive definite at the stationary point $\hat{\gamma}$ for a minimum to be guaranteed, but because a closed form for $\hat{\gamma}$ cannot be obtained, this condition cannot be checked explicitly here. It can be noted, however, that since $\hat{\gamma}$ converges in probability to $\mathbb{E}(\gamma)$ as $\alpha \to 0$ that $\frac{\partial^2 Q}{\partial \gamma \gamma^T}$ converges in probability to $2G$ and this matrix is
positive definite. Of course in any application the definiteness of 
\( \delta^2 Q/\delta \gamma \delta \gamma^T \) evaluated at the stationary point \( \hat{\gamma} \) can be checked by calculation of the eigenvalues.

If the matrix of weights \( W \) is not a function of \( \gamma \) then the equation,

\[
\frac{\partial Q}{\partial \gamma^T} = -2[\mathbb{I} - \sigma(\gamma)]^T W \frac{\partial \sigma(\gamma)}{\partial \gamma^T} = 0^T
\]

may be solved numerically using the following as starting values for \( a_1, a_2, \theta \) and \( F \), respectively:

\[
z_1 + w_1 + \frac{1}{2}(r_1), \quad z_2 + w_2 + \frac{1}{2}(r_2),
\]

\[
\hat{\theta} = \frac{(z_1 + z_2)}{(z_1 + z_2 + w_1 + w_2 + \frac{1}{2}(r_1 + r_2))},
\]

and

\[
\frac{(z_1 + z_2 + w_1 + w_2 - \frac{1}{2}(r_1 + r_2))}{(z_1 + z_2 + w_1 + w_2 + \frac{1}{2}(r_1 + r_2))}.
\]

The obtained weighted least squares estimator will be denoted by \( \hat{\gamma}_W \).

Following Browne (1974) several asymptotic properties \((a + \infty)\) of
the estimators \( \hat{\gamma} \) and \( \hat{\gamma}_W \) are

(i) \( \hat{\gamma} \) and \( \hat{\gamma}_W \) are consistent estimators of \( \gamma \).

(ii) \( \hat{\gamma} \) has an asymptotic distribution which is multivariate normal
with mean \( \gamma \) and variance-covariance matrix,

\[
G^{-1} \frac{\partial \sigma(\gamma)}{\partial \gamma} V \frac{\partial \sigma(\gamma)}{\partial \gamma^T} G^{-1}.
\]

(iii) \( \hat{\gamma}_W \) where \( W \) is a fixed positive definite matrix has an
asymptotic distribution which is multivariate normal with mean \( \gamma \) and
variance-covariance matrix,
\[ U_W = \left( \frac{\partial \gamma}{\partial \gamma_T} \right)^{-1} \frac{\partial \gamma}{\partial \gamma_T} \frac{\partial \gamma}{\partial \gamma_T} \left( \frac{\partial \gamma}{\partial \gamma_T} \right)^{-1}. \]

(iv) \( \hat{\gamma}_W \), where \( W \) is a consistent estimator of \( V^{-1} \), has an asymptotic distribution which is multivariate normal with mean \( \gamma \) and variance-covariance matrix,

\[ U = \left( \frac{\partial \gamma}{\partial \gamma_T} V^{-1} \frac{\partial \gamma}{\partial \gamma_T} \right)^{-1}. \]

Moreover this estimator is the best weighted least squares estimator in the sense that

\[ U_W - U \]

is positive semidefinite.

(v) If \( W \) is a consistent estimator of \( V^{-1} \) then the asymptotic distribution of

\[ [\bar{s} - \bar{\gamma}(\hat{\gamma}_W)]^T W [\bar{s} - \bar{\gamma}(\hat{\gamma}_W)] \]

is the central chi-square with 2 degrees of freedom.

Proofs of these properties, which rely on the regularity conditions mentioned earlier, mimic those in Browne (1974) and need not be presented here.

3.1.4 Large Sample Variances of Some Estimators of \( \theta \)

Consider the estimator \( \tilde{\theta} = (z_1 + z_2)/(z_1 + z_2 + w_1 + w_2 + \frac{1}{2}(r_1 + r_2)) \), presented in equation (3.1.6), which is also used as a starting value for both ordinary and weighted least squares estimation. Now \( \tilde{\theta} \) is a scalar function of the vector random variable \( \bar{s} \) which is asymptotically normal with mean \( \bar{s}(\gamma) \) and variance-covariance matrix \( V \), so by an application
of the theorem concerning functions of asymptotically normal vectors [see, for example, Serfling (1980, page 122)], \( \hat{\theta} \) is asymptotically normal with mean \( \hat{\theta} \) and variance

\[
(a_1 + a_2)^{-2} \begin{bmatrix}
1-\theta,1-\theta,1-\theta,-\frac{1}{2} \theta,-\frac{1}{2} \theta
\end{bmatrix} V^{-1} \begin{bmatrix}
1-\theta \\
1-\theta \\
-\theta \\
-\theta \\
-\frac{1}{2} \theta \\
-\frac{1}{2} \theta
\end{bmatrix}
\]  

(3.1.17)

Similarly the estimator

\[
\hat{\gamma} = \frac{1}{2} \left( \frac{z_1}{z_1 + w_1 + \frac{1}{2} r_1} + \frac{z_2}{z_2 + w_2 + \frac{1}{2} r_2} \right)
\]

presented in equation (3.1.5) is asymptotically normal with mean \( \theta \) and variance \( \hat{\gamma}^T V \hat{\gamma} \) where,

\[
\hat{\gamma}^T = \frac{1}{2} \left[ \frac{1-\theta}{a_1}, \frac{1-\theta}{a_2}, \frac{-\theta}{a_1}, \frac{-\theta}{a_2}, \frac{-\theta}{2a_1}, \frac{-\theta}{2a_2} \right].
\]

The ordinary least squares estimator \( \hat{\theta} = \hat{\varepsilon}_3^T \hat{\lambda} \) has asymptotic variance,

\[
\hat{\varepsilon}_3^T G^{-1} \frac{\partial \hat{\varepsilon}}{\partial \lambda} \frac{\partial \hat{\varepsilon}}{\partial \lambda} G^{-1} \hat{\varepsilon}_3
\]

while the weighted least squares estimator \( \hat{\theta}_w = \hat{\varepsilon}_3^T \hat{\lambda}_w \), when \( w \) is a consistent estimator of \( V^{-1} \), has asymptotic variance,

\[
\hat{\varepsilon}_3^T \hat{\varepsilon}_3 = \hat{\varepsilon}_3^T \left( \frac{\partial \hat{\varepsilon}}{\partial \lambda} V^{-1} \frac{\partial \hat{\varepsilon}}{\partial \lambda} \right)^{-1} \hat{\varepsilon}_3
\]
Of course $\hat{\theta}_w$, when $W$ is a consistent estimator of $V^{-1}$, has smaller asymptotic variance than any other weighted least squares estimator and $\hat{\theta}$. A direct comparison of $\sum_{j} e_j^T u e_j$ and the asymptotic variance of $\hat{\theta}$ does not appear to be possible as the derivation of a closed form for $V^{-1}$ for any mating system would appear to be intractable.

3.2 Many Loci Each with Two Codominant Alleles

The extension of this approach to many loci each with two codominant alleles does not require any new theory. Suppose there are $m$ loci. For the $k^{th}$ haplotype of the $j^{th}$ individual in the $i^{th}$ population, the basic random vector of indicator variables analogous to equation (3.1.1) is now $m$-dimensional:

$$\begin{pmatrix} x_{i,j,k,1} \\ x_{i,j,k,2} \\ \vdots \\ x_{i,j,k,m} \end{pmatrix} \quad (3.2.1)$$

where

$$x_{i,j,k,l} = \begin{cases} 1 & \text{if haplotype } k \text{ in individual } j \text{ in population } i \text{ carries} \\ & \text{allele } A_{l} \text{ at the } l^{th} \text{ locus,} \\ 0 & \text{otherwise} \end{cases}.$$ 

The multivariate analysis of variance for $x_{i,j,k}$ has the same structure as before (see Table 3.1). The variance-covariance component matrices are now $m \times m$ instead of $2 \times 2$ and have the form:
\[ \Sigma_a = \begin{bmatrix} \alpha_1 & D_{12} & \cdots & D_{1m} \\ D_{12} & \alpha_2 & \cdots & D_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ D_{1m} & D_{2m} & \cdots & \alpha_m \end{bmatrix}, \]

\[ \Sigma_b = \begin{bmatrix} \left( \frac{1+F}{2} - \theta \right) \alpha_1 & D_{b12} & \cdots & D_{b1m} \\ D_{b12} & \left( \frac{1+F}{2} - \theta \right) \alpha_2 & \cdots & D_{b2m} \\ \vdots & \vdots & \ddots & \vdots \\ D_{b1m} & D_{b2m} & \cdots & \left( \frac{1+F}{2} - \theta \right) \alpha_m \end{bmatrix}, \]

\[ \Sigma_c = \begin{bmatrix} (1-F)\alpha_1 & D_{c12} & \cdots & D_{c1m} \\ D_{c12} & (1-F)\alpha_2 & \cdots & D_{c2m} \\ \vdots & \vdots & \ddots & \vdots \\ D_{c1m} & D_{c2m} & \cdots & (1-F)\alpha_m \end{bmatrix}, \]

where

\[ \alpha_\lambda = p_\lambda (1-p_\lambda), \]

\[ D_{a\lambda\lambda'} = \frac{1}{\theta} A_{\lambda\lambda'} A_{\lambda\lambda'}, \]

\[ D_{b\lambda\lambda'} = \frac{1}{2} \left( F^1 + \frac{\bar{F}}{1-F} \right) A_{\lambda\lambda'} A_{\lambda\lambda'}, \]

and

\[ D_{c\lambda\lambda'} = (\bar{F}^1 - \frac{\bar{F}}{1-F}) A_{\lambda\lambda'} A_{\lambda\lambda'}. \]
and where $p_\lambda$ is the frequency of allele $\lambda$ in the initial population and $\Delta_{\lambda\lambda'}$ is the linkage disequilibrium between allele $\lambda$ and allele $\lambda'$ in the initial population.

The analog of equation (3.1.7) is now

$$s^T = \left[ z_1, z_2, \ldots, z_m, w_1, w_2, \ldots, w_m, r_1, r_2, \ldots, r_m \right], \quad (3.2.2)$$

so that

$$E(s^T) = \bar{s}^T(\gamma) = \left[ \theta a_1, \theta a_2, \ldots, \theta a_m, \left( \frac{1+F}{2} - \theta \right) a_1, \ldots, \left( \frac{1+F}{2} - \theta \right) a_m, \right. \left. (1-F)a_1, \ldots, (1-F)a_m \right], \quad (3.2.3)$$

where

$$\bar{\gamma}^T = [a_1, a_2, \ldots, a_m, \theta, F].$$

The estimation problem now consists in fitting the vector $\bar{s}(\gamma)$ which depends on $m+2$ parameters to the observed vector $s$ which is comprised of $3m$ observations. The variance-covariance matrix of $s$ has elements which are obvious analogs of those elements of the variance-covariance matrix given in section 3.1.1 and all other results concerning the asymptotic distribution of $s$, the estimation of $\gamma$ and large sample variances of estimators of $\theta$ are readily adapted from sections 3.1.2, 3.1.3 and 3.1.4, respectively.

Note, however, that since the two-locus descent measures depend on the amount of linkage between the two specified loci, the analog of a term such as $K(1,2)$ in the variance-covariance matrix of $s$ is

$$K(\lambda, \lambda') = [\Delta_{\lambda\lambda'}(\lambda, \lambda') - \left( \frac{1}{2} \right) a_\lambda a_{\lambda'}]$$

where $\lambda_{\lambda'}$ is the linkage parameter for loci $\lambda$ and $\lambda'$. This means that the variance-covariance matrix of $s$ depends on a large number of parameters. For example, if all $m(m-1)/2$ pairs of loci have different
linkage parameters then the $3m \times 3m$ variance-covariance matrix is composed of nonlinear functions of $5m^2 - 4m + 12$ parameters, that is, 12 one-locus descent measures

$m \alpha_\lambda$'s, and

$10m(m-1)/2$ two-locus descent measures.

One way of reducing the number of parameters is to consider the $m$ loci to be a random sample from the genome in the sense that the given $m$ loci, each with a specified $\alpha_\lambda$, have $m(m-1)/2$ linkage parameters which are a random sample of size $m(m-1)/2$ from some distribution with mean $\lambda$ and a range from 0 to 1. Taking expectations over all possible samples of size $m(m-1)/2$ of the linkage parameters might allow the replacement of the $10m(m-1)/2$ two-locus descent measures by just 10 two-locus measures so that, for example,

$$K(\ell, \ell') = [\Delta^2(\lambda) - \pi^2] \alpha_\lambda \alpha_{\lambda'} .$$

Such a device is presented in Weir, Avery and Hill (1980) and is applicable to univariate analyses where loci enter as a classification variable or label in the analysis of variance and one is arguing to change the status of loci from that of a fixed effect to a random effect (see also Cockerham, Weir, Reynolds, 1981). In the present situation, each locus is a measurement variable and is accorded a dimension in the measurement space. Thus the present analysis is localized to a fixed set of loci, and their pairwise linkage parameters, so that in the sequel the dependence of the two-locus descent measures on the amount of linkage is not suppressed in the notation.
3.3 A Special Case—Random Mating

Consider isolated finite monoecious populations, each of size N, which mate completely at random, including self-fertilization. In this case the number of one-locus gene identity measures reduces to four (Cockerham, 1971) in the manner:

$$\theta = F,$$
$$\gamma = \gamma_{XY} = \gamma_{XYZ},$$
$$\Delta = \Delta_{X \cdot Y} = \Delta_{X+Y} = \Delta_{X \cdot YZ} = \Delta_{X+YZ} = \Delta_{XY \cdot ZW},$$
$$\delta = \delta_{XY} = \delta_{XYZ} = \delta_{XYZW}.$$

The two-locus nonidentity measures reduce to three for each pair of loci (Weir, Avery and Hill, 1980) in the manner:

$$\theta = \theta_1 = \theta_2,$$
$$\gamma = \Gamma_1 = \Gamma_2 = \Gamma_3,$$
$$\Delta^* = \Delta^*_1 = \Delta^*_2 = \Delta^*_3 = \Delta^*_4 = \Delta^*_5.$$

These three two-locus measures will hereafter be subscripted in the manner $\theta_{\ell \ell'} \Gamma_{\ell \ell'}$ and $\Delta^*_{\ell \ell'}$, to indicate the pair of loci being referred to.

Suppose $m$ loci each with two codominant alleles are scored so that the expected value of the observed vector of variance components is

$$E(z) = E \begin{bmatrix} z_1 \\ \vdots \\ z_m \\ w_1 \\ \vdots \\ w_m \\ r_1 \\ \vdots \\ r_m \end{bmatrix} = E \begin{bmatrix} z \\ \vdots \\ w \\ r \end{bmatrix}.$$
\[ \sigma(\gamma) = \begin{bmatrix}
2\alpha_1 \\
\vdots \\
2\alpha_m \\
\frac{1}{2} (1-\theta)\alpha_1 \\
\vdots \\
\frac{1}{2} (1-\theta)\alpha_m \\
(1-\theta)\alpha_1 \\
\vdots \\
(1-\theta)\alpha_m
\end{bmatrix} = \begin{bmatrix}
\theta\alpha \\
\frac{1}{2} (1-\theta)\alpha \\
(1-\theta)\alpha
\end{bmatrix} \]

where \( \gamma^T = [\alpha_1, \ldots, \alpha_m, \theta]^T = [\alpha^T, \theta] \). The component matrices of the variance-covariance matrix of \( \mathbf{s} \) given in equation (3.1.9) have the form

\[ \Phi = \begin{bmatrix}
2\theta^2 [\text{diag}(\alpha^T \alpha)] & m_{2m} \\
2m_{m} & 2m_{2m}
\end{bmatrix} \]

\[ \Psi = \begin{bmatrix}
2\theta (1-\theta) [\text{diag}(\alpha^T \alpha)] & m_{2m} \\
2m_{m} & 2m_{2m}
\end{bmatrix} \]

\[ \Omega = \begin{bmatrix}
\frac{1}{2} (1-\theta)^2 [\text{diag}(\alpha^T \alpha)] & m_{2m} \\
2m_{m} & 2m_{2m}
\end{bmatrix} \]

\[ \Sigma = \begin{bmatrix}
\frac{1}{2} A_m & -A_m \\
-A_m & A_m \\
2m_{m} & m_{2m} \\
0 & m_{m}
\end{bmatrix} \]
where
\[ A_{m m} = (\theta - 2\gamma + \delta) \text{diag}(a) + (1 - 6\theta + 8\gamma + 3\Delta - 6\delta) \text{diag}(a^* a) \]
\[ + \left[ \Theta_{L^* - 2\Gamma_{L^*} + \Delta_{L^*} L^*} \right] a a^T - \text{diag}(a^* a) \]

and where the operation * denotes the Hadamard product of two matrices, \( \text{diag}(a) \) denotes the diagonal matrix with diagonal elements equal to those of the vector \( a \), and \( [\cdot]_{L^*} \) denotes an \( m \times m \) matrix with \( L, L^* \) element given by the term enclosed in the brackets.

The matrices \( \Gamma \) and \( \Xi \) may be partitioned into \( m \times m \) submatrices in the manner

\[ \Lambda = \begin{bmatrix} K & \frac{1}{2} R & R \\ \frac{1}{2} R & \frac{1}{4} M & \frac{1}{2} M \\ R & \frac{1}{2} M & M \end{bmatrix} \]

\[ \Xi = \begin{bmatrix} B & \frac{1}{2} G & G \\ \frac{1}{2} G & \frac{1}{4} D & H \\ G & H & D \end{bmatrix} \]

The submatrices are presented in Table 3.5.

As before, \( z \) has an asymptotic distribution \( (a \to \infty) \) which is multivariate normal with mean \( \mu(y) \) and variance-covariance matrix

\[ V = \frac{1}{a} A + \frac{1}{an} + \frac{1}{an(n-1)} T \]
Table 3.5 The submatrices of $\Lambda$ and $\Xi$ for the special case of random mating.

<table>
<thead>
<tr>
<th>Submatrix</th>
<th>$\text{diag}(q)$</th>
<th>Coefficients of $\text{diag}(q^\alpha)$</th>
<th>$\Theta_T^\top - \text{diag}(q^\alpha)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$</td>
<td>$\delta$</td>
<td>$(3\delta-6\delta-\delta^2)$</td>
<td>$\delta^*_{\lambda\lambda'} - (1-\delta)^2$</td>
</tr>
<tr>
<td>$M$</td>
<td>$\theta-2\gamma+\delta$</td>
<td>$(-4\theta+3\delta-6\delta-\delta^2)$</td>
<td>$\delta^*_{\lambda\lambda'} - (1-\delta)^2$</td>
</tr>
<tr>
<td>$R$</td>
<td>$\gamma - \delta$</td>
<td>$(-4\gamma - 3\delta + 6\delta + \delta^2)$</td>
<td>$- [\delta^*_{\lambda\lambda'} - (1-\delta)^2]$</td>
</tr>
<tr>
<td>$B$</td>
<td>$2(\gamma - \delta)$</td>
<td>$2(\theta - 4\gamma - 3\delta + 6\delta)$</td>
<td>$2(\Gamma_{\lambda\lambda'} - \delta^*_{\lambda\lambda'})$</td>
</tr>
<tr>
<td>$D$</td>
<td>$\frac{1}{2}(1-3\theta+4\gamma-2\delta)$</td>
<td>$(-1+6\theta-8\gamma-3\delta + 6\delta)$</td>
<td>$\Theta_{\lambda\lambda'} - \delta^*_{\lambda\lambda'}$</td>
</tr>
<tr>
<td>$G$</td>
<td>$\theta-3\gamma+2\delta$</td>
<td>$2(-3\theta+6\gamma+3\delta-6\delta)$</td>
<td>$-2(\Gamma_{\lambda\lambda'} - \delta^*_{\lambda\lambda'})$</td>
</tr>
<tr>
<td>$H$</td>
<td>$\frac{1}{4}(1-7\theta+12\gamma-6\delta)$</td>
<td>$\frac{3}{2}(-1+6\theta-8\gamma-3\delta+6\delta)$</td>
<td>$- \frac{1}{2}(\Theta_{\lambda\lambda'} - 4\Gamma_{\lambda\lambda'} + 3\delta^*_{\lambda\lambda'})$</td>
</tr>
</tbody>
</table>

1 The coefficients in this column are the $\lambda, \lambda'$ elements of a coefficient matrix whose Hadamard product with $\Theta_T^\top - \text{diag}(q^\alpha)$ is the component of a submatrix.

For example,

\[ K = \delta \text{diag}(q) + (3\delta-6\delta-\delta^2) \text{diag}(q^\alpha) + [\delta^*_{\lambda\lambda'} - (1-\delta)^2]_{\lambda\lambda'} * [\Theta_T^\top - \text{diag}(q^\alpha)] \]

so that the $\lambda, \lambda'$ element of $K$ for $\lambda \neq \lambda'$ is simply

\[ [\delta^*_{\lambda\lambda'} - (1-\delta)^2]_{\lambda\lambda'} \]

Now

\[
\frac{\partial \Theta_T}{\partial \gamma} = \begin{bmatrix}
\theta & 0 & 0 \\
\frac{1}{2}(1-\theta) & 1 & -\frac{1}{2}
\end{bmatrix}
\begin{bmatrix}
\alpha & -\alpha
\end{bmatrix}
\]
\[
\frac{\partial \sigma}{\partial y} \frac{\partial \sigma}{\partial \gamma}^T = \begin{bmatrix}
\left[ \theta^2 + \frac{5}{4} (1-\theta)^2 \right] m_m & \left( \frac{9}{4} \theta - \frac{5}{4} \right) \alpha \\
\left( \frac{9}{4} \theta - \frac{5}{4} \right) \alpha^T & \frac{9}{4} \alpha^T \alpha
\end{bmatrix}
\]

So that
\[
\left( \frac{\partial \sigma}{\partial y} \frac{\partial \sigma}{\partial \gamma}^T \right)^{-1} = \frac{1}{9\theta^2 - 10\theta + 5} \begin{bmatrix}
4 I_m + \left( \frac{9}{4} \theta - \frac{5}{4} \right) \alpha \alpha^T & -\left( \frac{9}{4} \theta - \frac{5}{4} \right) \alpha \\
\frac{9}{4} \alpha^T \alpha & \frac{9}{4} \alpha^T \alpha
\end{bmatrix}
\]

Ordinary least squares estimators \( \hat{\alpha} \) and \( \hat{\beta} \) must satisfy the two equations:

\[
\left( z - \frac{1}{2} \omega - r \right)^T \alpha = \left( \frac{9}{4} \theta - \frac{5}{4} \right) \alpha^T \alpha \]  
(3.3.1)

\[
\alpha = \frac{4}{9\theta^2 - 10\theta + 5} \left[ \frac{1}{2} \omega + r + \theta (z - \frac{1}{2} \omega - r) \right]. \]  
(3.3.2)

Substituting equation (3.3.2) into equation (3.3.1) gives a quadratic equation in \( \theta \):

\[
\begin{bmatrix}
5z + 4 \left( \frac{1}{2} \omega + r \right)
\end{bmatrix}^T \left( -z + \frac{1}{2} \omega + r \right) \theta^2 + \begin{bmatrix}
5z^T z - 10z^T \frac{1}{2} \omega + r \\
- 4 \left( \frac{1}{2} \omega + r \right) \frac{1}{2} \omega + r
\end{bmatrix} \theta + 5z^T \frac{1}{2} \omega + r = 0.
\]

So that \( \hat{\theta} \) is given by

\[
\hat{\theta} = \frac{10b - (5a-4c) \pm \sqrt{80b^2 + (5a-4c)^2}}{2b - 2(5a-4c)}, \]  
(3.3.3)

where \( a = z^T z \), \( b = z^T \left( \frac{1}{2} \omega + r \right) \) and \( c = \left( \frac{1}{2} \omega + r \right) \left( \frac{1}{2} \omega + r \right) \). To check which solution in equation (3.3.3) provides the absolute minimum, the definiteness of
\[
\frac{\partial^2 Q}{\partial \gamma \partial \gamma^T} = \begin{bmatrix}
\frac{1}{2} (9\theta^2 - 10\theta + 5) m \alpha m \\
(9\theta-5) \alpha^T - 2(z - \frac{1}{2} \bar{w} - \bar{r}) \\
(9\theta-5) \bar{z}^T - 2(z - \frac{1}{2} \bar{w} - \bar{r})^T \\
\frac{9}{2} \alpha^T \alpha
\end{bmatrix}
\]

should be checked and the residual sum of squares, \([s - g(\hat{y})]^T [s - g(\hat{y})]\)
should be calculated for each solution.

For a symmetric matrix of weights \(W\) that does not depend on \(\gamma\) the
weighted least squares estimator \(\hat{\gamma}_W\) is of course a solution to the
equation,

\[
\frac{\partial^2 Q_W}{\partial \gamma} = -2 [s - g(\gamma)]^T W \frac{\partial g(\gamma)}{\partial \gamma^T} = 0^T
\]

for which the Hessian \(\frac{\partial^2 Q_W}{\partial \gamma \partial \gamma^T}\) is positive definite. Partitioning \(W\)
into \(m \times m\) submatrices in the manner

\[
W = \begin{bmatrix}
W_1 & W_4 & W_6 \\
W_4^T & W_2 & W_5 \\
W_6^T & W_5^T & W_3
\end{bmatrix}
\]

the Hessian can be written in the form

\[
\frac{\partial^2 Q_W}{\partial \gamma \partial \gamma^T} = \begin{bmatrix}
m^c m & m^{-1} \\
\frac{d}{d^T} m^{-1} & 1^f 1
\end{bmatrix}
\]

where
\[ C = 2 \left\{ \theta^2 W_1 + \theta (1-\theta) \left[ \frac{1}{2} (W_4 + W_4^T) + W_6 + W_6^T \right] \right. \\
+ \left(1-\theta\right)^2 \left[ \frac{1}{4} W_2 + W_3 + \frac{1}{2} (W_5 + W_5^T) \right] \right\}, \]
\[ d_T = -2 \left\{ z^T (W_1 - \frac{1}{2} W_4^T W_6) + \frac{1}{2} W_4^T (W_4 - \frac{1}{2} W_2^T W_5) + \frac{1}{2} (W_6 - \frac{1}{2} W_5^T W_3) \right\} \]
\[ \left. - a^T \left[ 2 \theta W_1 + (1-\theta) (-\frac{1}{2} W_2 - 2 W_3 - W_5 - W_5^T) \right] + (1-2\theta) \left( \frac{1}{2} (W_4 + W_4^T) + W_6 + W_6^T \right) \right\}, \]
and
\[ f = 2 a^T \left[ W_1 + \frac{1}{4} W_2 + W_3 - \frac{1}{2} (W_4 + W_4^T) + \frac{1}{2} (W_5 + W_5^T) - (W_6 + W_6^T) \right] a . \]

Utilizing this expression for the Hessian a weighted least squares estimator can be computed by the Newton-Raphson method. Possible starting values for this iterative procedure include

\[ \mathbf{v} = z + w + \frac{1}{2} \mathbf{1} , \]
and
\[ \mathbf{y} = \frac{1}{m} \sum_{z=1}^{m} \left( \frac{z_z}{z_z + w_z + \frac{1}{2} \mathbf{1}} \right) , \]
or
\[ \tilde{\theta} = \frac{\mathbf{1}^T \mathbf{z}}{\mathbf{1}^T (z+w + \frac{1}{2} \mathbf{1})} \]

The asymptotic variance of \( \tilde{\theta} \) as \( n \to \infty \) is

\[ \varepsilon_{m+1}^T \left( \frac{\partial \sigma}{\partial y} \frac{\partial \sigma}{\partial y^T} \right)^{-1} \frac{\partial \sigma}{\partial y} \varepsilon \frac{\partial \sigma}{\partial y^T} \left( \frac{\partial \sigma}{\partial y} \frac{\partial \sigma}{\partial y^T} \right)^{-1} \varepsilon_{m+1} \]

\[ = (\mathbf{v}^T \mathbf{v})^{-2} \left( (1-\theta) \mathbf{v}^T - \frac{2}{5} \theta \mathbf{v}^T - \frac{4}{5} \theta \mathbf{v}^T \right) \varepsilon_{m+1} \]

\[ = \begin{bmatrix} (1-\theta) \mathbf{v}^T \\
- \frac{2}{5} \theta \mathbf{v}^T \\
- \frac{4}{5} \theta \mathbf{v}^T \end{bmatrix} \]
\[
= (\alpha^T \alpha)^{-2} \alpha^T \left\{ \frac{1}{a} \left[ (1-\theta)^2 K - 2\theta(1-\theta) R + \Theta^2 M \right] + \frac{1}{an} \left[ (1-\theta)^2 B - 2\theta(1-\theta) G + \frac{\theta^2}{25} (17D+16H) \right] \right. \\
+ \left. \frac{1}{2an(n-1)} \left( 1 - \frac{3}{5} \theta \right)^2 A \right\} \alpha
\]

\[
= \left( \sum_{t=1}^{m} \alpha_t^2 \right)^{-2} \left\{ \frac{1}{a} \left[ (\delta-2\Theta_\gamma+\Theta^3) \sum_{t=1}^{m} \alpha_t^3 + (3\Delta-6\delta+8\Theta_\gamma-\theta^2-4\Theta^3) \sum_{t=1}^{m} \alpha_t^4 \right. \\
+ \sum_{t \neq t'} \left( \Delta_{tt'}^* - (1-\theta)^2 \right) \alpha_t^2 \alpha_{t'}^2 \left. \right\} \\
+ \frac{1}{an} \left[ (2(\gamma-\delta) + 2\Theta_\gamma - 3 \frac{\theta^2}{2} + \Theta^2 \frac{(-7\Theta-36\gamma+18\delta)}{50} \right] \sum_{t=1}^{m} \alpha_t^3 \\
+ \left. \left( 2(\Theta-4\gamma-3\Delta+6\delta) - 8\Theta_\gamma + 6\Theta^2 \right. \right. \\
+ \left. \Theta^2 \left[ 9-4\Theta+9(8\gamma+3\Delta-6\delta) \right] \right] \sum_{t=1}^{m} \alpha_t^4 \\
+ \sum_{t \neq t'} \left( \begin{array}{c} 2 \left( \Gamma_{tt'} - \Delta_{tt'}^* \right) \\
+ 9\Theta^2 \left( \frac{\Theta_{tt'} - 2\Gamma_{tt'} + \Delta_{tt'}^*}{25} \right) \end{array} \right) \alpha_t^2 \alpha_{t'}^2 \left. \right\} \\
+ \frac{1}{2an(n-1)} \left[ (\Theta-2\gamma+6) \sum_{t=1}^{m} \alpha_t^3 + (1-6\Theta+8\Theta_\gamma+3\Delta-6\delta) \sum_{t=1}^{m} \alpha_t^4 \right. \\
+ \left. \sum_{t \neq t'} \left( \Theta_{tt'} - 2\Gamma_{tt'} + \Delta_{tt'}^* \right) \alpha_t^2 \alpha_{t'}^2 \right] \right\} \\
(3.3.4)
\]

For large sized samples within each population, that is \( n \) about the same order of magnitude as \( a \), equation (3.3.4) becomes,

\[
a^{-1} \left( \sum \alpha_t^2 \right)^{-2} \left\{ (\delta-2\Theta_\gamma+\Theta^3) \sum \alpha_t^3 + (3\Delta-6\delta+8\Theta_\gamma-\theta^2-4\Theta^3) \sum \alpha_t^4 \right. \\
+ \sum_{t \neq t'} \left( \Delta_{tt'}^* - (1-\theta)^2 \right) \alpha_t^2 \alpha_{t'}^2 \left. \right\} \\
(3.3.5)
\]
and this expression is equal to

\[
\frac{1}{\alpha} \left\{ \frac{(\delta - 2\theta + \theta^2)}{m^\alpha} + \frac{(3\Delta - 6\delta + 8\theta - \theta^2 - 4\theta^3)}{m} + \sum_{l \neq l'} \frac{[\Delta_{l,l'}^* - (1-\theta)^2]}{m^2} \right\}
\]

when all the \( \alpha \)'s are equal \( (\alpha = \bar{\alpha}, \forall \alpha) \). For large population sizes, \( N \), the contrasts in one-locus measures may be replaced by polynomials in \( \phi = e^{-t/2N} \), where \( t \) indexes the number of generations, in the manner

\[
\delta - 2\theta = \frac{1}{5} \phi - \phi^3 + \phi^4 - \frac{1}{5} \phi^6
\]

\[
3\Delta - 6\delta + 8\theta - \theta^2 - 4\theta^3 = -\phi^2 + 4\phi^3 - 4\phi^4 + \phi^6.
\]

The arguments which permit the replacement of one-locus measures by polynomials in \( \phi \) can be found in Chevalet, Gillois and Nassar (1977) and Cockerham, Weir and Reynolds (1981).

For the weighted least squares estimator \( \hat{x}_W \), with \( W \) a consistent estimator of \( V^{-1} \), the asymptotic variance of \( \hat{x}_W \) as \( a \to \infty \) is of course

\[
e_{m+1} \left( \frac{\partial^2}{\partial y^2} V^{-1} \frac{\partial^2}{\partial y^T} T \right)^{-1} \text{ e}_{m+1} \quad \text{(3.3.6)}
\]

but again a closed form for this variance in terms of descent measures and gene frequencies would appear to be intractable. However this asymptotic variance is certainly smaller than the asymptotic variance of \( \hat{\theta} \).

An expression for the asymptotic variance of \( \hat{\theta} \) using the formula in Kendall and Stuart (1977, page 247) for the approximate variance of a ratio of random variables, or the equivalent expression in equation (3.1.18), is
\[
\text{Var}(\tilde{\theta}) = (1^T \Sigma a)^{-2} \left\{ \text{Var}(\frac{1}{1^T Z}) - 2 \theta \text{Cov}\left[\frac{1}{1^T Z}, \frac{1}{1^T (Z+\tilde{w}+\frac{1}{2^T Z})}\right] + \theta^2 \text{Var}\left[\frac{1}{1^T (Z+\tilde{w}+\frac{1}{2^T Z})}\right] \right\} + o(a^{-1})
\]
\[
= (1^T \Sigma a)^{-2} \frac{1}{1^T a} \left\{ (1-\theta)^2 K - 2\theta(1-\theta) R + \theta^2 M \right\} \\
+ \frac{1}{an} \left\{ (1-\theta)^2 B - 2\theta(1-\theta) G + \theta^2 (H+\frac{1}{2^T D}) \right\} \\
+ \frac{1}{2 an (n-1)^A} \frac{1}{1^T a} + o(a^{-1}) \\
= \left( \sum_{\ell=1}^{m} \alpha_\ell \right)^{-2} \left\{ \frac{1}{a} \left[ (1-2\theta)(1+3^T \theta) \right] \sum_{\ell=1}^{m} \alpha_\ell \\
+ (3\Delta - 6\delta + 8\theta \gamma - \theta^2 - 4\theta^3) \sum_{\ell=1}^{m} \alpha_\ell^2 + \sum_{\ell \neq \ell'} \left[ \Delta_{\ell \ell'}^* - (1-\theta)^2 \right] \alpha_\ell \alpha_{\ell'} \right\} \\
+ \frac{1}{an} \left\{ (2(\gamma - \delta) + 2\theta \gamma - \frac{3\theta^2}{2} - \frac{3\theta^3}{2}) \sum_{\ell=1}^{m} \alpha_\ell \\
+ [2(\theta - 4\theta + 3\Delta - 6\delta) - 8\theta \gamma + 6\theta^2 + 2\theta^3] \sum_{\ell=1}^{m} \alpha_\ell^2 \\
+ 2 \sum_{\ell \neq \ell'} (\Gamma_{\ell \ell'} - \Delta_{\ell \ell'}^*) \alpha_\ell \alpha_{\ell'} \right\} \\
+ \frac{1}{2an(n-1)} \left\{ (\theta - 2\gamma + \delta) \sum_{\ell=1}^{m} \alpha_\ell \\
+ (1-6\delta + 8\theta \gamma - 3\Delta - 6\delta) \sum_{\ell=1}^{m} \alpha_\ell^2 \right\} \\
+ \sum_{\ell \neq \ell'} \left( \theta_{\ell \ell'} - \frac{2\Gamma_{\ell \ell'} - \Delta_{\ell \ell'}^*}{\Delta_{\ell \ell'}^* - (1-\theta)^2} \right) \right\} + o(a^{-1}) \
(3.3.7)
\]

For large sized samples within each population, that is \(n\) about the same order of magnitude as \(a\), the variance of \(\tilde{\theta}\) is approximately,
\[
\frac{1}{a^{-1} \left( \sum_{\ell=1}^{m} \alpha_\ell \right)^{-2}} \left\{ (\delta - 2\theta \gamma + \theta^3) \sum_{\ell=1}^{m} \alpha_\ell \\
+ (3\Delta - 6\delta + 8\theta \gamma - \theta^2 - 4\theta^3) \sum_{\ell=1}^{m} \alpha_\ell^2 \\
+ \sum_{\ell \neq \ell'} \left[ \Delta_{\ell \ell'}^* - (1-\theta)^2 \right] \alpha_\ell \alpha_{\ell'} \right\} 
\]
\[
(3.3.8)
\]

and this expression is equal to
\[
\frac{1}{2} \left\{ \frac{(\delta - 2\theta \gamma + \theta^3)}{m^2} \sum_{\ell=1}^{m} \alpha_\ell^2 + \frac{(3\Delta - 6\delta + 8\theta \gamma - \theta^2 - 4\theta^3)}{m} \sum_{\ell \neq \ell'} \frac{\Delta_{\ell \ell'}^* - (1-\theta)^2}{m^2} \right\} 
\]
when $\alpha_l = \bar{\alpha}$ for all $l$.

For equal initial gene frequency functions $\alpha_l$ at each locus and large sized samples within each population the asymptotic variances of $\hat{\sigma}$ and $\tilde{\sigma}$ are clearly the same, but for different $\alpha_l$ the relationship between expressions (3.3.5) and (3.3.8) is unclear. Certainly, for $m$ loci,

$$\frac{\sum_l \alpha_l^3}{(\sum_l \alpha_l^2)^2} > \frac{\sum_l \alpha_l}{(\sum_l \alpha_l)^2}$$

and

$$\frac{\sum_l \alpha_l^4}{(\sum_l \alpha_l^2)^2} > \frac{\sum_l \alpha_l^2}{(\sum_l \alpha_l)^2}$$

but for two loci

$$\frac{\alpha_1^2 \alpha_2^2}{(\alpha_1^2 + \alpha_2^2)^2} < \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_2)^2}$$

For more than two loci, it is possible for

$$\sum_{i \neq l'} \left[ \sum_{i, l'} \left( \Delta_{i,l'}^* - (1-\theta)^2 \right) \alpha_i^2 \alpha_{l'}^2 \right]$$

$$\left( \sum_l \alpha_l^2 \right)^2$$

to be either less than or greater than

$$\sum_{i \neq l', l''} \left[ \sum_{i, l', l''} \left( \Delta_{i,l', l''}^* - (1-\theta)^2 \right) \alpha_i \alpha_{l'} \alpha_{l''} \right]$$

$$\left( \sum_l \alpha_l \right)^2$$
However, it is the first two terms (the terms with coefficients of \( \delta - 2\theta \gamma + \theta^3 \) and \( 3\Delta - 6\delta + 8\theta \gamma - 6\theta^2 - 4\theta^3 \)) in expressions (3.3.5) and (3.3.8) that are likely to dominate these asymptotic variances, in which case the asymptotic variance of \( \hat{\gamma} \) will be greater than the asymptotic variance of \( \tilde{\gamma} \). That this is so, should not be too surprising as the ordinary least squares estimator ignores the variance-covariance structure of \( \gamma \) in the fitting of \( s(\gamma) \) to \( s \). Whether the biases and variances of these estimators differ substantially in small samples (i.e., small numbers of populations sampled) is another matter.

An approximate expression for the variance of \( \hat{\delta} \) is given by

\[
\text{Var}(\hat{\delta}) = \frac{1}{m} \left[ \frac{1}{\alpha} \left( \frac{(\delta - 2\theta \gamma + \theta^3)}{m} \sum_{l} \left( \frac{1}{\alpha_l} \right) + (3\Delta - 6\delta + 8\theta \gamma - 6\theta^2 - 4\theta^3) \right) \right]

+ \sum_{l \neq l'} \left[ \frac{\Delta^*_{ll'} - (1 - \theta)^2}{m} \right] + \frac{1}{an} \left[ \frac{(2(\gamma - \delta) + 2\theta \gamma - \frac{3\theta^2}{2} - \frac{\theta^3}{2})}{m} \sum_{l} \left( \frac{1}{\alpha_l} \right) \right]

+ \left( 2(\theta - 4\gamma + 3\Delta + 6\delta) - 8\theta \gamma + 6\theta^2 + 2\theta^3 \right) + \sum_{l \neq l'} \left( \frac{\Gamma_{ll'} - \Delta^*_{ll'}}{m} \right)

+ \frac{1}{2an(n-1)} \left[ \frac{(\theta - 2\gamma + \delta)}{m} \sum_{l} \left( \frac{1}{\alpha_l} \right) + (1 - 6\theta + 8\gamma + 3\Delta - 6\delta) \right]

+ \sum_{l \neq l'} \left( \frac{\Theta_{ll'} - 2\Gamma_{ll'} + \Delta^*_{ll'}}{m} \right) \right] + o(a^{-1}) \quad (3.3.9)

If \( n \) is of the same order of magnitude as \( a \), the variance of \( \hat{\delta} \) is approximately,

\[
a^{-1} \left\{ (\delta - 2\theta \gamma + \theta^3)m^{-2} \sum_{l} \left( \alpha^{-1}_l \right) + (3\Delta - 6\delta - 8\theta \gamma - 6\theta^2 - 4\theta^3)m^{-1} \right.

+ \left. m^{-2} \sum_{l \neq l'} \left[ \Delta^*_{ll'} - (1 - \theta)^2 \right] \right\} \quad (3.3.10)
\]
Comparing expression (3.3.10) with the corresponding expression (3.3.8) for $\tilde{\theta}$ it can be seen that

$$m^{-2} \sum_{\ell} \left( \alpha_{\ell}^{-1} \right) > \frac{\sum \alpha_{\ell}}{\left( \sum \alpha_{\ell} \right)^2},$$

but that,

$$\frac{1}{m} < \frac{\sum \alpha_{\ell}^2}{\left( \sum \alpha_{\ell} \right)^2}$$

while the relationship between $m^{-2}$ and $\alpha_{\ell}^{-1} \left( \sum \alpha_{\ell} \right)^{-2}$ can only be determined when $\alpha_{\ell}$, $\sum \alpha_{\ell}$ and $\sum \alpha_{\ell}^2$ are known. However each of the expressions (3.3.8) and (3.3.10) are likely to be dominated by their first terms, as time increases, in which case the asymptotic variance of $\tilde{\theta}$ will be smaller than the asymptotic variance of $\hat{\theta}$. Asymptotic variances of $\hat{\theta}$, $\bar{\theta}$, $\hat{\theta}$ and $\hat{\theta}_w$ are compared for the special cases of a single locus and two loci in the next two sections.

As an aside, an interesting special case alluded to earlier is that of large population size, $N$, and long time, $t$. In this case the one-locus measures may be replaced by polynomials in $\phi = e^{-t/2N}$. Moreover since the contrasts in the two-locus nonidentity measures which appear in the variance-covariance matrix of $\mathbf{s}$ are likely to be close to zero for all values of the linkage parameter $\lambda$ (examples are given in Table 3.6), the variance-covariance matrix of $\mathbf{s}$ may be approximated by a matrix which is a function of the parameter vector,

$$\gamma^* = \begin{bmatrix} \alpha \\ \phi \end{bmatrix}.$$
Table 3.6 Values of contrasts in two locus nonidentity measures for a monoecious population size $N = 1000$, with linkage parameter $\lambda$, at various times $(t)$.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>$t$</th>
<th>$\Delta^*-(1-\theta)^2$</th>
<th>$\Gamma-\Delta^*$</th>
<th>$\theta-2\Gamma+\Delta^*$</th>
<th>$\theta-4\Gamma+3\Delta^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
<td>$&lt;10^{-5}$</td>
<td>$&lt;10^{-5}$</td>
<td>.00033</td>
<td>.00033</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>$&lt;10^{-5}$</td>
<td>$&lt;10^{-5}$</td>
<td>.00032</td>
<td>.00032</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>$&lt;10^{-5}$</td>
<td>$&lt;10^{-5}$</td>
<td>.00030</td>
<td>.00030</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>$&lt;10^{-5}$</td>
<td>$&lt;10^{-5}$</td>
<td>.00020</td>
<td>.00020</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>$&lt;10^{-5}$</td>
<td>$&lt;10^{-5}$</td>
<td>.00012</td>
<td>.00012</td>
</tr>
<tr>
<td>0.9</td>
<td>10</td>
<td>$&lt;10^{-5}$</td>
<td>$&lt;10^{-5}$</td>
<td>.0029</td>
<td>.0029</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>$&lt;10^{-5}$</td>
<td>.00003</td>
<td>.0044</td>
<td>.0043</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>$&lt;10^{-5}$</td>
<td>.00004</td>
<td>.0042</td>
<td>.0041</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>.00003</td>
<td>.00003</td>
<td>.0028</td>
<td>.0028</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>.00003</td>
<td>.00002</td>
<td>.0017</td>
<td>.0017</td>
</tr>
</tbody>
</table>

This means that the estimation problem reduces to that of estimating $\phi$ and consequently $\theta = 1-\phi$ by fitting $s$ to $c(\gamma^*)$ where $s$ has an approximate asymptotic variance-covariance matrix given by $V(\gamma^*)$, the variance-covariance matrix $V$ with all one-locus descent measures replaced by polynomials in $\phi$ and all off-diagonal elements in the component submatrices of $V$, given in Table 3.5, set equal to zero. A consistent estimator of $V(\gamma^*)$ is easily obtained by replacing $\gamma^*$ by, say,

$$
\begin{bmatrix}
\gamma_x \\
1-\theta
\end{bmatrix}
$$

in $V(\gamma^*)$ and an iterative weighted least squares procedure based on equation (3.1.12) but with reestimation of $V(\gamma^*)$ at each step may be used to estimate $\gamma^*$. 
3.3.1 One Locus

For the case of one locus,

\[ \mathbf{\bar{z}}^T = [z, w, r] , \]

and

\[ \mathbf{\bar{z}}^T(y) = [\theta \alpha, \frac{1}{2}(1-\theta)\alpha, (1-\theta)\alpha] \]

where

\[ \mathbf{\bar{y}}^T = [\alpha, \theta] . \]

Estimators of \( \theta \) include

(i) \( \hat{\theta} = \bar{y} = z/(z+w+\frac{1}{2}r) \), presented in Cockerham (1969),

(ii) \( \hat{\theta} \) given by equation (3.3.3) with \( a = z^2 \), \( b = z(\frac{1}{2}w+r) \) and

\[ c = (\frac{1}{2}w+r)^2 , \]

(iii) \( \hat{\theta}_w \), a weighted least squares estimator where the matrix of weights \( W \) is a consistent estimator of \( V^{-1} \).

Now the asymptotic variance of \( \bar{\theta} \) as \( a \to \infty \) is simply

\[
\text{Var}(\bar{\theta}) = \frac{1}{a} \left[ \left( \frac{\delta - 2\Theta \gamma + \Theta^3}{\alpha} \right) + (3\Delta - 6\delta + 8\Theta \gamma - \Theta^2 - 4\Theta^3) \right] \\
+ \frac{1}{2an(n-1)} \left[ \frac{(\Theta - 2\gamma + \delta)}{\alpha} + (1 - 6\Theta + 8\gamma + 3\Delta - 6\delta) \right] \\
+ o(a^{-1}) .
\]  

The asymptotic variance of \( \hat{\theta} \) as \( a \to \infty \) is
\[
\text{Var} (\hat{\theta}) = \frac{1}{a} \left[ \frac{(3-2\theta+\theta^3)}{\alpha} + \left(3\Delta-6\delta+8\theta-6\theta^2-4\theta^3\right) \right] \\
+ \frac{1}{\text{an}} \left[ \frac{2(\gamma-\delta)+2\theta-5\delta}{\alpha} + \theta^2 \left(6+2\gamma+18\delta\right) \right] \\
+ \left[ \frac{2(\theta-4\gamma-3\Delta+6\delta)-8\theta+6\theta^2+\theta^3 (9-4\theta+9(8\gamma+3\Delta-6\delta))}{\alpha} \right] \\
+ \frac{(1-3\theta)^2}{2\text{an}(n-1)} \left[ \frac{(\theta-2\gamma+\delta)}{\alpha} + (1-6\theta+8\gamma+3\Delta-6\delta) \right] \\
+ o(a^{-1}) . \hspace{1cm} (3.3.12)
\]

The difference between \(\text{Var} (\hat{\theta})\) and \(\text{Var} (\hat{\theta})\) is likely to be miniscule as

\[
\text{Var} (\hat{\theta}) - \text{Var} (\hat{\theta}) = \frac{3\theta}{25\text{an}(n-1)} \left[ 3 \left( n - \frac{1}{2} \right) \theta - 5 \left( \frac{\theta-2\gamma+\delta}{\alpha} + (1-6\theta+8\gamma+3\Delta-6\delta) \right) \right] \\
+ o(a^{-1}) \\
= \frac{3\phi(1-\phi)}{25\text{an}(n-1)} \left[ 3 \left( n - \frac{1}{2} \right)(1-\phi) - 5 \right] \left[ \frac{(1-\phi)}{5\alpha} + \phi \right]
\]

where the approximate expression in \(\phi = e^{-t/2N}\) is for large sized populations and long time, \(t\).

The asymptotic variance of \(\hat{\theta}_w\) as \(a \to \infty\) is given by expression (3.3.6) with \(m\) set equal to 1, namely,

\[
\text{Var} (\hat{\theta}_w) = \frac{(\xi_1 - 2\xi_2 + \xi_3)\theta^2 + 2(\xi_2 - \xi_3)\theta + \xi_3}{(\xi_1\xi_3 - \xi_2^2)\alpha^2} , \hspace{1cm} (3.3.13)
\]

where

\[
\xi_1 = \frac{1}{\det(V)} \left( \begin{array}{cc} V_{22} & V_{23} \\ V_{33} & V_{22} \end{array} \right) , \\
\xi_2 = \frac{1}{\det(V)} \left( \begin{array}{ccc} V_{12} & V_{23} & -\frac{1}{2} V_{33} \end{array} \right) , \\
\xi_3 = \frac{1}{\det(V)} \left( \begin{array}{ccc} V_{11} & V_{22} & -\frac{1}{4} V_{11} \end{array} \right) ,
\]

\[
V_{12} = V_{22} V_{23} - \frac{1}{2} V_{33} V_{12} + \frac{1}{2} V_{13} V_{23} , \\
V_{11} = V_{11} V_{22} - V_{12} V_{23} , \\
V_{13} = \frac{1}{4} V_{11} V_{33} + V_{11} V_{22} - V_{12} V_{23} - \frac{1}{4} V_{12}^2 - V_{13}^2 + V_{12} V_{13} .
\]
and where \( \text{det}(V) \) denotes the determinant of \( V \) and \( v_{ij} \) a typical element of \( V \).

Some numerical comparisons of asymptotic standard deviations of \( \tilde{\theta} \) and \( \hat{\theta}_W \), using equations (3.3.11) and (3.3.13), are presented in Table 3.7 for subdivided monoecious populations of size 100. It is apparent from Table 3.7 that there is little difference in the asymptotic standard deviations of the two estimators except in the cases where the sample size within each population (\( n \)) is much less than the number of populations sampled (\( a \)) and this case is rarely found in practice.

3.3.2 Two Loci

For the two-locus case, asymptotic variances of the estimators \( \mathbf{y}^2 \), \( \tilde{\theta} \), \( \hat{\theta} \) and \( \hat{\theta}_W \) may be calculated by setting \( m = 2 \) in equations (3.3.9), (3.3.7), (3.3.4) and (3.3.6), respectively. In Tables 3.8 and 3.9 asymptotic standard deviations of the estimators have been calculated using these equations for samples of 30 replicate monoecious populations of size 100. The computation of these asymptotic standard deviations for various generations, \( t \), utilized the closed form solutions for the one-locus descent measures given in Cockerham, Weir and Reynolds (1981) and the transition equations for the two-locus nonidentity descent measures given in Weir, Avery and Hill (1980).

As expected, for equal initial gene frequencies at both loci, the asymptotic standard deviation of \( \mathbf{y}^2 \) is equal to the asymptotic standard deviation of \( \tilde{\theta} \). For equal initial gene frequencies and \( n \geq a \), the asymptotic standard deviation of \( \hat{\theta} \) is very close to that of \( \tilde{\theta} \) and of course \( \mathbf{y}^2 \). In all cases the asymptotic standard deviation of \( \hat{\theta}_W \) is less than or equal to the asymptotic standard deviations of the other
Table 3.7  Asymptotic standard deviations (sd) of $\hat{\theta}$ and $\hat{\theta}_W$
for subdivided monoecious populations of size $N = 100$ at various generation times ($t$).

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\theta$</th>
<th>$a$</th>
<th>$n$</th>
<th>$sd(\hat{\theta})$</th>
<th>$sd(\hat{\theta}_w)$</th>
<th>$sd(\hat{\theta})$</th>
<th>$sd(\hat{\theta}_w)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.09539</td>
<td>30</td>
<td>10</td>
<td>0.03926</td>
<td>0.03910</td>
<td>0.03379</td>
<td>0.03367</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30</td>
<td>10</td>
<td>0.02912</td>
<td>0.02712</td>
<td>0.02610</td>
<td>0.02609</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>50</td>
<td>10</td>
<td>0.02789</td>
<td>0.02789</td>
<td>0.02463</td>
<td>0.02463</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>30</td>
<td>10</td>
<td>0.03041</td>
<td>0.03029</td>
<td>0.02617</td>
<td>0.02608</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>50</td>
<td>10</td>
<td>0.02302</td>
<td>0.02301</td>
<td>0.02022</td>
<td>0.02021</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>50</td>
<td>10</td>
<td>0.02160</td>
<td>0.02160</td>
<td>0.01908</td>
<td>0.01908</td>
</tr>
<tr>
<td>50</td>
<td>0.22169</td>
<td>30</td>
<td>10</td>
<td>0.07111</td>
<td>0.07102</td>
<td>0.05454</td>
<td>0.05448</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30</td>
<td>10</td>
<td>0.06222</td>
<td>0.06221</td>
<td>0.04853</td>
<td>0.04852</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>50</td>
<td>10</td>
<td>0.06051</td>
<td>0.06051</td>
<td>0.04738</td>
<td>0.04738</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>10</td>
<td>10</td>
<td>0.05508</td>
<td>0.05501</td>
<td>0.04225</td>
<td>0.04220</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>30</td>
<td>10</td>
<td>0.04819</td>
<td>0.04819</td>
<td>0.03759</td>
<td>0.03759</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>50</td>
<td>10</td>
<td>0.04687</td>
<td>0.04687</td>
<td>0.03670</td>
<td>0.03670</td>
</tr>
<tr>
<td>100</td>
<td>0.39423</td>
<td>30</td>
<td>10</td>
<td>0.1051</td>
<td>0.1051</td>
<td>0.07252</td>
<td>0.07249</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30</td>
<td>10</td>
<td>0.09804</td>
<td>0.09804</td>
<td>0.06827</td>
<td>0.06827</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>50</td>
<td>10</td>
<td>0.09668</td>
<td>0.09668</td>
<td>0.06746</td>
<td>0.06746</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>10</td>
<td>10</td>
<td>0.08144</td>
<td>0.08140</td>
<td>0.05618</td>
<td>0.05615</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>30</td>
<td>10</td>
<td>0.07594</td>
<td>0.07594</td>
<td>0.05288</td>
<td>0.05288</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>50</td>
<td>10</td>
<td>0.07489</td>
<td>0.07489</td>
<td>0.05226</td>
<td>0.05226</td>
</tr>
<tr>
<td>200</td>
<td>0.63304</td>
<td>30</td>
<td>10</td>
<td>0.1231</td>
<td>0.1231</td>
<td>0.07823</td>
<td>0.07821</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>30</td>
<td>10</td>
<td>0.1188</td>
<td>0.1188</td>
<td>0.07569</td>
<td>0.07569</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>50</td>
<td>10</td>
<td>0.1180</td>
<td>0.1180</td>
<td>0.07522</td>
<td>0.07522</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>10</td>
<td>10</td>
<td>0.09534</td>
<td>0.09531</td>
<td>0.06059</td>
<td>0.06058</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>30</td>
<td>10</td>
<td>0.09203</td>
<td>0.09203</td>
<td>0.05863</td>
<td>0.05863</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>50</td>
<td>10</td>
<td>0.09140</td>
<td>0.09140</td>
<td>0.05826</td>
<td>0.05826</td>
</tr>
</tbody>
</table>
Table 3.8  Asymptotic standard deviations (sd) of $\tilde{\gamma}$, $\tilde{\theta}$, $\hat{\theta}$ and $\hat{\theta}_W$ for 30 subdivided monoecious populations of size $N = 100$ with linkage parameter $\lambda = 0.0$.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\delta$</th>
<th>$n$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>sd($\tilde{\theta}$)</th>
<th>sd($\tilde{\gamma}$)</th>
<th>sd($\hat{\theta}$)</th>
<th>sd($\hat{\theta}_W$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.09539</td>
<td>10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.02844</td>
<td>0.02844</td>
<td>0.02833</td>
<td>0.02833</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.09539</td>
<td>0.24</td>
<td>0.24</td>
<td>0.02627</td>
<td>0.02692</td>
<td>0.02993</td>
<td>0.02578</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.09539</td>
<td>0.24</td>
<td>0.24</td>
<td>0.02390</td>
<td>0.02390</td>
<td>0.02382</td>
<td>0.02382</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.09539</td>
<td>0.09</td>
<td>0.09</td>
<td>0.02147</td>
<td>0.02147</td>
<td>0.02149</td>
<td>0.02146</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.09539</td>
<td>0.24</td>
<td>0.24</td>
<td>0.02002</td>
<td>0.02071</td>
<td>0.02321</td>
<td>0.01979</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.09539</td>
<td>0.24</td>
<td>0.24</td>
<td>0.01846</td>
<td>0.01846</td>
<td>0.01846</td>
<td>0.01846</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.09539</td>
<td>0.09</td>
<td>0.09</td>
<td>0.02013</td>
<td>0.02013</td>
<td>0.02013</td>
<td>0.02013</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.09539</td>
<td>0.24</td>
<td>0.24</td>
<td>0.01882</td>
<td>0.01953</td>
<td>0.02190</td>
<td>0.01863</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.09539</td>
<td>0.24</td>
<td>0.24</td>
<td>0.01742</td>
<td>0.01742</td>
<td>0.01744</td>
<td>0.01742</td>
</tr>
<tr>
<td>50</td>
<td>0.22169</td>
<td>10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.05222</td>
<td>0.05222</td>
<td>0.05231</td>
<td>0.05216</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.22169</td>
<td>0.24</td>
<td>0.24</td>
<td>0.04591</td>
<td>0.04449</td>
<td>0.04875</td>
<td>0.04383</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.22169</td>
<td>0.24</td>
<td>0.24</td>
<td>0.03857</td>
<td>0.03857</td>
<td>0.03863</td>
<td>0.03853</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.22169</td>
<td>0.09</td>
<td>0.09</td>
<td>0.04561</td>
<td>0.04561</td>
<td>0.04573</td>
<td>0.04561</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.22169</td>
<td>0.24</td>
<td>0.24</td>
<td>0.04036</td>
<td>0.03944</td>
<td>0.04338</td>
<td>0.03878</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.22169</td>
<td>0.24</td>
<td>0.24</td>
<td>0.03432</td>
<td>0.03432</td>
<td>0.03439</td>
<td>0.03432</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.22169</td>
<td>0.09</td>
<td>0.09</td>
<td>0.04434</td>
<td>0.04434</td>
<td>0.04442</td>
<td>0.04434</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.22169</td>
<td>0.24</td>
<td>0.24</td>
<td>0.03930</td>
<td>0.03847</td>
<td>0.04232</td>
<td>0.03781</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.22169</td>
<td>0.24</td>
<td>0.24</td>
<td>0.03351</td>
<td>0.03351</td>
<td>0.03356</td>
<td>0.03351</td>
</tr>
<tr>
<td>100</td>
<td>0.39423</td>
<td>10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.07797</td>
<td>0.07797</td>
<td>0.07842</td>
<td>0.07793</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.39423</td>
<td>0.24</td>
<td>0.24</td>
<td>0.06599</td>
<td>0.06072</td>
<td>0.06537</td>
<td>0.06057</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.39423</td>
<td>0.24</td>
<td>0.24</td>
<td>0.05129</td>
<td>0.05129</td>
<td>0.05156</td>
<td>0.05126</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.39423</td>
<td>0.09</td>
<td>0.09</td>
<td>0.07265</td>
<td>0.07265</td>
<td>0.07289</td>
<td>0.07265</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.39423</td>
<td>0.24</td>
<td>0.24</td>
<td>0.06168</td>
<td>0.05701</td>
<td>0.06136</td>
<td>0.05686</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.39423</td>
<td>0.24</td>
<td>0.24</td>
<td>0.04828</td>
<td>0.04828</td>
<td>0.04842</td>
<td>0.04827</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.39423</td>
<td>0.09</td>
<td>0.09</td>
<td>0.07163</td>
<td>0.07163</td>
<td>0.07179</td>
<td>0.07163</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.39423</td>
<td>0.24</td>
<td>0.24</td>
<td>0.06086</td>
<td>0.05631</td>
<td>0.06057</td>
<td>0.05615</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.39423</td>
<td>0.24</td>
<td>0.24</td>
<td>0.04771</td>
<td>0.04771</td>
<td>0.04780</td>
<td>0.04771</td>
</tr>
</tbody>
</table>
Table 3.9 Asymptotic standard deviations (sd) of $\tilde{y}$, $\hat{\theta}$, $\hat{\theta}$ and $\hat{\theta}_{W}$ for 30 subdivided monoeccious populations of size $N = 100$ with linkage parameter $\lambda = 1.0$.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\hat{\theta}$</th>
<th>$n$</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
<th>sd($\tilde{y}$)</th>
<th>sd($\hat{\theta}$)</th>
<th>sd($\hat{\theta}$)</th>
<th>sd($\hat{\theta}_{W}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>.09539</td>
<td>10</td>
<td>.09</td>
<td>.09</td>
<td>.02890</td>
<td>.02890</td>
<td>.02878</td>
<td>.02878</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>.09</td>
<td>.24</td>
<td>.02677</td>
<td>.02730</td>
<td>.03012</td>
<td>.02626</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>.24</td>
<td>.24</td>
<td>.02445</td>
<td>.02445</td>
<td>.02435</td>
<td>.02435</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.09</td>
<td>.09</td>
<td>.02175</td>
<td>.02175</td>
<td>.02178</td>
<td>.02175</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.09</td>
<td>.24</td>
<td>.02032</td>
<td>.02094</td>
<td>.02332</td>
<td>.02009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.24</td>
<td>.24</td>
<td>.01879</td>
<td>.01879</td>
<td>.01881</td>
<td>.01878</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.09</td>
<td>.09</td>
<td>.02038</td>
<td>.02038</td>
<td>.02041</td>
<td>.02038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.09</td>
<td>.24</td>
<td>.01909</td>
<td>.01973</td>
<td>.02200</td>
<td>.01889</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.24</td>
<td>.24</td>
<td>.01771</td>
<td>.01771</td>
<td>.01773</td>
<td>.01771</td>
</tr>
<tr>
<td>50</td>
<td>.22169</td>
<td>10</td>
<td>.09</td>
<td>.09</td>
<td>.05356</td>
<td>.05356</td>
<td>.05366</td>
<td>.05349</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>.09</td>
<td>.24</td>
<td>.04742</td>
<td>.04574</td>
<td>.04938</td>
<td>.04526</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>.24</td>
<td>.24</td>
<td>.04036</td>
<td>.04036</td>
<td>.04043</td>
<td>.04031</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.09</td>
<td>.09</td>
<td>.04668</td>
<td>.04668</td>
<td>.04681</td>
<td>.04668</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.09</td>
<td>.24</td>
<td>.04157</td>
<td>.04042</td>
<td>.04387</td>
<td>.03993</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.24</td>
<td>.24</td>
<td>.03573</td>
<td>.03573</td>
<td>.03582</td>
<td>.03573</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.09</td>
<td>.09</td>
<td>.04537</td>
<td>.04537</td>
<td>.04546</td>
<td>.04537</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.09</td>
<td>.24</td>
<td>.04045</td>
<td>.03941</td>
<td>.04279</td>
<td>.03891</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.24</td>
<td>.24</td>
<td>.03485</td>
<td>.03485</td>
<td>.03491</td>
<td>.03485</td>
</tr>
<tr>
<td>100</td>
<td>.39423</td>
<td>10</td>
<td>.09</td>
<td>.09</td>
<td>.08074</td>
<td>.08074</td>
<td>.08123</td>
<td>.08070</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>.09</td>
<td>.24</td>
<td>.06924</td>
<td>.06352</td>
<td>.06683</td>
<td>.06348</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>.24</td>
<td>.24</td>
<td>.05540</td>
<td>.05540</td>
<td>.05574</td>
<td>.05537</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.09</td>
<td>.09</td>
<td>.07513</td>
<td>.07513</td>
<td>.07539</td>
<td>.07513</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.09</td>
<td>.24</td>
<td>.06458</td>
<td>.05951</td>
<td>.06265</td>
<td>.05949</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>.24</td>
<td>.24</td>
<td>.05193</td>
<td>.05193</td>
<td>.05211</td>
<td>.05193</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.09</td>
<td>.09</td>
<td>.07406</td>
<td>.07406</td>
<td>.07423</td>
<td>.07406</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.09</td>
<td>.24</td>
<td>.06369</td>
<td>.05875</td>
<td>.06183</td>
<td>.05872</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>.24</td>
<td>.24</td>
<td>.05128</td>
<td>.05128</td>
<td>.05139</td>
<td>.05128</td>
</tr>
</tbody>
</table>
estimators. In the preparation of Tables 3.8 and 3.9, calculations were performed for values of \( t \) ranging from 10 to 200 in steps of 10 generations. For the one case of unequal initial gene frequencies at the two loci that was studied, the behavior of the asymptotic standard deviations of \( \hat{\theta}, \tilde{\theta}, \) and \( \hat{\theta} \) exhibited the following pattern. For \( \lambda = 0.0 \), that is in the construction of Table 3.8, and for \( t \leq 30 \),

\[
\text{sd}(\hat{\theta}) < \text{sd}(\tilde{\theta}) < \text{sd}(\hat{\theta}) ,
\]

but for \( 40 \leq t \leq 90 \),

\[
\text{sd}(\tilde{\theta}) < \text{sd}(\hat{\theta}) < \text{sd}(\hat{\theta}) ,
\]

while for \( 100 \leq t \leq 200 \),

\[
\text{sd}(\tilde{\theta}) < \text{sd}(\hat{\theta}) < \text{sd}(\hat{\theta}) .
\]

A similar pattern was observed for \( \lambda = 1.0 \) in the construction of Table 3.9. This behavior may be deduced from the relevant equations for the asymptotic variances and can be attributed to the relationships among the coefficients of the functions of the one-locus descent measures, namely,

\[
\frac{1}{\alpha_1 + \alpha_2} \leq \frac{\alpha_1^2 + \alpha_2^3}{(\alpha_1 + \alpha_2)^2} \leq \frac{1}{4} \left( \frac{1}{\alpha_1} + \frac{1}{\alpha_2} \right)
\]

and

\[
\frac{1}{2} \leq \frac{\alpha_1^2 + \alpha_2^2}{(\alpha_1 + \alpha_2)^2} \leq \frac{\alpha_1^4 + \alpha_2^4}{(\alpha_1^2 + \alpha_2^2)^2} ,
\]

and the changes over time in the relative magnitude of the functions of the one-locus descent measures, for example, for \( N = 100 \) and for changes in \( t \) from 20 to 50 to 100, the corresponding changes in
\[
\frac{\delta - 2\theta \gamma + \theta^3}{3\Delta - 6\theta + 8\theta \gamma - \theta^2 - 4\theta^3}
\]

are from 0.054 to 0.186 to 0.695.

Another example of this change in the ranking of the asymptotic standard deviations of the estimators over time is provided in Figure 3.1 where asymptotic coefficients of variation of each estimator are graphed on a logarithmic time scale for samples of 50 individuals from 20 replicate monoecious populations of size 1000 with quite different initial gene frequencies at two tightly linked loci. In this situation the weighted least squares estimator, with weight matrix \( W \), a consistent estimator of \( V^{-1} \), is uniformly best among the four estimators. However for long times, say \( t > 500 \), there is little difference between the asymptotic coefficients of variation and asymptotic standard deviations of \( \hat{\theta}, \tilde{\theta}, \text{ and } \hat{\theta}_W \), so that the preferred estimator on the grounds of computational ease might be \( \tilde{\theta} \). The estimator \( \hat{\theta} \) which is an unweighted average across loci performs poorly when \( t \) is large. However, when \( t \) is small, \( \hat{\theta} \) performs almost as well as \( \hat{\theta}_W \) and certainly better than \( \tilde{\theta} \) and \( \hat{\theta} \).

Returning to Tables 3.8 and 3.9 it is apparent that linkage increases the asymptotic standard deviations of the estimators but does not appear to change the ranking of the estimators.

In summary it might appear from these limited numerical examples that when \( \theta \) is to be estimated from data gathered on two loci from a large number of isolated monoecious populations, and it cannot be assumed that initial gene frequencies are the same at both loci, that

(i) \( \hat{\theta} \) is the "appropriate" estimator when it is suspected that the populations have been isolated for a long time;
Figure 3.1  Asymptotic coefficients of variation (CV) of $\hat{\theta}$, $\hat{\theta}$, $\hat{\theta}$ and $\hat{\theta}^W$ for samples of size 50 from 20 monoeocious populations of size $N = 1000$ with $\alpha_1 = 0.0196$, $\alpha_2 = 0.2496$ and $\lambda = 0.9$.

(ii) $\hat{\theta}$ is the "appropriate" estimator when it is suspected that the populations have been isolated for a short time;

(iii) $\hat{\theta}^W$ is the "appropriate" estimator when no auxiliary information regarding time is available and when a consistent and positive definite estimator of $\nu^{-1}$ is able to be constructed.
4. A SIMULATION STUDY

In the previous chapter some large sample properties \((a \to \infty)\) of four estimators of the coancestry coefficient were presented. In this chapter an attempt to elucidate the small sample properties of the estimators via a small simulation study is described. The case of two replicate monoecious populations \((a = 2)\) practicing random mating, including selfing, with two loci scored, was simulated.

4.1 Discussion of Programming

The program for the simulation study was written in the Fortran IV language and was run on computers at the Triangle Universities Computation Center (TUCC). Pairs of monoecious random mating populations of 100 individuals each were constructed by sampling without replacement from an initial reference population of infinite size. Four kinds of initial reference populations, all in Hardy-Weinberg equilibrium at each of two loci and all in linkage equilibrium, were utilized. These reference populations had the following parameters:

\[
\begin{align*}
\alpha_1 &= .25, \quad \alpha_2 = .25, \quad \lambda = 0 \\
\alpha_1 &= .25, \quad \alpha_2 = .25, \quad \lambda = 0.9 \\
\alpha_1 &= .24, \quad \alpha_2 = .09, \quad \lambda = 0 \\
\text{and} \quad \alpha_1 &= .24, \quad \alpha_2 = .09, \quad \lambda = 0.9 .
\end{align*}
\]

A replication of the experiment consisted of a pair of populations "monitored" over 100 generations of random mating and 100 replicates were generated for each of the four kinds of initial reference populations. Sampling of 100 individuals from the reference population to form one of a pair of founder populations and sampling of 200 gametes
from an infinite pool of gametes to form a next generation of a population was achieved by the usual method of obtaining random numbers from a particular discrete distribution utilizing pseudo-random uniform (0,1) deviates. The IMSL (1979) subroutine GGUBS (a multiplicative generator) was used for this purpose.

The estimators \( \hat{\theta}, \tilde{\theta}, \hat{\delta} \) and \( \hat{\delta}_W \) were computed for each generation for each replicate pair of populations with the consequence that sample size \( n \) equaled population size \( N \). If a replicate pair of populations became fixed for the same alleles at both loci, that replicate was ignored in subsequent generations. This happened to one of the replicates in the case

\[
\alpha_1 = .24, \quad \alpha_2 = .09, \quad \lambda = 0
\]

at generation 91 and to one of the replicates in the case

\[
\alpha_1 = .24, \quad \alpha_2 = .09, \quad \lambda = 0.9
\]

at generation 81. If a pair of populations became fixed for the same allele at a single locus, the estimation of \( \hat{\theta} \) for that replicate was overlooked in the incidental generation and subsequent generations. The other three estimators were calculated. By generation 100, fixation of the same allele at a single locus had occurred in four replicates for the case \( \alpha_1 = \alpha_2 = .25, \lambda = 0 \) and \( \lambda = 0.9 \), in 41 replicates for the case \( \alpha_1 = .24, \alpha_2 = .09, \lambda = 0 \), and in 44 replicates for the case \( \alpha_1 = .24, \alpha_2 = .09, \lambda = 0.9 \).

Computation of the estimators \( \hat{\theta}, \tilde{\theta} \) and \( \hat{\delta} \) posed no problems as closed forms exist for these estimators (see equations 3.1.5, 3.1.6 and 3.3.3). The computation of \( \hat{\delta}_W \) was accomplished by way of damped
Newton-Raphson iteration. The main features of the algorithm to compute $\hat{\theta}_W$ are sketched below.

1. An initial consistent estimate of generation time $t$ was recovered from $\tilde{\theta}$ by the application of the formula

$$t = \log_e (1-\tilde{\theta})/\log_e (1- \frac{1}{2N})$$

If the application of this formula yielded a negative value for $t$, $t$ was set equal to zero. This value of $t$ was used to generate, via the appropriate transition equations, consistent estimates of the higher order one-locus descent measures ($\gamma$, $\Delta$ and $\delta$) and the two-locus measures ($\Theta$, $\Gamma$ and $\Delta^*$). The calculation of the two-locus measures presupposed that the linkage parameter $\lambda$ was known exactly.

2. The estimated total variances for locus 1 ($z_1 + w_1 + \frac{1}{2} r_1$) and locus 2 ($z_2 + w_2 + \frac{1}{2} r_2$) were used as initial estimates of $\alpha_1$ and $\alpha_2$, respectively. If either of these was greater than 0.25 it was set equal to 0.25. If one of these was zero the locus was dropped from the analysis.

3. A consistent estimate of the variance-covariance matrix of the observed variance components was formed from the results of (1) and (2). Since $a = 2$, terms with coefficients $1/a(a-1)$ were not ignored in the construction of this matrix.

4. The weight matrix was formed by inversion of the estimated variance-covariance matrix. Since the estimated variance-covariance matrix is positive definite and symmetric, the IMSL (1979) subroutine LINV3P was used for the inversion.

5. Damped Newton-Raphson iteration using $z_1 + w_1 + \frac{1}{2} r_1$, $z_2 + w_2 + \frac{1}{2} r_2$ and $\tilde{\theta}$ as initial estimates of $\alpha_1$, $\alpha_2$ and $\theta$, respectively, was carried out. The IMSL (1979) subroutine LINV3P was used to invert the
Hessian. If this subroutine returned an error message indicating that
the Hessian was algorithmically not positive definite, iteration was
stopped and the current value of the iterate was reported as the
estimate. This type of error occurred in less than 2% of the minimiza-
tions and seemed to occur when the estimated value of \( t \) was far from
the actual value of \( t \), for example an estimated \( t \) of 921 and an actual
\( t \) of 97, or an estimated \( t \) equal to zero and an actual \( t \) of 52.

The damping of the Newton step consisted in reducing the length of
the Newton step by a factor of one-half until the current value of the
objective function \( Q_w(k-1) \) was bettered. If a reduction in the objec-
tive function had not occurred for reduction in the Newton step by a
factor of \( 2^{-32} \), then a series of damped steps in the direction of the
gradient vector were taken until the current value of the objective
function was reduced. If such a reduction was not achieved for a
step of \( 2^{-32} \) times the gradient vector, the current iterate \( \hat{\gamma}(k-1) \) was
reported as the estimate. This type of occurrence happened in less
than 0.5% of the minimizations and generally after four or five
iterations.

6. Convergence was deemed to occur when both

\[
|Q_w(k-1) - Q_w(k)| < 10^{-5} \left[ Q_w(k-1) + 10^{-3} \right]
\]

and

\[
||\hat{y}_w(k-1) - \hat{y}_w(k)|| < 10^{-5} [||\hat{y}_w(k-1)|| + 10^{-3}]
\]

were satisfied or when the number of iterations \( k \) exceeded 50,
although this last criterion did not have to be invoked in any of the
minimizations.
Realized mean square errors and variances of the estimators for each of the 100 generations were computed by averaging over the number of replicates \( r \) in the manner,

\[
\frac{1}{r} \sum_{i=1}^{r} (\theta_i - \bar{\theta})^2 = \frac{r-1}{r} \left[ \frac{1}{r-1} \sum_{i=1}^{r} (\theta_i - \bar{\theta})^2 \right] + (\bar{\theta} - \theta)^2
\]

where \( \theta_i \) denotes an estimate of the true value \( \theta \) for the \( i \)th replicate and \( \bar{\theta} \) is the average of these estimates over replicates.

### 4.2 Results

Selected results for generations 5, 10, 20, 50 and 100 for the four types of initial reference population are presented in Tables 4.1, 4.2, 4.3 and 4.4. In each of these tables the four rows nested within each time, \( t \), refer to the four estimators,

\[ \hat{\theta} \]
\[ \tilde{\theta} \]
\[ \check{\theta} \]
\[ \hat{\theta}_w \]

respectively, while the column labeled Avar contains numerical quantities obtained by evaluating the asymptotic variance formulae, presented in section 3.3, for \( a = 2 \).

In all four tables there is a trend for realized coefficients of variation for each estimator to decrease as \( t \) increases. When there is no linkage (Tables 4.1 and 4.3), there is a tendency for \( \hat{\theta} \) to be the least biased and \( \tilde{\theta} \) to be the most biased of the four estimators in later generations. This can be perceived by comparing the mean square errors
Table 4.1 Realized means, standard deviations, mean square errors and variances for four estimators of $\theta$ for the case of two monoecious populations diverging from an initial reference population with parameters $a_1 = a_2 = .25$ and $\lambda = 0$.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\theta$</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>MSE x 10</th>
<th>Var x 10</th>
<th>Avar x 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.0248</td>
<td>.0257</td>
<td>.0257</td>
<td>.0066</td>
<td>.0066</td>
<td>.0042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0261</td>
<td>.0263</td>
<td>.0069</td>
<td>.0069</td>
<td>.0042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0259</td>
<td>.0257</td>
<td>.0066</td>
<td>.0066</td>
<td>.0042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0259</td>
<td>.0261</td>
<td>.0067</td>
<td>.0068</td>
<td>.0041</td>
</tr>
<tr>
<td>10</td>
<td>.0489</td>
<td>.0441</td>
<td>.0441</td>
<td>.0195</td>
<td>.0194</td>
<td>.0131</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0451</td>
<td>.0457</td>
<td>.0208</td>
<td>.0209</td>
<td>.0131</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0443</td>
<td>.0445</td>
<td>.0198</td>
<td>.0198</td>
<td>.0131</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0446</td>
<td>.0450</td>
<td>.0203</td>
<td>.0203</td>
<td>.0145</td>
</tr>
<tr>
<td>20</td>
<td>.0954</td>
<td>.0899</td>
<td>.0852</td>
<td>.0722</td>
<td>.0726</td>
<td>.0413</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0945</td>
<td>.0938</td>
<td>.0871</td>
<td>.0880</td>
<td>.0413</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0937</td>
<td>.0969</td>
<td>.0930</td>
<td>.0939</td>
<td>.0413</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0922</td>
<td>.0905</td>
<td>.0811</td>
<td>.0818</td>
<td>.0406</td>
</tr>
<tr>
<td>50</td>
<td>.222</td>
<td>.189*</td>
<td>.148</td>
<td>.228</td>
<td>.219</td>
<td>.160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.200</td>
<td>.160</td>
<td>.257</td>
<td>.255</td>
<td>.160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.200</td>
<td>.168</td>
<td>.283</td>
<td>.282</td>
<td>.160</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.193</td>
<td>.153</td>
<td>.240</td>
<td>.234</td>
<td>.135</td>
</tr>
<tr>
<td>100</td>
<td>.394</td>
<td>.307*</td>
<td>.201</td>
<td>.474</td>
<td>.403</td>
<td>.325</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.343*</td>
<td>.225</td>
<td>.530</td>
<td>.508</td>
<td>.325</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.363</td>
<td>.261</td>
<td>.682</td>
<td>.679</td>
<td>.326</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.331*</td>
<td>.224</td>
<td>.538</td>
<td>.503</td>
<td>.348</td>
</tr>
</tbody>
</table>
Table 4.2  Realized means, standard deviations, mean square errors and variances for four estimators of $\theta$ for the case of two monoecious populations diverging from an initial reference population with parameters $\alpha_1 = \alpha_2 = .25$ and $\lambda = 0.9$.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\theta$</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>MSE x 10</th>
<th>Var x 10</th>
<th>Avar x 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.0248</td>
<td>.0173*</td>
<td>.0196</td>
<td>.0044</td>
<td>.0039</td>
<td>.0042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0174*</td>
<td>.0198</td>
<td>.0044</td>
<td>.0039</td>
<td>.0042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0174*</td>
<td>.0197</td>
<td>.0044</td>
<td>.0039</td>
<td>.0042</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0174*</td>
<td>.0198</td>
<td>.0044</td>
<td>.0039</td>
<td>.0041</td>
</tr>
<tr>
<td>10</td>
<td>.0489</td>
<td>.0467</td>
<td>.0450</td>
<td>.0201</td>
<td>.0203</td>
<td>.0132</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0477</td>
<td>.0467</td>
<td>.0216</td>
<td>.0218</td>
<td>.0132</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0467</td>
<td>.0451</td>
<td>.0202</td>
<td>.0204</td>
<td>.0132</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0474</td>
<td>.0464</td>
<td>.0214</td>
<td>.0216</td>
<td>.0142</td>
</tr>
<tr>
<td>20</td>
<td>.0954</td>
<td>.0780*</td>
<td>.0791</td>
<td>.0649</td>
<td>.0625</td>
<td>.0419</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0807</td>
<td>.0831</td>
<td>.0706</td>
<td>.0691</td>
<td>.0419</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0780*</td>
<td>.0808</td>
<td>.0675</td>
<td>.0654</td>
<td>.0419</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0795</td>
<td>.0819</td>
<td>.0690</td>
<td>.0671</td>
<td>.0429</td>
</tr>
<tr>
<td>50</td>
<td>.222</td>
<td>.164*</td>
<td>.127</td>
<td>.192</td>
<td>.160</td>
<td>.162</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.174*</td>
<td>.137</td>
<td>.209</td>
<td>.188</td>
<td>.162</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.175*</td>
<td>.145</td>
<td>.230</td>
<td>.210</td>
<td>.163</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.167*</td>
<td>.131</td>
<td>.198</td>
<td>.170</td>
<td>.112</td>
</tr>
<tr>
<td>100</td>
<td>.394</td>
<td>.281*</td>
<td>.188</td>
<td>.478</td>
<td>.354</td>
<td>.330</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.326*</td>
<td>.239</td>
<td>.612</td>
<td>.570</td>
<td>.331</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.298*</td>
<td>.210</td>
<td>.530</td>
<td>.441</td>
<td>.327</td>
</tr>
</tbody>
</table>
Table 4.3  Realized means, standard deviations, mean square errors and variances for four estimators of $\theta$ for the case of two monoecious populations diverging from an initial reference population with parameters $a_1 = .24$, $a_2 = .09$ and $\lambda = 0$.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\theta$</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>MSE x 10</th>
<th>Var x 10</th>
<th>Avar x 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>.0248</td>
<td>.0188*</td>
<td>.0202</td>
<td>.0044</td>
<td>.0041</td>
<td>.0031</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0194</td>
<td>.0281</td>
<td>.0081</td>
<td>.0079</td>
<td>.0036</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0193</td>
<td>.0327</td>
<td>.0109</td>
<td>.0107</td>
<td>.0047</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0193*</td>
<td>.0218</td>
<td>.0050</td>
<td>.0048</td>
<td>.0031</td>
</tr>
<tr>
<td>10</td>
<td>.0489</td>
<td>.0374*</td>
<td>.0356</td>
<td>.0139</td>
<td>.0127</td>
<td>.0142</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0390*</td>
<td>.0428</td>
<td>.0191</td>
<td>.0184</td>
<td>.0162</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0385*</td>
<td>.0480</td>
<td>.0238</td>
<td>.0230</td>
<td>.0207</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0378*</td>
<td>.0374</td>
<td>.0151</td>
<td>.0140</td>
<td>.0141</td>
</tr>
<tr>
<td>20</td>
<td>.0954</td>
<td>.0798*</td>
<td>.0668</td>
<td>.0466</td>
<td>.0446</td>
<td>.0483</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0921</td>
<td>.0910</td>
<td>.0820</td>
<td>.0827</td>
<td>.0522</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0950</td>
<td>.102</td>
<td>.102</td>
<td>.103</td>
<td>.0657</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0848</td>
<td>.0770</td>
<td>.0598</td>
<td>.0593</td>
<td>.0473</td>
</tr>
<tr>
<td>50</td>
<td>.222</td>
<td>.184*</td>
<td>.131</td>
<td>.184</td>
<td>.172</td>
<td>.223</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.214</td>
<td>.184</td>
<td>.335</td>
<td>.337</td>
<td>.214</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.221</td>
<td>.201</td>
<td>.400</td>
<td>.404</td>
<td>.259</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.204</td>
<td>.175</td>
<td>.307</td>
<td>.307</td>
<td>.206</td>
</tr>
<tr>
<td>100</td>
<td>.394</td>
<td>.280*</td>
<td>.167</td>
<td>.405</td>
<td>.280</td>
<td>.545</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.332*</td>
<td>.255</td>
<td>.680</td>
<td>.648</td>
<td>.467</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.336*</td>
<td>.264</td>
<td>.724</td>
<td>.697</td>
<td>.540</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.326*</td>
<td>.257</td>
<td>.699</td>
<td>.660</td>
<td>.464</td>
</tr>
</tbody>
</table>
Table 4.4  Realized means, standard deviations, mean square errors and variances for four estimators of \( \theta \) for the case of two monoecious populations diverging from an initial reference population with parameters \( \alpha_1 = .24, \alpha_2 = .09 \) and \( \lambda = 0.9 \).

<table>
<thead>
<tr>
<th>( t )</th>
<th>( \theta )</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>MSE x 10</th>
<th>Var x 10</th>
<th>Avar x 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>(.248)</td>
<td>.0207</td>
<td>.0233</td>
<td>.0056</td>
<td>.0055</td>
<td>.0044</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0201</td>
<td>.0248</td>
<td>.0063</td>
<td>.0062</td>
<td>.0051</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0193*</td>
<td>.0272</td>
<td>.0076</td>
<td>.0074</td>
<td>.0066</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0206</td>
<td>.0231</td>
<td>.0055</td>
<td>.0053</td>
<td>.0044</td>
</tr>
<tr>
<td>10</td>
<td>(.0489)</td>
<td>.0415</td>
<td>.0444</td>
<td>.0201</td>
<td>.0197</td>
<td>.0144</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0410</td>
<td>.0482</td>
<td>.0236</td>
<td>.0232</td>
<td>.0163</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0396</td>
<td>.0517</td>
<td>.0273</td>
<td>.0267</td>
<td>.0208</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0416</td>
<td>.0447</td>
<td>.0203</td>
<td>.0200</td>
<td>.0143</td>
</tr>
<tr>
<td>20</td>
<td>(.0954)</td>
<td>.0732*</td>
<td>.0634</td>
<td>.0447</td>
<td>.0402</td>
<td>.0489</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0787*</td>
<td>.0780</td>
<td>.0631</td>
<td>.0609</td>
<td>.0526</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0777*</td>
<td>.0847</td>
<td>.0742</td>
<td>.0718</td>
<td>.0659</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.0760*</td>
<td>.0696</td>
<td>.0517</td>
<td>.0484</td>
<td>.0479</td>
</tr>
<tr>
<td>50</td>
<td>(.222)</td>
<td>.151*</td>
<td>.116</td>
<td>.183</td>
<td>.135</td>
<td>.225</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.180*</td>
<td>.153</td>
<td>.251</td>
<td>.235</td>
<td>.216</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.186*</td>
<td>.170</td>
<td>.298</td>
<td>.288</td>
<td>.260</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.168*</td>
<td>.142</td>
<td>.227</td>
<td>.200</td>
<td>.209</td>
</tr>
<tr>
<td>100</td>
<td>(.394)</td>
<td>.250*</td>
<td>.158</td>
<td>.453</td>
<td>.250</td>
<td>.549</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.287*</td>
<td>.218</td>
<td>.587</td>
<td>.477</td>
<td>.471</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.296*</td>
<td>.233</td>
<td>.635</td>
<td>.543</td>
<td>.542</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.277*</td>
<td>.212</td>
<td>.581</td>
<td>.449</td>
<td>.468</td>
</tr>
</tbody>
</table>
and variances of the estimators for the later generations and by noting
the asterisks in the column labeled "Mean." For a particular time $t$ a
one-sample $t$-test can be carried out to test whether the realized mean
of an estimator differs from the true value of $\theta$. The asterisk indicates
rejection of the null hypothesis that the realized mean of an estimator
is equal to the true value at the 5% significance level. Note, however,
that since the sample of replicates at time $t$ is related to the sample
of replicates at time $t + k$ the outcomes of the $t$-tests for each
generation will be related so that comparisons of performance of an
estimator for different times should be resisted. For tight linkage
(Tables 4.2 and 4.4) all four estimators appear to be badly biased but
again $\hat{\theta}$ is the least biased in later generations. The extent and direc-
tion of some of these biases can be predicted from the following formulae
for the means of $\hat{\theta}$ and $\tilde{\theta}$:

\[
E(\hat{\theta}) = \theta + \frac{1}{m} \sum_{t=1}^{m} \left( \frac{1}{\alpha_t} - 4 \right) \left\{ \frac{1}{n} \left( \theta^2 - \gamma \right) + \frac{1}{an} \left[ \frac{1}{2} \theta (1 - \theta) - (\theta - \gamma) \right] \right\} + o(a^{-1})
\]

\[
E(\tilde{\theta}) = \theta + \left( \frac{\sum_{t=1}^{m} \alpha_t}{m} \right) \left( \frac{1}{\alpha_t} - 4 \right) \left\{ \frac{1}{n} \left( \theta^2 - \gamma \right) + \frac{1}{an} \left[ \frac{1}{2} \theta (1 - \theta) - (\theta - \gamma) \right] \right\} + o(a^{-1}) \quad (4.2.2)
\]

Notice that if $\alpha_t = .25$ for all $t$ that the approximate bias in $\hat{\theta}$ and $\tilde{\theta}$
should be zero. This appears to agree with the situation in early
generations in Table 4.1. However terms involving the linkage parameter
$\lambda$ do not enter the approximate formulae in equations (4.2.1) and (4.2.2)
so that the large biases in Tables 4.2 and 4.4 are not predicted. For Table 4.3 realized biases in \( \hat{\theta} \) for times 10 and 50 are -0.0115 and -0.038, while the approximate biases calculated from equation (4.2.1) are -0.0023 and -0.035. Similarly realized biases in \( \tilde{\theta} \) for times 10 and 50 are -0.0099 and -0.008 while the approximate biases calculated from equation (4.2.2) are -0.0006 and -0.010.

In all four tables the estimator \( \hat{\theta} \) generally has the smallest realized mean square error and variance but its bias might prevent an unequivocal recommendation for its routine use in the case of two sampled populations (\( a = 2 \)). There is some agreement between the realized variances and asymptotic variances calculated for \( a = 2 \) particularly for the later generations in Table 4.4 (the case where gene frequencies at the two loci in the initial reference population are quite different and linkage is tight). Routine calculation of these asymptotic variance formulae even for the case of \( a = 2 \) may provide some insight into the variance of an estimate of \( \theta \).
5. DOMINANCE, MULTIPLE ALLELES AND OTHER EXTENSIONS

In this chapter, the effects of failures of some assumptions such as codominance, two alleles per locus, no migration and no mutation, on the methodology described in Chapters 3 and 4 are briefly described. This discussion is certainly not complete and should be regarded as merely an indication of areas worthy of future research.

5.1 Dominance

In previous chapters, alleles at each locus have been assumed to be codominant; that is, genotypes at each locus have been assumed to be identifiable. Consider, now, two loci each with two allelic classes, A and A at the first locus with A completely recessive to A and B and B at the second locus with B completely recessive to B. Let the indicator random variable $z_{ij}$ be such that

$$
\begin{bmatrix}
z_{ij1} \\
\end{bmatrix}
= \begin{bmatrix}
1 \\
1
\end{bmatrix}, \text{ if individual } j \text{ in population } i \text{ is AABB}
$$

$$
\begin{bmatrix}
z_{ij2} \\
\end{bmatrix}
= \begin{bmatrix}
1 \\
0
\end{bmatrix}, \text{ if individual } j \text{ in population } i \text{ is AABB or ABB}
$$

$$
\begin{bmatrix}
0 \\
1
\end{bmatrix}, \text{ if individual } j \text{ in population } i \text{ is ABB or ABBB}
$$

$$
\begin{bmatrix}
0 \\
0
\end{bmatrix}, \text{ otherwise.}
$$

The values of this random variable, for one locus, correspond to the assigned phenotypic values of Robertson (1952) and in this brief discussion an exact treatment of the effect of inbreeding on the variation due to recessive genes is provided. The expectations of various products of the elements of these random vectors over the two sampling processes (a
populations from an infinite number of replicate populations and n individuals from an infinite progeny array in each population) are as follows:

\[
E(z_{ij} z_{i'j}) = \begin{bmatrix}
\frac{P^A}{A} & \frac{P^{AB}}{AB} \\
\frac{P^A}{A} & \frac{P^{AB}}{AB} \\
\frac{P^{AB}}{AB} & \frac{P^B}{B}
\end{bmatrix},
\]

\[
E(z_{ij} z_{i'j'}) = \begin{bmatrix}
\frac{P^A}{A} & \frac{P^{AB}}{AB} \\
\frac{P^A}{A} & \frac{P^{AB}}{AB} \\
\frac{P^{AB}}{AB} & \frac{P^B}{B}
\end{bmatrix},
\]

\[
E(z_{ij} z_{i'j'}) = \begin{bmatrix}
\left(P^A\right)^2 & \frac{P^AP^B}{A} \\
\frac{P^AP^B}{A} & \left(P^B\right)^2
\end{bmatrix}.
\]

A bivariate analysis of variance with labels for among and within populations can be constructed. The among-populations variance-covariance component matrix \( \Sigma_a \) can be written in the form

\[
\Sigma_a = \begin{bmatrix}
\frac{P^A}{A} & \left(P^A\right)^2 & \frac{P^AP^B}{A} \\
\frac{P^AP^B}{A} & \frac{P^AP^B}{AB} \\
\frac{P^AP^B}{A} & \frac{P^AP^B}{AB} & \frac{P^B}{B}
\end{bmatrix}.
\]

and the within-populations variance-covariance component matrix \( \Sigma_b \) can be written in the form

\[
\Sigma_b = \begin{bmatrix}
\frac{P^A}{A} - \frac{P^A}{A} & \frac{P^AB}{AB} - \frac{P^AP^B}{A} \\
\frac{P^AB}{AB} - \frac{P^AP^B}{A} & \frac{P^B}{B} - \frac{P^B}{B}
\end{bmatrix}.
\]
The translations of the elements of these matrices to functions of
descent measures and gene frequencies (assuming linkage equilibrium in
the initial reference population) are

\[
\begin{align*}
\frac{p^A_A|A}{p^A_A} - (\frac{p^A_A}{p_A})^2 &= \delta_{XY} p_A (1-p_A) \\
&+ (4\gamma_{XY} + \Delta_{X,Y} + 2\Delta_{X,Y} - 6\delta_{XY} - F^2) p_A^2 (1-p_A^2) \\
&+ [4\Theta - 12\gamma_{XY} - 2(\Delta_{X,Y} + 2\Delta_{X,Y} + 12\delta_{XY} + 2F^2)] p_A^3 (1-p_A),
\end{align*}
\]

\[
\begin{align*}
\frac{p^A_B}{p^A_B} - \frac{p^A_B}{p^A_B} &= [\Delta_2^* - (1-F)^2] p_A (1-p_A) p_B (1-p_B),
\end{align*}
\]

\[
\begin{align*}
\frac{p^A}{p^A} - \frac{p^A}{p^A} &= (F - \delta_{XY}) p_A (1-p_A) \\
&+ [1 - 4\gamma_{XY} - (\Delta_{X,Y} + 2\Delta_{X,Y}) + 6\delta_{XY}] p_A^2 (1-p_A^2) \\
&- [2F + 4\Theta - 12\gamma_{XY} - 2(\Delta_{X,Y} + 2\Delta_{X,Y} + 12\delta_{XY})] p_A^3 (1-p_A),
\end{align*}
\]

\[
\begin{align*}
\frac{p_{AB}}{p_{AB}} - \frac{p_{AB}}{p_{AB}} &= (\Theta - \Delta_2^*) p_A (1-p_A) p_B (1-p_B).
\end{align*}
\]

It is clear that \(\Theta\) cannot be recovered from the expected values of
the among and within variance components for each locus. If, however,
the recessive alleles, \(A\) at the first locus and \(B\) at the second locus,
are rare in the initial reference population, that is \(p_A\) and \(p_B\) are
close to zero, then

\[
\Sigma_a \approx \begin{bmatrix}
\delta_{XY} p_A & 0 \\
0 & \delta_{XY} p_B
\end{bmatrix},
\]

and

\[
\Sigma_b \approx \begin{bmatrix}
(F - \delta_{XY}) p_A & 0 \\
0 & (F - \delta_{XY}) p_B
\end{bmatrix}.
\]
The parametric function $\delta_{XY}/F$, which increases from 0 to 1 as time increases from generation 1, is identifiable and can be estimated by the methods discussed in Chapters 3 and 4.

As an aside, exact expressions for the among and within variances for the cases of full sib mating and random mating monoecious populations with selfing, considered by Robertson (1952), can be obtained by substituting the exact solutions for the descent measures presented in Cockerham, Weir and Reynolds (1981) into equations (5.1.1) and (5.1.2).

5.2 Multiple Alleles

Consider two loci, each with a series of codominant allelic classes. Let the first locus have $r+1$ alleles $A_1, A_2, \ldots, A_{r+1}$ and the second locus have $s+1$ alleles $B_1, B_2, \ldots, B_{s+1}$. The vector of indicator random variables analogous to equation (3.1.1) and based on a scoring system proposed by Stanton (1960) is

$$ x_{ijk} = \left[ \begin{array}{c} x_{ijk11} \\ \vdots \\ \vdots \\ x_{ijklr} \\ x_{ijk21} \\ \vdots \\ \vdots \\ x_{ijk2s} \end{array} \right] \tag{5.2.1} $$

where $x_{ijklu} = 1$ if haplotype $k$ in individual $j$ in population $i$ carries allele $A_u$, and zero otherwise (the index $u$ runs from 1 to $r$),
and where,

\[ x_{ijk2v} = 1 \text{ if haplotype } k \text{ in individual } j \text{ in population } i \text{ carries } \]
\[ \text{allele } B_v, \text{ and zero otherwise (the index } v \text{ runs from 1 to } s). \]

A multivariate analysis of variance for \( x_{ijk} \) with the same structure as that in Table 3.1 can be derived. The \((r+s) \times (r+s)\) variance-covariance component matrix for among populations, \( \Sigma_a \), can be written in the form

\[
\Sigma_a = \begin{bmatrix}
A_{11} & A_{12} \\
A_{12}^T & A_{22}
\end{bmatrix}
\]

where \( A_{11} \), which is \( r \times r \), has typical diagonal element,

\[
\frac{\hat{p}_u - p_{Au}}{\hat{p}_{Au}^2} = \theta \hat{p}_{Au} (1 - p_{Au}); \quad u = 1, \ldots, r
\]

and typical off-diagonal element,

\[
\frac{\hat{p}_u - p_{Au}}{\hat{p}_{Au}^2} - \hat{p}_{Au} \hat{p}_{Au}^\prime = -\theta \hat{p}_{Au} \hat{p}_{Au}^\prime; \quad u \neq u^\prime.
\]

The matrix \( A_{22} \) is the \( s \times s \) \( B \)-locus analog of \( A_{11} \) and \( A_{12} \), which is \( r \times s \), has typical element,

\[
\frac{p_u^{B_v} - \hat{p}_u p_{B_v}}{\hat{p}_u p_{B_v}} = \Delta_{AuB_v}; \quad u = 1, \ldots, r; \quad v = 1, \ldots, s;
\]

where \( \Delta_{AuB_v} \) is the disequilibrium for alleles \( A_u \) and \( B_v \) in the initial reference population.

The individuals-within-populations variance-covariance component matrix has the form
\[ \Sigma_b = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{bmatrix} \]

where \( B_{11} \), which is \( r \times r \), has typical diagonal element,

\[ \frac{1}{2} \left( p_{A_u}^u + p_{A_u}^u - 2p_{A_u}^u \right) = \left( \frac{1+F}{2} - \theta \right) p_{A_u} (1-p_{A_u}) ; \quad u = 1, \ldots, r \]

and typical off-diagonal elements,

\[ \frac{1}{2} \left( p_{A_u}^u - p_{A_u}^u \right) = -\left( \frac{1+F}{2} - \theta \right) p_{A_u} p_{A_{u'}} ; \quad u \neq u' . \]

The matrix \( B_{22} \) is the \( s \times s \) B-locus analog of \( B_{11} \) and \( B_{12} \) is the \( r \times s \) matrix with typical element,

\[ \frac{1}{2} \left( p_{A_u}^u + p_{A_u}^u - 2p_{A_u}^u \right) = \frac{1}{2} \left( \frac{F}{2} + \frac{1-F}{2} \theta \right) \Delta_{A_u B_v} . \]

The variance-covariance component matrix for haplotypes or gametes within individuals is,

\[ \Sigma_c = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^T & C_{22} \end{bmatrix} \]

where the \( r \times r \) matrix \( C_{11} \) has typical diagonal element,

\[ p_{A_u}^u - p_{A_u}^u = (1-F) p_{A_u} (1-p_{A_u}) ; \quad u = 1, \ldots, r \]

and typical off-diagonal element,

\[ -p_{A_u}^u = -(1-F) p_{A_u} p_{A_{u'}} ; \quad u \neq u' . \]
The matrix $C_{22}$ is the $s \times s$ B-locus analog of $C_{11}$ and $C_{12}$ which is $r \times s$ has typical element

$$A \begin{bmatrix} u \\ v \end{bmatrix} - A \begin{bmatrix} u \\ v \end{bmatrix} = (\hat{F}^{-1} \hat{F}) A \begin{bmatrix} u \\ v \end{bmatrix}.$$

The problem of estimating $\theta$ can now be viewed as the fitting of a vector $s$ which is $\frac{3}{2}[r(r+1) + s(s+1)] \times 1$ to a vector $\sigma(\gamma)$ where $s$ is given by

$$s = \begin{bmatrix}
\text{vech}(\hat{A}_{11}) \\
\text{vech}(\hat{A}_{22}) \\
\text{vech}(\hat{B}_{11}) \\
\text{vech}(\hat{B}_{22}) \\
\text{vech}(\hat{C}_{11}) \\
\text{vech}(\hat{C}_{22})
\end{bmatrix}$$

and where carets denote unbiased estimators, from the multivariate analysis of variance, of the submatrices of the variance-covariance component matrices and the vech operator, to be found in Henderson and Searle (1979), for example, stacks only those parts of columns of matrices on or below the diagonal. The vector $\sigma(\gamma)$ is the expectation of $s$ over the three sampling processes, and

$$\gamma^T = [p_{A_1}, \ldots, p_{A_r}, p_{B_1}, \ldots, p_{B_s}, F, \theta].$$

Various estimators of $\gamma$ and hence of $\theta$ are available. As in Chapters 3 and 4 various norms between $s$ and $\sigma(\gamma)$ may be minimized with respect to $\gamma$. Ratio estimators of $\theta$ similar to $\tilde{\theta}$ and $\hat{\theta}$ may be constructed in various ways with the analogs of $\tilde{\gamma}$ and $\hat{\gamma}$ being
\[
\tilde{\theta} = \frac{\mathbf{f}^T [\text{vech}(\hat{A}_{11})] + \mathbf{g}^T [\text{vech}(\hat{A}_{22})]}{\mathbf{f}^T [\text{vech}(\hat{A}_{11} + \hat{B}_{11} + \frac{1}{2} \hat{C}_{11})] + \mathbf{g}^T [\text{vech}(\hat{A}_{22} + \hat{B}_{22} + \frac{1}{2} \hat{C}_{22})]}
\]

and

\[
\tilde{\theta} = \frac{\mathbf{f}^T [\text{vech}(\hat{A}_{11})] + \mathbf{g}^T [\text{vech}(\hat{A}_{22})]}{\mathbf{f}^T [\text{vech}(\hat{A}_{11} + \hat{B}_{11} + \frac{1}{2} \hat{C}_{11})] + \mathbf{g}^T [\text{vech}(\hat{A}_{22} + \hat{B}_{22} + \frac{1}{2} \hat{C}_{22})]}
\]

where \( \mathbf{f} \) is some \( r(r+1) \times 1 \) vector and \( \mathbf{g} \) is some \( s(s+1) \times 1 \) vector.

Whether asymptotic variances of these estimators differ much for different choices of \( \mathbf{f} \) and \( \mathbf{g} \) and whether an optimal choice of \( \mathbf{f} \) and \( \mathbf{g} \) can be routinely chosen is another matter. Elucidation of the asymptotic variances of these estimators requires the derivation, in the manner of Appendix 9.1 of the variances of the off-diagonal terms in \( \hat{A}_{11}, \hat{A}_{22}, \hat{B}_{11}, \hat{B}_{22}, \hat{C}_{11} \) and \( \hat{C}_{22} \), and the covariances of these off-diagonal terms with all other terms. If the results in Chapters 3 and 4 are any indication there may well be little penalty to be paid (in terms of asymptotic variance) in using the computationally simple estimators \( \tilde{\theta} \) and \( \tilde{\hat{\theta}} \) with \( \mathbf{f} = \mathbf{1} \) and \( \mathbf{g} = \mathbf{1} \). However, the small sample properties of these estimators may well be quite different.

5.3 Mutation and Migration

In the development of the genetic distances for short-term evolution in this thesis a fixed allele model has been utilized; that is, the expectations of gene frequencies over the relevant sampling processes have been assumed to be constant and undisturbed by mutation or any other
force. As an introduction to the problem of accounting for both finite population size and mutation, consider an infinite-alleles mutation model. Any allele which is present in at least two of the sampled replicate populations must have been present in the initial reference population. An allele which is present in only one of the sampled replicate populations is either a mutant which has arisen since divergence or is an allele that was actually present in the initial population but was lost from the other sampled replicate populations or missed in the sampling of individuals from these other replicate populations. Confining attention to alleles that are present in more than one sampled replicate population allows score vectors similar to that in equation (5.2.1) to be set up. A multivariate analysis of variance produces similar variance-covariance component matrices to those in section 5.2, the only differences being that the gene frequencies, $F$ and $\theta$, are functions of the mutation rate. For example, if allele $u$ at locus $A$ appears in samples from more than one replicate line, then the among populations variance-covariance matrix will have a diagonal term equal to

$$\theta \left[ p_{A_u} (1-\nu)^t \right] \left[ 1 - p_{A_u} (1-\nu)^t \right]$$

where $\nu$ is the mutation rate at the $A$-locus. For monoecy with random selfing, $\theta$ is given by

$$\frac{(1-\lambda)^t (1-\nu)^2}{2N(1-\lambda)}$$

where $\lambda = (1-\nu)^2 \frac{(2N-1)}{2N}$. The recovery of divergence time from an estimate of $\theta$ requires a separate estimate of $\nu$, and of course, knowledge of $N$, as $\nu$ is not identifiable. The expected gene frequencies
\[ p_{Au}^* = p_{Au} (1-\nu)^t \]

are identifiable.

The effect of the incorporation of various migration models into the fixed allele model and the joint effect of mutation and migration on measures of genetic distance or more generally on descent measures raises several problems. The general approach that one must follow is to derive the transition equations for the various one-locus and two-locus descent measures under the models. The next step is to find suitable estimators of the one-locus descent measures such as \( \Theta \) that may be candidate genetic distances. If divergence time is to be recovered (note that it may not be possible to speak of divergence if migration among the replicate lines is substantial) from an estimate of a parameter such as \( \Theta \) care must be taken to ensure that the parameter is a monotonic function of time. Estimates of the parameters in the concurrent mutation and migration models, which are required if \( t \) is to be recovered, may require samples of individuals to be taken at more than one time.
6. DISCUSSION

The use of a measure of genetic relationship, \( \theta \), as a genetic distance has been proposed and methods for estimating such a measure have been developed. As a starting point the hierarchical analysis of variance structure discussed by Cockerham (1969, 1973) has been used to elucidate the various sampling processes inherent in the model of evolutionary divergence. The genetic distance \( \theta \) arises from this structure in a natural way as a parameter and the task is to estimate this parameter. In contrast, the treatments by previous authors of such distances have blurred the distinction between statistics and parameters (Latter, 1973; Nei and Chakravarti, 1977).

The distance measures suggested by Nei (1972, 1973) are all functions of the quantity \( I \) presented in equation (2.1.1). The suggested estimator for such a quantity, for a pair of populations \( X, Y \) and for single-locus data is

\[
\bar{I} = \frac{\sum \hat{p}_{X_i} \hat{p}_{Y_i}}{\sqrt{\sum \hat{p}_{X_i}^2 \sum \hat{p}_{Y_i}^2}}
\]

\[I^* = \frac{\sum \hat{p}_{X_i} \hat{p}_{Y_i}}{\sum \hat{p}_{X_i}^2 + \sum \hat{p}_{Y_i}^2} \]

where \( \hat{p}_{X_i} \) is the frequency of allele \( i \) in the sample from population \( X \). This estimator can be approximated by
If the statistic in equation (6.2) is computed for the case of drift and no mutation, its approximate expectation is (Cockerham, 1979)

\[
\frac{\sum_{i} p_{i}^2}{\sum_{i} p_{i}^2 + \theta (1 - \sum_{i} p_{i}^2)}
\]

where \( p_{i} \) is the frequency of allele \( i \) in the initial reference population. Clearly, divergence time cannot be recovered from any function of this statistic without knowledge of \( \sum_{i} p_{i}^2 \), and of course population size \( N \).

While Nei (1972, 1973, 1978a) does not claim that his measures are designed to account for drift alone, the routine calculation of statistics such as \( \bar{\bar{v}} \) and \( \bar{\bar{v}}^* \) as "kernels" for measures of genetic distance for short-term evolution is clearly inappropriate.

In order to investigate how to combine information over several loci in the estimation of the genetic distance \( \theta \), a multivariate analysis of variance framework was set up with the loci as separate "response" variables. It was shown that only the diagonal elements of the estimated variance-covariance component matrices provide information about the parameter \( \theta \). Variances and covariances of these estimated variance components which comprise the statistics from which estimators of the genetic distance are constructed were derived in section 3.1.1 and Appendix 9.1. These variances and covariances may be expressed in terms of frequencies of gene combinations. The effect of different mating systems on these variances and covariances is illuminated by translation of these frequencies of gene combinations to functions of initial gene frequencies and descent measures. The resulting expressions are rather complicated (see Tables 3.3 and 3.4) but simplify for the case of monoecy with random sampling (Table 3.5). An important
feature is that there is covariation between the statistics for separate loci. This covariation is generated by the sampling of gametes in the population to form successive generations and it depends on the mating system, linkage and population size. In the case of large monoecious populations with random selfing that have been in existence for a long time such covariation may be negligible (Table 3.6). The other important feature is that these variances and covariances are at most of order \( a^{-1} \), see equation (3.1.9). Since gametes or haplotypes within individuals, and individuals within replicate populations are related by virtue of the mating system and the finite population size, the only independent samples are the replicate lines themselves so that a central limit theorem for the estimated variance components is obtained only when the number of replicate lines \((a)\) tends to infinity. This means that asymptotic variances of estimators of distances are for large \( a \) (a situation rarely met in practice).

In Chapters 3 and 4, four types of estimators were introduced: an unweighted average over loci of ratio estimators, designated by "\( \bar{v} \)"; a weighted average over loci of ratio estimators designated by a tilde, an ordinary least squares estimator designated by a caret, and a weighted least squares estimator designated by a caret and subscripted with a "\( w \)" to indicate its dependence on a matrix of weights, \( W \). The choice of the weight matrix for the fourth estimator has been overlooked somewhat in the previous chapters. In order to obtain a "best" weighted least squares estimator, \( W \) must be a consistent estimator of the inverse of the variance-covariance matrix of the estimated variance components. One approach is to recover a consistent estimate of time from an initial consistent estimate of \( \theta \), say \( \hat{\theta} \) or \( \tilde{\theta} \), then use this
estimate of time to construct, via the respective transition equations for the system of mating, the necessary higher-order one-locus descent measures and two-locus nonidentity descent measures. Elements of a consistent estimator of the variance-covariance matrix may be obtained by combining these calculated and consistent estimates of the descent measures with initial consistent estimates of \( p_i(1-p_i) \). Such an approach which suggests iterative updating of the estimate of the variance-covariance matrix merits further investigation. A similar method would be used to estimate the "asymptotic standard errors" of estimates of \( \theta \). In the special case of large monoecious populations isolated for a long time, the construction of a consistent estimator of the variance-covariance matrix was seen to involve less computation.

Only one type of mating system has been studied in this thesis, namely monoecy with random selfing, although the necessary machinery to study asymptotic variances of the estimators of \( \theta \) for any regular system of mating has been established. Transition equations for the one-locus identity measures and two-locus nonidentity descent measures for other mating systems, namely monoecious populations excluding selfing and dioecious populations with random or hierarchical pairing are available (Weir, Avery and Hill, 1980). There is some interest in establishing the transition equations for a double-first-cousin mating system as a first step towards approximating the mating pattern of some so-called primitive tribes (Neel, 1978). This would allow the recovery of divergence time for these tribes from some common ancestral population.

Some authors (Smouse and Neel, 1977; Chakraborty, 1980) essentially average the score vectors \( x_{ijk} \) defined in equations (3.1.1) and (3.2.1) over gametes in the manner,
\[ y_{ij} = \frac{1}{2}(x_{ij1} + x_{ij2}). \]

Previous discussions of score vectors similar to \( y_{ij} \) have been given by Wright (1922, 1965), Stanton (1960) and Cockerham (1969). Except in the case of random mating monoecious populations, when \( F = 0 \), it can be shown that the coancestry coefficient \( \theta \) is no longer identifiable when a multivariate analysis of variance for \( y_{ij} \) is constructed but that the parametric function \( \rho = 2\theta/(1+F) \) which is related to Wright's coefficient of relationship is identifiable. The parameter \( \rho \), however, is not a particularly well behaved distance measure as it asymptotes earlier than \( \theta \). This has been pointed out elsewhere (Cockerham, 1978).

Perhaps the greatest weakness in the present study is the lack of an extensive investigation of the small sample properties of the various estimators of \( \theta \). One system of mating, monoecy with selfing, has been the object of a small simulation study. It was found that for \( a = 2 \) (the smallest number of replicate populations that allows the estimation of \( \theta \)) and if two tightly linked loci are scored, the suggested estimators of \( \theta \) are apt to show substantial bias. This is particularly distressing in the case of the estimator \( \hat{\theta}_W \) which was designed to take into account covariation between the scored loci by utilizing a stochastic weight matrix. Future research may need to be directed toward the construction of a simpler, and possibly deterministic, weight matrix. Other research might proceed in the direction of accepting the existing estimators and finding suitable corrections for bias. Equations (3.1.3) and (3.1.4) may provide bias corrections and since they arise as scalar functions of bivariate analogs of the intraclass correlation further research in this area may be fruitful.
As they stand, the four estimators of $\theta$ considered in this thesis may be negative. This occurs most often when divergence time is small and one or more of the observed among variance components, $z_{l}$, are negative. One suggestion (Cockerham, 1979) to obtain non-negative estimates of $\theta$ is to replace the observed among variance components, $z_{l}$, in the numerators of the formulae for $\hat{\gamma}$ and $\hat{\theta}$ by the observed among mean squares for each locus divided by $2n$. This will also give the numerators of these modified ratio estimators a positive bias [at most of order $1/(2n-1)$] and this may help correct the overall negative bias of the ratio estimators.
7. SUMMARY

The use of the coancestry coefficient as a measure of genetic
distance has been espoused. This distance has its genesis in ratios of
among population variation to total variation, hence its propriety as a
measure of genetic diversity among populations. Also, this distance
when viewed as a function of measures of identity by descent is a
monotonic function of divergence time so that estimates of divergence
time can be obtained in either closed form or by graphical or tabular
means from estimates of the distance.

The estimation of this distance using multilocus data was studied.
Variances and covariances of the statistics with which these estimators
are constructed were derived and expressed in terms of higher order
one-locus measures of identity by descent and a set of measures of non-
identity for pairs of loci, under the assumption of linkage equilibrium
at all pairs of scored loci in the initial reference population.
General expressions for any regular mating system for the asymptotic
variances (as the number of replicate populations tends to infinity)
of four types of estimator of this distance were obtained. Asymptotic
normality of these estimators was established, again for the case of a
large number of replicate populations so that hypothesis tests concerning
the values of the genetic distance and the construction of confidence
intervals for the genetic distance and divergence time are routine.

The special case of monoecious random mating populations with
selfing was discussed in some detail and it was found that when $\theta$ is to
be estimated from data gathered at two loci, with two alleles per locus,
from a large number of such isolated populations, that
(i) If initial gene frequencies are the same at both loci, the computationally simple unweighted average of the ratio estimators \( \tilde{\theta} \) is as good as any of the others, and

(ii) If initial gene frequencies are different at both loci, the weighted least squares estimator with weight matrix a consistent estimator of the inverse of the variance-covariance matrix of the constituent statistics appears to have the uniformly smallest asymptotic variance of the four estimators. The computationally simple unweighted average of the ratio estimators \( \hat{\theta} \) is as good when the populations have only recently diverged while the weighted average of the ratio estimators \( \tilde{\theta} \) is as good when the populations have been isolated for some time.

The mean square error properties of the four types of estimator for situations in which only a small number of replicate populations are available must await an extensive simulation study.
8. LIST OF REFERENCES


9. APPENDICES
9.1 Derivation of the Variance-Covariance Matrix of the Vector \( \mathbf{s} \)

The variance-covariance matrix of \( \mathbf{s} \) is,

\[
\text{Var}(\mathbf{s}) = \delta(\mathbf{s}_s^T) - \sigma(\mathbf{y})\sigma^T(\mathbf{y}),
\]

where \( \delta(\mathbf{s}_s^T) \) is the symmetric matrix,

\[
\delta(\mathbf{s}_s^T) = \delta \begin{bmatrix}
    z_1^2 & z_1 z_2 & z_1 \omega_1 & z_1 \omega_2 & z_1 r_1 & z_1 r_2 \\
    z_1 z_2 & z_2^2 & z_2 \omega_1 & z_2 \omega_2 & z_2 r_1 & z_2 r_2 \\
    z_1 \omega_1 & z_2 \omega_1 & \omega_1^2 & \omega_1 \omega_2 & \omega_1 r_1 & \omega_1 r_2 \\
    z_1 \omega_2 & z_2 \omega_2 & \omega_1 \omega_2 & \omega_2^2 & \omega_2 r_1 & \omega_2 r_2 \\
    z_1 r_1 & z_2 r_1 & \omega_1 r_1 & \omega_2 r_1 & r_1^2 & r_1 r_2 \\
    z_1 r_2 & z_2 r_2 & \omega_1 r_2 & \omega_2 r_2 & r_1 r_2 & r_2^2
\end{bmatrix}
\]

(9.1.1)

9.1.1 \( \delta(r^2) \)

\[
\delta(r^2) = \delta \left[ \frac{1}{2a^2} \left( \sum_{i,j,k} x_{ijk} \right)^2 - \frac{2}{4a^2} \left[ \sum_{i,j,k} \sum_{i,j,k} x_{ijk} x_{ijk} \right] \right] 
\]

Now,

\[
\delta \left( \sum_{i,j,k} x_{ijk} \right)^2 = \delta \sum_{i,j,k} x_{ijk} + \sum_{i,j,k} \sum_{i,j,k} x_{ijk} x_{ijk} + \sum_{i,j,k} \sum_{i,j,k} x_{ijk} x_{ijk} 
\]

\[
= an \left( 2p_A + 2p_A^A \right) + an(n-1)(4p_A^A) + a(a-1)n^2(2p_A)^2 
\]

\[= 2an \left[ p_A + p_A + 2(n-1)p_A^A \right] + 4a(a-1)n^2 p_A^2. \]
\[ \delta \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \right]^2 = \delta \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^4 + \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \left( \sum_{i,j,k} x_{ijkl} \right)^2 \right] \]

\[ + \sum_{i,j,k} \sum_{i,j,k} \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \left( \sum_{i,j,k} x_{ijkl} \right)^2 \]

\[ = an \left( 2p_A + 14p_A^4 \right) + an(n-1) \left( 4p_A^4 + 8p_A^4 A^A + 4p_A^4 A \right) \]

\[ + a(a-1)n^2 \left( 2p_A + 2p_A^4 \right)^2 \]

\[ = 2an \left[ p_A + 7p_A^4 + 2(n-1) \left( p_A^4 + 2p_A^4 A^A + p_A^4 A \right) + 2(a-1) \left( p_A + p_A^4 \right) \right] \]

so that,

\[ \delta(r_1^2) = \left( p_A - p_A^4 \right)^2 + \frac{1}{a} \left[ \left( p_A^4 - 2p_A^4 A^A \right) + p_A^4 A - \left( p_A - p_A^4 \right)^2 \right] \]

\[ + \frac{1}{an} \left[ \frac{1}{2} \left( p_A - p_A^4 \right) - \left( p_A^4 - 2p_A^4 A^A \right) + p_A^4 A \right] \]

Of course \( \delta(r_2^2) \) is just the analog of this with \( A \)'s replaced by \( B \)'s.

9.1.2 \( \delta(r_1^2) \)

\[ \delta(r_1^2) = \delta \left( \frac{1}{2} \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \right) - \frac{1}{2} \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \]

\[ \left( \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \right) \]

\[ + \frac{1}{4a^2} \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \right] \]

\[ + \frac{1}{2} \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \left( \sum_{i,j,k} x_{ijkl} \right)^2 \]

\[ + \frac{1}{2} \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijkl} \right)^2 \left( \sum_{i,j,k} x_{ijkl} \right)^2 \]. \]
Now,

\[
\delta \left[ \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \right] = \delta \left[ \sum_{i,j,k} \epsilon_{ijk} \left( \sum_{i,j,k} \epsilon_{ijk} \right) \right] + \sum_{i',j',k} \epsilon_{i'j'k} \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right)
\]

\[+ \sum_{i',j',k} \epsilon_{i'j'k} \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \]

\[= 2an \left[ p_A^B + p_A^{AB} + 2(n-1)p_A^{B|B} + 2(a-1)np_Ap_B \right],
\]

\[
\delta \left[ \left( \sum_{i,j,k} \epsilon_{ijk} \right)^2 \right] \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) = 2an \left[ p_A^B + p_A^{AB} + 2p_A^{AB} \right]
\]

\[+ 2(n-1)(p_A^{B|B} + p_A)|B|B \]

\[+ 2(a-1)np_A(p_B + p_A^B) \right],
\]

\[
\delta \left[ \left( \sum_{i,j,k} \epsilon_{ijk} \right)^2 \right] \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) = 2an \left[ p_A^B + p_A^{AB} + 2p_A^{AB} \right]
\]

\[+ 2(n-1)(p_A^{B|B} + p_A)|B|B \]

\[+ 2(a-1)np_B(p_A + p_A^A) \right],
\]

(9.1.2)

\[
\delta \left[ \left( \sum_{i,j,k} \epsilon_{ijk} \right)^2 \right] \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) = \delta \left[ \left( \sum_{i,j,k} \epsilon_{ijk} \right)^2 \right] \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right)
\]

\[+ \sum_{i',j',k} \epsilon_{i'j'k} \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \]

\[+ \sum_{i',j',k} \epsilon_{i'j'k} \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \left( \sum_{i,j,k} \epsilon_{ijk} \right) \]

\[= 2an \left[ p_A^B + p_A^{AB} + 2(n-1)p_A^{B|B} + 2(a-1)np_Ap_B \right],
\]
\[ = 2an \left[ \frac{1}{2} (p_A^{AB} + p_A^B) + 2 \left( p_A^{AB} + p_A^B \right) + 2p_A^{AB} \right] \\
+ 2(n-1) \left( p_A^B + p_A^{AB} + p_A^B + p_A^{AB} \right) + 2(n-1) \left( p_A^B + p_A^{AB} \right) \]  
(9.1.3) \\
+ 2(n-1)(p_A^B + p_A^{AB}) 
\]

so that,
\[
\delta(r_1, r_2) = (p_A - p_A^A)(p_B - p_B^B) + \frac{1}{a} \left[ p_A^B - (p_A^A + p_A^B) + p_A^{AB} - (p_A - p_A^A)(p_B - p_B^B) \right] \\
+ \frac{1}{a(n-1)} \left[ \frac{1}{2} (p_A^{AB} + p_A^B) + p_A^{AB} - (p_A^A + p_A^{AB}) + p_A^{AB} \right] . 
\]

As a check, it can be seen that \( \delta(r_1, r_2) \) reduces to \( \delta(r_1^2) \) when \( B \)'s are replaced by \( A \)'s and when redundancies in the notation are eliminated in the manner:

- \( p_A^A \) corresponds to \( p_A^A \),
- \( p_A^{AA} \) corresponds to \( p_A \),
- \( p_A^{AA} \), \( p_A^{AA} \), \( p_A^{AA} \) and \( p_A^{AA} \) correspond to \( p_A^A \).

Utilizing this fact, that expectations of cross products involving two loci reduce to expectations of squares or cross products involving one locus, only the expectations of the off-diagonal elements in the submatrices in equation (9.1.1) are presented in the following sections. One locus expectations are recovered by substituting say \( B \)'s for \( A \)'s if a result for the first locus is required and by using Table (9.1) to eliminate redundancies in the notation.
Table 9.1 Translations of redundant two-locus notation to one-locus notation

<table>
<thead>
<tr>
<th>Redundant notation</th>
<th>One-locus notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p^{AA}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{A/A}, p^{AA}, p^{AA}, p^{AA}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{AA}, p^{AA}, p^{AA}, p^{AA}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{A/A}, p^{A/A}, p^{A/A}, p^{A/A}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{A/A}, p^{A/A}, p^{A/A}, p^{A/A}, p^{A/A}, p^{A/A}, p^{A/A}, p^{A/A}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{A/A}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{A/A}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{A/A}$</td>
<td>$P_A$</td>
</tr>
<tr>
<td>$p^{A/A}$</td>
<td>$P_A$</td>
</tr>
</tbody>
</table>

9.1.3 $\delta(w_1, w_2)$

$$
\delta(w_1, w_2) = \frac{1}{16a^2(n-1)^2} \{ \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk1} \right)^2 \right] \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk2} \right)^2 \right] \\
- \frac{1}{n} \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk1} \right)^2 \right] \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk2} \right)^2 \right] \\
- \frac{1}{n} \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk1} \right)^2 \right] \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk2} \right)^2 \right] \\
+ \frac{1}{n} \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk1} \right)^2 \right] \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk2} \right)^2 \right] \}
$$

where

$$
\delta \left\{ \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk1} \right)^2 \right] \left[ \sum_{i,j,k} \left( \sum_{i,j,k} x_{ijk2} \right)^2 \right] \right\} \text{ was given in equation (9.1.3),}
$$
\[ \delta \left\{ \left[ \sum_{i,j,k} \left( x_{ijk1} \right)^2 \right] \left[ \sum_{i,j,k} \left( x_{ijk2} \right)^2 \right] \right\} = 2an \left\{ \left( p^A + p^B \right) + 2 \left( p^A A + p^B B \right) + 2p^A AB \right. \\
+ 2(n-1) \left( p^A \left| B \right. + p^A \left| A \right. \right) + p^B \left| B \right. \right) \\
+ 4(n-1) \left( p^A AB + p^A / B + 2p^B B \right) \\
+ 4(n-1) \left( p^A / B + p^B \right) + 2a-1) \left( p^A A + p^B + p^B + 2(n-1)F^B B \right) \right\}, \]

(9.1.4)

is the analog of equation (9.1.4)

with A and B interchanged,

\[ \delta \left\{ \left[ \sum_{i,j,k} \left( x_{ijk1} \right)^2 \right] \left[ \sum_{i,j,k} \left( x_{ijk2} \right)^2 \right] \right\} = 2an \left\{ \left( p^A + p^B \right) + 2 \left( p^A A + p^B B \right) + 2p^A AB \right. \\
+ 2(n-1) \left( p^A \left| B \right. + p^A \left| A \right. \right) + p^B \left| B \right. \right) \\
+ 4(n-1) \left( p^A AB + p^A / B + p^B \right) + 2 \left( p^A B + p^A B \right) \\
+ 4(n-1) \left( p^A AB + p^A / B + p^B \right) + 2a-1) \left( p^A A + p^B + p^B + 2(n-1)F^B B \right) \right\}, \]

(9.1.5)
so that,

\[ \delta(\omega_{1}, \omega_{2}) = \frac{1}{4} \left( p_{A} + p_{A}^{A} - 2p_{A}^{A} \right) \left( p_{B} + p_{B}^{A} - 2p_{B}^{A} \right) \]

\[ + \frac{1}{4} \sum_{a} \left( \left( p_{A}^{a} \right)^{B} + p_{A}^{a} \right) \left( p_{B}^{a} \right)^{B} + p_{A}^{a} \left( p_{B}^{a} \right)^{B} \right) - \frac{1}{2} \left( p_{A}^{a} \right)^{B} + \left( p_{A}^{a} \right)^{B} + p_{A}^{a} \left( p_{B}^{a} \right)^{B} + p_{A}^{a} \left( p_{B}^{a} \right)^{B} \right) \]

\[ - \frac{1}{4} \left( p_{A} + p_{A}^{A} - 2p_{A}^{A} \right) \left( p_{B} + p_{B}^{A} - 2p_{B}^{A} \right) + p_{A}^{B} \right] \]

\[ + \frac{1}{4} \left( \sum_{ab} \left( p_{A}^{a} \right)^{B} + p_{A}^{a} \right) \left( p_{A}^{b} \right)^{B} + 2 \left( p_{A}^{a} \right)^{B} + p_{A}^{a} \left( p_{A}^{b} \right)^{B} \right) - \frac{1}{4} \left( p_{A}^{a} \right)^{B} + p_{A}^{a} \left( p_{A}^{b} \right)^{B} + p_{A}^{a} \left( p_{A}^{b} \right)^{B} \right) \]

\[ - \frac{1}{2} \left( p_{A} \right)^{B} + p_{A}^{B} + p_{A}^{B} + p_{A}^{B} + 2 \left( p_{A} \right)^{B} + p_{A}^{B} \right] \]

\[ + \left( p_{A} \right)^{B} + p_{A}^{B} + p_{A}^{B} + p_{A}^{B} \left( p_{A} \right)^{B} + 2 \left( p_{A} \right)^{B} + p_{A}^{B} \left( p_{A} \right)^{B} \right) \]

\[ + \frac{1}{2} \left( p_{A} \right)^{B} + 2p_{A}^{B} + p_{A}^{B} + 2 \left( p_{A} \right)^{B} + p_{A}^{B} \right) - \frac{1}{2} \left( p_{A} \right)^{B} + p_{A}^{B} + p_{A}^{B} \right] . \]

9.1.4 \( \delta(\omega_{1}, \omega_{2}) \)

\[ \delta(\omega_{1}, \omega_{2}) = \frac{1}{4a^{n}(n-1)} \delta \left\{ \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \right. \]

\[ - \frac{1}{n} \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \right. \]

\[ - \frac{1}{2} \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \right. \]

\[ + \frac{1}{2n} \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \right. \}

where

\[ \delta \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \left( \sum_{i=1}^{n} \left( \sum_{j=1}^{k} x_{i,j,k} \right)^{2} \right) \]

was given in equation (9.1.2).
\[
\delta \left\{ \left[ \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right] \left[ \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right] \right\} = 2an \left( p_{AB}^{A/B} + p_{A/B}^{A/B} + 2p_{A}^{AB} + 2(n-1) \left( p_{A}^{A/B} + p_{A}^{A/B} \right) \right) \\
+ 4(n-1) \left( p_{AB}^{A/B} + p_{A/B}^{A/B} \right) + 4(n-1)(n-2)p_{A}^{AB} \left( p_{B}^{A/B} + p_{A/B}^{A/B} \right) \\
+ 2(n-1)n \left[ p_{A} + p_{A}^{A/B} + 2(n-1)p_{A}^{A/B} \right] p_{B}^j, \quad (9.1.6)
\]

\[
\delta \left\{ \left[ \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right] \left[ \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right] \right\} \text{ was given in equation (9.1.3)},
\]

\[
\delta \left\{ \left[ \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right] \left[ \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right] \right\} \text{ is the analog of equation (9.1.4)}
\]

with A and B interchanged, so that,

\[
\delta (w_{1} r_{2}) = \frac{1}{2} \left( p_{A}^{A/B} - 2p_{A}^{A/B} \right) (p_{B} - p_{B}^{B/B}) \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) \\
+ \frac{1}{4(a-1)n} \delta \left\{ \left[ \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right] w_{2} - \left[ \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right] w_{1} \right\} \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left[ \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) w_{1} \right] \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left[ \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) w_{1} \right] \\
+ \frac{1}{n} \delta \left\{ w_{1} w_{2} \right\}, \quad (9.1.7)
\]

9.1.5 \( \delta (z_{1} z_{2}) \)

\[
\delta (z_{1} z_{2}) = \frac{1}{16(a-1)n} \delta \left\{ \left[ \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right] \left[ \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right] \right\} \\
- \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) \\
- \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) \\
+ \frac{1}{4(a-1)n} \delta \left\{ \left[ \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right] w_{2} - \left[ \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right] w_{1} \right\} \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left[ \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) w_{1} \right] \\
+ \frac{1}{a} \left( \sum_{i} \left( \sum_{j} x_{ijk1} \right)^{2} \right) \left[ \left( \sum_{i} \left( \sum_{j} x_{ijk2} \right)^{2} \right) w_{1} \right] \\
+ \frac{1}{n} \delta \left\{ w_{1} w_{2} \right\},
\]
where

\[
\begin{align*}
\delta \left\{ \left[ \sum_{i,j,k} (x_{i,j,k})^2 \right] \left[ \sum_{i,j,k} (y_{i,j,k})^2 \right] \right\} &= \text{equation (9.1.5)} + \\
& \quad 8a(a-1)n^2 p_{BL}^A + p_{BL}^B + 2p_{AB}^A + 2(n-1) \\
& \quad \times \left( p_{A}^A \cdot + p_{A}^B \right) \\
& \quad + 4(n-1) \left( p_{A}^A + p_{A}^B \right) \\
& \quad + 4(n-1)(n-2) p_{A}^B \\
& \quad + 8a(a-1)(a-2)n p_{BL}^A + p_{A}^A + 2(n-1)p_{A}^B \\
& \quad + 8a(a-1)(a-2)n^2 p_{BL}^A + p_{A}^A + 2(n-1)p_{A}^B \\
& \quad (9.1.8)
\end{align*}
\]

\[
\begin{align*}
\delta \left\{ \left( \sum_{i,j,k} (x_{i,j,k})^2 \right) \left( \sum_{i,j,k} (y_{i,j,k})^2 \right) \right\} &= \text{equation (9.1.8)} + \\
& \quad 8a(a-1)n^2 \left( p_{AL}^A \right) + 2p_{AB}^A + 2(n-1) \\
& \quad \times \left( p_{A}^A \cdot + p_{A}^B \right) \\
& \quad + 4(n-1) \left( p_{AB}^A \cdot + p_{AB}^B \right) \\
& \quad + 4(n-1)(n-2) p_{A}^B \\
& \quad + \left[ p_{A}^A \cdot + p_{A}^B + 2(n-1)p_{A}^B \right]^2 \\
& \quad + (a-2)n p_{A}^A \left( p_{B}^A + p_{B}^B + 2(n-1)p_{B}^B \right) \\
& \quad + 4(n-2) p_{A}^B \left( p_{B}^A \cdot + p_{A}^B \right) \\
& \quad + 2(n-1)p_{A}^B \\
& \quad + 2(a-2)(a-3)n^2 p_{A}^B \right)^2 \\
& \quad (9.1.8)
\end{align*}
\]
\[ 0 \left( \sum_{i \neq j} \sum_{k} x_{1jk}^2 \right)^2 \left( \sum_{i \neq j} \sum_{k} x_{1jk}^2 \right)^2 \] 

\[ = 2an \left( p_{A}^{AB} + p_{A}^{AB} + 2(p_{A}^{AB} + p_{A}^{AB}) + 2p_{A}^{AB} \right) 
+ 4(n-1) \left( p_{A}^{AB} + 2p_{A}^{AB} + p_{A}^{AB} \right) 
+ 2(n-1) \left( p_{A}^{AB} + p_{A}^{AB} + p_{A}^{AB} + p_{A}^{AB} \right) 
+ 4(n-1)(n-2) \left( p_{A}^{AB} + p_{A}^{AB} \right) 
+ 2(a-1)n \left( p_{A}^{AB} + 2(n-1)p_{A}^{AB} \right) \left( p_{B} + p_{B}^{B} \right) 
+ 4(a-1)np_{A} \left( p_{B}^{AB} + p_{B}^{AB} + 2p_{B}^{AB} \right) 
+ 2(n-1) \left( p_{B}^{AB} + p_{B}^{AB} \right) 
+ 4(a-1)(a-2)n^2 \left( p_{B}^{AB} + p_{B}^{AB} \right) \].

The other terms in equation (9.1.7) have appeared in previous sections.

Thus,

\[ \delta(x_{1}x_{2}) = (p_{A}^{AB} - p_{A}^{AB})(p_{B}^{AB} - p_{B}^{AB}) \]

\[ + \frac{1}{a} \left[ p_{A}^{AB} - p_{B}^{AB} + 2p_{A}^{AB} - 2p_{A}^{AB} + 4p_{A}^{AB} \right] \]

\[ + \frac{2}{a(n-1)} \left( p_{A}^{AB} - p_{B}^{AB} \right)^2 \]

\[ + \frac{1}{a(n-1)} \left( p_{A}^{AB} + p_{B}^{AB} - 2p_{A}^{AB} \right) \left( p_{A}^{AB} + p_{B}^{AB} - 2p_{A}^{AB} \right) \]

\[ + \frac{1}{2an(n-1)} \left[ p_{A}^{AB} + 2p_{A}^{AB} + p_{A}^{AB} - 4(p_{A}^{AB} + p_{A}^{AB} + p_{A}^{AB} + p_{A}^{AB}) \right] 
+ 4p_{A}^{AB} \right) \]

\[ + \frac{1}{2a(a-1)n^2} \left( p_{A}^{AB} + p_{B}^{AB} - 2p_{A}^{AB} \right)^2 \].
\[ \delta(x_1^2) \]

\[
\delta(x_1^2) = \frac{1}{4a(a-1)n^2} \delta \left\{ \left[ \sum_{i} \left( \sum_{j,k} x_{ijk1} \right)^2 \right] \left( \sum_{i} \sum_{j,k} x_{ijk2} \right) \right.
\]

\[
- \frac{1}{2} \left[ \sum_{i} \left( \sum_{j,k} x_{ijk1} \right)^2 \right] \left[ \sum_{i} \left( \sum_{j,k} x_{ijk2} \right)^2 \right]
\]

\[
- \frac{1}{a} \left( \sum_{j,k} x_{ijk1} \right)^2 \left( \sum_{i} \sum_{j,k} x_{ijk2} \right)
\]

\[
+ \frac{1}{2a} \left( \sum_{j,k} x_{ijk1} \right)^2 \left[ \sum_{i} \left( \sum_{j,k} x_{ijk2} \right)^2 \right]
\]

\[
\left. \right\} - \frac{1}{n} \delta(x_1^2), \]

where,

\[
\delta \left( \left( \sum_{i} \sum_{j,k} x_{ijk1} \right)^2 \left( \sum_{i} \sum_{j,k} x_{ijk2} \right) \right) = \text{equation (9.1.6)} +
\]

\[
+ 8a(a-1)n^2 p_{\text{A}}^\text{B} p_{\text{A}} + p_{\text{A}} / p_{\text{B}} + 2(n-1)p_{\text{A}} / p_{\text{B}}
\]

\[
+ (a-2)np_{\text{A}} / p_{\text{B}} \}
\]

and where other terms have appeared in previous sections, so that,

\[
\delta(x_1^2) = \left( p_{\text{A}} / p_{\text{B}} - p_{\text{B}} / p_{\text{B}} \right)
\]

\[
+ \frac{1}{a} \left\{ p_{\text{A}} / p_{\text{A}} - p_{\text{B}} / p_{\text{A}} \right\} + (2p_{\text{B}} - p_{\text{A}}) / p_{\text{B}} - 2p_{\text{A}} / p_{\text{B}} - 2p_{\text{A}} / p_{\text{A}}
\]

\[
+ \frac{1}{an} \left( p_{\text{A}} / p_{\text{B}} - 2p_{\text{B}} / p_{\text{A}} + p_{\text{A}} / p_{\text{B}} - 2\left( p_{\text{A}} / p_{\text{A}} \right) - p_{\text{A}} / p_{\text{B}} \right) - p_{\text{A}} / p_{\text{B}} + p_{\text{A}} / p_{\text{B}} - 2p_{\text{A}} / p_{\text{B}}
\]

\[
+ 2\left( p_{\text{A}} / p_{\text{B}} - p_{\text{A}} / p_{\text{B}} \right) \}
\]
\[ \delta(\alpha_1, \alpha_2) = \delta \left( \frac{1}{4(a-1)n} \sum_{j,k} (\Sigma \Sigma x_{i,j,k})^2 \right) w_2 \]

\[ + \frac{1}{4a(a-1)n^2} (\Sigma \Sigma x_{i,j,k})^2 w_2 - \frac{w_1 w_2}{n} \]

\[ = (\frac{p_A^2}{A} - \frac{p_A^2}{B}) \frac{1}{2} \left( p_B + \frac{p_B^2}{B} - 2\frac{p_B^2}{B} \right) \]

\[ + \frac{1}{2a} \left[ p_{A|B} + p_{A|B} - 2p_{A|B} - 2p_{A|B} \right] \]

\[ + \frac{1}{2an} \left[ p_{A/B} + 2p_{A/B} + p_{A/B} - 2(p_{A/B} + p_{A/B}) \right] \]

\[ - \frac{1}{2an} \left[ p_{A|B} + p_{A|B} - 2p_{A|B} - 2p_{A|B} \right] \]

\[ - \frac{1}{2an(n-1)} \left[ p_{A/B} + 2p_{A/B} + p_{A/B} - 4(p_{A/B} + p_{A/B}) + 4p_{A/B} \right]. \]
9.2 Asymptotic Multivariate Normality of $\mathbf{s}$ as $a \to \infty$

The vector $\mathbf{s}$ in equation (3.1.7) is, as $a \to \infty$, asymptotically multivariate normal with mean vector $\bar{\mathbf{s}}(\mathbf{y})$ and variance-covariance matrix

$$
V = \frac{1}{a} \Lambda + \frac{1}{an} \Xi + \frac{1}{an(n-1)} \mathbf{Y}.
$$

Proof Outline:

The vector $\mathbf{s}$ may be written in the form,

$$
\mathbf{s} = \begin{bmatrix}
  z_1 \\
  z_2 \\
  w_1 \\
  w_2 \\
  r_1 \\
  r_2
\end{bmatrix} = \frac{1}{a} \sum_{i=1}^{a} \begin{bmatrix}
  \frac{a}{a-1} (\zeta_{11} - \bar{\zeta}_{11})^2 - \frac{1}{n} w_{11} \\
  \frac{a}{a-1} (\zeta_{12} - \bar{\zeta}_{12})^2 - \frac{1}{n} w_{12} \\
  w_{11} \\
  w_{12} \\
  \rho_{11} \\
  \rho_{12}
\end{bmatrix}
$$

where

$$
\zeta_{11} - \bar{\zeta}_{11} = \frac{1}{2n} (\sum_{j \neq k} x_{ijk1} - \frac{1}{a} \sum_{i \neq j \neq k} x_{ijk1})
$$

$$
\zeta_{12} - \bar{\zeta}_{12} = \frac{1}{2n} (\sum_{j \neq k} x_{ijk2} - \frac{1}{a} \sum_{i \neq j \neq k} x_{ijk2})
$$

$$
w_{11} = \frac{1}{4(n-1)} \left[ \sum_{j \neq k} \left( x_{ijk1} \right)^2 - \frac{1}{n} \left( \sum_{j \neq k} x_{ijk1} \right)^2 \right]
$$

$$
w_{12} = \frac{1}{4(n-1)} \left[ \sum_{j \neq k} \left( x_{ijk2} \right)^2 - \frac{1}{n} \left( \sum_{j \neq k} x_{ijk2} \right)^2 \right]
$$

$$
\rho_{11} = \frac{1}{n} \left[ \sum_{j \neq k} x_{ijk1} - \frac{1}{2} \left( \sum_{j \neq k} x_{ijk1} \right)^2 \right]
$$

$$
\rho_{12} = \frac{1}{n} \left[ \sum_{j \neq k} x_{ijk2} - \frac{1}{2} \left( \sum_{j \neq k} x_{ijk2} \right)^2 \right].
$$
Now the term

\[ \frac{1}{a} \sum_{i=1}^{a} \left( \frac{a}{a-1} \right) (\zeta_{i1} - \bar{\zeta}_{.1})^2 \]

is not an average of independent random variables, but,

\[ \frac{1}{a} \sum_{i=1}^{a} \left( \frac{a}{a-1} \right) (\zeta_{i1} - \bar{\zeta}_{.1})^2 = \frac{1}{a} \sum_{i=1}^{a} (\zeta_{i1} - p_A)^2 + O_p \left( \frac{1}{a-1} \right), \quad a \to \infty \]

where use has been made of the lemma given by Serfling (1980, page 72) and given earlier by Cramér (1946, page 365). So by Slutsky's Theorem [see, for example, Cramér (1946, page 255)] if the sequence of random variables

\[ \frac{1}{a} \sum_{i=1}^{a} (\zeta_{i1} - p_A)^2 \]

converges in distribution to some random variable then the sequence of random variables

\[ \frac{1}{a} \sum_{i=1}^{a} \left( \frac{a}{a-1} \right) (\zeta_{i1} - \bar{\zeta}_{.1})^2 \]

converges in distribution to the same random variable. A similar result holds for the term

\[ \frac{1}{a} \sum_{i=1}^{a} \left( \frac{a}{a-1} \right) (\zeta_{i2} - \bar{\zeta}_{.2})^2 \]

So the asymptotic distribution of \( s \) is the same as the asymptotic distribution of

\[ s^* = \frac{1}{a} \sum_{i=1}^{a} \begin{bmatrix} (\zeta_{i1} - p_A)^2 - \frac{1}{n} w_{i1} \\ (\zeta_{i2} - p_B)^2 - \frac{1}{n} w_{i2} \\ w_{i1} \\ w_{i2} \\ \rho_{i1} \\ \rho_{i2} \end{bmatrix} \]
Now the vectors
\[
\begin{bmatrix}
(C_{11} - p_A)^2 - \frac{1}{n} w_{11} \\
(C_{12} - p_B)^2 - \frac{1}{n} w_{12} \\
w_{11}
w_{12}
p_{11}
p_{12}
\end{bmatrix}
\]
are i.i.d. with mean vector $\bar{g}(\gamma)$ and variance-covariance matrix given by
\[
\Lambda + \frac{1}{n} \Xi + \frac{1}{n(n-1)} \Psi.
\]
So by the multivariate version of the Lindeberg-Lévy version of the central limit theorem [see, for example, Cramér (1946, page 316) or Serfling (1980, page 28)], $\bar{g}^*$ is asymptotically multivariate normal with mean vector $\bar{g}(\gamma)$ and variance-covariance matrix $\Psi$. 