Role of sexlinked genes in quantitative inheritance

Neeti Ranjan Bohidar

Iowa State University
ROLE OF SEXLINKED GENES IN QUANTITATIVE INHERITANCE

by

Neeti Ranjan Bohidar

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subjects: Statistics
Animal Breeding

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Signature was redacted for privacy.

Heads of Major Departments

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1960
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>TITLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II.</td>
<td>REVIEW OF LITERATURE</td>
<td>15</td>
</tr>
<tr>
<td>III.</td>
<td>COMPLETE SEXLINKED GENES</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>General Considerations</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Random Mating</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Inbreeding</td>
<td>69</td>
</tr>
<tr>
<td>IV.</td>
<td>PARTIALLY SEXLINKED GENES</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>General Considerations</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Quantitative Inheritance</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Random mating</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Inbreeding</td>
<td>112</td>
</tr>
<tr>
<td>V.</td>
<td>GENE EFFECTS DEPENDING ON SEX</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>General Considerations</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>Biological Evidence</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>Mathematical Consequences</td>
<td>138</td>
</tr>
<tr>
<td>VI.</td>
<td>THE ESTIMATION OF GENETIC PARAMETERS</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>General Considerations</td>
<td>147</td>
</tr>
<tr>
<td></td>
<td>Characteristics Expressed By Males</td>
<td>149</td>
</tr>
<tr>
<td></td>
<td>Characteristics Expressed By Females</td>
<td>152</td>
</tr>
<tr>
<td></td>
<td>Characteristics Expressed By Both Sexes</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td>Inbreeding Followed By Random Mating</td>
<td>168</td>
</tr>
<tr>
<td></td>
<td>Non-orthogonal Situation</td>
<td>169</td>
</tr>
<tr>
<td>VII.</td>
<td>THE MONTE CARLO APPROACH</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>General Consideration</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>Simulation Procedure</td>
<td>172</td>
</tr>
<tr>
<td>VIII.</td>
<td>THE PROBLEM AND THE PROCEDURE</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>The Problem</td>
<td>180</td>
</tr>
<tr>
<td></td>
<td>The Procedure</td>
<td>185</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>TITLE</td>
<td>Page</td>
</tr>
<tr>
<td>-----------------</td>
<td>---------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>IX.</td>
<td>RESULTS AND DISCUSSION</td>
<td>194</td>
</tr>
<tr>
<td></td>
<td>General Consideration</td>
<td>194</td>
</tr>
<tr>
<td></td>
<td>Results</td>
<td>196</td>
</tr>
<tr>
<td></td>
<td>Discussion</td>
<td>227</td>
</tr>
<tr>
<td>X.</td>
<td>SUMMARY</td>
<td>235</td>
</tr>
<tr>
<td>XI.</td>
<td>LITERATURE CITED</td>
<td>241</td>
</tr>
<tr>
<td>XII.</td>
<td>ACKNOWLEDGEMENTS</td>
<td>243</td>
</tr>
<tr>
<td>XIII.</td>
<td>APPENDIX</td>
<td>244</td>
</tr>
</tbody>
</table>
CHAPTER I. INTRODUCTION

A qualitative trait is characterized by its mutually exclusive categories and thus follows a discrete distribution such as the binomial or the multinomial distribution. Mendelian principles were derived from the behaviour of such characters. In contrast, there are many characteristics which show subtle variations from one individual to another and offer no basis for classification into a small number of mutually exclusive groups. The variation in these instances is expressed in terms of some continuous scale of measurement like inches, pounds or bushels and the characters are called quantitative characters. These characters can not be identified with one or two genes but rather depend upon the action and interaction of several genes. Examples of such characters are stature in man, milk production in dairy cattle, yield in corn. In economic species generally most, but not all, characters of great practical importance are quantitative. The distribution of such a trait approaches a continuous distribution as measurement techniques are refined.

From a strict logical point of view even qualitative characters can be quantified and expressed as quantitative characters because an ingenious worker can devise a continuous measurement which is related closely to a supposed discrete factor as for instance, color. Several theories have been extended to explain the mode of inheritance of a quantitative character of which the more prominent are East and Nilsson-Ehle's
multiple gene hypothesis, and Mather's polygenic hypothesis. It is not the scope of this work to go into details of this matter but one can visualise the possibility that the phenotypic effect usually ascribed to the operation of a particular multiple gene complex may be more or less duplicated by consequences of the action of a single major gene. The normal continuous range of stature in man, for instance, seems almost certainly to depend upon multiple gene effects. The stature of a very short person may usually be accounted for by the combined influences of many different genes. However, a single gene for dwarf stature may give rise to the same end result. There can be a possible role of major genes in the multiple gene system itself. Genes could serve a dual capacity by simultaneously affecting both quantitative and qualitative characters. One can envisage in relation to a particular quantitative character, a spectrum of gene effects, some gene pairs having small effects, others intermediate effects and probably a few with major, or in some cases catastrophic, effects. Most classical genetics consists of looking for genes with major effects and ignoring genes with minor effects on the same attribute.

The basis of qualitative genetics is the discovery of identifiable phenotypes and hence the evaluation of their frequencies with different kinds of mating. This approach is not possible with quantitative genetics since there is no information other than the numerical information on the phenotype. Therefore statistical methods must be used to obtain some understanding of the genetic determination. By statistical procedures one studies the central tendency, variability and covariability of populations and examines consequences of various breeding plans. Since the only
real proof of a genetic explanation is whether one can induce predictable changes by genetic means. The simplest such case is the use of the regression of offspring on parent to estimate the genetic advance per generation, which one then compares with the actual gain after some generations of artificial selection. Also the statistical method deals with partitioning the total genotypic variance into meaningful components to understand the various influences the population is subjected to. To exemplify this, consider a single locus situation, and let the genotypes AA, Aa and aa have genotypic values of 2, 1 and 0 respectively. The sole aim here is to show an example of partitioning the total genotypic variance into components of interest in a random mating population. The value $\sigma_G^2 = 2pq$ is the total genotypic variance and it can be subdivided into components using the family as the criterion for grouping the individuals. Table 1 gives the family structure under random mating. The average value of variance "within" families weighted by frequency of family structure in the population is $pq$. The variance between the mean values of families, also weighted by corresponding frequencies is found to be $pq$. In summary we have,

$$\text{variance within families} = pq = \frac{1}{2} \sigma_G^2$$

$$\text{variance between families} = pq = \frac{1}{2} \sigma_G^2$$

$$\text{total variance in the population} = 2pq = \sigma_G^2.$$
Table 1. Partition of genotypic variance in a random mating population

<table>
<thead>
<tr>
<th>Source of the families</th>
<th>Frequencies</th>
<th>Structure of the family</th>
<th>Family mean</th>
<th>Variance within family</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AA</td>
<td>Aa</td>
<td>aa</td>
</tr>
<tr>
<td>1. AA x AA</td>
<td>$p^4$</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2. AA x Aa</td>
<td>$4p^3q$</td>
<td>0.50</td>
<td>0.50</td>
<td>0.00</td>
</tr>
<tr>
<td>3. Aa x Aa</td>
<td>$4p^2q^2$</td>
<td>0.25</td>
<td>0.50</td>
<td>0.25</td>
</tr>
<tr>
<td>4. AA x aa</td>
<td>$2p^2q^2$</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5. Aa x aa</td>
<td>$4pq^3$</td>
<td>0.00</td>
<td>0.50</td>
<td>0.50</td>
</tr>
<tr>
<td>6. aa x aa</td>
<td>$q^4$</td>
<td>0.00</td>
<td>0.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Using similar tables we can evaluate the correlations between relatives. We shall, of course, be dealing with more general cases in this study. The general aim is to study some aspects of the inheritance of quantitative characters, though we should note that a qualitative factor can be represented as a quantitative factor, so the results are true for all Mendelian cases.

Before we embark on the main theme of this work, it is appropriate at this point to recapitulate briefly some of the elementary biological facts on which this mathematical study is based.

Living organisms are composed of functioning cells each of which possesses a governing structure appropriately called "nucleus". This nucleus contains the hereditary material called "chromosomes" and "genes". The inference of the existence of "genes" actually originated from Mendel's law of heredity. On the origin of the "gene" it is relevant
to paraphrase Johannsen (1911): The word gene is an applicable little word easily combined with others, which may be useful as an expression for the "unit factors", "elements" or "allelomorphs" in the gametes, demonstrated by modern Mendelian researches. Inheritance consists mostly the transmission of particulate units, the genes, located on larger units, the chromosomes. The existence of a specific gene is based on the recognition of allelic alternative forms which generally act upon the same aspect of the phenotype, but in different degree. Each gene is a physical entity which (i) is duplicated and passed intact from generation to generations (ii) has a specific function in the production of the phenotype and (iii) can mutate to another entity satisfying (i) and (ii). It should be noted that there may exist also in the cytoplasm particulate units which have all the properties of genes except that they are not attached to chromosomes and hence do not show Mendelian segregation. In this study the term "gene" will refer only to chromosomal entities which segregate according to the laws of Mendel.

The chromosomal theory of sex-determination will be the next consideration. Chromosomes usually appear in pairs called homologous chromosomes. In many animals and some plants, males and females are found to differ not in the presence or absence of one whole chromosome, but in the presence in one sex of a chromosome that is unlike its mate (homologue) and unlike any chromosome in the opposite sex. In the Drosophila species the female has four pairs of chromosomes, the members of each pair being alike. In the male there is only one of the straight rod-like chromosomes, the place of the other member of this
pair being taken by a rod with a hook-shaped or bent end. The rod-like member is called the sex-chromosome or X-chromosome and the unlike member is called the Y-chromosome. The other chromosomes are called autosomes. Sex-linked genes are genes which are located in X-chromosome. For example, color blindness in human beings is a sex-linked character, and the following diagram shows the mode of inheritance and the determination of sex at the same time.

![Diagram showing the inheritance and determination of sex with X and Y chromosomes.]

The diagram above shows that a carrier mother and a color-blind father will have half the daughters (1) carrier, half the daughters (2) color-blind, half the sons (3) normal and half the sons (4) color-blind. In general the XY sex is the male and the XX is the female; the opposite holds for birds and Lepidoptera. At this point a genetic fact will be introduced to which reference will be made in later chapters. Between the X and Y chromosomes and between the sexes which they determine, an important discontinuity arises, in that there is some restriction on recombination, that is, some limitation on crossing over. The genetic system has methods at its disposal for limiting crossing over. A genotypic localization of chiasmata or structural inversion of a segment of chromosome will suppress crossing over. Available evidences suggest that genotypic control is primarily responsible, structural
changes secondarily. Thus the XY sex always has lower crossing over than the XX sex, in the other chromosomes as well as in the sex-chromosomes. The extreme example of genotypic control is found in Drosophila, where crossing over does not occur in the XY sex, that is in the male. This genotypic restriction of recombination can be of great consequences in quantitative genetics.

It will be of interest to introduce at this point the concept of genic balance and its impact on determination of sex. A living organism is the resultant of thousands of genes acting in concert. Its phenotype results from the blended, balanced and synchronized activities of the products of all genes in the genome. On this point of view all genes interact with one another through their products and create a modifying system. In order for the modifying system to act harmoniously to produce a functional organism the system must be in balance and this is what is known as genic balance. To achieve a balanced system, it is, in general, not only necessary to have a full complement of genes but a complement of genes present in the proper dosage relations. The concept of genic balance receives its support from all experiments which show the dependence of a specific character upon more than a single gene, but its impact on the determination of sex is rather significant and the Drosophila species offers the best example in this respect. In this species, as in nearly all organisms with sexual dimorphism, sex is determined by the chromosomes. There is a mechanism which acts to produce an approximately equal distribution of sexes, and in Drosophila it appears to be associated with a balance between the genes on the autosomes versus those on the X-chromosome.
Bridges (1939) has described the effects on sex of changing the normal chromosome complement of Drosophila melanogaster, which consists of (2X's and 2 pairs of each autosomal type in the female, and XY with 2 pairs of each autosomal type in the male. Some of the results are given in Table 2.

Table 2. Determination of sex in Drosophila melanogaster

<table>
<thead>
<tr>
<th>Sex type</th>
<th>X:A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Super male</td>
<td>1:3</td>
</tr>
<tr>
<td>2. Male</td>
<td>2:4</td>
</tr>
<tr>
<td>3. Male</td>
<td>1:2</td>
</tr>
<tr>
<td>4. Intersex</td>
<td>2:3</td>
</tr>
<tr>
<td>5. Intersex</td>
<td>3:4</td>
</tr>
<tr>
<td>6. Female</td>
<td>2:2</td>
</tr>
<tr>
<td>7. Female</td>
<td>3:3</td>
</tr>
<tr>
<td>8. Female</td>
<td>4:4</td>
</tr>
<tr>
<td>9. Super female</td>
<td>3:2</td>
</tr>
</tbody>
</table>

It will be seen from the table that females are produced whenever the ratio of X-chromosomes to autosomes in each of the homologous groups is 1:1. Exceptional females, the so-called "super females" result from a ratio of 3X:2A. If now the number of X-chromosome is reduced relative to the number of autosomes, there results a normal male e.g. IX:2A or 2X:4A, and then with further reduction the super male IX:3A. Individuals with both male and female secondary sex
characteristics, the so-called "intersexes", arise as a result of ratios intermediate between IX:2A and IX:1A. Thus 3X:4A and 2X:3A individuals are intersexes. The Y-chromosome normally present in the male apparently has no role to play in sex determination, because XXY:2A individuals are typical females and XO:2A individuals are definitely male in appearance although sterile. In this study the individuals under consideration are assumed to be "normal" and to have achieved genic balance. The males possess one X-chromosome and females have two homologous X-chromosomes. There is "normal" autosomal representation in both sexes.

So far we have considered broadly the internal "environment" influencing the expression of the phenotype. Quantitative characters are subject to considerable phenotypic modification by external environment, for instance, milk production by dairy cattle varies from farm to farm and year to year, and with different systems of management and so on. The essential difference is that the influences of external environment on the phenotype are not inherited unless they cause genetic material to mutate. The "normal" phenotype is that obtained under "normal" environmental conditions under the control of a "normal" genotype. The extent of phenotypic change is dependent, of course, upon two factors, the degree of environmental change and the response of the organism. The relationship can be expressed as follows,

\[ P = G + E + GE \]

where \( P \) is the phenotypic response of the organism.
G is the genotypic contribution
E is the environmental contribution
GE is the contribution due to interaction between environment and genotype.

This simple relationship may not be true in actuality, and one has to develop experiments to understand the true relationships among the forces. To make the genotypic contribution approach the phenotypic response in the model we have to minimize variability due to E and GE. This can be partially achieved by controlled experimentation, or imposing statistical adjustments or by both. In this study we will be dealing with genotypic values themselves and the environmental influence is considered to be negligible. The theoretical results on variances and covariances will remain true in the presence of strictly random environmental effects and interactions, providing variances are increased by the amount of variance of this origin.

In this discussion we will be concerned with the contribution of sex-linked genes to quantitative characteristics. These characteristics are influenced by the effects of sexlinked genes, autosomal genes and the interactions of these effects. The representation of sexlinked genes in an organism may be modest, but the magnitude of their actions can be large particularly when a great number of autosomal genes may interact with the sexlinked genes. It has long been established that a significant portion of the genetic variability is caused by sexlinkage. The reduction of recombinations in the XY sex is itself one of the brilliant examples of the impact of sexlinked genes on the nature of inheritance.
Partially sexlinked genes will also be considered in this study. The definitions, mode of transmission, behaviour and so on will be treated separately in Chapter IV.

In this study two methods of approach are followed viz. the mathematical approach and the Monte Carlo approach. The mathematical approach gives algebraic solutions to problems under suitable restrictions and conditions. Initial simplification of real problems is a necessity to enable the mathematical approach to be at all effective. A mathematical solution to an idealized problem has, however, the advantage that the evaluation of a particular case of interest is achieved by assigning appropriate numerical values to the algebraic elements. Deductive inferences based on the results so obtained are valid only under the assumptions and the conditions defined, and there is always considerable subjective judgement involved in deciding whether the mathematically limited problem has relevance to the actual scientific problem which led to its formulation.

The Monte Carlo method, on the other hand, is an entirely different approach. It is based on simulation of stochastic processes by the use of random or pseudo-random numbers. The simulation operation involves hundreds of thousands of arithmetical steps. The algebraic aspects of life-processes, e.g. biological reproduction, segregation, evaluation of phenotypic values, selection, etc. are simulated and a complete experiment with a well-defined objective is conducted in a high speed computer and direct numerical results are obtained. Consider, for example, a
problem in population dynamics: Suppose one desires to study the approach to equilibrium of a population mating at random. In a single locus situation we have,

\[ n_1 \text{ AA} + n_2 \text{ Aa} + n_3 \text{ aa} \longrightarrow p^2 \text{ AA} + 2pq \text{ Aa} + q^2 \text{ aa} \]

after one generation of random mating. This is a simple algebraic consequence for a single locus with two alleles. The approach to equilibrium for two loci with an arbitrary number of alleles with or without linkage involves some notational and algebraic difficulties which are manageable. But if one attempts to work with more loci with different linkage relationships among the linked loci, the mathematics become cumbersome, complex and indeed unmanageable. The Monte Carlo method offers a satisfactory solution to such problems. An arbitrary population composition is conceived and by passing the individuals in this population through a reproduction subroutine one can generate a large population, the size of which is determined by the number of genetic parameters involved. The frequency of each genotype which is influenced by the linkage relationships and the random choice of mates is accurately recorded for every generation. This sequence of events is repeated in the machine and the actual point of asymptotic convergence is determined numerically. The mathematical approach, of course, has to a certain extent, the capacity of handling the whole realm of possibilities at one time and thus offers deeper insight into the problem and this is one of the reasons why this method is complementary and antecedent to the Monte Carlo approach. At the moment both have their limitations. But certainly one can recognize the fact that quantitative genetics (or population genetics or biometric genetics) is passing through a period of
transition. With the rapid growth of high speed digital computers and the
demand on the part of the experimenter to seek facts more close to
reality, the Monte Carlo approach will gain more popularity and will
supplement other methods to an increasing extent, if only because
numerical simulation can be done with little mathematical manipulation
for problems which are quite intractable to the highest mathe-
matical abilities available in the world. This would not, of course, be a
factor unless there were developed machines capable of doing elementary
numerical operations at fantastic speeds and of following a sequence of
related numerical operation at similar speeds. The numerical work
described later in this thesis would require an amount of work with ordi-
nary desk computers which is measurable in terms of man-years, as
compared to a few hours on the digital computer used. The algebraic
representation of selection with more than one locus, generation after
generation, in a finite population is an example where one can easily see
that algebraic treatment is impossible with currently available mathe-
matical tools, but numerical simulation is relatively simple to develop.

Since this work is produced in the transitional period, both the
approaches have been taken advantage of to solve different aspects of the
same problem. The partition of total genotypic variances into its appro-
priate constituents and the construction of the structure of covariances
between individuals have been derived algebraically. More complex
problems like the effect of sexlinkage and autosomal linkage on the
efficiency of artificial selection have been examined by taking recourse
to the Monte Carlo method. The latter method will be described in de-
tail in later chapter of this thesis.

After a review of literature in the ensuing chapter we will examine the complete sexlinked system in Chapter III and the partially sexlinked system in Chapter IV. Some deductions for the case when gene effects depend on sex will be presented in Chapter IV. In Chapter V we will examine estimation problems.

In the last five chapters of this thesis, the Monte Carlo approach to problems in general will be introduced and the procedure adopted in this study will be explained in detail. Results, discussion and summary will form the last part of the presentation.
CHAPTER II. REVIEW OF LITERATURE

The facts concerning sex determination were discovered early in the current century by Wilson, Sutton, Stevens, Montgomery and others. Many questions were, however, left unsettled. Morgan's (1910) investigations of sex-linked inheritance showed that many sex-linked genes are carried in the X-chromosome, whereas Y behaved as though it were "empty" of genes. The critical evidence was supplied by Bridges (1913) in a classical paper characteristically entitled "Non-disjunction as proof of the chromosome theory of heredity". Since then several workers have been involved in solving various aspects of the inheritance of sex-linked genes. It is impossible to do justice to their great contributions in this short survey of history.

In the biometrical area the investigation of genotypic correlations between relatives under the assumption of autosomal transmission started as early as 1900. This problem was examined by Pearson (1904), Yule (1904), Snow (1910), Brownlee (1910), Weinberg (1908, 1910), Jennings (1916, 1917). Fisher's 1918 and Wright's 1921 papers are the classical papers in this area. Fisher (1918) obtained many ancestral and collateral correlations under conditions of random mating, with the presence of dominance, and then of epistasis among pairs of loci. Some cases were worked out for two alleles at a locus and a few for many alleles at a locus. He also considered a population in which mating was assortative and gene effects were small and equilibrium existed. Wright (1921) obtained correlations among relatives for very many types of relationship in population under random or assortive mating or in which inbreeding was
occurring providing the effects of genes were additive, i.e., there was no dominance or epistacy. Robbins (1918), in the area of population dynamics, examined the consequences of the initial genotypic frequencies for both male and female after $n$ generations of random mating under the assumption of sexlinked transmission. He extended this study to the case of assortive mating, brother-sister mating, parent offspring mating and to some other special cases of interest. Hogben (1932) considered the filial and fraternal correlations arising from the sexlinked transmission for the case involving a single gene substitution in a system of random mating. He calculated the numerical values of the correlation coefficients of few special cases under the assumption of intermediate heterozygote, complete dominance and equal frequencies of allelomorphic genes. Charles (1933) extended the study of Hogben (1932) to the case of collateral and ancestral correlations for sexlinked transmission irrespective of sex. In this study he considered the correlation between relatives without regard to the sex of the individuals in each pair or to the sex of their common ancestor and concluded that sib correlations might be higher than parent offspring correlations when there was dominance with respect to autosomal genes or there was contribution of sexlinked genes irrespective of dominance effect. Wright (1933) evaluated the correlation between uniting gametes and between zygotes in terms of gene frequencies and finite population numbers of males and females to study the progress of heterozygosis. Wright (1935) dealt with simple correlations of relatives with a particular epistatic model. The formulae for inbreeding coefficient ($F$) and coefficient of relationship
under the assumption of sexlinked inheritance were derived by Wright (1951) using the path coefficient analysis. Lush (1948) presented a formula for the correlation between parent and offspring under the assumption of \( t \) autosomal genes and \( s \) sexlinked genes using the method of path coefficient. He obtained results for a special case with \( t = 19 \) and \( s = 1 \) and showed that sexlinked genes have slight effect on the biometric relationships. Inter and intra allelic interactions were not considered in his study. Malecot (1948) attacked the problem with probability arguments and derived the following formula in terms of additive and dominance components for the covariance of relatives under the assumptions of no linkage, no epistacy, no non-autosomal effects:

\[
\frac{1}{2} (\emptyset + \emptyset^1) \sigma_A^2 + (\emptyset \emptyset^1) \sigma_D^2
\]

Kempthorne (1954) solved the general problem of partitioning the total genotypic variance and obtaining the genotypic covariances between relatives with arbitrary number of loci, arbitrary numbers of alleles, arbitrary dominance and epistasis under the assumption of autosomal inheritance, by using the probability method of approach. This generalization has been developed for both random mating and inbreeding but does not take account of linkage. The probability approach has the virtue of achieving a considerable degree of algebraic generalization; for example, the covariance between two individuals \( X \) and \( Y \) is given by Kempthorne (1954) as

\[
\text{Cov}(X, Y) = \frac{1}{2} (\emptyset + \emptyset^1) \sigma_A^2 + (\emptyset \emptyset^1) \sigma_D^2 + \sum_{i,j}^{rs} (\emptyset + \emptyset^1)^i (\emptyset \emptyset^1)^j \left( \frac{1}{2} \right)^i \sigma_x^2
\]
where $\phi$ is the probability that they will receive the same gene by way of the Sire.

$\phi'$ is the probability that they will receive the same gene by way of the Dam.

$r$ is the number of additive effects involved in the interaction.

$s$ is the number of dominance deviation involved in the interaction.

$s^2_x$ is the component of variance for $x$, where $x$ includes 1 to $r$ additive effects and 1 to $s$ dominance deviation.

Cockerham (1956) stated a few simple effects of autosomal linkage on covariances among relatives.

The work of Mather (1949) which is very valuable for plant quantitative genetics, is not relevant to this study and is therefore not included.
CHAPTER III. COMPLETE SEXLINKED GENES

General Considerations

In this chapter the algebraic formulation will be presented in step-by-step detail to make apparent the assumptions and the procedure. In the ensuing section we will deal with the case of a completely sexlinked locus. We shall then develop multiple loci theory with simultaneous consideration of both sexlinked and autosomal loci. The chief objective in the following section will be to obtain the general partition of the total genotypic variances for males and females into their respective constituents and to construct the general structures of the covariances between (i) male and male, (ii) female and male and, (iii) female and female. The development refers to bisexual diploid interbreeding organisms with males as heterogametic sex. All populations are assumed to be infinite and the initial population is assumed to have achieved equilibrium under random mating. Selection, maternal effects, mutation and linkage are assumed absent for the derivation. A more complete treatment of the topic would include these factors also. Random mating and inbreeding will be dealt with separately. The genotypes are characterized by genotypic values which are by definition independent of environment. The genotypic value of a genotype is defined as the average phenotype realized over the relevant population of environments. The genotypic value will be partitioned additively into components called additive effects, dominance effects and epistatic effects. Probability arguments will be used to derive the algebraic expressions for the covariances among relatives.
Random Mating

Single locus theory

Partition of genotypic variance Consider a population of males with following genotypic array

\[
\sum_{i=1}^{t} p_i S_i
\]

and a population of females with the genotypic array

\[
\sum_{j=1}^{t} \sum_{k=1}^{t} p_j p_k S_j S_k \quad \text{or} \quad \sum_{j=1}^{t} \sum_{j \neq k}^{t} p_j p_k S_j S_k
\]

where \( S_i \) is the \( j^{th} \) sex-linked allele, \( j = 1, 2, \ldots, t \) and \( p_i \) is the frequency of the \( S_i \) allele.

It will be assumed that the expressions of the possible genes in the two sexes are represented by the following genotypic values, measured around the respective means:

- Genotypic value of male \( S_i \) = \( u_i \)
- Genotypic value of female \( S_j S_k \) = \( v_{jk} \)

Hence

\[
E(u_i) = \sum_g p_g u_g = 0
\]

\[
E(v_{jk}) = \sum_g \sum_h p_g p_h v_{gh} = 0
\]

Now define,

\[
V(u_i) = \sum_g p_g u_g^2 = \sigma^2_{G_m}
\]

\[
V(v_{jk}) = \sum_g \sum_h p_g p_h v_{gh}^2 = \sigma^2_{G_f}
\]

where \( E \) and \( V \) stand for expected value and variance, and \( \sigma^2_{G_m} \) and \( \sigma^2_{G_f} \)
\( \sigma^2_{G_f} \) indicate the total genotypic variance of male and female respectively. The subscripts \( m \) and \( f \) will always pertain to males and females respectively unless otherwise stated.

Since the male sex is heterogametic, no partition of variance is involved.

In the case of females, we can express the genotypic values by an additive model, so that for individual of genotype \( S_j S_k \)

\[
v_{jk} = \beta_j + \beta_k + \pi_{jk}
\]

where

- \( \beta_j \) is the effect of gene \( S_j \),
- \( (\beta_j + \beta_k) \) is the total of the additive gene effects for the individual
- \( \pi_{jk} \) is the dominance deviation.

The quantities \( \beta_j \) are defined by the application of least squares to the model \( v_{jk} = \beta_j + \beta_k \). We take the general multiple allelic situation and denote the average effect of the sex-linked genes \( S_1, S_2, \ldots, S_t \) by \( \beta_1, \beta_2, \ldots, \beta_t \). Then the values for \( \beta_1, \beta_2, \ldots, \beta_t \) are obtained by minimizing the expression,

\[
Q = \sum_{gh} \frac{1}{p_g p_h} (v_{gh} - \beta_g - \beta_h)^2
\]

The normal equations are

\[
\frac{d}{d\beta_r} \left( \frac{1}{2} Q \right) = 0 \quad r = 1, 2, \ldots, t
\]
or,

\[ 2 \sum_{g} p_g \beta_g + 2 \sum_{g} p_g \beta_r = 2 \sum_{g} p_g \nu_{rg} \]

or,

\[ 2 \beta_r (\sum_{g} p_g \beta_g) + 2 \beta_r = 2 \beta_r \sum_{g} p_g \nu_{rg} \]

Adding over \( t \), we find that \( \sum_{g} p_g \beta_g = 0 \)

So that

\[ \beta_r = \sum_{g} p_g \nu_{rg} \]

The additive genetic variance is defined to be the part of the total variance of the population which is removed by the fitting of the additive gene effects, one for each allele. It is the sum of products of the estimates and right-hand sides of the normal equations. The right-hand side of the \( r^{th} \) normal equation is

\[ 2 \beta_r \sum_{g} p_g \nu_{gr} = 2 \beta_r \]

The additive genetic variance is therefore given by the equation

\[ \sigma_{A_f}^2 = 2 \sum_{r} \beta_r^2, \quad (1) \]

and the dominance variance by the equation

\[ \sigma_{D_f}^2 = \sum_{gh} p_{gh} \nu_{gh}^2 - 2 \sum_{gh} \beta_r^2 = \sum_{gh} p_{gh} \pi_{gh}^2 \quad (2) \]

We are concerned with the effects and variances arising from the fact that the individuals considered are random members of the population, subject to different relationships. Any gene possessed by an individual has probability \( p_g \) of having the \( g^{th} \) allele. The average effect associated with this gene is random, because the individual and gene in
the individual are random and if we denote the gene by \( j \) and its effect by \( \beta_j \) then

\[
E(\beta_j) = \Sigma p_g \beta_g = 0
\]  

(3)

Similarly if the other gene is denoted by \( k \) with effect \( \beta_k \) then

\[
E(\beta_k) = \Sigma p_h \beta_h = 0
\]

and since \( E(v_{jk}) = 0 \) (because the \( v_{jk} \) are defined as deviations around the population),

\[
E(\pi_{jk}) = E(v_{jk} - \beta_j - \beta_k) = 0
\]  

(4)

Now we proceed to partition the genotypic variance of a female as follows,

\[
V(v_{jk}) = V(\beta_j + \beta_k + \pi_{jk}) = V(\beta_j) + V(\beta_k) + V(\pi_{jk})
\]

\[
+ 2 \text{Cov}(\beta_j, \beta_k) + 2 \text{Cov}(\beta_j, \pi_{jk})
\]

\[
+ 2 \text{Cov}(\beta_k, \pi_{jk}).
\]

The probability that a gene of an individual is the \( g \)th allele is \( p_g \), so it follows from (1) that,

\[
V(\beta_j) = \sum_{g=1}^{t} p_g \beta_g^2 = \frac{1}{2} \sigma^2_A
\]

The range of summation of dummy subscripts like \( g \) will be from 1 to \( t \) and we shall not state this every time.

Similarly,

\[
V(\beta_k) = \sum p_g \beta_g^2 = \frac{1}{2} \sigma^2_A
\]

\[
V(\pi_{jk}) = \sum p_g p_h \pi_{gh} = \sigma^2_D
\]
\[
\text{Cov}(\beta_j, \beta_k) = \sum_{gh} p_g p_h \beta_g \beta_h = (\sum p_g \beta_g)(\sum p_h \beta_h) = 0
\]

\[
\text{Cov}(\beta_j, \pi_{jk}) = \sum_{gh} p_g p_h (v_{gh} - \beta_g - \beta_h) = 0
\]

Thus we have achieved the partition of both the sexes

\[
\sigma^2_{G_m} = \sigma^2_A_m
\]

\[
\sigma^2_{G_f} = \sigma^2_A_f + \sigma^2_D_f
\]

Covariance between relatives

Before we attempt to get a general formula for the single locus case, it will be appropriate to examine some of the special cases to get an idea of the subsequent developments.

Special case

Unilineal

Genotypic covariance of Sire and son

\[
\begin{align*}
\text{Sire} & \quad \text{Dam} \\
S_i & \quad S_j S_k \\
\text{Son} & \quad \frac{1}{2} S_j + \frac{1}{2} S_k
\end{align*}
\]

\[
\text{Cov}(\text{Sire, son}) = \text{Cov}(u_i, \frac{1}{2} u_j + \frac{1}{2} u_k)
\]

\[
= \frac{1}{2} E(u_i u_j) + \frac{1}{2} E(u_i u_k)
\]

\[
= 0
\]

In the autosomal case this would be \( \frac{1}{2} \sigma^2_A \).
Genotypic covariance of sire and daughter

\[ \text{Sire} \quad S_i \quad \text{Daughter} \quad S_i \left( \frac{1}{2} S_j + \frac{1}{2} S_k \right) \]

\[ \text{Cov}(\text{Sire, Daughter}) = \text{Cov} \left( \frac{1}{2} v_{ij} \right) \]

\[ = \frac{1}{2} \text{E} (u_i v_{ij}) + \frac{1}{2} \text{E} (u_i v_{ik}) \]

\[ = \frac{1}{2} \sum p_i u_i \beta_i + \frac{1}{2} \sum p_i u_i \beta_i \]

\[ = \sum p_i u_i \beta_i \]

\[ = \frac{1}{2} \text{CA}_p \]

which is the average value of the product of the effect of a gene in the males and the average effect of the gene in the females.

In the autosomal case this would be \( \frac{1}{2} \sigma^2_A \).

Genotypic covariance of full brothers

\[ \text{Sire} \quad S_k \quad \text{Dam} \quad S_i S_j \]

\[ \text{Sons} \quad \frac{1}{2} S_i + \frac{1}{2} S_j \quad \frac{1}{2} S_i + \frac{1}{2} S_j \]
Cov (full brothers) = \text{Cov} \left( \frac{1}{2} u_i + \frac{1}{2} u_j, \frac{1}{2} u_i + \frac{1}{2} u_j \right) \\
= \frac{1}{2} \sum p_i u_i^2 \\
= \frac{1}{2} \sigma^2_A \\
In the autosomal case this would be $\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_D$

\underline{Bilineal cases}

**Genotypic covariance of sire and maternal grand daughter**

\begin{center}
\begin{tikzpicture}
  \node (sire) at (0,0) {Sire $S_i$};
  \node (dam) at (1,0) {Dam $S_j S_k$};
  \node (daug) at (0.5,0.5) {Daughter $S_i \left( \frac{1}{2} S_j + \frac{1}{2} S_k \right)$};
  \node (grand_daug) at (0.5,1.5) {Grand daughter $S_m \left( \frac{1}{2} S_i + \frac{1}{4} S_j + \frac{1}{4} S_k \right)$};
  \node (sm) at (0.5,2.5) {Sm (male)};

  \draw (sire) -- (daug);
  \draw (dam) -- (daug);
  \draw (daug) -- (grand_daug);
  \draw (daug) -- (sm);
\end{tikzpicture}
\end{center}

\begin{align*}
\text{Cov (Sire, maternal grand daughter)} &= \text{Cov} \left( u_i, \frac{1}{2} v_{im} + \frac{1}{4} v_{jm} + \frac{1}{4} v_{mk} \right) \\
&= \frac{1}{2} \sum p_i u_i \beta_i \\
&= \frac{1}{4} C_A s
\end{align*}

In the autosomal case it would be $\left( \frac{1}{2} \right)^2 \sigma^2_A$
Genotypic covariance of dam and paternal grand daughter

\[ \text{Sire} \quad \text{Son} \quad \text{Dam} \]
\[ S_i \quad \frac{1}{2} S_j + \frac{1}{2} S_k \quad S_j S_k \]
\[ \text{Grand daughter} \quad \left( \frac{1}{2} S_m + \frac{1}{2} S_n \right) \left( \frac{1}{2} S_j + \frac{1}{2} S_k \right) \]

\[ \text{Cov(Dam, paternal grand daughter)} = \text{Cov}(v_{jk'}, \frac{1}{4} v_m + \frac{1}{4} v_m + \frac{1}{4} v_n + \frac{1}{4} v_n) \]
\[ = \frac{1}{4} \sum_j p_j \beta_j^2 + \frac{1}{4} \sum_k p_k \beta_k^2 + \frac{1}{4} \sum_j p_j \beta_j^2 \]
\[ + \frac{1}{4} \sum_k p_k \beta_k^2 \]
\[ = \frac{1}{2} \sigma_f^2 \]

In the autosomal case it would be \( \left( \frac{1}{2} \right)^2 \sigma_A^2 \)

**General cases**

**Cov(male, male)**

To obtain the covariance of two males, in general, we shall have to find the expectation of the product of two effects say, \( u_i \) and \( u_j \) one for each male. Now there are two possibilities, either the corresponding genes are identical by descent, written briefly as \( i = j \) or they are not
identical by descent written as \( i \neq j \), so then,

\[
\text{Cov}(X_m, Y_m) = E(u_i^m u_j^m)
= P(i = j) E u_i^2 + P(i \neq j) E u_i u_j
= P(i = j) \sigma_A^2
= \nu_{XY} \sigma_A^2 \tag{5}
\]

because if \( i \) and \( j \) are not identical by descent they are statistically independent so that

\[
E(u_i u_j) = (E u_i)(E u_j) = 0.
\]

The quantity \( \nu_{XY} \) is defined as the probability that a random sexlinked gene in male \( X \) is identical by descent with a random sexlinked gene in male \( Y \) and \( \sigma_A^2 \) is the variance due to additive effects of sexlinked genes in a male.

\[
\text{Cov(female, female)}
\]

The procedure of deriving the covariance between two members of the homogametic sex is more elaborate than that for the heterogametic sex. The expectation of the product of the genotypic values of the two members is decomposed into the expectation of the nine terms, which, in this case, fall into three categories namely, terms involving (i) two additive effects (ii) one additive effect and one dominance deviation and (iii) two dominance deviations. For each category we construct a table of mutually exclusive possibilities and evaluate their conditional probabilities of genes being identical by descent. The sum of the probabilities over all terms constitute the algebraic expression of the coefficient of the variance.
components in the covariance under consideration. Only the relevant possibilities have been recorded in the Table 3 below; it will be of interest to note that thirteen such possibilities exists for the third category. Symbolically the scheme is represented as follows:

$$\text{Cov}(X_f, Y_f) = E(v_{ij} v_{kl})$$

$$= E(\beta_i + \beta_j + \pi_{ij})(\beta_k + \beta_l + \pi_{kl})$$

$$= E(\beta_i \beta_k) + E(\beta_i \beta_l) + \ldots$$

$$+ E(\beta_j \beta_k) + \ldots$$

$$+ E(\beta_i \pi_{kl}) + \ldots + E(\pi_{ij} \pi_{kl}).$$

(i) $$E(\beta_i \beta_k) = \sum_r \beta_r \beta_r \sum_p(i = k) \beta_p + \sum_p(i \neq k) \sum_k \beta_k$$

$$= P(i = k) \sum_r \beta_r^2$$

$$= P(i = k) \frac{1}{2} \sigma_A^2_f$$

Now combining all the expectations of the terms involving the additive effects, we have

$$\sum P(i = k) + P(i \neq 1) + P(j = k) + P(j = 1) \sum \frac{1}{2} \sigma_A^2_f$$

$$= 2 X_Y \sigma_A^2_f$$

(ii) Now consider terms like $$E(\beta_i \pi_{kl}).$$
Table 3. Mutually exclusive possibilities for (ii)

<table>
<thead>
<tr>
<th>k = 1</th>
<th>k ≠ 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>i = k</td>
<td>i ≠ k</td>
</tr>
</tbody>
</table>

i = 1  (1) possible  (2) possible
i ≠ 1  (3) impossible  (4) possible

We have three mutually exclusive possibilities as indicated in the table and the evaluation of their probabilities are as follows:

(1) \( P(k = 1, i = 1, i = k) = 0 \), since this can only arise in inbreeding.

(2) \( P(k ≠ 1, i ≠ k, i = 1) \sum p_{ik} p_k \beta_i \pi_{ki} = \sum p_i \beta_i (\sum p_k \pi_{ik}) = 0 \)

(4) \( P(k ≠ 1, i ≠ k, i ≠ 1) \sum p_{ikl} p_k p_l \beta_i \pi_{kl} = \sum p_i \beta_i (\sum p_k p_l \pi_{kl}) = 0 \)

so \( E(\beta_i \pi_{kl}) = 0 \)

(iii) Lastly consider terms involving \( \pi_{ij} \pi_{kl} \)

Assume \( P(i = j) = P(k = 1) = 0 \), since there is no inbreeding. One set of the possibilities as regards pairs of the genes being identical or non-identical is presented in Table 4. We consider only the cases which have an asterisk mark since the expectations for the other possibilities are easily seen to be zero.
Table 4. Mutually exclusive possibilities for (iii)

<table>
<thead>
<tr>
<th>i ≠ j</th>
<th>i = k</th>
<th>i ≠ k</th>
<th>k ≠ 1</th>
<th>i = 1</th>
<th>i ≠ 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>j = k</td>
<td>impossible</td>
<td>possible</td>
<td>possible*</td>
<td>possible</td>
<td>possible</td>
</tr>
<tr>
<td>j ≠ k</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
</tr>
<tr>
<td>j = 1</td>
<td>possible*</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
</tr>
<tr>
<td>j ≠ 1</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
<td>possible</td>
</tr>
</tbody>
</table>

We have

\[ P(i ≠ j, i = k, j = 1) = \sum_{ij} p_i p_j \pi_{ij} \pi_{ij} \]

\[ = P(i ≠ j, i = k, j = 1) \sigma^2_{Df} \]

and

\[ P(k ≠ 1, i = 1, j = k) = \sum_{ij} p_i p_j \pi_{ij} \pi_{ij} \]

\[ = P(k ≠ 1, i = 1, j = k) \sigma^2_{Df} \]

The condition \( k ≠ 1 \) can be omitted, since we do not allow individuals to be inbred.

Hence,

\[ E(\pi_{ij}, \pi_{kl}) = (P(i = k, j = 1) + P(i = 1, j = k)) \sigma^2_{Df} \]

The result is therefore

\[ \text{Cov}(X_f, Y_f) = 2 r_{XY}^* \sigma^2_{A_f} + u_{XY}^* \sigma^2_{D_f} \] (6)

where \( r_{XY}^* \) is defined to be the probability that a random sexlinked gene from female \( X \) is identical by descent with a random sexlinked gene from female \( Y \), and \( u_{XY}^* \) is defined to be the probability that one of the
random genes of $X$ is identical by descent with a random gene of $Y$ and the other gene of $X$ is identical by descent with the other gene of $Y$.

Cov(female, male)

The derivation of the covariance between a member of the homogametic sex and a member of the heterogametic sex follows the procedures described above. There are only five relevant possibilities recorded in the Table 5. The new feature in the algebraic expression of the covariance is that it has a covariance component instead of variance component unlike the previous two cases.

In the following we have the symbolical representations of the sexes, the genotypes and their corresponding genotypic values.

<table>
<thead>
<tr>
<th>Sex</th>
<th>genotype</th>
<th>genotypic value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_m$</td>
<td>$S_k$</td>
<td>$u_k$</td>
</tr>
<tr>
<td>$Y_f$</td>
<td>$S_i S_j$</td>
<td>$v_{ij}$</td>
</tr>
</tbody>
</table>

$$\text{Cov}(Y_f, X_m) = E(v_{ij} u_k)$$

The possible cases are tabulated below.
Table 5. Mutually exclusive possibilities for (iii)

<table>
<thead>
<tr>
<th></th>
<th>( i = j )</th>
<th>( i \neq j )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( k = j )</td>
<td>possible</td>
<td>impossible</td>
</tr>
<tr>
<td>( k \neq j )</td>
<td>impossible</td>
<td>possible</td>
</tr>
</tbody>
</table>

The possibilities (1) and (2) can only arise due to inbreeding, and the evaluation of the probabilities of the rest of the possibilities is as follows,

\[
\begin{align*}
(3) & \quad P(i \neq j, k = j, k \neq i) \sum_k p_k u_k \beta_k \\
(4) & \quad P(k = i, k \neq j, i \neq j) \sum_k p_k u_k \beta_k \\
(5) & \quad P(k \neq i, k \neq j, i \neq j) \sum_{ijk} p_i p_j p_k u_k v_{ij} = 0
\end{align*}
\]

So we have,

\[
\text{Cov}(X_{m}, Y_{f}) = \sum_k p_k u_k \beta_k 
\]

\[
= w_{XY} \frac{1}{2} C_{A_s} 
\]

where \( w_{XY} \) is defined to be the probability that a random sexlinked gene in male \( X \) is identical by descent with a random sexlinked gene in female \( Y \), and \( C_{A_s} \) will symbolise the covariance between the additive effects of genes in the male and female sexes.

The coefficients of variance and covariance components for some special cases are presented in Table 6.
Table 6. The coefficients of the variance and covariance components under single sexlinked locus theory

<table>
<thead>
<tr>
<th>Relationship</th>
<th>$\sigma^2_{A_m}$</th>
<th>$\sigma^2_{A_f}$</th>
<th>$\sigma^2_{D_f}$</th>
<th>$C_{A_s}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sire and son</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sire and daughter</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>Dam and son</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>Dam and daughter</td>
<td>-</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Brothers</td>
<td>$\frac{1}{2}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sisters</td>
<td>-</td>
<td>$\frac{3}{4}$</td>
<td>$\frac{1}{2}$</td>
<td>-</td>
</tr>
<tr>
<td>Brother - sister</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>Paternal half brother</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Paternal half sister</td>
<td>-</td>
<td>$\frac{1}{2}$</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Paternal half brother - sister</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Maternal half brothers</td>
<td>$\frac{1}{2}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Maternal half sisters</td>
<td>-</td>
<td>$\frac{1}{4}$</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Maternal half brother - sister</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>Paternal uncle - nephew</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Paternal uncle - niece</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>Paternal aunt - nephew</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Paternal aunt - niece</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Maternal uncle - nephew</td>
<td>$\frac{1}{8}$</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Maternal uncle - niece</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$\frac{1}{8}$</td>
</tr>
</tbody>
</table>
Table 6. (Continued)

<table>
<thead>
<tr>
<th>Relationship</th>
<th>$\sigma^2_{A_m}$</th>
<th>$\sigma^2_{A_f}$</th>
<th>$\sigma^2_{D_f}$</th>
<th>$C_{A_s}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal aunt - nephew</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3/8</td>
</tr>
<tr>
<td>Maternal aunt - niece</td>
<td>-</td>
<td>$1/4$</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>Sire and maternal grand daughter</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$1/4$</td>
</tr>
<tr>
<td>Dam and paternal grand daughter</td>
<td>-</td>
<td>$1/2$</td>
<td>0</td>
<td>-</td>
</tr>
</tbody>
</table>

Multiple loci theory

Partition of genotypic variance A quantitative trait is an aggregate of actions and interactions of both autosomal genes and sexlinked genes. So a simultaneous consideration of both types of genes is appropriate in developing multiple loci theory. The derivation of the formulation is based on the following assumptions,

1. no linkage
2. no differential viability
3. no mutation, no maternal effects and no selection
4. arbitrary number of alleles and loci
5. arbitrary dominance and epistasis
6. males are heterogametic
7. population is infinite and at equilibrium under random mating
8. diploid, interbreeding bisexual population

Since we will be dealing with a number of genetic parameters, the derivation and the symbols will necessarily be complex. So at the outset,
a brief description of the procedure involved will be appropriate.

For the evaluation of the genotypic values, an arbitrary genotype is assigned to each member of the sex. A genotypic value represents the postulated "true" response envisaged to depend upon the actions and interactions of a finite number of entities, the genes. The "true" response can be decomposed into arbitrary quantities and when expressed in the form of an equation, we generate a population identity. Some meaningful interpretations can be assigned to these arbitrary quantities depending upon the field of interest. An example of the above procedure might be useful in clarifying the situation.

Consider $Y_{ij}$ as the "true" response which depended upon two factors, namely A and B. Then we can decompose $Y_{ij}$ as follows:

$$Y_{ij} = \bar{Y} + (Y_{i.} - \bar{Y}) + (Y_{.j} - \bar{Y}) + (Y_{ij} - Y_{i.} - Y_{.j} + \bar{Y})$$

(1) \hspace{1cm} (2) \hspace{1cm} (3) \hspace{1cm} (4)

Now it can be verified that the algebraic sum of the quantities in the right-hand side is equal to the left-hand side and this expression represents an algebraic identity. If we are dealing with population structures, then this is called population identity. We can assign statistical interpretation to these quantities as follows,

(1) as the mean
(2) as the main effect of factor A
(3) as the main effect of factor B
(4) as the interaction of factor A and B
So then we have

\[ Y_{ij} = \mu + a_i + \beta_j + (a\beta)_{ij} \]

In this case the decomposition is on the "true" response. The observed response can take the following structure suitable for statistical analysis,

\[ Y_{ijk} = Y_{ij} + \epsilon_{ijk} \]

where \( k \) is the subscript for replication and \( \epsilon \) stands for measurement error.

After assigning meaningful interpretation to each of the components in the algebraic identity of the genotypic value, we proceed to prove the orthogonality of each component and obtain the variances of each orthogonal component and hence achieve the partition of the total variance.

Before we take up the construction of a population identity for a genotypic value with arbitrary number of alleles, arbitrary number of loci, arbitrary dominance and arbitrary epistacy, it will be appropriate to go through briefly a simple case involving one autosomal locus and one sexlinked locus for each sex.

Let \( A_jA_kS_q \) represent the genotype of a male and \( A_jA_kS_rS_q \) the genotype of a female. \( A_jA_k \) belongs to a autosomal locus and \( S_q \) to a sexlinked locus. Let the same symbols be used for a genotypic values. Then we have in case of males

\[
A_jA_kS_q = \sum (\Sigma_{p}p_A_s)(\Sigma_{p}p_A_s)+(A_j - \Sigma_{p}p_A_s)(\Sigma_{p}p_A_s)+(A_k - \Sigma_{p}p_A_s)(\Sigma_{p}p_A_s)
\]

\[ + (A_j - \Sigma_{p}p_A_s)(A_k - \Sigma_{p}p_A_s) \sum p_oS_o + (S_q - \Sigma_{p}p_oS_o) \sum \]

(8)
and in case of female,
\[ A_j A_k S_q S_r = \mathcal{L}(\Sigma p_s A_s)(\Sigma p_s A_s) + (A_j - \Sigma p_s A_s)(\Sigma p_s A_s) + (A_k - \Sigma p_s A_s)(\Sigma p_s A_s) \]
\[ + (A_j - \Sigma p_s A_s)(\Sigma p_s A_s) + (A_k - \Sigma p_s A_s)(\Sigma p_s A_s) \]
\[ \mathcal{L}(\Sigma p_o S_o)(\Sigma p_o S_o) + (S_q - \Sigma p_o S_o)(\Sigma p_o S_o) \]
\[ + (S_r - \Sigma p_o S_o)(\Sigma p_o S_o) + (S_q - \Sigma p_o S_o)(S_r - \Sigma p_o S_o) \]
(9)

where \( p_s \) and \( p_o \) are the gene frequencies of the autosomal alleles and the sexlinked alleles respectively.

The right-hand sides of the algebraic identities \( \mathcal{L}(8) \) and \( \mathcal{L}(9) \) can be expanded and symbols like \( A_p A_q \) or \( S_p S_q \) can be replaced by their corresponding genotypic values and thus we obtain the following linear models.

For males we have,
\[ y_{jkq} = \mu + a_j + a_k + d_{jk} + \beta_q + (a\beta)_{jq} + (a\beta)_{kq} + (d\beta)_{jkq} \]
and for females,
\[ y_{jkqr} = \mu + a_j + a_k + d_{jk} + \beta_q + \beta_r + \pi_{qr} + (a\beta)_{jq} + (a\beta)_{jr} + (a\beta)_{kq} + (a\beta)_{kr} + (a\pi)_{jqr} + (a\pi)_{kqr} + (d\pi)_{jkqr} + (d\beta)_{jkq} + (d\beta)_{jkr} \]

where, \( a's \) and \( \beta's \) are autosomal and sexlinked additive gene effects, \( d_{jk} \) and \( \pi_{qr} \) are autosomal and sexlinked dominance deviations and terms involving parenthesis are corresponding interactions.

Now we present the case of male. Let the genotype of the male be
\[ \begin{array}{cccc}
 n_1 & A_j^c & A_k^c & n_2 \\
 \pi & j_c & k_c & \pi & S_q^a \\
c=1 & a=1 & & q_a
\end{array} \]
where the superscript \( c \) denotes the autosomal loci, \( c = 1, 2, \ldots, n_1 \),
the superscript \( a \) denotes the sexlinked loci, \( a = 1, 2, \ldots, n_2 \),

\[
A_j^c \text{ symbolizes the } j^{th} \text{ allele of the } c^{th} \text{ locus,}
\]

\[
S_{q_a}^a \text{ symbolizes the } q^{th} \text{ allele of the } a^{th} \text{ locus,}
\]

and \( \pi \) stands for consecutive arrangement and not for products.

Let the genotypic symbol represent the genotypic value. Then we have

the population identity:

\[
\pi \left( \sum_{c=1}^{n_1} A_j^c A_k^c \right) \pi \left( \sum_{a=1}^{n_2} S_{q_a}^a \right) = \pi \left( \sum_{c=1}^{n_1} \left( \sum_{c} A_j^c A_k^c \right) \left( \sum_{c} S_{q_a}^a \right) + \left( A_j^c - \sum_{c} A_j^c \right) \left( \sum_{c} S_{q_a}^a \right) \right)
\]

\[+ \left( \sum_{a=1}^{n_2} \left( \sum_{a} S_{q_a}^a \right) \right) \left( \sum_{a} S_{q_a}^a \right) + \left( A_k^c - \sum_{c} A_k^c \right) \left( \sum_{c} S_{q_a}^a \right) \left( \sum_{c} S_{q_a}^a \right) \right]
\]

To minimize the notational burden we shall have the following formal definitions, as follows,

\[
\sum_{c} A_j^c A_k^c = u^c
\]

\[
A_j^c - \sum_{c} A_j^c = \theta^c
\]

\[
\sum_{c} S_{q_a}^a = \rho^a
\]

then we have,

\[
A_j^c = \theta_j^c + u^c \tag{10}
\]
40

= 0^

+

.

Now s u b s t i t u t i n g (lO) a n d (11) i n t h e f o r m u l a f o r g e n o t y p i c v a l u e a n d
e x p a n d i n g t h e r e s u l t a n t e x p r e s s i o n we h a v e ,

*1
*2
*1
*2
C + u°) TT
7T A° A, C TT S a = TT (Q C + UC )(9
(0 & +
1
1 k
1
1
k
,
q
1
J
1
n
c-1 c
c a=l a
c=l
c
c
a= 1 q a

)

n
cc
a , „ „c
c
c1 c '
TT U u IT p + 2 0 . U ITU u i r p
c=l
a
c ^c
cT
a
, ^ nc
+ 2 9^
c c

c
c' c!
a
u tt u u tt p
c 1
a

+ 2 0. 0
TT U
c "'c
c c1

u

TT p
a

On C,!
c
U IT p C

+ 2 O'T U^ 0T u^ IT U^
c c ' ^c
^c'
c11
, ^

c
k

c

A
+ 2 0.

,
cc -

a

c1
c!
c'1 c"
a
UTTU U
TT Pr
k .1
c
c"
a

n
U 0,

c

, ^ nc
c a c! c'
c11 c»
a
+ 2 9. u Gk
u TT u
u
up
c c,i1 "jc.
c1
c"
a
^ ncr c1 Ac
c
c" cn
a
+ 20. u 9, u tt u
u
tt p
c c 1 ^c '
c
c"
a

+

c c1 c?„cn c11
cin cm
a
nc
2
TTU
U
TT p
Jcn
c c ' c " ^c
^c 1
cm
a

+

c nc! cTncn c"
cm cm
a
nc
2 9 1 U 0.
U 9,
U
TT U
U
TT p +
cc'c" c
c'
c"
c"'
a


\[ + \sum_{a=1}^{n_2} \theta^a_{q_a a} \pi \rho^a \pi u^c c \]

\[ + \sum_{a a'} \theta^a_{q_a a} \theta^{a'}_{q_{a'} a'} \pi \rho^{a'} \pi u^c c \]

\[ + \sum \theta^a_{q_a a} \theta^{a'}_{q_{a'} a'} \theta^{a''}_{q_{a''} a''} \pi \rho^{a''} \pi u^c c + \ldots \]

\[ + \sum \sum \theta^c_{q_c c} \theta^c_{q_{c'} c'} \theta^{c''}_{q_{c''} c''} \pi \rho^{c''} \pi u^c c \]

\[ + \sum \sum \theta^c_{q_c c} \theta^c_{q_{c'} c'} \theta^{c''}_{q_{c''} c''} \pi \rho^{c''} \pi u^c c \]

\[ + \sum \sum \theta^c_{q_c c} \theta^c_{q_{c'} c'} \theta^{c''}_{q_{c''} c''} \pi \rho^{c''} \pi u^c c \]

\[ + \sum \sum \theta^c_{q_c c} \theta^c_{q_{c'} c'} \theta^{c''}_{q_{c''} c''} \pi \rho^{c''} \pi u^c c \]

and so on, where \( c \neq c' \neq c'' \neq c''' \) and \( a \neq a' \neq a'' \).

Each of the quantities identified in the population model for males can be given the following genetic interpretations:

\[ \mu = \pi u^c c \pi \rho^a = \text{over all genotypic mean of the population.} \]
\[ a_{jc}^c = \theta_{jc}^c \pi_{jc} u^c \pi_{jc} u^c \pi_{ja} \]

= additive effect of autosomal gene \( A_{jc}^c \)

\[ d_{jk}^{c, c} = \theta_{jc}^c \theta_{kc}^c \pi_{jc} u^c \pi_{jc} u^c \pi_{ja} \]

= autosomal dominance deviation associated with gene \( A_{jc}^c \) and \( A_{kc}^c \) at the \( c^th \) locus.

\[ (a^c a^c)_{jc}^{k, c} = \theta_{jc}^c \theta_{kc}^c \pi_{jc} u^c \pi_{jc} u^c \pi_{ja} \]

= additive (autosomal) x additive (autosomal) deviation associated with genes \( A_{jc}^c \) and \( A_{kc}^c \)

\[ (a^c a^c)_{jc}^{k, c'} = \theta_{jc}^c \theta_{kc}^c \pi_{jc} u^c \pi_{jc} u^c \pi_{ja} \]

= additive (autosomal) x additive (autosomal) deviation associated with gene \( A_{jc}^c \) and \( A_{kc}^{c'} \) at the loci \( c \) and \( c' \).

\[ (a^a a^a)_{qa} = \phi_{qa}^a \pi_{qa} u^a \]

= additive effect of sexlinked gene \( S_{qa}^a \) at the \( a^th \) locus.

\[ (\beta^a \beta^a)_{qa}^{qa} = \phi_{qa}^a \phi_{qa}^{qa} \pi_{qa} u^a \]

= additive (sexlinked) x additive (sexlinked) deviation associated with \( S_{qa}^a \) and \( S_{qa}^{a} \) at the \( a^th \) and \( a'^th \) loci.

\[ (a^c a^a)_{jc}^{qa} = \theta_{jc}^c \theta_{qa}^a \pi_{ja} u^c \]

= additive (autosomal) x additive (sexlinked) deviation associated with \( A_{jc}^c \) gene at the \( c^th \) autosomal locus and
with \( S_{a}^{a} \) gene at the \( a^{th} \) sexlinked locus.

\[
(d_{j}^{c} \rho_{k}^{a})_{j_{c} k_{c} q_{a}}^{a} = \phi_{j_{c}}^{c} \phi_{k_{c}}^{c} \phi_{q_{a}}^{a} \pi_{a^{t}} \rho_{a^{t}} \pi_{u^{t}} u^{t} c^{t}
\]

= dominance (autosomal) x additive (sexlinked) deviation

associated with \( A_{j_{c}}^{c} \) and \( A_{k_{c}}^{c} \) at the \( c^{th} \) locus and

\( S_{a}^{a} \) at the sexlinked locus.

\[
(d_{c}^{r} d_{c}^{r})_{j_{c} k_{c} j_{c} k_{c} r_{c} r_{c}}^{a} = \phi_{j_{c}}^{c} \phi_{k_{c}}^{c} \phi_{j_{c}}^{r} \phi_{k_{c}}^{r} \pi_{c^{t}} u^{t} c^{t} \pi_{a^{t}} a^{t}
\]

= dominance (autosomal) x dominance (autosomal)

deviation associated with \( A_{j_{c}}^{c} \) and \( A_{k_{c}}^{c} \) at the \( c^{th} \)
locus and \( A_{j_{c}^{c}}^{c} \) and \( A_{k_{c}^{c}}^{c} \) at the \( c^{th} \) locus.

Now we can express explicitly the genotypic value of males in terms of
gene effects and interactions.

\[
\pi_{1} A_{j_{c}}^{c} A_{k_{c}}^{c} \pi_{S_{a}^{a}} = \mu + \Sigma a_{j_{c}}^{c} + \Sigma a_{k_{c}}^{c} \pi_{c=1}^{n_{1}} + \Sigma a_{j_{c}}^{c} \pi_{c=1}^{n_{1}} + \Sigma a_{k_{c}}^{c} \pi_{c=1}^{n_{1}}
\]

\[
+ \Sigma d_{j_{c}}^{c} \pi c=1^{n_{1}} + \Sigma \beta_{a}^{a} \pi c=1^{n_{1}} + \Sigma \beta_{a}^{a} \pi c=1^{n_{1}} + \Sigma \beta_{a}^{a} \pi c=1^{n_{1}} + \ldots
\]
The genotypic value of a male consists of following components.

The total number of terms are \(2n_1 + n_2\) and they are as follows,

- number of means = 1,
- number of additive autosomal gene effects = \(2n_1\),
- number of additive sexlinked gene effects = \(n_2\),
- number of dominance deviations = \(n_1\),
- number of additive (autosomal) x additive (autosomal) interaction deviations = \(2n_1(n_1 - 1)\),
- number of additive (sexlinked) x additive (sexlinked) interaction deviations = \(\frac{1}{2}n_2(n_2 - 1)\),
- number of additive (autosomal) x additive (sexlinked) interactions = \(2n_1n_2\),
- number of additive (autosomal) x dominance deviation interactions = \(2n_1(n_1 - 1)\),
- number of additive (sexlinked) x dominance deviation interactions = \(\frac{1}{2}n_2(n_2 - 1)\),

and so on.

Now we present the case of female. Let the genotype of the female be

\[
\pi_{c=1}^{n_1} A_{j_c}^c A_{k_c}^c \pi_{a=1}^{n_2} S_a^a S_r^a
\]

where the subscript \(c\) denotes the autosomal loci, \(c = 1, 2, \ldots, n_1\),
the superscript \(a\) denotes the sexlinked loci, \(a = 1, 2, \ldots, n_2\),

\(A_{j_c}^c\) symbolizes the \(j^{th}\) allele of the \(c^{th}\) locus,
$S_r^a$ symbolizes the $r^{th}$ allele of the $a^{th}$ locus,

and $\pi$ stands for consecutive arrangement and not for products.

Let the genotypic symbol represent the genotypic value. Then we have the population identity:

$$n_1 \pi A^c_k A^c_j \pi S_r^a S_r^a = \sum_{c=1}^{n_1} \left( \sum_{c' \neq c} p_{A_{c'}} A_{c'\bar{c}} \right) \left( \sum_{c' \neq c} p_{S_{c'}} S_{c'\bar{c}} \right) + \left( A^c_j - \sum_{c' \neq c} p_{A_{c'}} A_{c'\bar{c}} \right) \left( \sum_{c' \neq c} p_{S_{c'}} S_{c'\bar{c}} \right) + \left( A^c_k - \sum_{c' \neq c} p_{A_{c'}} A_{c'\bar{c}} \right) \left( \sum_{c' \neq c} p_{S_{c'}} S_{c'\bar{c}} \right)$$

Now we make some new formal definitions to minimize the notational burden as follows,

$$\sum_{c=1}^{n_1} p_{A_{c'}} A_{c'\bar{c}} = u^c \quad \sum_{c=1}^{n_1} p_{S_{c'}} S_{c'\bar{c}} = v^c \quad \sum_{c=1}^{n_1} p_{S_{c'}} S_{c'\bar{c}} = w^c \quad \sum_{c=1}^{n_1} p_{S_{c'}} S_{c'\bar{c}} = x^c \quad \sum_{c=1}^{n_1} p_{S_{c'}} S_{c'\bar{c}} = y^c$$
then we have

\[ A_{j_c}^c = \theta_{j_c}^c + u^c \]  
\[ J_{c} = \theta_{q_a}^a + \rho^a \]  

Now substituting (12) and (13) in the formula for genotypic value and expanding the resultant expression, we have

\[
\pi \sum_{c=1}^{n_1} A_{j_c}^c A_{k_c}^c \pi \sum_{a=1}^{n_2} S_{a}^a S_{r_a}^a = 
\]
\[
\sum_{c=1}^{n_1} \left( (u^c + \theta_{j_c}^c)(u^c + \theta_{k_c}^c) \right) \sum_{a=1}^{n_2} \left( (\rho^a + \theta_{q_a}^a)(\rho^a + \theta_{r_a}^a) \right)
\]

\[
= \pi \sum_{c=1}^{n_1} u^c \pi \sum_{a=1}^{n_2} \rho^a + \sum_{c=1}^{n_1} \sum_{c' = 1}^{n_1} \theta_{j_c}^c \theta_{j_{c'}}^c \pi \sum_{a=1}^{n_2} \rho^a \pi \sum_{a=1}^{n_2} \rho^a + \ldots 
\]

\[ + \sum_{c=1}^{n_1} \theta_{j_c}^c \theta_{j_{c'}}^c \pi \sum_{a=1}^{n_2} \rho^a \pi \sum_{a=1}^{n_2} \rho^a + \ldots 
\]

\[ + \sum_{c=1}^{n_1} \theta_{j_c}^c \theta_{j_{c'}}^c \theta_{j_{c''}}^c \pi \sum_{a=1}^{n_2} \rho^a \pi \sum_{a=1}^{n_2} \rho^a + \ldots 
\]

\[ + \sum_{c=1}^{n_1} \theta_{j_c}^c \theta_{j_{c'}}^c \theta_{j_{c''}}^c \theta_{j_{c'''}}^c \pi \sum_{a=1}^{n_2} \rho^a \pi \sum_{a=1}^{n_2} \rho^a + \ldots 
\]

\[ + \sum_{c=1}^{n_1} \theta_{j_c}^c \theta_{j_{c'}}^c \theta_{j_{c''}}^c \theta_{j_{c'''}}^c \theta_{j_{c''''}}^c \pi \sum_{a=1}^{n_2} \rho^a \pi \sum_{a=1}^{n_2} \rho^a + \ldots 
\]
\[ + \sum_{j_c} \theta_c^c \rho a \rho a' \pi u c' u' c' \]
\[ + \sum_{k_c} \theta_c^c \rho a \rho a' \pi u c' u' c' \]
\[ + \sum_{j_c} \theta_c^c \rho a \rho a' \pi u c' u' c' \]
\[ + \sum_{v_c} \theta_c^c \rho a \rho a' \pi u c' u' c' \]

+ ... and so on,

where \( c \neq c' \neq c'' \) and \( a \neq a' \neq a'' \)

Each of the quantities identified in the population model for females can be given the following genetic interpretations.

\[ \mu = \pi u c u \pi a a \]
\[ c=1 \]
\[ a=1 \]

= overall genotypic mean of the population.

\[ a_{c}^{j_c} = \theta_c^c \theta_c^c \pi c' u' c' \pi a a \]
\[ a_{c}^{j_c} = \theta_c^c \theta_c^c \pi c' u' c' \pi a a \]

= additive effect of autosomal gene \( A_{c}^{j_c} \)

\[ d_{j_c}^{k_c} = \theta_c^c \theta_c^c \pi c' u' c' \pi a a \]
\[ d_{j_c}^{k_c} = \theta_c^c \theta_c^c \pi c' u' c' \pi a a \]

= autosomal dominance deviation associated with genes \( A_{c}^{j_c} \) and \( A_{c}^{k_c} \) at the \( c \)th locus.
\[ (a^c a^c c^c c^c)_{c^c c^c} = \theta^c_{c^c} u^c c^c c^c c^c \pi^c_{c^c} u^c c^c c^c c^c \pi^c_a \rho^c_a \]

= additive (autosomal) x additive (autosomal)

deviation associated with genes \( A^c_{c^c} \) and \( A^c_{c^c} \) at the \( c^c \)th and \( c^c \)th loci.

\[ p^a_{q^a} = \theta^a_{q^a} \rho^a_{q^a} \pi^a_{q^a} a^a_{q^a} \pi^a_{q^a} u^c_{q^a} c^c \]

= additive effect of sexlinked gene \( S^a_{q^a} \).

\[ \pi^a_{q^a r^a} = \theta^a_{q^a} \pi^a_{q^a} a^a_{q^a} \rho^a_{q^a} \pi^a_{q^a} u^c_{q^a} c^c \]

= sexlinked dominance deviation associated with

sexlinked genes \( S^a_{q^a} \) and \( S^a_{r^a} \)

\[ (p^a p^a a^a)_{q^a a^a} = \theta^a_{q^a} \rho^a_{q^a} p^a_{q^a} a^a_{q^a} \pi^a_{q^a} p^a_{q^a} a^a_{q^a} \pi^a_{q^a} u^c_{q^a} c^c \]

= additive (sexlinked) x additive (sexlinked)

deviation associated with sexlinked genes \( S^a_{q^a} \)
at locus \( a \) and \( S^a_{q^a} \) gene at \( a^a \).

\[ (a^c p^a)_{j^c q^a} = \theta^c_{j^c} u^c a^c p^a_{j^c} \pi^c_{j^c} a^a_{j^c} \pi^c_{j^c} u^c_{j^c} c^c \]

= additive (autosomal) x additive (sexlinked)

deviation associated with gene \( A^c_{j^c} \) and \( S^a_{q^a} \)
at the loci \( c \) and \( a \) respectively.
\[ (d^c{p^a})_{j,k,c}q_a = \theta^c_j \theta^c_k \phi^a_c \rho^a \pi^a \rho^a \pi^a \pi^c u^c \]

= dominance (autosomal) x additive (sexlinked) deviation associated with gene \( A^c_j \) and \( A^c_k \) and \( S^a_q \).

\[ (a^c{p^a})_{j,c}q_a r_a = \theta^c_j \phi^a_c \rho^a \pi^a \rho^a \pi^a \pi^c u^c \]

= additive (autosomal) x dominance (sexlinked) deviation associated with genes \( A^c_j \) at locus \( c \)
and \( S^a_q \) at the \( a^{th} \) locus.

\[ (d^c{p^a})_{j,k,c}q_a r_a = \theta^c_j \theta^c_k \phi^a_c \rho^a \pi^a \rho^a \pi^a \pi^c u^c \]

= dominance (autosomal) x dominance (sexlinked) deviation associated with gene \( A^c_j \) and \( A^c_k \) at the \( c^{th} \) locus and \( S^a_q \) and \( S^a_r \) at the \( a^{th} \) locus.

and so on.

Now we can express explicitly the genotypic value of females in terms of gene effects and interactions.

\[
\begin{align*}
\pi_{n_1} A^c_j A^c_k & = \sigma_{n_1} \pi S^a_q S^a_r \\
\pi_{c=1} & = \mu + \sum_{c=1}^{n_1} a^c_j + \sum_{c=1}^{n_1} a^c_k + \sum_{c=1}^{n_1} d^c_j k^c_k \\
& + \sum_{a=1}^{n_2} \beta^c_q + \sum_{a=1}^{n_2} \beta^c_r + \sum_{a=1}^{n_2} \pi^a_q a^r_a \\
& + \sum_{a=1}^{n_2} \pi^a q^r_a
\end{align*}
\]
The genotypic value of a female consists of following components,

\[ 2(n_1 + n_2) \]

The total number of terms are \( 2(n_1 + n_2) \) and they are as follows

- number of means = 1,
- number of additive autosomal gene effects = \( 2n_1 \),
- number of additive sexlinked gene effects = \( 2n_2 \),
- number of additive (sexlinked) x additive (sexlinked) gene interaction = \( 2n_2(n_2 + 1) \),
- number of dominance deviations (autosomal) = \( n_1 \),
- number of dominance deviations (sexlinked) = \( n_2 \),
- number of additive (autosomal) x additive (sexlinked) gene interaction = \( 4n_1n_2 \),
- number of additive (autosomal) x dominance (sexlinked) = \( 2n_1(n_2 - 1) \),
- number of (sexlinked) x dominance (autosomal) = \( 2n_2(n_1 - 1) \),

and so on.

Now we proceed to prove the orthogonality of each component in the general models. Let the probability that \( A_{jc}^c \) is a particular gene be \( p_r^c \), where \( c \) ranges from 1 to \( n_1 \) and let the probability that \( S_{qa}^a \) is a particular gene be \( p_t^a \), where \( a \) ranges from 1 to \( n_2 \). Symbolically we can represent these relations as follows,
Since the individual under consideration is a random member of the population, the genes possessed by the individual are random with probabilities noted above. In the following we shall derive the expectations of the effects and interactions of the random genes. Since the derivations are similar for the cases of male and female, they will not be treated separately.

\[ E(a_j^c) = \sum p_j^c a_j^c \]

\[ = \sum p_j^c (A_j^c - \sum p_s^c A_s^c) \sum p_s^c A_s^c \pi \nu^{c'} \nu^{c'} \pi \rho^a \]

\[ = (\mu - \mu) \pi \nu^{c'} \nu^{c'} \pi \rho^a \]

\[ = 0 \]

where \( \sum p_j^c = 1 \)

This shows that the average value of an additive effect of a random autosomal gene over the whole population is zero.
\[
E(d_{j\rightarrow k}) = \sum \sum p^c_g p^c_h q^c_g q^c_h \pi \ u^{c^1} u^{c^1} \pi \ \rho^a_{g \rightarrow h}
\]

\[
= \sum \sum p^c_g p^c_h \sum (A^c_g - \sum p^c_{s_g} A^c_s) (A^c_h - \sum p^c_{s_h} A^c_s) \pi \ u^{c^1} u^{c^1} \pi \ \rho^a_{g \rightarrow h}
\]

\[
= \sum \sum p^c_g p^c_h A^c_g A^c_h - \sum \sum p^c_g p^c_h A^c_g A^c_h
\]

\[
- \sum \sum p^c_s p^c_h A^c_s A^c_h + \sum \sum p^c_s p^c_h A^c_s A^c_h
\]

\[
= (\mu - \mu - \mu + \mu) \pi \ u^{c^1} u^{c^1} \pi \ \rho^a_{g \rightarrow h}
\]

\[
= 0
\]

where $\sum p^c_g = 1$ and $\sum p^c_h = 1$

This shows that the average value of a dominance deviation arising from two autosomal genes in a locus over the whole population is zero.

\[
E(\rho^a_{q\rightarrow a}) = \sum p^a_g \rho^a_{g \rightarrow a}
\]

\[
= \sum p^a_g \rho^a_{g \rightarrow a} \pi \ a^f \pi \ u^c u^c
\]

\[
= \sum p^a_g (S^a_g - \sum p^a_{s_g} S^a_s) \pi \ a^f \pi \ u^c u^c
\]

\[
= \sum \sum p^a_g p^a_{s_g} S^a_g S^a_s \pi \ a^f \pi \ u^c u^c
\]

\[
= (\mu - \mu) \pi \ a^f \pi \ u^c u^c
\]

\[
= 0
\]
This shows that the average value of an additive effect of a random sex-linked gene in male over the whole population is zero.

$$E(n_{q_a r}) = \sum \sum p_a p_h g_{a_a} g_{h_a} \phi_{a_a} \phi_{h_a} \pi p^a p^h \pi u^c u^c$$

$$= \sum \sum p_a p_h g_{a_a} g_{h_a} (S_{a_a} - \Sigma p_o S_{a_a}) (S_{a_a} - \Sigma p_o S_{a_a}) \pi p^a p^h \pi u^c u^c$$

$$= \sum \sum p_a p_h g_{a_a} g_{h_a} S_{a_a}^2 - \sum \sum p_a p_h g_{a_a} g_{h_a} S_{a_a} S_{a_a} - \sum \sum p_a p_h g_{a_a} g_{h_a} S_{a_a} S_{a_a} + \sum \sum p_a p_h g_{a_a} g_{h_a} S_{a_a} S_{a_a}$$

$$= (\mu - \mu - \mu + \mu) \pi p^a p^h \pi u^c u^c$$

$$= 0$$

where $\sum p_a = 1$ and $\sum p_h = 1$

This shows that the average value of a dominance deviation of two sex-linked genes in a locus in females over the whole population is zero. We showed above that the expected value of an additive effect or dominance deviation is zero. Now let us examine the expectation of an epistatic effect.

$$E(a_{c q_a r}^c) = \sum \sum p_c p_a q_{c c} q_{c c} \phi_{c c} \phi_{c c} \pi p^a p^c \pi a^c a^c$$

$$= \sum \sum p_c p_a q_{c c} q_{c c} (S_{c c} - \Sigma p_o S_{c c}) (S_{c c} - \Sigma p_o S_{c c}) \pi p^a p^c \pi a^c a^c$$

$$= \sum \sum p_c p_a q_{c c} q_{c c} S_{c c}^2 - \sum \sum p_c p_a q_{c c} q_{c c} S_{c c} S_{c c} - \sum \sum p_c p_a q_{c c} q_{c c} S_{c c} S_{c c} + \sum \sum p_c p_a q_{c c} q_{c c} S_{c c} S_{c c}$$

$$= (\mu - \mu - \mu + \mu) \pi p^a p^c \pi a^c a^c$$

$$= 0$$

where $\sum p_c = 1$ and $\sum p_a = 1$
Now then we can prove that the expectation of an epistatic effect is zero. It can also be easily seen that the covariance between any two terms in the genotypic value is zero. So all the terms in the genotypic value are uncorrelated and hence we have,

\[ V(\text{genotypic value}) = \sum_{\text{over all terms}} \mathbb{E}(\text{each term})^2 \]

Now collecting the terms of similar nature we define, for the case of male,

\[ \mathbb{E}(a_{jc}^2) = \frac{1}{2} \sigma_A^2 \text{ occurs} \]

= half the additive (autosomal) variance due to locus \( c \).

\[ \mathbb{E}(d_{jc}^2) = \sigma_D^2 \text{ occurs} \]

= dominance (autosomal) variance due to locus \( c \).

\[ \mathbb{E}(a_{jc} a_{jc'}^2) = \frac{1}{4} \sigma_{A\text{A}}^2 \text{ occurs} \]

= \( \frac{1}{4} \) additive (autosomal) \( \times \) additive (autosomal) variance due to loci \( c \) and \( c' \).

\[ \mathbb{E}(\beta_{qa}^2) = \sigma_{A_m}^2 \text{ occurs} \]

= additive (sexlinked) variance due to locus \( a \).

\[ = (\sum_{g_c} \theta_{jc}^c \sigma_{ja}^c) (\sum_{h_a} \rho_{ja}^a \pi_{ja}^a) \pi_{ja}^a \pi_{ja}^a \pi_{jc}^c \sigma_{ja}^c \]

\[ = \mathbb{E}(\eta_{jc}^c) \mathbb{E}(\beta_{qa}^a) \]

\[ = 0 \]
\[ E(\beta^a_q \beta^{a'}_q)^2 = \sigma^2_{A_m A_m}(aa') = \text{additive (sexlinked) x additive (sexlinked) variance due to loci } a \text{ and } a'. \]

\[ E(a^c_j \beta^a_c)^2 = \frac{1}{2} \sigma^2_{A A_m}(ca) = \text{additive (autosomal) x additive (sexlinked) variance due to } c^\text{th} \text{ autosomal locus and } a^\text{th} \text{ sexlinked locus.} \]

\[ E(d^c_j k_c \beta^a_c)^2 = \sigma^2_{D A_m}(ca) = \text{dominance (autosomal) x additive (sexlinked) variance due to } c^\text{th} \text{ autosomal locus and } a^\text{th} \text{ sexlinked locus.} \]

and so on.

Now in general, if \( X_m \) involves \( t_1 \) autosomal additive effects, \( t_2 \) dominance deviations, \( t_3 \) additive (sexlinked) effects in males, then,

\[ EX^2_m = \left(\frac{1}{2}\right)^{t_1} \sigma^2_{A_1 D_2 A_3} = \left(\frac{1}{2}\right)^{t_1} \sigma^2_{A A A} \ldots \ldots \ldots A \text{ terms} \]

Now we proceed to evaluate the total variances for males, by summing any component over all possible sets of loci.

\[ \sigma^2_A = \sum_{c=1}^{n_1} \sigma^2_A(c) = \text{total additive (autosomal) variance in the population.} \]
\[ \sigma_D^2 = \sum_{c=1}^{n_2} \sigma_D^2(c) \] = total dominance (autosomal) variance in the population.

\[ \sigma_{AA}^2 = \sum_{cc'} \sigma_{AA}^2(cc') \] = total additive (autosomal) x additive (autosomal) variance in the population.

\[ \sigma_{Am}^2 = \sum_a \sigma_{Am}^2(a) \] = total additive (sexlinked) variance in the population.

\[ \sigma_{DAm}^2 = \sum_c \sum_a \sigma_{DAm}^2(ca) \] = total dominance (autosomal) x additive (sexlinked) variance in the population.

\[ \sigma_{AAm}^2 = \sum_c \sum_a \sigma_{AAm}^2(ca) \] = total additive (autosomal) x additive (sexlinked) variance in the population.

and so on.

In case of male, the genotypic variance is partitioned into the following constituents,

\[ \sigma_{Gm}^2 = \sigma_A^2 + \sigma_D^2 + \sigma_{AA}^2 + \sigma_{DD} + \ldots + \sigma_{Am}^2 + \sigma_{AAm}^2 \]

\[ + \sigma_{DAm}^2 + \sigma_{AAm}^2 + \sigma_{DAm}^2 + \sigma_{AAm}^2 + \ldots \text{ and so on.} \]

In case of females we collect the terms of similar nature and define,

\[ E(p_{qa}^2) = \sigma_{Aq}^2(a) \] = additive (sexlinked) variance due to locus a.
additive (sexlinked) x additive (sexlinked) variance due to loci a and a'.

\[ E(\beta_{q_a}^{a} \beta_{q_a}^{a'})^2 = \sigma^2_{A_f A_f}(aa') \]

additive (autosomal) x additive (sexlinked) variance due to loci c and a.

\[ E(\alpha_{j}^{c} \beta_{q_a}^{a})^2 = \frac{1}{2} \sigma^2_{AA_f}(ca) \]

dominance (autosomal) x additive (sexlinked) variance due to loci c and a.

\[ E(d_{j\text{C}}^{c} \beta_{q_a}^{a})^2 = \sigma^2_{DA_f}(ca) \]

additive (autosomal) x dominance (sexlinked) variance due to loci c and a.

\[ E(\alpha_{j\text{C}}^{c} \alpha_{r_a}^{a})^2 = \sigma^2_{AD_f}(ca) \]

dominance (autosomal) x dominance (sexlinked) variance due to loci c and a.

\[ E(d_{j\text{C}}^{c} \pi_{q_a}^{a})^2 = \sigma^2_{DD_f}(ca) \]

and so on.

Now in general if \( X_f \) involves \( t_1 \) autosomal additive effects, \( t_2 \) dominance (autosomal) deviations, \( t_3 \) additive (sexlinked) effects and \( t_4 \) dominance (sexlinked) deviations in females, then,

\[ E X_f^2 = \left( \frac{1}{2} \right)^{t_1} \sigma^2 A A \ldots A \quad D D \ldots D \quad A_f A_f \ldots A_f \quad D_f D_f \ldots D_f \]

\[ t_1 \text{ terms} \quad t_2 \text{ terms} \quad t_3 \text{ terms} \quad t_4 \text{ terms} \]

\[ = \left( \frac{1}{2} \right)^{t_1} \sigma^2 \quad A^t_1 D_2 A_f^t_3 D_f^t_4 \]
Now we sum any component over all possible sets of loci to evaluate the total variances in females.

\[ \sigma^2_{A^f} = \sum_a \sigma^2_{A_A^f(a)} \]

= total additive (sexlinked) variance in the population.

\[ \sigma^2_{AA^f} = \sum_c \sum_a \sigma^2_{AA^f(ca)} \]

= total additive (autosomal) x additive (sexlinked) variance in the population.

\[ \sigma^2_{DA^f} = \sum_c \sum_a \sigma^2_{DA^f(ca)} \]

= total additive (autosomal) x dominance (sexlinked) variance in the population.

\[ \sigma^2_{DD^f} = \sum_c \sum_a \sigma^2_{DD^f(ca)} \]

= total dominance (autosomal) x dominance (sexlinked) variance in the population.

and so on.

In case of female, the genotypic variance is partitioned into the following components,

\[ \sigma^2_{G_f} = \sigma^2_A + \sigma^2_D + \sigma^2_{AA} + \sigma^2_{DD} + \ldots + \sigma^2_{A^f} + \sigma^2_{AA^f} + \sigma^2_{D^f} + \sigma^2_{D^f_A^f} + \sigma^2_{D^f_A^f} + \ldots \]

and so on.

The genotypic covariance between relatives We shall now develop the general structures for the covariances between (i) male and male (ii) female and female (iii) female and male, with an arbitrary degree of relationship under the assumption of both autosomal and sexlinked transmission. There will be two restrictions on the development (1) assumption of no linkage, that is, the genes segregate independently and
(2) assumption of no inbreeding, that is, there is zero probability that
two genes possessed by an individual at any locus are identical by descent.
Since this topic has been elaborated in detail under single locus theory,
we shall give only a brief description of the procedure involved, and then
present the results.

We shall obtain the covariance between the genotypic values of an
individual \( \mathbf{X} \) who is a random member of the population and an indi­
vidual \( \mathbf{Y} \) who is in particular relationship to \( \mathbf{X} \). Later we will intro­
duce the subscripts \( m \) and \( f \) on \( \mathbf{X} \) and \( \mathbf{Y} \) to indicate their sexes. Let
the genes of \( \mathbf{X} \) at the \( a \) locus be \( i_a, j_a \) and of \( \mathbf{Y} \) be \( k_a, l_a \). The
quantities \( v_{XY}, r_{XY}^*, u_{XY}^*, w_{XY} \) will retain their original defi­
\( \equiv \)

First we shall consider the covariance of autosomal additive effects.
At the \( a^{th} \) locus \( \mathbf{X} \) has two genes and \( \mathbf{Y} \) has two genes and at the \( b^{th} \)
locus \( \mathbf{X} \) has two genes and \( \mathbf{Y} \) has two genes and the covariance of the
effects of these genes can be represented symbolically as follows,

\[
\sum \mathbb{P}(i_a = k_a) + \mathbb{P}(i_a = l_a) + \mathbb{P}(j_a = k_a) + \mathbb{P}(j_a = l_a) \frac{1}{2} \sigma_A^2(a) \]
\[
+ \sum \mathbb{P}(i_b = k_b) + \mathbb{P}(i_b = l_b) + \mathbb{P}(j_b = k_b) + \mathbb{P}(j_b = l_b) \frac{1}{2} \sigma_A^2(b) \]
\[
+ \ldots \text{ and so on.}
\]
Now the sum of these components over the number of loci involved yields the covariance of additive (autosomal) effects. So then we have

\[
\frac{1}{2} 4 r_{XY} \sigma^2_{A(a)} + \sigma^2_{A(b)} + \ldots \]

\[
= \frac{1}{2} r_{XY} \sum_c \sigma^2_{A(c)}
\]

\[
= 2 r_{XY} \sigma^2_A
\]

Now consider the covariances of autosomal dominance deviations; under the same argument above we have,

\[
\sum P(i_a = k_a, j_a = 1_a) + P(i_a = 1_a, k_a = j_a) \sigma^2_{D(a)}
\]

\[
+ \sum P(i_b = k_b, j_b = 1_b) + P(i_b = 1_b, k_b = j_b) \sigma^2_{D(b)}
\]

\[
+ \ldots \quad \text{and so on.}
\]

Now the sum of these components over the number of loci involved gives the covariances of autosomal dominance deviations. Symbolically we have,

\[
u_{XY} \sum \sigma^2_{D(a)} + \sigma^2_{D(b)} + \ldots \]

\[
= u_{XY} \sum_c \sigma^2_{D(c)}
\]

\[
= u_{XY} \sigma^2_D
\]

Considering the covariances of male sexlinked loci, we have,

\[
P(i_a = j_a) \sigma^2_{A_m(a)}
\]

\[
+ P(i_b = j_b) \sigma^2_{A_m(b)} + \ldots \quad \text{and so on.}
\]
Now the sum of these components over the number of sexlinked loci involved gives the required covariance. Symbolically we have,

\[ \sum_{X} \left[ \sigma_{m}^{2}(a) + \sigma_{m}^{2}(b) + \ldots \right] \]

\[ = \sum_{a} \sigma_{m}^{2}(a) \]

\[ = \sum_{X} \sigma_{m}^{2} \]

Considering the covariance of female additive sexlinked loci and following the same argument as the autosomal case we have,

\[ \sum_{Y} \left[ P(i = k) + P(i = 1) + P(j = k) + P(j = 1) \right] \frac{1}{2} \sigma_{t}^{2}(a) \]

\[ + \sum_{Y} \left[ P(i = k) + P(i = 1) + P(j = k) + P(j = 1) \right] \frac{1}{2} \sigma_{t}^{2}(b) \]

\[ + \ldots \text{ and so on.} \]

Now summing these expressions we have,

\[ \frac{1}{2} \sum_{Y} \sigma_{t}^{2} \]

\[ = 2 \sum_{Y} \sigma_{t}^{2} \]

In case of covariance of dominance (sexlinked) deviations in females we have,

\[ \sum_{Y} \left[ P(i = k, j = 1) + P(i = 1, j = k) \right] \sigma_{D}^{2}(a) \]

\[ + \sum_{Y} \left[ P(i = k, j = 1) + P(i = 1, j = k) \right] \sigma_{D}^{2}(b) \]

\[ + \ldots \text{ and so on.} \]

Now summing we have,
\[ u_{XY} \sum_c \sigma_D^2(c) = u_{XY} \sigma_D^2 \]

Considering the covariance of a male sexlinked loci and a female sexlinked loci, we have,

\[ \sum P(i_a = j_a) + P(i_a = k_a) \frac{1}{2} C_{A_s}(a) \]

\[ + \sum P(i_b = j_b) + P(i_b = k_b) \frac{1}{2} C_{A_s}(b) \]

\[ + \ldots \text{ and so on.} \]

Then summing we have,

\[ w_{XY} \frac{1}{2} \left( C_{A_s}(a) + C_{A_s}(b) + \ldots \right) \]

\[ = w_{XY} \frac{1}{2} \sum_c C_{A_s}(c) \]

\[ = w_{XY} \frac{1}{2} C_{A_s} \]

where \( X \) is the male with genes \( i_a, i_b, \) and so on, and \( Y \) is the female with genes \( j_a k_a, j_b k_b \) and so on.

Now briefly let us consider the covariance of effects or deviations involving interactions. Since the loci are not linked and independent conditions hold at loci other than the loci under consideration, we will have a covariance of \( \left( \frac{1}{2} \right)^{t_1} \sigma_A^2 \) for the additive interactions of a set of \( t_1 \) autosomal loci, if \( i_a = k_a \) for every locus (of the set).

Now there are three other cases which can arise namely, \( j_a = 1_a, i_a = 1_a \) and \( j_a = k_a \). With independent assortment the total covariance for the
additive autosomal effects, we have

\[(4r_{XY})^t_1 (\frac{1}{2})^t_1 \sigma^2 \]

\[A^t_1 \]

With regard to dominance deviation at \( t_2 \) loci under independent conditions we have

\[(\frac{1}{2})^t_1 (4r_{XY})^t_1 \left( (u_{XY})^t_2 \sigma^2 \right) \]

\[A^t_1 D^t_2 \]

If the interaction also involves \( t_3 \) sexlinked additive effects in male, we have,

\[(\frac{1}{2})^t_1 (4r_{XY})^t_1 \left( t_2^t (u_{XY})^t_3 \sigma^2 \right) \]

\[A^t_1 D^t_2 A^t_3 \]

If the consideration is extended to the \( t_4 \) sexlinked additive effects and \( t_5 \) sexlinked dominance deviation in female then we have,

\[(\frac{1}{2})^t_1 (4r_{XY})^t_1 \left( (u_{XY})^t_2 (r_{XY})^t_4 \right) \]

\[A^t_1 D^t_2 A^t_4 D^t_5 \]

and so on.

In this development two conditions have to be remembered namely (i) in evaluating the covariance between \( X \) and \( Y \), a term gives zero covariance unless the subscripts are identical over and above the fact of being alike by chance, and (ii) the covariance from different sets of loci are additive under our assumptions.

Now then we enumerate the results as follows,

(1) Cov(male, male)
\[ \text{Cov}(X_m, Y_m) = 2r_{XY} \sigma^2_A + u_{XY} \sigma^2_D + \nu_{XY} \sigma^2_{A_m} + \sum_{t_1} \sum_{t_2} \sum_{t_3} (2r_{XY})^{t_1} (u_{XY})^{t_2} (\nu_{XY})^{t_3} \sigma^2_{A_1} \sigma^2_{D_2} \sigma^2_{A_m} \] (14)

where \( t_1 + t_2 + t_3 > 1 \)

where \( t_1 + t_2 = n_1 \), the number of autosomal loci
\( t_3 = n_2 \), the number of sexlinked loci
\( t_1 \) is the number of autosomal additive effects
\( t_2 \) is the number of autosomal dominance deviations
\( t_3 \) is the number of sexlinked additive effects in male

\[ \text{Cov}(\text{female, female}) = \text{Cov}(X_f, Y_f) \]

\[ = 2r_{XY} \sigma^2_A + u_{XY} \sigma^2_D + 2r_{XY} (u_{XY})^{t_1} (2r_{XY})^{t_2} (u_{XY})^{t_4} \sigma^2_{A_f} \sigma^2_{D_f} \] (15)

where \( t_1 + t_2 = n_1 \), the number of autosomal loci
\( t_4 + t_5 = n_2 \), the number of sexlinked loci
\( t_1 \) is the number of autosomal additive effects
\( t_2 \) is the number of autosomal dominance deviations
\( t_4 \) is the number of sexlinked additive effects
\( t_5 \) is the number of sexlinked dominance deviations
(3) \( \text{Cov(male, female)} \)

\[
\text{Cov}(X_m, Y_f) = 2 \sigma_{XY}^2 \sigma_A^2 + u_{XY}^2 \sigma_D^2 + w_{XY}^2 C_{AS}^2 \\
+ \sum_{t_1} \sum_{t_2} \left( \frac{2 r_{XY}}{n_1} \right)^{t_1} \left( u_{XY} \right)^{t_2} \sigma^2_{AD} \\\n+ \sum_{t_1} \sum_{t_2} \sum_{t_6} \left( \frac{2 r_{XY}}{n_1} \right)^{t_1} \left( u_{XY} \right)^{t_2} \left( w_{XY} \right)^{t_6} C_{ADAS} \\
\text{where } t_1 + t_2 + t_6 > 1
\]

where, the number of autosomal loci

- \( t_1 \) is the number of autosomal loci
- \( t_2 \) is the number of autosomal dominance deviations
- \( t_6 \) is the number of sexlinked additive effects

If a particular interaction involves \( t_1 \) autosomal additive effects, \( t_2 \) dominance deviations and \( t_6 \) sexlinked additive effects, the covariance of such interactions summed over the relevant number of loci is

\[
(2 r_{XY})^{t_1} (u_{XY})^{t_2} (w_{XY})^{t_6} C_{ADAS} \\
\text{and when } t_1 = t_2 = t_6 = 1, \text{ the "variance" component is } C_{ADAS}, \text{ which is short for autosomal additive by autosomal dominance by sexlinked additive covariance.}

If we consider only the multiple sexlinked loci, then the genotypic covariance between individuals is as follows,
The symbols above have already been defined.

The genotypic covariances of some of the special cases which are likely to be of interest have been worked out under the supposition of both sexlinked and autosomal inheritance. The effects of the interactions involving more than two factors have been assumed to be negligible in these examples. The numerical values of the coefficients of the various variance and covariance components contained in the genotypic covariances are presented in Table 7.
Table 7. Values of the coefficients of the variance and covariance components under the supposition of multiple loci inheritance

<table>
<thead>
<tr>
<th>Relations</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sire and son</td>
<td>$\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_{AA}$</td>
</tr>
<tr>
<td>Sire - daughter</td>
<td>$\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_{AA} + \frac{1}{2} C_{As} + \frac{1}{4} C_{As} A_s$</td>
</tr>
<tr>
<td>Dam - son</td>
<td>$\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_{AA} + \frac{1}{2} C_{As} + \frac{1}{4} C_{As} A_s$</td>
</tr>
<tr>
<td>Dam - daughter</td>
<td>$\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_{AA} + \frac{1}{2} \sigma^2_f + \frac{1}{4} \sigma^2_{AAf} + \frac{1}{4} \sigma^2_{AfAf}$</td>
</tr>
<tr>
<td>Brothers</td>
<td>$\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_D + \frac{1}{2} \sigma^2_{Am} + \frac{1}{4} \sigma^2_{AA} + \frac{1}{8} \sigma^2_{AD} + \frac{1}{16} \sigma^2_{DD}$</td>
</tr>
<tr>
<td></td>
<td>$+ \frac{1}{4} \sigma^2_{AAm} + \frac{1}{8} \sigma^2_{DAm} + \frac{1}{4} \sigma^2_{AmAm}$</td>
</tr>
<tr>
<td>Sisters</td>
<td>$\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_D + \frac{3}{4} \sigma^2_f + \frac{1}{2} \sigma^2_{Af} + \frac{1}{4} \sigma^2_{AA} + \frac{1}{8} \sigma^2_{AD}$</td>
</tr>
<tr>
<td></td>
<td>$+ \frac{1}{16} \sigma^2_{DD} + \frac{3}{8} \sigma^2_{AAf} + \frac{1}{4} \sigma^2_{ADf} + \frac{3}{16} \sigma^2_{DAf} + \frac{1}{8} \sigma^2_{DDf}$</td>
</tr>
<tr>
<td></td>
<td>$+ \frac{9}{16} \sigma^2_{AfAf} + \frac{1}{4} \sigma^2_{DfDf}$</td>
</tr>
<tr>
<td>Brother and sister</td>
<td>$\frac{1}{2} \sigma^2_A + \frac{1}{4} \sigma^2_D + \frac{1}{4} C_{As} + \frac{1}{4} \sigma^2_{AA} + \frac{1}{8} \sigma^2_{AD} + \frac{1}{16} \sigma^2_{DD}$</td>
</tr>
<tr>
<td></td>
<td>$+ \frac{1}{8} C_{AA} + \frac{1}{16} C_{DA} + \frac{1}{16} C_{AsAs}$</td>
</tr>
<tr>
<td>Paternal half brothers</td>
<td>$\frac{1}{4} \sigma^2_A + \frac{1}{16} \sigma^2_{AA}$</td>
</tr>
<tr>
<td>Paternal half sisters</td>
<td>$\frac{1}{4} \sigma^2_A + \frac{1}{2} \sigma^2_{Af} + \frac{1}{16} \sigma^2_{AA} + \frac{1}{8} \sigma^2_{AAf} + \frac{1}{4} \sigma^2_{AfAf}$</td>
</tr>
<tr>
<td>Paternal half brother and half sister</td>
<td>$\frac{1}{4} \sigma^2_A + \frac{1}{16} \sigma^2_{AA}$</td>
</tr>
</tbody>
</table>
Table 7. (Continued)

<table>
<thead>
<tr>
<th>Relations</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal half brothers</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{2} \sigma^2 A_m + \frac{1}{16} \sigma^2 AA_m + \frac{1}{8} \sigma^2 AA_m + \frac{1}{4} \sigma^2 A_m A_m$</td>
</tr>
<tr>
<td>Maternal half sisters</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{4} \sigma^2 A_f + \frac{1}{16} \sigma^2 AA_f + \frac{1}{16} \sigma^2 AA_f + \frac{1}{16} \sigma^2 A_f A_f$</td>
</tr>
<tr>
<td>Maternal half brother and half sister</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{4} C_{A_s} + \frac{1}{16} \sigma^2 AA_s + \frac{1}{16} C_{AA_s} + \frac{1}{16} C_{A_s A_s}$</td>
</tr>
<tr>
<td>Paternal uncle nephew</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{16} \sigma^2 AA$</td>
</tr>
<tr>
<td>Paternal uncle niece</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{4} \sigma^2 A_f + \frac{1}{16} \sigma^2 AA_f + \frac{1}{16} \sigma^2 AA_f + \frac{1}{16} \sigma^2 A_f A_f$</td>
</tr>
<tr>
<td>Paternal aunt nephew</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{16} \sigma^2 AA$</td>
</tr>
<tr>
<td>Paternal aunt niece</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{4} \sigma^2 A_f + \frac{1}{16} \sigma^2 AA_f + \frac{1}{16} \sigma^2 AA_f + \frac{1}{16} \sigma^2 A_f A_f$</td>
</tr>
<tr>
<td>Maternal uncle nephew</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{8} \sigma^2 A_m + \frac{1}{16} \sigma^2 AA_m + \frac{1}{32} \sigma^2 AA_m + \frac{1}{64} \sigma^2 A_m A_m$</td>
</tr>
<tr>
<td>Maternal uncle niece</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{8} C_{A_s} + \frac{1}{16} \sigma^2 AA_s + \frac{1}{32} C_{AA_s} + \frac{1}{64} C_{A_s A_s}$</td>
</tr>
<tr>
<td>Maternal aunt nephew</td>
<td>$\frac{1}{4} \sigma^2 + \frac{3}{8} C_{A_s} + \frac{1}{16} \sigma^2 AA_s + \frac{3}{32} C_{AA_s} + \frac{9}{64} C_{A_s A_s}$</td>
</tr>
<tr>
<td>Maternal aunt niece</td>
<td>$\frac{1}{4} \sigma^2 + \frac{3}{4} \sigma^2 A_f + \frac{1}{16} \sigma^2 AA_f + \frac{3}{16} \sigma^2 AA_f + \frac{3}{16} \sigma^2 A_f A_f$</td>
</tr>
<tr>
<td>Sire and maternal grand daughter</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{4} C_{A_s} + \frac{1}{16} \sigma^2 AA_s + \frac{1}{16} C_{AA_s} + \frac{1}{16} C_{A_s A_s}$</td>
</tr>
<tr>
<td>Dam and paternal grand daughter</td>
<td>$\frac{1}{4} \sigma^2 + \frac{1}{2} \sigma^2 A_f + \frac{1}{16} \sigma^2 AA_f + \frac{1}{8} \sigma^2 AA_f + \frac{1}{4} \sigma^2 A_f A_f$</td>
</tr>
</tbody>
</table>
Inbreeding

In the last section we have examined the cases when there is zero probability that the two genes possessed by an individual are identical by descent. This case was approached first because the results obtained under the assumptions outlined there are not only easily comprehensible but they form a base line from which one can compare or contrast the results obtained under different assumptions. With this objective in mind, we have elaborated the random mating case in step-by-step detail. In this section we shall relax the assumption noted above to the case when there is non-zero probability that the two genes possessed by an individual are identical by descent. This condition is easily achieved by inbreeding, that is, by mating related individuals. The consequences of this condition are manifold and of great practical interest but unfortunately it has been difficult to achieve complete algebraic generalization even for simple cases.

Wright (1921) introduced the coefficient of inbreeding (F) and examined the progress of panmictic index (P) under the assumption of autosomal inheritance. Malécot (1948) attacked the problem with probability arguments and simplified the concept to a certain degree. Generation matrix methodology which is a standard tool of elementary stochastic process theory was applied to genetic systems first by Bartlett and Haldane (1935). Fisher (1949) has developed this approach considerably as regards the mathematical structure of inbreeding systems. A fairly complete discussion of the role of the approach based on the probability that a pair of genes are identical by descent, which in
simple cases at least leads to difference equations with constant co-
efficients and of the generation matrix method is given by Kempthorne
(1957). A general summary statement of the situation is that both are
facets of the same situation, that it is somewhat a matter of taste whether
one follows one approach or the other for simpler problems and that the
generation matrix methodology does give a more complete picture of the
situation. Whether the additional feature of the situation yielded by the
generation matrix methodology in problem susceptible to both approaches
have much scientific value may be debated but there are very definitely
real scientific problems which require generation matrix methodology, or
even more complicated mathematical tools. The work of Kempthorne
(1955, 1956, 1957) and Horner (1956) are noteworthy. Kempthorne
solved the long-standing problem of deriving the correlations between
relatives under selfing and fullsibbing. Horner obtained similar results
for the case of parent-offspring mating. The results are, as expected,
complex even for the single locus situation with regular inbreeding. It is
expected that the joint consideration of autosomal genes and sexlinked
genes in studying the theory of inbreeding and its impact on quantitative
genetics with arbitrary number of loci with arbitrary epistasis under any
system of inbreeding (regular or irregular) will lead to formidably
complex results. In this study no attempt will be made even to initiate a
modest approach to the problem, but we will be concerned with some
other problem of practical interest involving inbreeding.

Malécot has presented the basic notion of two genes being alike in the
population for two different reasons:
(a) They may be alike because one is a copy of the other, that is, each is a copy of one gene occurring in the ancestors of both and this is called identical by descent.

(b) They may be alike because they happen to be copies of different genes in the base population, the different genes being indistinguishable from each other. This is called identical by state, but not by descent.

He has also introduced the concept of coefficient of parentage, subsequently called coefficient of parentage by Kempthorne (1957), exclusively applicable to autosomal genes. In the following we will be defining certain quantities applicable to only sex-linked genes and name them as coefficients of genic relationship. The justification of introducing such a terminology will be given in chapter IV.

The quantity $v_{xy}$ is defined as the probability that a random sex-linked gene in male $X$ is identical by descent with a random sex-linked gene in male $Y$.

The quantity $r_{xy}^s$ is defined as the probability that a random sex-linked gene in female $X$ is identical by descent with a random sex-linked gene in female $Y$.

The quantity $w_{xy}^s$ is defined as the probability that a random sex-linked gene in female $X$ is identical by descent with a random sex-linked gene in male $Y$.

Now we shall examine the relationships among the three coefficients of genic relationship under the assumption of inbreeding and sexlinked inheritance. We have presented a picture below, in which $X$ is the dam
with two genes (ab), Y is the Sire with one gene (c), \( Z_m \) is a son with one gene (d) and \( Z_f \) is a daughter with two genes (ef).

The relationships among the genes are given by the following probability arrays,

\[
d = \frac{1}{2} a + \frac{1}{2} b \\
e = \frac{1}{2} a + \frac{1}{2} b \\
f = c
\]

A word of explanation is necessary to spell out the exact meaning of these equations noted above. In the first equation it really does not mean that gene \( d \) is algebraically equal to sum of half of gene \( a \) and half of gene \( b \). We have already defined in the introduction that genes will always refer in this study, to indivisible entities. The equation actually is interpreted as follows:

The probability that \( d \) is equal to \( a \) is 0.5 and the probability that \( d \) is equal to \( b \) is 0.5. The equation

\[
d = \frac{1}{2} a + \frac{1}{2} b
\]

is a conventional way of expressing these facts and is a probability array for \( d \).

Here we shall introduce a convention that will be followed throughout
the thesis. This relates to the specification of sexes of individuals, the 
evaluation of whose coefficient of genic relationship is under consideration. 
This will be explained with simple examples. Suppose one writes $v_{XY}$ 
and wants to find out the sexes of individual $X$ and individual $Y$, it is 
obvious that $X$ and $Y$ are both males because $v_{XY}$ stands for the 
coefficient of genic relationship of two males. So also is the case of 
$r_{AB}^*$ in which $A$ and $B$ are both females because $r_{AB}^*$ stands for the 
coefficient of genic relationship between two females. So in these two 
respects the sexes of the individuals concerned need not be specified.

Now consider $w_{CD}$. From the symbol only we cannot know if $C$ is a 
males or a female or, if $D$ is a male or a female, because $w_{CD}$ stands 
for the coefficient of genic relationship between a male and a female.

Now if we introduce a convention that we shall always place the female 
first and male next in the subscripts, then, in this example, obviously 
$C$ is the female and $D$ is the male. Also in the convention we include 
the condition that $w_{DC}$ is not permissible, if $D$ is a male and $C$ is a 
female. Now if we follow the above noted convention we do not have to 
specify the sexes of the individuals the determination of whose coefficient 
of genic relationship is under consideration.

Now let us consider the algebraic relationships among the coefficients 
of genic relationship with reference to the pedigree given in page 72. In 
the following whenever there is a reference to Sire, Dam, Son and 
Daughter, it will mean $X$, $Y$, $Z_m$ and $Z_f$ respectively in the pedigree.
(1) The coefficient of genic relationship between Sire and Son is equal to the coefficient of genic relationship between their parents. Symbolically, we have

\[ v_{YZ} = P(c = d) \]
\[ = \frac{1}{2} P(a = c) + \frac{1}{2} P(b = c) \]
\[ = w_{XY} \]

(2) The coefficient of genic relationship between Dam and Son is equal to half of one plus the coefficient of inbreeding of the Dam. Symbolically we have,

\[ w_{XZ} = \frac{1}{2} \left( P(a = d) + P(b = d) \right) \]
\[ = \frac{1}{2} \left( 1 + F_X \right) \]

where \( F_X \) is the inbreeding coefficient of \( X \) and is equal to \( P(a = b) \), that is,

\[ F_X = P(a = b) \]

(3) The coefficient of genic relationship between Dam and Daughter is equal to half the sum of the coefficients of genic relationship between the Daughter and herself and her parents. Symbolically,

\[ r_{XZ} = \frac{1}{4} \left( P(a = e) + P(a = f) + P(b = e) + P(b = f) \right) \]
\[ = \frac{1}{2} \left( r_{XX}^* + w_{XY} \right) \]

where \( r_{XX}^* \) is defined as the coefficient of genic relationship between \( X \) and herself.
(4) The coefficient of genic relationship between Sire and Daughter is equal to half of one plus the coefficient of genic relationship between the parents.

\[ w_{ZY} = \frac{1}{2} \left[ P(e = c) + P(f = c) \right] \]

Now we plan to examine the coefficient of genic relationship between a member of the pedigree in page 72 and any male or any female in the population, so that the results thus obtained can be applied to a wide variety of cases. Let \( N \) represent the male with gene (g) and \( N \) represent the female with gene (kl), then we have the following relationship among the coefficients of genic relationship with reference to the pedigree presented in page 72.

(5) The coefficient of genic relationship between the Son and male \( N \) is equal to the coefficient of genic relationship between the Dam and male \( N \). Symbolically,

\[ v_{NZ} = v_{N(XY)} = P(d = g) \]

\[ = \frac{1}{2} \left[ P(a = g) + P(b = g) \right] \]

\[ = w_{XN} \]

(6) The coefficient of genic relationship between the Son and female \( N \) is equal to the coefficient of genic relationship between the Dam and female \( N \).

\[ w_{NZ} = w_{N(XY)} \]

\[ = \frac{1}{2} \left[ P(k = d) + P(l = d) \right] \]
\[= \left[ P(a = k) + P(a = l) + P(b = k) + P(b = l) \right] \]

\[= r_{NX}^* \]

In (5) and (6) and in subsequent relationships XY in parenthesis stands for the parents of Z. The female parent is placed first and the male parent next.

(7) The coefficient of genie relationship between the Daughter and male N is equal to half of the sum of the coefficients of genie relationship between male N and the parents of the Daughter. Symbolically,

\[w_{ZN} = w_{(XY)N} = \frac{1}{2} \left[ P(e = g) + P(f = g) \right] \]

\[= \frac{1}{2} \left[ w_{XN} + v_{YN} \right] \]

(8) The coefficient of genie relationship between the Daughter and female N is equal to half of the sum of the coefficients of genie relationship between female N and the parents of the Daughter. Symbolically,

\[r_{NZ}^* = \frac{1}{4} \left[ P(e = k) + P(e = l) + P(f = k) + P(f = l) \right] \]

\[= \frac{1}{2} \left[ r_{NX}^* + w_{NY} \right] \]

(9) The coefficient of genie relationship between a female and herself is equal to half the sum of one and the coefficient of inbreeding of the female, that is,

\[r_{XX}^* = \frac{1}{2} \left( 1 + F_X \right) \]
(10) The coefficient of genic relationship between a male and himself is equal to one, that is,

\[ v_{yy} = 1 \]

Now to summarize the above results, we have the following basic formulae,

(a) \( v_{NZ} = v_{N(XY)} = w_{XN} \)

(b) \( w_{NZ} = w_{N(XY)} = r_{NX}^{x} \)

(c) \( w_{ZN} = w_{(XY)N} = \frac{1}{2} \sum w_{XN} + v_{YN} \) \( \sum \)

(d) \( r_{NZ}^{x} = r_{N(XY)}^{x} = \frac{1}{2} \sum r_{NX}^{x} + w_{NY} \) \( \sum \)

(e) \( r_{XX}^{x} = \frac{1}{2} \sum 1 + F_{X} \) \( \sum \)

(f) \( v_{yy} = 1 \)

where \( X \) and \( Y \) are the parents of \( Z \) and \( F_{X} \) is the coefficient of inbreeding of \( X \).

Now, for the moment, let us examine what the "F" does in a female population of the following composition,

\[ (\sum p_{i} s_{i})^{2} \]

If we choose individuals at random from the population under consideration, the probability that any sexlinked gene is \( S_{i} \) is \( p_{i} \) and if \( F \) is the probability that two genes in an individual are alike by descent, the probability that they are both \( S_{i} \) is \( F p_{i} \). The probability that they are unlike by descent is \( (1 - F) \) and the probability that two ordered genes in the population are \( S_{i} \) and \( S_{j} \) is \( p_{i} p_{j} \). So after inbreeding the population to an extent measured by \( F \), the resulting population has the
It is of interest to examine the proportion of heterozygous individuals in the inbred population compared to the proportion in the original population. In this particular case the original proportion of heterozygotes is \( \sum p_i p_j \). After inbreeding to the extent measured by \( F \), the proportion is \( (1 - F) \sum p_i p_j \), so that the proportion of heterozygotes is reduced from the original proportion by a relative amount \( F \). Now it will be of interest to record the progressive reduction of the proportion of heterozygous individuals from generation to generation in a regular system of inbreeding, particularly when the coefficient of inbreeding of the heterogametic sex is predetermined, that is zero. The basic formulae of page 77 will be used to determine this. In the beginning we shall exemplify this with a few cases of regular systems of inbreeding, namely, full sibbing, parent-offspring, and double first cousin and only the results will be given for some other cases.

In the case of a diploid individual \( X \) with genes \( a \) and \( b \) at a locus, the coefficient of inbreeding has been defined by

\[
F_X = P(a = b)
\]

where the equality sign in the parentheses stands for "identical by descent". We need also the relationship

\[
P = (1 - F)
\]

or \( F = (1 - P) \)
where $P$ is Wright's panmictic index.

I. Full sibbing

Let two full sibs in generation $n$ be $X_f$ and $Y_m$. The picture below shows the pedigree.

![Pedigree diagram]

Let the coefficient of inbreeding in generation $n$ be $F_n$ so that

$$F_X = F_n$$

$$F_A = w_{XY} = F_{n+1}$$

Then

$$F_C = w_{AB} = w_{(XY)B}$$

$$= \frac{1}{2} [w_{XB} + v_{YB}]$$

$$= \frac{1}{2} [w_{XB} + v_{Y(XX)}]$$

$$= \frac{1}{2} [w_{XB} + w_{XY}]$$

$$= \frac{1}{2} [r_{XX} + w_{XY}]$$

$$= \frac{1}{2} \left[ \frac{1}{2} (1 + F_X) + F_A \right]$$

$$= \frac{1}{4} + \frac{1}{4} F_X + \frac{1}{2} F_A$$
Since \( F_C = F_{n+2} \), we have

\[
F_{n+2} = \frac{1}{2} F_{n+1} + \frac{1}{4} F_n + \frac{1}{4}
\]

\[
P_{n+2} = \frac{1}{2} P_{n+1} + \frac{1}{4} P_n
\]

This result is the same as that given by Wright (1933) and Haldane (1937).

2. Parent-offspring

There can be many different types of parent-offspring system of inbreeding. We will be dealing with one type shown below,

\[
\begin{align*}
F_{n+1} &= F_A = w_{XY} \\
F_C &= w_{AB} = w_{(XY)B} = \frac{1}{2} \mathcal{L} w_{XB} + v_{YB} \mathcal{J} \\
&= \frac{1}{2} \mathcal{L} \frac{\mathcal{J}}{2} + w_{XY} \mathcal{J} \\
&= \frac{1}{2} \mathcal{L} \frac{1}{2} (1 + F_X) + F_A \mathcal{J}
\end{align*}
\]

or,

\[
F_{n+2} = \frac{1}{2} F_{n+1} + \frac{1}{4} F_n + \frac{1}{4}
\]

Since \( F_C = F_{n+2} \)

or,

\[
P_{n+2} = \frac{1}{2} P_{n+1} + \frac{1}{4} P_n
\]
3. Double first cousin

There can be many different types of this system of inbreeding and we are considering only the type, the picture of which is given below.

Let

\[ F_n = F_A = F_C \]
\[ F_{n+1} = F_E = F_G \]
\[ F_{n+2} = w_{EF} = w_{GH} \]

Then we have

\[ F_{n+3} = w_{IJ} = w_{(EF)(GH)} = \frac{1}{2} r_{eG} + w_{GF} \]

where

\[ r_{EG} = \frac{3}{8} + \frac{1}{8} F_n + \frac{1}{2} F_{n+1} \]

and

\[ w_{GF} = w_{EF} = F_{n+2} \]

So,

\[ F_{n+3} = \frac{3}{16} + \frac{1}{16} F_n + \frac{1}{4} F_{n+1} + \frac{1}{2} F_{n+2} \]

or

\[ P_{n+3} = \frac{1}{2} P_{n+2} + \frac{1}{4} P_{n+1} + \frac{1}{16} P_n \].
We summarize the recurrence relations of the panmictic index over generations in the systems of inbreeding detailed below. The regular system of inbreeding for which the relation holds is only that particular one given in the text. The pedigrees for octuple third cousin is the same as one given by Wright (1921).

<table>
<thead>
<tr>
<th>System</th>
<th>Recurrence relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parent-offspring:</td>
<td>( P_{n+2} = \frac{1}{2} P_{n+1} + \frac{1}{4} P_n )</td>
</tr>
<tr>
<td>Full sibbing:</td>
<td>( P_{n+2} = \frac{1}{2} P_{n+1} + \frac{1}{4} P_n )</td>
</tr>
<tr>
<td>Double first cousin:</td>
<td>( P_{n+3} = \frac{1}{2} P_{n+2} + \frac{1}{4} P_{n+1} + \frac{1}{16} P_n )</td>
</tr>
<tr>
<td>Quadruple second cousin:</td>
<td>( P_{n+4} = \frac{1}{2} P_{n+3} + \frac{1}{4} P_{n+2} + \frac{1}{16} P_{n+1} + \frac{5}{64} P_n )</td>
</tr>
<tr>
<td>Octuple third cousin:</td>
<td>( P_{n+5} = \frac{1}{2} P_{n+4} + \frac{1}{4} P_{n+3} + \frac{1}{16} P_{n+2} + \frac{5}{64} P_{n+1} + \frac{9}{256} P_n )</td>
</tr>
</tbody>
</table>

A summary statement taking into account both autosomal loci and sex-linked loci is that the males become homozygous faster than the females because they have less segregating loci, and the sex-linked loci of the females become homozygous faster than the autosomal loci.

These conclusions of course hold only in the absence of other disturbances such as mutation, natural or artificial selection, random fluctuation due to finite population size and so on.

**Covariance between relatives** The population we are concerned with here is generated by random mating of the individuals of a population which resulted by inbreeding a random mating population to the extent measured by \( F \). The knowledge of the covariances between relatives in such a population is of great practical interest. In the single sexlinked locus situation the partition of the genotypic variance yields the same
components that we have already described in preceding section. Under
the assumption of the additive effects and the dominance deviations being
uncorrelated under any circumstances, we have the same general
structure of the covariances between individuals as in the previous
section that is,

\[ \text{Cov}(X_m, Y_m) = v_{XY} \sigma^2_{A_m} \]
\[ \text{Cov}(X_f, Y_f) = 2r_{XY} \sigma^2_{A_f} + u_{XY} \sigma^2_{D_f} \]
\[ \text{Cov}(X_m', Y_f) = w_{XY} C_{A_s} \]

Now we shall show the application of the above formulae to some of the
special cases of interest. We consider the case of full sibs first. Let
\( X(a) \) and \( Y(bc) \) be the Sire and Dam of \( A(d) \) and \( B(e) \), two of their
male offspring as shown in the picture below,

\[ \begin{align*}
X(a) & \quad Y(bc) \\
A(d) & \quad B(e)
\end{align*} \]

then the relationships among the genes are given by the following
probability arrays

\[ d = \frac{1}{2} b + \frac{1}{2} c \]
\[ e = \frac{1}{2} b + \frac{1}{2} c \]
\[ F = P(b = c) \]

and

\[ v_{AB} = P(d = e) \]
= \frac{1}{2} (1 + F)

and \( \text{Cov}(A, B) = \frac{1}{2} (1 + F) \sigma^2_{A_{m}} \),

that is the covariance between two brothers is \( \frac{1}{2} (1 + F) \sigma^2_{A_{m}} \).

Now if we let \( B \) possess (ef) genes in the picture above where we have the probability array,

\[
f = a
\]

then

\[
w_{BA} = P(d = e) + P(d = f)
\]

\[
= \frac{1}{2} (1 + F)
\]

So the covariance between brother and sister is \( \frac{1}{4} (1 + F) C_{A_{s}} \).

Now if in the same picture we let \( A \) possess genes (dg), where we have the probability array,

\[
g = a
\]

then

\[
r_{AB}^{g} = \frac{1}{4} \left[ P(d = e) + P(d = f) + P(g = e) + P(g = f) \right]
\]

\[
= \frac{1}{8} (3 + F)
\]

\[
u_{AB}^{g} = \left[ P(d = e, g = f) + P(d = f, g = e) \right]
\]

\[
= \frac{1}{2} (1 + F)
\]

So then the covariance between two sisters is

\[
\frac{1}{4} (3 + F) \sigma^2_{A_{g}} + \frac{1}{2} (1 + F) \sigma^2_{D_{g}}
\]

Now in the parent-offspring situation, the covariances between father and son and father and daughter do not involve the inbreeding coefficient \( F \) and are equal to zero and \( \frac{1}{2} C_{A_{s}} \) respectively.
The covariance between mother and daughter is equal to $\frac{1}{2}(1 + F)\sigma^2_{A_f}$ if we make the assumption that the additive effects are uncorrelated with dominance deviations under any circumstances.

We can extend this situation to the general case, that is, multiple loci case. The partition of the genotypic variance can be accomplished by the same procedure as described in the preceding section. The general structure of the covariances between individuals remains the same as in random mating situation, that is,

$$\text{Cov}(X_m, Y_m)$$

$$= 2 r_{XY} \sigma_A^2 + u_{XY} \sigma_D^2 + v_{XY} \sigma_{A_m}^2 + \sum_{t_1 t_2 t_3} (2 r_{XY})^{t_1} (u_{XY})^{t_2} (v_{XY})^{t_3} \sigma_{A_f} A^{t_1} D^{t_2} A_m^{t_3}$$

$$\geq t_1 + t_2 + t_3 > 1$$

$$\text{Cov}(X_f, Y_f)$$

$$= 2 r_{XY} \sigma_A^2 + u_{XY} \sigma_D^2 + 2 r_{XY} \sigma_{A_f}^2 + u_{XY} \sigma_{D_f}^2$$

$$+ \sum_{t_1 t_2 t_3 t_4 t_5} (2 r_{XY})^{t_1} (u_{XY})^{t_2} (2 r_{XY})^{t_3} (u_{XY})^{t_4} (u_{XY})^{t_5} \sigma_{A_f} A^{t_1} D^{t_2} A_f^{t_3} D_f^{t_4}$$

$$t_1 + t_2 + t_3 + t_4 + t_5 \geq 1$$

$$\text{Cov}(X_m, Y_f)$$

$$= 2 r_{XY} \sigma_A^2 + u_{XY} \sigma_D^2 + w_{XY} \sigma_{A_s}$$
The genotypic covariances of some of the special cases which are likely to be of interest have been worked out with the assumption of both sexlinked, autosomal inheritance and inbreeding. The additive effects, dominance effects and only the first order epistatic effects have been considered important. The numerical values of the coefficients of the various variance and covariance components belonging to the genotypic covariances are presented in Table 8. It is also assumed that the inbreeding coefficient (F) is the same for males and females and for both autosomal and sexlinked loci. If F is not the same for both sexes and for autosomal and sexlinked loci, it is not difficult to make the appropriate modifications in the coefficients presented in Table 8.
Table 8. Values of the coefficients of the variance and covariance components under inbreeding and complete sex-linked inheritance

<table>
<thead>
<tr>
<th>Relations</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sire and son</td>
<td>(\frac{1}{2}(1+F)\sigma_A^2 + \frac{1}{4}(1+F)^2\sigma_{AA}^2)</td>
</tr>
<tr>
<td>Sire and daughter</td>
<td>(\frac{1}{2}(1+F)\sigma_A^2 + \frac{1}{2}C_{As} + \frac{1}{4}(1+F)^2\sigma_{AA}^2 + \frac{1}{4}(1+F)C_{AA_s})</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{4}C_{As})</td>
</tr>
<tr>
<td>Dam and son</td>
<td>(\frac{1}{2}(1+F)\sigma_A^2 + \frac{1}{2}(1+F)C_{As} + \frac{1}{4}(1+F)^2\sigma_{AA}^2)</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{4}(1+F)^2C_{AA_s} + \frac{1}{4}(1+F)^2C_{As})</td>
</tr>
<tr>
<td>Dam and daughter</td>
<td>(\frac{1}{2}(1+F)\sigma_A^2 + \frac{1}{2}(1+F)\sigma_{Af}^2 + \frac{1}{4}(1+F)^2\sigma_{AA}^2)</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{4}(1+F)^2\sigma_{Af}^2 + \frac{1}{4}(1+F)^2\sigma_{Af})</td>
</tr>
<tr>
<td>Brothers</td>
<td>(\frac{1}{2}(1+F)\sigma_A^2 + \frac{1}{4}(1+F)^2\sigma_{D} + \frac{1}{2}(1+F)\sigma_{Am}^2)</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{4}(1+F)^2\sigma_{AA}^2 + \frac{1}{8}(1+F)^3\sigma_{AD} + \frac{1}{16}(1+F)^4\sigma_{DD}^2)</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{4}(1+F)^2\sigma_{AA_m} + \frac{1}{8}(1+F)^3\sigma_{DA_m}^2)</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{4}(1+F)^2\sigma_{Am}^2)</td>
</tr>
<tr>
<td>Sisters</td>
<td>(\frac{1}{2}(1+F)\sigma_A^2 + \frac{1}{4}(1+F)^2\sigma_{Af}^2 + \frac{1}{4}(3+F)\sigma_{Af}^2)</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{2}(1+F)\sigma_{Df}^2 + \frac{1}{4}(1+F)^2\sigma_{AA}^2 + \frac{1}{8}(1+F)^3\sigma_{AD}^2)</td>
</tr>
<tr>
<td></td>
<td>+ (\frac{1}{16}(1+F)^4\sigma_{DD}^2 + \frac{1}{8}(1+F)(3+F)\sigma_{AA}^2)</td>
</tr>
</tbody>
</table>
Table 8. (Continued)

<table>
<thead>
<tr>
<th>Relations</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\frac{1}{4}(1+F)^2\sigma^2_{AD_f} + \frac{1}{16}(1+F)^2(3+F)\sigma^2_{DA_f}$</td>
</tr>
<tr>
<td></td>
<td>$\frac{1}{8}(1+F)^3\sigma^2_{DD_f} + \frac{1}{16}(3+F)^2\sigma^2_{A_fA_f}$</td>
</tr>
<tr>
<td></td>
<td>$\frac{1}{4}(1+F)^2\sigma^2_{D_fD_f}$</td>
</tr>
<tr>
<td>Brother and sister</td>
<td>$\frac{1}{2}(1+F)\sigma^2_A + \frac{1}{4}(1+F)^2\sigma^2_D + \frac{1}{4}(1+F)C_{A_s}$</td>
</tr>
<tr>
<td></td>
<td>$+\frac{1}{8}(1+F)^3\sigma^2_{AD} + \frac{1}{16}(1+F)^4\sigma^2_{DD}$</td>
</tr>
<tr>
<td></td>
<td>$+\frac{1}{8}(1+F)^2C_{A_sA_s} + \frac{1}{16}(1+F)^3C_{DA_s}$</td>
</tr>
<tr>
<td></td>
<td>$+\frac{1}{16}(1+F)^2C_{A_sA_s}$</td>
</tr>
<tr>
<td>Paternal half brothers</td>
<td>$\frac{1}{4}(1+F)\sigma^2_A + \frac{1}{16}(1+F)^2\sigma^2_{AA}$</td>
</tr>
<tr>
<td>Paternal half sisters</td>
<td>$\frac{1}{4}(1+F)\sigma^2_A + \frac{1}{2}\sigma^2_{A_f} + \frac{1}{16}(1+F)^2\sigma^2_{AA}$</td>
</tr>
<tr>
<td></td>
<td>$+\frac{1}{8}(1+F)^3\sigma^2_{AA_f} + \frac{1}{4}\sigma^2_{A_fA_f}$</td>
</tr>
<tr>
<td>Paternal half brothers</td>
<td>$\frac{1}{4}(1+F)\sigma^2_A + \frac{1}{16}(1+F)^2\sigma^2_{AA}$</td>
</tr>
<tr>
<td>and half sisters</td>
<td>$\frac{1}{4}(1+F)^2\sigma^2_{AA_f} + \frac{1}{2}(1+F)\sigma^2_{A_m} + \frac{1}{16}(1+F)^2\sigma^2_{AA}$</td>
</tr>
<tr>
<td>Maternal half brother</td>
<td>$\frac{1}{4}(1+F)^2\sigma^2_{AA_f} + \frac{1}{16}(1+F)^2\sigma^2_{AA}$</td>
</tr>
<tr>
<td></td>
<td>$+\frac{1}{16}(1+F)^2\sigma^2_{AA_m} + \frac{1}{16}(1+F)^2\sigma^2_{A_mA_m}$</td>
</tr>
<tr>
<td>Maternal half sisters</td>
<td>$\frac{1}{4}(1+F)^2\sigma^2_{AA_f} + \frac{1}{4}(1+F)^2\sigma^2_{AA_m}$</td>
</tr>
</tbody>
</table>
Table 8. (Continued)

<table>
<thead>
<tr>
<th>Relations</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal half brothers and half sisters</td>
<td>[ + \frac{1}{16} (1+F)^2 \sigma^2_{AA_f} + \frac{1}{16} (1+F)^2 \sigma^2_{D_f D_f} ]</td>
</tr>
<tr>
<td>Paternal uncle and nephew</td>
<td>[ \frac{1}{4} (1+F) \sigma^2_A + \frac{1}{16} (1+F)^2 \sigma^2_{AA} ]</td>
</tr>
<tr>
<td>Paternal uncle niece</td>
<td>[ \frac{1}{4} (1+F) \sigma^2_A + \frac{1}{4} (1+F) C_{A_s} + \frac{1}{16} (1+F)^2 \sigma^2_{AA} ]</td>
</tr>
<tr>
<td>Paternal aunt nephew</td>
<td>[ \frac{1}{4} (1+F) \sigma^2_A + \frac{1}{16} (1+F)^2 \sigma^2_{AA} ]</td>
</tr>
<tr>
<td>Paternal aunt niece</td>
<td>[ \frac{1}{4} (1+F) \sigma^2_A + \frac{1}{4} (1+F) C_{A_s} + \frac{1}{16} (1+F)^2 \sigma^2_{AA} ]</td>
</tr>
<tr>
<td>Maternal uncle nephew</td>
<td>[ \frac{1}{4} (1+F) \sigma^2_A + \frac{1}{8} (1+F) \sigma^2_{A_m} + \frac{1}{16} (1+F)^2 \sigma^2_{AA} ]</td>
</tr>
<tr>
<td>Maternal uncle niece</td>
<td>[ \frac{1}{4} (1+F) \sigma^2_A + \frac{1}{8} (1+F) C_{A_s} + \frac{1}{16} (1+F)^2 \sigma^2_{AA} ]</td>
</tr>
<tr>
<td>Maternal aunt nephew</td>
<td>[ \frac{1}{4} (1+F) \sigma^2_A + \frac{1}{8} (3+F) C_{A_s} + \frac{1}{16} (1+F)^2 \sigma^2_{AA} ]</td>
</tr>
</tbody>
</table>
### Table 8. (Continued)

<table>
<thead>
<tr>
<th>Relations</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal aunt niece</td>
<td>[ + \frac{1}{32}(1+F)(3+F) C_{AA_s} + \frac{1}{64}(3+F)^2 C_{A_s A_s} ]</td>
</tr>
<tr>
<td>Sire and maternal grand daughter</td>
<td>[ \frac{1}{4}(1+F) \sigma_A^2 + \frac{1}{8}(3+F) \sigma_{A_A}^2 + \frac{1}{16}(1+F)^2 \sigma_{AA}^2 ]</td>
</tr>
<tr>
<td></td>
<td>[ + \frac{1}{32}(1+F)(3+F) \sigma_{AA_f}^2 + \frac{1}{64}(3+F)^2 \sigma_{A_f A_f}^2 ]</td>
</tr>
<tr>
<td>Dam and paternal grand daughter</td>
<td>[ \frac{1}{4}(1+F) \sigma_A^2 + \frac{1}{2}(1+F) \sigma_{A_f}^2 + \frac{1}{16}(1+F)^2 \sigma_{AA}^2 ]</td>
</tr>
<tr>
<td></td>
<td>[ + \frac{1}{8}(1+F)^2 \sigma_{AA_f}^2 + \frac{1}{4}(1+F)^2 \sigma_{A_f A_f}^2 ]</td>
</tr>
</tbody>
</table>
CHAPTER IV. PARTIALLY SEXLINKED GENES

General Considerations

Until 1921, autosomal inheritance and complete sexlinkage constituted the only type of chromosomal inheritance in animals and men. But a third type of inheritance was discovered in experiments with small fresh-water fish (Aplocheilus laptipes) in 1921 by Aida. The traits studied in the fish consisted of various color patterns and were transmitted in a modified sexlinked fashion. The genetic behaviour of these traits was explicable on the theory that a pair of alleles is located in a pair of homologous loci in the sex-chromosome of both females and males. In the female, with its two X-chromosomes, this situation is not different from X-linked genes in the other organisms. In the male, however, the theory proposes that both the X- and Y- chromosomes contain alleles for such traits. Moreover, the facts of transmission had to be explained by the assumption that the X- and Y- chromosomes in the male can exchange their alleles so that the X-linked allele becomes Y-linked, and the Y-linked allele X-linked.

Cytologically it has been observed that during meiosis the X- and Y-chromosomes pair, but only in a limited region of each. Since the pairing of chromosomes is apparently caused by specific attractions between homologous parts, this cytological finding was interpreted to mean that the pairing regions in the X- and Y- chromosomes were genetically homologous and the non-pairing regions genetically non-homologous. The existence of homologous, paired regions fits the genetic theory of homologous loci in X- and Y- chromosomes and leads to the assumption
that homologous loci are present in the paired regions of these chromosomes. The nonpairing, non-homologous regions, also called the differential segments, had to be regarded as the seat of the typical, hemizygous sexlinked genes, the differential segment of the Y-chromosome being occupied by absolutely Y-linked genes, and that of the X-chromosome by complete X-linked genes.

Following the cytological discovery of the pairing segments in X- and Y-chromosomes, Haldane (1936) searched for evidence of incomplete sexlinkage among published human pedigrees. He was successful in establishing, with various degrees of certainty, that the transmission of five characters follows this hitherto unknown type of inheritance in man.

Incomplete sexlinkage is, by definition, easily differentiated from complete X- or Y-linkage. It is more difficult to distinguish incomplete sexlinkage from ordinary autosomal inheritance. The search for incomplete sexlinkage was, accordingly, made among those traits which had, so far, been considered autosomal. If R represents a rare dominant gene located in an autosome, then the offspring from marriages of the type Rr x rr will yield equal numbers of Rr and rr children with random distribution of the two genotypes among sons and daughters. This will be so regardless of whether the mother is Rr and the father rr, or, whether, reciprocally, the mother is rr and the father Rr. If R is an incompletely sexlinked gene and the marriage is Rr female x rr male, the expectation of the equality of Rr and rr among daughters and sons remains unchanged, since the R and the r carrying X-chromosomes of the mother have an equal chance of segregating into eggs which, after fertilization, may develop into either female or male as can be seen in
In the picture above (1) and (2) are daughters and (3) and (4) are sons. If, however, the Rr parent is the male, that is, rr female X Rr male, an association between the sex of the off-spring and the genotypes Rr and rr is to be expected.

The closeness of this association should vary inversely with the frequency with which a crossover interchanges the R and r alleles between the homologous segments of the X- and Y-chromosome of the father. If, in a group of such male parents, R happens to be located in the X-chromosome and r in the Y-chromosome as in picture (I), and if in 70 percent of the sperm cells, R remains in the X-chromosome (X^R) and r in the Y(Y^r), then 70 percent of all the daughters (1) would inherit R and 70 percent of all sons r (3) from their father. The remaining 30 percent of the sperm cells would carry, as a result of crossing over,

the picture below:
either an $X^r$ or a $Y^R$ chromosome so that 30 percent of all the daughters (2) would inherit $r$ and 30 percent of all sons (4) $R$ from their father. If the frequency of crossing over is less than 30 percent, a still higher number of daughters and lower number of sons would inherit $R$, thus approaching the limiting case of complete X-linkage of $R$. If crossover occurs more often than 30 percent, a lower number of daughters and higher number of sons will inherit $R$. Since not more than 50 percent recombinations occur between two loci (at the present state of our knowledge) it may also be assumed that, for the frequency of recombinations between a gene $R$ and the differential end point of the homologous segment, 50 percent represents the upper limit. At this limit, equal numbers of $X^R$ and $X^r$, $Y^R$ and $Y^r$ sperm would be formed, resulting in equal numbers of $Rr$ and $rr$ daughters and sons. Without additional data, such a case would be indistinguishable from autosomal inheritance, even though the $R$, $r$ alleles were carried on the sex chromosomes. Males heterozygous for a partially sexlinked locus occur in two different genotypes. One of these, $X^R Y^r$, has been dealt with above. The other genotype has the constitution $X^r Y^R$, with recessive allele in the $X$- and the dominant allele in the $Y$-chromosome (II). Assuming as before 30 percent of crossing over, the latter genotype would give rise to 70 percent of $X^r$ and $Y^R$ sperm and 30 percent of $X^R$ and $Y^r$ sperm, resulting in 70 percent of $rr$ daughters (5) and $Rr$ sons (7), and in 30 percent of $Rr$ daughters (6) and $rr$ sons (8) - a result opposite in its frequency relations to that expected from first named $X^R Y^r$ genotype of the male parent.
Haldane (1936) examined the data for retinitis pigmentosa and found that about 40 percent of the genes responsible for dominant retinitis pigmentosa are partially sexlinked with a crossover value of about 33 percent. He obtained the mathematical relationship among proportion of crossovers in the total data, frequency of the autosomal and sexlinked genes and the frequency of the crossover value in the latter under certain assumptions and derived a formula to estimate the genetic parameters of interest. Fisher (1936) applied statistical methods to Haldane's data and derived the analysis of variance of the functions of certain statistics, called u-statistics and studied the related tests of significance.

Quantitative Inheritance

The importance of partially sexlinked genes arises from the fact that a quantitative character is influenced not only by the direct effect of these genes but also by the interactions of these genes with the other genes such as autosomal and sexlinked genes. It can be easily visualized that the magnitude of the interaction deviations can be large and a significant portion of the epistatic variability can be explained by these genes in some of the metrical traits of practical interest. The contribution of the partially sexlinked genes to a quantitative character has not been explored so far, probably because of mathematical difficulty involved in handling the mode of transmission of genes particularly their asymmetric frequency relations. Tukey (1954) examined the progress of the panmictic index under brother-sister mating applying the path-coefficient method. He obtained the fundamental difference equation to study the
progress of heterozygosis and derived the transitional mean \((4 \times 4)\) matrix and variance-covariance \((4 \times 4)\) matrix in terms of \(\theta\), where

\[ \theta = \frac{r}{r-1} \quad (r \text{ is the crossover value}) \]

Kempthorne (1957) obtained a \((6 \times 6)\) generation matrix in terms of \(r\) to study the progress of panmictic index under brother-sister mating. He also showed that Malecôt's coefficient of parentage, exclusively meant for autosomal genes, cannot be used in this case. This is what led the author to pursue the matter to obtain the requisite solutions.

In this chapter we shall partition the total genotypic variance and obtain the general structure of the covariances between relatives under the assumption of partially sexlinked transmission. Single locus theory and multiple loci theory will be dealt for random mating and inbreeding and recurrence relations over generations will be obtained to study the progress of panmictic index under regular systems of inbreeding.

Random Mating

Now we proceed to partition the genotypic variance of male and female under the assumption of partially sexlinked inheritance. Let each of the genotypes \(P_1P_j\) and \(P_kP_l\) be assigned to a male and a female respectively. We do not intend to repeat the elaborate procedure of achieving the partition, since the details of this have already been described in the preceding chapter. But in the following, we shall briefly note down the various steps necessary to obtain the required partition,

(a) assignment of genotypic values to male and female,

(b) construction of the algebraic identity,
(c) assignment of meaningful genetic interpretation to each of the terms of the population identity,

(d) obtaining the derived model.

We have the following model, for males,

\[ v_{ij} = \beta_i + \beta_j + \pi_{ij} \]

and for females,

\[ v'_{kl} = \beta'_k + \beta'_l + \pi'_{kl} \]

where \( \beta \neq \beta', \, v \neq v' \) and \( \pi \neq \pi' \).

Note that the gene effects are not assumed to be the same in both sexes. Now we fit \( v_{ij} = \beta_i + \beta_j \) and \( v'_{kl} = \beta'_k + \beta'_l \) by least squares procedure and after solving the normal equations, we have

\[ \beta_g = \sum p_i v_{gi} \]

and

\[ \beta'_h = \sum p_k v'_{hk} \]

We can follow the usual method of proving the orthogonality of the components in the models and then we obtain the variances of the genotypic values and achieve the partitions as follows,

in males

\[ \sigma^2_G_{pm} = \sigma^2_A_{pm} + \sigma^2_D_{pm} \]

and in females

\[ \sigma^2_G_{pf} = \sigma^2_A_{pf} + \sigma^2_D_{pf} \]

Now we proceed to derive the genotypic covariance between relatives in a single locus situation under random mating.


Definition of coefficient of genic relationship

Suppose a random gene from a diploid individual is associated with a random gene from another diploid individual, there are four possible associations, which constitute a set of mutually exclusive events. The probability of each event depends upon the frequency relation existant between the genes at a locus. The autosomal genes in both sexes and the completely sexlinked genes in females exhibit symmetrical probability relations between the two genes at any locus. More general cases like linked autosomal genes and partially sexlinked genes have asymmetrical frequency relation between the genes at a locus. When these associations pertain exclusively to autosomal genes with symmetric frequency relations between the genes at a locus, we define Malecot's coefficient of parentage $r_{XY}$. The name 'Coefficient of genic relationship' is introduced to embrace the cases of autosomal, linked autosomal, completely sexlinked and partially sexlinked genes with both symmetrical and asymmetrical frequency relations between the genes at a locus. Then, the coefficient of genic relationship is defined as the probability that a random gene (autosomal, completely sexlinked, partially sexlinked and linked autosomal genes with either symmetrical or asymmetrical frequency relations) from a diploid individual is identical by descent with a random gene from another diploid individual.

Now we shall examine the following pedigree in which Dam A is in possession of two genes (ab), Sire B has two genes (cd), son C possesses two genes (ef) and the two genes (gh) are possessed by the daughter D.
A general expression of the frequency relation between genes is given in the following probability arrays

\[ e = r_1^1 a + r_2^1 b \]
\[ f = r_1^1 c + r_2^2 d \]
\[ g = r_2^1 a + r_1^1 b \]
\[ h = r_2^2 c + r_1^1 d \]

where

\[ r_1^1 + r_2^2 = 1 \]
and

\[ r_1^1 + r_2^2 = 1 \]

A word of explanation is necessary to exactly understand the quantities \( r_1 \) and \( r_1^1 \). The probability that the Dam gene \( e \) is equal to gene \( a \) is \( r_1^1 \) and that equal to gene \( b \) is \( r_2^2 \); likewise, the probability of the Sire gene \( f \) being equal to gene \( c \) is \( r_1^1 \) and being equal to gene \( d \) is \( r_2^2 \).

If we have,

\[ r_1^1 = r_2^2 = r_1 = r_2 = \frac{1}{2} \]

then, the probability arrays correspond to those in the autosomal case.

If we have,
then, the arrays correspond to completely sexlinked genes. And if we have
\[ r_1' = \frac{1}{2} \]
the probability arrays correspond to partially sexlinked genes. Now if we do not assign any specific values to \( r_1' \) and \( r_2' \), then the probability arrays correspond to the case of autosomal genes with asymmetrical frequency relations of genes at locus caused by gametic preference, random drift, linkage and so on. It is beyond the scope of this study to present a general treatment of the autosomal case and so we shall confine ourselves to the case of partially sexlinked genes only.

Now let \( A_f (= ab) \) and \( B_f (= cd) \) be two females. Then the coefficient of genic relationship between \( A_f \) and \( B_f \) is the probability that a random gene from \( A_f \) and a random gene from \( B_f \) are identical by descent. Symbolically we define,
\[
\text{r}_{AB} = \frac{1}{4} \left[ P(a = c) + P(a = d) + P(b = c) + P(b = d) \right]
\]

Now let \( A_m (= ab) \) and \( B_f (= cd) \) represent a male and a female respectively. Then the coefficient of genic relationship between \( A_m \) and \( B_f \) is the probability that a random gene from \( A_m \) and a random gene from \( B_f \) are identical by descent. Symbolically we define,
\[
\text{w}_{AB} = \frac{1}{4} \left[ P(a = c) + P(a = d) + P(b = c) + P(b = d) \right]
\]

Now let \( A_m (= ab) \) and \( B_m (= cd) \) be two males. Then we define the coefficient of genic relationship of \( A_m \) and \( B_m \) to be the probability
that a random gene from \( A \) and a random gene from \( B \) are identical by descent. Symbolically we define,

\[
v_{AB}^{\ddagger} = \frac{1}{4} \sum P(a = c) + P(a = d) + P(b = c) + P(b = d)
\]

In short then, the symbols \( r^{\ddagger} \), \( w^{\ddagger} \) and \( v^{\ddagger} \) stand for coefficient of genic relationship between two females, female and male and two males respectively. As can be seen, they are the same symbols as used for completely sexlinked genes, except that they have now been primed to identify them when they enter into the general structure of the genotypic covariances. The equality sign in the parenthesis stands for "identical by descent".

We now proceed to derive the general structures of the covariances between relatives under the assumption of random mating at a single locus. The details of the procedure have already been elaborated in connection with completely sexlinked genes, and here we shall briefly follow the same argument and present only the essential mechanics involved. Let \( A \) and \( B \) be two females and suppose we express their genotypic values in terms of additive and dominance effects as follows,

\[
A = \beta_{a_s} + \beta_{a_d} + \pi_{a_s a_d}
\]
\[
B = \beta_{b_s} + \beta_{b_d} + \pi_{b_s b_d}
\]

Then the covariance of \( A \) and \( B \) is given in the following expression,

\[
\frac{1}{4} \sum P(a_s = b_s) + P(a_s = b_d) + P(a_d = b_s) + P(a_d = b_d) \sigma_{A_{\rho f}}^2
\]

\[
+ \sum P(a_s = b_s', a_d = b_d') + P(a_s = b_d', a_d = b_s) \sigma_{D_{\rho f}}^2
\]
Next we consider a male and a female expressed in terms of components of their genotypic values,

\[ A = \beta^I_a s + \beta^I_a d + \pi^I_a s d \]
\[ B = \beta^I_b s + \beta^I_b d + \pi^I_b s d \]

To obtain the covariance between A and B we have to evaluate similar probability expressions. The case of two males in which genotypic values are expressed as follows,

\[ A = \beta^I_a s + \beta^I_a d + \pi^I_a s d \]
\[ B = \beta^I_b s + \beta^I_b d + \pi^I_b s d \]

involves a similar type of calculation. In general the mechanics involved are to express the individuals in terms of the components of their genotypic values and then evaluate the expectation of the product of the components. Three types of product are involved, consisting of the two \( \beta \) terms, or two \( \pi \) terms, or one \( \beta \) and one \( \pi \) term. Then we can construct the tables of possibilities and evaluate the probabilities, and are led to the coefficient of the various variance components involved in the genotypic covariances.

We now present the algebraic structure of the covariances under the assumption of

(a) random mating
(b) single locus segregating
(c) partially sexlinked inheritance.

They are,
\[
\text{Cov}(A_f, B_f) = 2r^*_{AB} \sigma^2_{Apf} + u^*_{AB} \sigma^2_{Dpf}, \\
\text{Cov}(A_f, B_m) = 2w^l_{AB} C_{A\rho} + u^*_{AB} C_{D\rho}, \\
\text{Cov}(A_m, B_m) = 2v^l_{AB} \sigma^2_{Am\rho} + u^*_{AB} \sigma^2_{Dm\rho}. 
\]

where,

- \(\sigma^2_{Apf}\) denotes the variance due to additive partially sexlinked gene effect in females,
- \(\sigma^2_{Dpf}\) denotes the variance due to dominance deviations of partially sexlinked genes in female,
- \(\sigma^2_{Am\rho}\) denotes the variance due to additive effects of partially sexlinked genes in male,
- \(\sigma^2_{Dm\rho}\) denotes the variance due to dominance deviation of partially sexlinked genes in male,
- \(C_{A\rho}\) denotes the covariance between the additive effects of partially sexlinked genes in male and female,
- \(C_{D\rho}\) denotes the covariance between the dominance deviations of partially sexlinked genes in male and female.

The definitions of \(r^*_{AB}\), \(w^l_{AB}\), \(v^l_{AB}\) and \(u^*_{AB}\) will be the same as before. The prime and asterisk on \(u_{AB}\) indicate that the coefficient belongs to the partially sexlinked case.

The genotypic covariance of Sire and daughter is equal to

\[
2w^l_{DB} C_{A\rho} + u^*_{AB} C_{D\rho}
\]
\[
= 2 \frac{1}{4} \left[ \sum P(c=g) + P(c=h) + P(d=h) + P(d=g) \right] C_A \rho \\
+ \sum P(c=g, d=h) + P(c=h, d=g) \right] C_D \rho \\
= 2 \frac{1}{4} \left( r_1 + r_2 \right) C_A \rho \\
= \frac{1}{2} C_A \rho
\]

The genotypic covariance of Dam and son is equal to
\[
2 w^t \frac{A}{C} C_A \rho + u^t \frac{A}{C} C_D \rho \\
= 2 \frac{1}{4} \left[ \sum P(a=e) + P(a=f) + P(b=e) + P(b=f) \right] C_A \rho \\
+ \sum P(a=e, b=f) + P(a=f, b=e) \right] C_D \rho \\
= \frac{1}{2} C_A \rho
\]

The genotypic covariance of Dam and daughter is equal to
\[
2 r^t \frac{A}{D} \frac{A}{D} + u^t \frac{A}{D} \frac{A}{D} \sigma^2 \rho_f \\
= 2 \frac{1}{4} \left[ \sum P(a=g) + P(a=h) + P(b=g) + P(b=h) \right] \sigma^2 \rho_f \\
+ \sum P(a=g, b=h) + P(a=h, b=g) \right] \sigma^2 \rho_f \\
= \frac{1}{2} \sigma^2 \rho_f
\]

Now we shall demonstrate the application of the above formulae by presenting the detailed derivation of the covariances of parent and offspring and full sibs. Let A (ab) be the Dam, B (cd) be the Sire, C (ef) be the son and D (gh) be the daughter. Diagramatically we have,
The probability arrays are

\[ e = \frac{1}{2} a + \frac{1}{2} b \]

\[ f = r_1 c + r_2 d \]

\[ g = \frac{1}{2} a + \frac{1}{2} b \]

\[ h = r_2 c + r_1 d \]

The genotypic covariance of Sire and son is equal to

\[ 2 \nu^t_{BC} \sigma^2_{A_{\rho m}} + u^t_{BC} \sigma^2_{D_{\rho m}} \]

\[ = 2 \cdot \frac{1}{4} \left( P(c = e) + P(c = f) + P(d = e) + P(d = f) \right) \sigma^2_{A_{\rho m}} \]

\[ + \left( P(c = e, d = f) + P(c = f, d = e) \right) \sigma^2_{D_{\rho m}} \]

\[ = 2 \cdot \frac{1}{4} \cdot (r_1 + r_2) \sigma^2_{A_{\rho m}} \]

\[ = \frac{1}{2} \sigma^2_{A_{\rho m}} \]

The genotypic covariance of brother and sister is equal to (the same pedigree is under reference)

\[ 2 \nu^t_{DC} C^t_{A_{\rho}} + u^t_{DC} C_{D_{\rho}} \]

\[ = 2 \cdot \frac{1}{4} \left( P(e = g) + P(e = h) + P(f = g) + P(f = h) \right) C_{A_{\rho}} \]
Now to evaluate the covariance of two brothers consider the following pedigree,

where A (ab) is the Dam, B (cd) is the Sire, and C (ef) and D (gh) are two sons with same Sire and Dam that is C and D are two brothers.

Then the probability arrays are

\[ e = \frac{1}{2} a + \frac{1}{2} b \]
\[ f = r_1 c + r_2 d \]
\[ g = \frac{1}{2} a + \frac{1}{2} b \]
\[ h = r_1 c + r_2 d \]

The genotypic covariance of two brothers is equal to

\[
\frac{1}{2} \left( \frac{1}{2} \right) A_{\rho m}^2 \sigma^2_A + r^2 \left( \frac{1}{2} \right) \sigma^2_D \rho m
\]

\[
= \frac{1}{2} \left( \frac{1}{2} \right) \sum P(e = g) + P(e = h) + P(f = g) + P(f = h) \sigma^2_A \rho m
\]

\[
+ \left( \frac{1}{2} \right) \sum P(e = g, f = g) + P(e = h, f = g) \sigma^2_D \rho m
\]

\[
= \frac{1}{2} \left( \frac{1}{2} \right) \left( r_1^2 + r_2^2 \right) \sigma^2_A \rho m + \frac{1}{2} \left( r_1^2 + r_2^2 \right) \sigma^2_D \rho m
\]
To evaluate the genotypic covariance of two sisters assume that C(ef) and D(gh) are two sisters in the pedigree above, and then we have the following probability arrays,

\[
e = \frac{1}{2}a + \frac{1}{2}b, \quad f = r_c + r_d
\]
\[
g = \frac{1}{2}a + \frac{1}{2}b, \quad h = r_c + r_d
\]

The genotypic covariance between two sisters is equal to

\[
2 r_{CD}^2 \sigma_{\rho_f}^2 + u_{CD}^2 \sigma_{\rho_f}^2
\]
\[
= \frac{1}{2} \left[ P(e = g) + P(e = h) + P(f = g) + P(f = h) \right] \sigma_{A_{\rho_f}}^2
\]
\[
+ \left[ P(e = g, f = h) + P(e = h, f = g) \right] \sigma_{D_{\rho_f}}^2
\]
\[
= \frac{1}{2} \left[ \frac{1}{2} + (r_c^2 + r_d^2) \right] \sigma_{A_{\rho_f}}^2 + \frac{1}{2} (r_c^2 + r_d^2) \sigma_{D_{\rho_f}}^2
\]

Multiple loci theory under random mating will be the next consideration. As usual, we shall make a simultaneous consideration of the autosomal genes and partially sexlinked genes. The general procedure of approaching the problem has been elaborated in the preceding chapter. Here we shall mention only the essential features.

Let the genotype of a male or a female be

\[
n_1 \pi A^c_j A^c_k \pi n_3 \pi P^b_c P^b_b
\]
\[
c = 1, j, \ldots, n_1
\]

where

\[
j \quad \text{and} \quad k \quad \text{represent two autosomal alleles at the locus} \quad c,
\]
\[
c = 1, 2, \ldots, n_1
\]
g and h represent two partially sexlinked alleles at the locus b, 
b = 1, 2, ..., n

P symbolizes the partially sexlinked gene
A symbolizes the autosomal gene
π stands for the consecutive arrangement

Now we shall note down the necessary steps required to partition the
genotypic variance under the assumption of both autosomal and partially
sexlinked inheritance,

(a) evaluation of the genotypic value
(b) construction of an algebraic identity
(c) assignment of meaningful genetic interpretation to 2 (n1 + n3)
terms
(d) proving the orthogonality of each term in the genotypic value
(e) defining the variances of each of the orthogonal components of
the genotypic value
(f) formation of total variances by summing over all the loci.

Now then the net result of these procedures is,

for males,

\[ \sigma^2_{Gm} = \sigma^2_A + \sigma^2_D + \sigma^2_{AA} + \sigma^2_{DD} + \sigma^2_{AD} + \ldots \]
\[ + \sigma^2_{A\rho m} + \sigma^2_{D\rho m} + \sigma^2_{A\rho m\rho m} + \sigma^2_{D\rho m\rho m} + \ldots \]
\[ + \sigma^2_{AA\rho m} + \sigma^2_{AD\rho m} + \sigma^2_{DD\rho m} + \ldots \]

and for females,
The development of the general structures of the covariances between (i) male and male, (ii) female and male and (iii) female and female will be the next consideration under the assumption of both autosomal and partially sexlinked inheritance. An arbitrary degree of relationship between the individuals concerned will be assumed. As usual, there will be two restrictions, (1) no linkage, that is, the loci segregate independently and (2) no inbreeding. The general procedure to be adopted and the arguments necessary to achieve the algebraic structures of the covariances in question are identical with that we have elaborated in the case of completely sexlinked genes. The peculiar situation involved in this particular case is that there is symmetrical frequency relationship between the two genes located in an autosomal locus but an asymmetric probability relation exists between the two genes at a partially sexlinked locus and this situation is taken care of by the coefficients of genic relationship defined earlier.

Now we shall present the results as follows;

(1) Cov(male and male)

\[ \sigma_{G_f}^2 = \sigma_A^2 + \sigma_D^2 + \sigma_{AA}^2 + \sigma_{DD}^2 + \sigma_{AD}^2 + \ldots \]

\[ + \sigma_{A\rho_f}^2 + \sigma_{D\rho_f}^2 + \sigma_{A\rho_f A\rho_f}^2 + \sigma_{D\rho_f D\rho_f}^2 + \ldots \]

\[ + \sigma_{AA\rho_f}^2 + \sigma_{AD\rho_f}^2 + \sigma_{DD\rho_f}^2 + \ldots \]

\[ = \text{Cov}(X_m, Y_m) \]

\[ = 2 r_{XY} \sigma_A^2 + u_{XY} \sigma_D^2 + 2 v_{XY} \sigma_{A \rho_f}^2 + u_{XY} \sigma_{D \rho_f}^2 \]
\[ + \sum \sum \sum \left(2r_{XY}\right)^{t_1}(u_{XY})^{t_2}(2v_{XY})^{t_7}(u_{XY}^t)^{t_8} \sigma^2_{XY} \right) \]
\[ t_1 > t_2 > t_7 > t_8 \]

where \( t_1 + t_2 = n_1 \), the number of autosomal loci

\( t_7 + t_8 = n_3 \), the number of partially sexlinked loci

\( t_1 \) is the number of autosomal additive effects

\( t_2 \) is the number of autosomal dominance deviations

\( t_7 \) is the number of partially sexlinked additive effects

\( t_8 \) is the number of partially sexlinked dominance deviations in male

(2) Cov(female and male)

\[ \text{Cov}(X_f, Y_m) \]
\[ = 2r_{XY} \sigma^2_A + u_{XY}^2 A + 2w_{XY} C_A + u_{XY}^T C_D + u_{XY}^T C_D \]
\[ + \sum \sum \left(2r_{XY}\right)^{t_1}(u_{XY})^{t_2} \sigma^2_{XY} \right) \]
\[ t_1 > t_2 \]

where \( t_1 \) is the number of autosomal additive effects

\( t_2 \) is the number of autosomal dominance deviations
\( t_7 \) is the number of partially sexlinked additive effects

\( t_8 \) is the number of partially sexlinked dominance deviations

(3) Cov(female and female)

\[
\text{Cov}(X^, Y^) = 2r_{XY} \sigma_A^2 + u_{XY} \sigma_D^2 + 2r_{XY} \sigma_A^2 + u_{XY} \sigma_D^2
\]

\[
+ \sum \sum \sum (2r_{XY})^{t_1} (u_{XY})^{t_2} (2r_{XY})^{t_3} (u_{XY})^{t_4} \sigma^2
\]

\[
A^1 D^1 A^2 D^2 A^3 D^3 A^4 D^4
\]

where \( t_7 \) and \( t_8 \) are the number of partially sexlinked additive and dominance effects in females respectively.

Now that we have completed the consideration of the contribution of the three types of genes viz., autosomal, completely sexlinked and partially sexlinked genes, to a quantitative character, we shall, in the following, approach a simultaneous consideration of all the three cases and establish the fundamental genotype, fundamental partition of the genotypic variance and the fundamental structures of the genotypic covariances between (i) male and male (ii) female and male and (iii) female and female in quantitative genetics.

Let the genotype of a male be

\[
\begin{pmatrix}
\pi_{1, a} A^c c & \pi_{2, a} A^c c & \pi_{3, a} A^c c \\
\pi_{4, a} A^c c & \pi_{5, a} A^c c & \pi_{6, a} A^c c \\
\pi_{7, a} A^c c & \pi_{8, a} A^c c & \pi_{9, a} A^c c
\end{pmatrix}
\]

and the female be

\[
\begin{pmatrix}
\pi_{1, a} A^c c & \pi_{2, a} A^c c & \pi_{3, a} A^c c \\
\pi_{4, a} A^c c & \pi_{5, a} A^c c & \pi_{6, a} A^c c \\
\pi_{7, a} A^c c & \pi_{8, a} A^c c & \pi_{9, a} A^c c
\end{pmatrix}
\]
where \( n_1 + n_2 + n_3 = n \), the total number of loci in the individual.

Then adopting the usual procedure for partitioning the genotypic variance, we have for the male

\[
\sigma^2_{G_m} = \sigma^2_A + \sigma^2_D + \sigma^2_{A m} + \sigma^2_{A \rho m} + \sigma^2_{D \rho m} + \ldots
\]

\[
+ \sigma^2_{AA} + \sigma^2_{AD} + \sigma^2_{DD} + \sigma^2_{A m} + \sigma^2_{A \rho m} + \ldots
\]

\[
+ \sigma^2_{A m D \rho m} + \sigma^2_{A m \rho m D \rho m} + \ldots
\]

for the female we have,

\[
\sigma^2_{G_f} = \sigma^2_A + \sigma^2_D + \sigma^2_{A f} + \sigma^2_{D f} + \sigma^2_{A \rho f} + \sigma^2_{D \rho f} + \ldots
\]

\[
+ \sigma^2_{AA} + \sigma^2_{AD} + \sigma^2_{DD} + \sigma^2_{A f} + \sigma^2_{D f} + \ldots
\]

\[
+ \sigma^2_{D f A \rho f} + \sigma^2_{D f D \rho f} + \sigma^2_{A \rho f D f A f D \rho f} + \ldots
\]

Now we shall enumerate the algebraic structure of the covariances,

(1) \( \text{Cov(male and male)} \)

\[
= \text{Cov}(X_m', Y_m)
\]

\[
= 2r_{XY} \sigma^2_A + u_{XY} \sigma^2_D + v_{XY} \sigma^2_{A m} + 2v^1_{XY} \sigma^2_{A \rho m} + \mu_{XY} \sigma^2_D \rho_m
\]

\[
+ \sum \sum \sum \sum (2r_{XY})^{t_1} (u_{XY})^{t_2} (v_{XY})^{t_3} (2v^1_{XY})^{t_4} (u_{XY}^*)^{t_7} (t_{XY})^{t_8} \sigma^2_{A m} + A_{D m} \rho m \rho m \rho m
\]

\[
t_1 t_2 t_3 t_7 t_8 \sigma^2_{A m} + A_{D m} \rho m \rho m \rho m
\]
(2) Cov(female and male)

\[ = \text{Cov}(X_f, Y_m) \]

\[ = 2r_{XY} \sigma^2 + u_{XY} \sigma^2_D + w_{XY} C_A + \]

\[ + 2w_{XY} C_A \rho + u_{XY} C_D \rho + \Sigma (2r_{XY})^t_1 t_1 (u_{XY})^t_1 t_2 \sigma^2_A \]

\[ \cdot A_1 D_2 t_1 + t_2 > 1 \]

\[ + \Sigma \Sigma \Sigma \Sigma (2r_{XY})^t_1 t_1 \cdot \Sigma \Sigma \Sigma \Sigma (2r_{XY})^t_1 t_1 (u_{XY})^t_1 t_2 \cdot t_6 (2w_{XY})^t_6 t_7 (u_{XY})^t_7 t_8 \cdot C_{t_1 t_2 t_6 t_7 t_8} \]

\[ A_1 D_2 A_s A_2 D_2 \rho \]

\[ t_1 + t_2 + t_6 + t_7 + t_8 > 1 \]

(3) Cov(female and female)

\[ = \text{Cov}(X_f, Y_f) \]

\[ = 2r_{XY} \sigma^2 + u_{XY} \sigma^2_D + 2r^{\text{af}}_{XY} \sigma^2_A_f + u^{\text{af}}_{XY} \sigma^2_D_f \]

\[ + 2r^{\text{af}}_{XY} \sigma^2_A_f + u^{\text{af}}_{XY} \sigma^2_D_f \rho_f \]

\[ + \Sigma \Sigma \Sigma \Sigma (2r_{XY})^t_1 t_1 (u_{XY})^t_1 t_2 (2r^{af}_{XY})^t_4 t_5 (u^{af}_{XY})^t_5 t_7 t_8 \]

\[ A_1 D_2 A_f D_f A_2 D_2 \rho_f \rho_f \]

\[ t_1 + t_2 + t_4 + t_5 + t_7 + t_8 > 1 \]

Next we consider the construction of the table of coefficients of the variance components and the covariance components with the assumption that the second order epistatic effects are negligible. Under this assumption the following component of variances and covariances will be involved,
and also the first-order epistatic variance. There are 66 such components and it will be easy to construct the table if the coefficients of the variance and covariance components associated with the partially sexlinked case are given. Table 9 contains the required coefficients, which are given as functions of $r_1$ and $r_2$ to make the results generally applicable.

Inbreeding

In the preceding chapter we have established certain relationships between the quantities $r_{XY}$, $v_{XY}$, and $w_{XY}$ under the assumption of inbreeding and completely sexlinked inheritance. These basic relationships were used to formulate the difference equations over generations to study the progress of the panmictic index in both regular and irregular systems of inbreeding. The advantage of such a procedure is that given a regular or irregular system of inbreeding by a pedigree and having a knowledge of the sexes of individuals in the pedigree and the paths of descent, we can apply the basic relationship of the quantities $r_{XY}$, $v_{XY}$, and $w_{XY}$ and achieve the recurrent relation of generations to study the progress of panmictic index. In other words, we do not have to deal with genes per se but we only deal with the individuals in a pedigree. This is convenient and faster than dealing with individual genes because in a large pedigree it will be a formidable job to keep track of all the genes involved.
Table 9. Genotypic covariances of some of the special cases of interest under the assumption of partially sexlinked inheritance

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sire and son</td>
<td>$\frac{1}{2} \sigma^2_{A\rho m}$</td>
</tr>
<tr>
<td>2. Sire and daughter</td>
<td>$\frac{1}{2} \sigma^2_{A\rho}$</td>
</tr>
<tr>
<td>3. Dam and son</td>
<td>$\frac{1}{2} \sigma^2_{A\rho}$</td>
</tr>
<tr>
<td>4. Dam and daughter</td>
<td>$\frac{1}{2} \sigma^2_{A\rho f}$</td>
</tr>
<tr>
<td>5. Full brothers</td>
<td>$\frac{1}{2} \sum_{n} (r_1^2 + r_2^2) \sigma^2_{A\rho m} + \frac{1}{2} (r_1^2 + r_2^2) \sigma^2_{D\rho m}$</td>
</tr>
<tr>
<td>6. Brother and sister</td>
<td>$\frac{1}{2} (\frac{1}{2} + 2 r_1 r_2) \sigma^2_{A\rho} + r_1 r_2 \sigma^2_{D\rho}$</td>
</tr>
<tr>
<td>7. Full sisters</td>
<td>$\frac{1}{2} \sum_{n} (r_1^2 + r_2^2) \sigma^2_{A\rho f} + \frac{1}{2} (r_1^2 + r_2^2) \sigma^2_{D\rho f}$</td>
</tr>
<tr>
<td>8. Paternal half brother</td>
<td>$\frac{1}{2} (r_1^2 + r_2^2) \sigma^2_{A\rho m}$</td>
</tr>
<tr>
<td>9. Paternal half brother-sister</td>
<td>$r_1 r_2 \sigma^2_{A\rho}$</td>
</tr>
<tr>
<td>10. Paternal half sisters</td>
<td>$\frac{1}{2} (r_1^2 + r_2^2) \sigma^2_{A\rho f}$</td>
</tr>
<tr>
<td>11. Maternal half brother</td>
<td>$\frac{1}{4} \sigma^2_{A\rho m}$</td>
</tr>
<tr>
<td>12. Maternal half sisters</td>
<td>$\frac{1}{4} \sigma^2_{A\rho f}$</td>
</tr>
<tr>
<td>13. Maternal half brother-sister</td>
<td>$\frac{1}{4} \sigma^2_{A\rho}$</td>
</tr>
<tr>
<td>14. Paternal uncle nephew</td>
<td>$\frac{1}{2} \sum_{n} r_1 + r_2 (r_1^2 + r_2^2) \sigma^2_{A\rho m}$</td>
</tr>
</tbody>
</table>
Table 9. (Continued)

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Covariances</th>
</tr>
</thead>
<tbody>
<tr>
<td>15. Paternal uncle niece</td>
<td>( \frac{1}{2} \left( \frac{1}{2} r_2 + r_1 (r_1^2 + r_2^2) \right) C_\rho )</td>
</tr>
<tr>
<td>16. Paternal aunt nephew</td>
<td>( \frac{1}{2} \left( \frac{1}{2} r_1 + 2 r_1 r_2 \right) C_\rho )</td>
</tr>
<tr>
<td>17. Paternal aunt niece</td>
<td>( \frac{1}{2} \left( \frac{1}{2} r_2 + 2 r_1 r_2 \right) \sigma^2 )</td>
</tr>
<tr>
<td>18. Maternal uncle nephew</td>
<td>( \frac{1}{2} \left( \frac{1}{4} + r_1 r_2 \right) C_\rho )</td>
</tr>
<tr>
<td>19. Maternal uncle niece</td>
<td>( \frac{1}{2} \left( \frac{1}{4} + r_1 r_2 \right) \sigma^2 )</td>
</tr>
<tr>
<td>20. Maternal aunt nephew</td>
<td>( \frac{1}{2} \left( \frac{1}{4} + \frac{1}{2} (r_1^2 + r_2^2) \right) C_\rho )</td>
</tr>
<tr>
<td>21. Maternal aunt niece</td>
<td>( \frac{1}{2} \left( \frac{1}{4} + \frac{1}{2} (r_1^2 + r_2^2) \right) \sigma^2 )</td>
</tr>
<tr>
<td>22. Sire and maternal grand daughter</td>
<td>( \frac{1}{4} C_\rho )</td>
</tr>
<tr>
<td>23. Dam and paternal grand daughter</td>
<td>( \frac{1}{2} r_2 \sigma^2 )</td>
</tr>
<tr>
<td>24. Sire and paternal grand daughter</td>
<td>( \frac{1}{2} r_1 C_\rho )</td>
</tr>
</tbody>
</table>
In the case of partially sexlinked genes, due to the asymmetrical frequency relations between the genes at a locus, difficulty is encountered in achieving such relationships among the quantities $r_{XY}^{1}$, $v_{XY}^{1}$ and $w_{XY}^{1}$, previously defined, to be applicable to a wide variety of inbreeding systems. In the absence of such formulae, we have to deal with the genes possessed by the individuals. This will be, of course, a cumbersome process particularly when we are involved with large pedigrees. Here we shall briefly go through the essential steps involved and then we shall exemplify the situation by dealing with certain types of parent-offspring inbreeding system.

Now given a pedigree, we must label the genes of the females as to whether they originate on the Sire side or the Dam side. In case of males, which are assumed to be heterogametic in this study, we know that the gene located in Y-chromosome comes from the Sire and that in the X-chromosome comes from the Dam. Now let two related individuals be denoted by X and Y. The genes possessed by X and the genes possessed by Y have four possible associations. We derive the probability associated with these associations and when we express them as linear equations, we have a system of linear equations when we consider all possible cases. The vector associated with the $n^{th}$ generation can be expressed in terms of the vector associated with the first generation through a transitional matrix. After deriving the canonical roots and the canonical vectors we can obtain a general formula for any particular component in the vector associated with $n^{th}$ generation. The mechanics of this procedure have been detailed by Kempthorne (1957). In the following we shall construct the transitional matrix for parent-offspring inbreeding system of the
following types

(a) parent-offspring recurrent male
(b) parent-offspring recurrent female
(c) parent-offspring alternating male and female

The parameter involved in the derivation is $r$, the cross-over value, so that the frequency of the non-cross-over is $1-r$. We shall designate the former as $r_1$ and the latter as $r_2$ where

$$r_1 + r_2 = 1$$

The quantities we need to calculate are,

(i) the probability of the Sire gene of $X$ being identical by descent with the Sire gene of $Y$,
(ii) the probability of the Dam gene of $X$ being identical by descent with the Sire gene of $Y$,
(iii) the probability of the Sire gene of $X$ being identical by descent with the Dam gene of $Y$,
(iv) the probability of the Dam gene of $X$ being identical by descent with the Dam gene of $Y$,
(v) the probability of two genes possessed by $X$ being identical by descent,
(vi) the probability of two genes possessed by $Y$ being identical by descent.

In the pedigrees below we shall assign two genes to each individual and we shall write the Sire gene on the left and the Dam gene on the right and this convention will be followed throughout.
Now let us consider the parent-offspring inbreeding where the Sire is recurrent. In the pedigree below we have Sire A (ab) mated with a random Dam B (c_1 d_1) and a female offspring C (c_2 d_2) is mated back to Sire A and the resultant offspring D possesses (c_3 d_3) genes, and so on.

Then we have the probability arrays

\[ c_2 = r_1 b + r_2 a, \]
\[ d_2 = \frac{1}{2} c_2 + \frac{1}{2} d_1, \]

so that

\[ P(a = c_2) = r_1 P(a = b) + r_2, \]
\[ P(a = d_2) = \frac{1}{2} P(a = c_1) + \frac{1}{2} P(a = d_1), \]
\[ P(b = c_2) = r_1 + r_2 P(a = b), \]
\[ P(b = d_2) = \frac{1}{2} P(b = c_1) + \frac{1}{2} P(b = d_1), \]
\[ P(c_2 = d_2) = \frac{1}{2} r_1 P(b = c_1) + \frac{1}{2} r_1 P(b = d_1) + \frac{1}{2} r_2 P(a = c_1) + \frac{1}{2} r_2 P(a = d_1). \]
Now if we express the above equations in matrix form, we have,

\[
\begin{bmatrix}
P(a = c_2) \\
P(a = d_2) \\
P(b = c_2) \\
P(b = d_2) \\
P(c_2 = d_2)
\end{bmatrix} =
\begin{bmatrix}
r_2 & r_1 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
r_1 & r_2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & \frac{1}{2}r_2 & \frac{1}{2}r_2 & \frac{1}{2}r_1 & \frac{1}{2}r_1
\end{bmatrix}
\begin{bmatrix}
1 \\
P(a = b) \\
P(a = c_1) \\
P(a = d_1) \\
P(b = c_1) \\
P(b = d_1)
\end{bmatrix}
\]

Also we can see that,

\[
\begin{bmatrix}
P(a = c_3) \\
P(a = d_3) \\
P(b = c_3) \\
P(b = d_3) \\
P(c_3 = d_3)
\end{bmatrix} =
\begin{bmatrix}
r_2 & r_1 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\
r_1 & r_2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & \frac{1}{2}r_2 & \frac{1}{2}r_2 & \frac{1}{2}r_1 & \frac{1}{2}r_1
\end{bmatrix}
\begin{bmatrix}
1 \\
P(a = b) \\
P(a = c_2) \\
P(a = d_2) \\
P(b = c_2) \\
P(b = d_2)
\end{bmatrix}
\]

Now define vector \( \theta_i \) where

\[
\theta_i = \begin{bmatrix}
1 \\
P(a = b) \\
P(a = c_1) \\
P(a = d_1) \\
P(b = c_1) \\
P(b = d_1) \\
P(c_1 = d_1)
\end{bmatrix}
\]
Then

\[
\begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
r_2 & r_1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 \\
r_1 & r_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0 \\
0 & 0 & 1/2r_2 & 1/2r_2 & 1/2r_1 & 1/2r_1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
P(a=b) \\
P(a=c) \\
P(a=d) \\
P(b=c) \\
P(b=d) \\
P(c_2=d_2)
\end{bmatrix}
= 
\begin{bmatrix}
P(a=b) \\
P(a=c) \\
P(a=d) \\
P(b=c) \\
P(b=d) \\
P(c_1=d_1)
\end{bmatrix}
\]

Now using matrix notation we have

\[
\theta_2 = T \theta_1
\]

or, \[\theta_2 = T^2 \theta_1\]

or, \[\theta_n = T^n \theta_1\]

where the matrix \(T\) is the transitional matrix. Now we can choose a range of values of \(r_1\) from 0 to \(1/2\) and for each such value of \(r_1\) we can obtain the canonical roots and canonical vectors and hence obtain a general formulation for \(\theta_n\) and solve for any particular component of \(\theta\) which is of interest. In this case we shall be interested in finding the solution for \(P(c_1 = d_1)\) since this defines the inbreeding coefficient of the female in the \(i^{th}\) generation.

Now we examine the consequences of having

\[r_1 = r_2 = \frac{1}{2}\]

in the transitional matrix.
Define,
\[ P(a = b) = F_n \]
\[ P(c_2 = d_2) = F_{n+1} \]
\[ P(c_3 = d_3) = F_{n+3} \]
in the pedigree above.

Clearly we see that,
\[ P(a = c_2^2) = P(b = c_2^2) = \frac{1}{2} (1 + F_n) \]

where
\[ c_2 = r_1 b + r_2 a \quad (1) \]
\[ F_{n+1} = P(c_2 = d_2) = \frac{1}{2} \sum P(a = d_2) + P(b = d_2) \quad \text{from (1)} \]
or,
\[ P(a = d_2^2) + P(b = d_2^2) = 2F_{n+1} \quad (2) \]
\[ F_{n+2} = P(c_3 = d_3) = \frac{1}{4} \sum P(a = c_2^2) + P(b = c_2^2) + P(a = d_2^2) + P(b = d_2^2) \]
\[ = \frac{1}{4} \sum 2P(a = c_2^2) + 2F_{n+1} \quad \text{from (2)} \]
\[ = \frac{1}{4} \sum 1F_n + 2F_{n+1} \]
or,
\[ F_{n+2} = \frac{1}{2} F_{n+1} + \frac{1}{4} F_n + \frac{1}{4} \]

This result, as expected, conforms to the recurrence relation of generations under the assumption of autosomal inheritance. So the whole body of difference equations can be "absorbed" and a simple recurrence relation can be established, when we have symmetrical relationship between \( r_1 \) and \( r_2 \).

The next case we consider is parent-offspring inbreeding with
recurrent female. In the pedigree \(A(a_1b_1)\) is the Sire and \(B(cd)\) is the Dam and the male offspring \(C\) has two genes \((a_2b_2)\) and it is mated back to the Dam \(B\) and the resultant female offspring is \(D(a_3b_3)\).

\[
\begin{align*}
A_m & \quad (a_1b_1) \\
C_m & \quad (a_2b_2) \\
B_f & \quad (cd) \\
D_m & \quad (a_3b_3)
\end{align*}
\]

The probability arrays are,

\[
a_2 = r_1a_1 + r_2b_1 \\
b_2 = \frac{1}{2}c + \frac{1}{2}d
\]

Then we have,

\[
\begin{align*}
P(c = a_2) &= r_1 P(c = a_1) + r_2 P(c = b_1) \\
P(c = b_2) &= \frac{1}{2} + \frac{1}{2} P(c = d) \\
P(d = a_2) &= r_1 P(d = a_1) + r_2 P(d = b_1) \\
P(d = b_2) &= \frac{1}{2} P(c = d) + \frac{1}{2}
\end{align*}
\]

\[
P(a_2 = b_2) = \frac{1}{2} r_1 P(c = a_1) + \frac{1}{2} r_2 P(c = b_1) + \frac{1}{2} r_1 P(d = a_1) + \frac{1}{2} r_2 P(d = b_1)
\]

Now expressing the above relations in matrix form, we have the recurrence relation,
\[
\begin{bmatrix}
P(c = a_2) \\
P(c = b_2) \\
P(d = a_2) \\
P(d = b_2) \\
P(a_2 = b_2)
\end{bmatrix} =
\begin{bmatrix}
0 & 0 & r_1 & r_2 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & r_1 & r_2 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}r_1 & \frac{1}{2}r_2 & \frac{1}{2}r_1 & \frac{1}{2}r_2
\end{bmatrix}
\begin{bmatrix}
1 \\
P(c = d) \\
P(c = a_1) \\
P(c = b_1) \\
P(d = a_1) \\
P(d = b_1) \\
P(c_1 = d_1)
\end{bmatrix}
\]

Now define,
\[
\Theta_2 =
\begin{bmatrix}
1 \\
P(c = d) \\
P(c = a_2) \\
P(c = b_2) \\
P(d = a_2) \\
P(d = b_2) \\
P(c_2 = d_2)
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & r_1 & r_2 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & r_1 & r_2 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \frac{1}{2}r_1 & \frac{1}{2}r_2 & \frac{1}{2}r_1 & \frac{1}{2}r_2 & 0
\end{bmatrix}
\begin{bmatrix}
1 \\
P(c = d) \\
P(c = a_1) \\
P(c = b_1) \\
P(d = a_1) \\
P(d = b_1) \\
P(c_1 = d_1)
\end{bmatrix}
\]
Now using the matrix notation we can express the recurrence relation as follows,

\[ \theta_2 = T \theta_1 \]

so \[ \theta_3 = T^2 \theta_2 \]

and \[ \theta_n = T^n \theta_1 \]

where the matrix \( T \) is the transitional matrix. After we get the canonical roots and canonical vectors by assigning certain numerical values to \( r_1 \), we can solve for any components of the vector \( \theta_n \). In this case we will be interested in obtaining solutions for \( P(c_2 = d_2) \) since, by definition, it gives the inbreeding coefficient of \( C_m \) that is \( F_{C_m} \).

When we have

\[ r_1 = r_2 = \frac{1}{2} \]

the result conforms to the result under autosomal inheritance. The details have been worked out in the preceding example. Finally we consider the case of parent-offspring inbreeding with alternating male and female. In the pedigree below male \( A(a_1, b_1) \) mates with female \( B(c_1, d_1) \) and their female offspring is \( C(c_2, d_2) \) and then \( A \) mates with \( C \) and the resultant male offspring is \( Da_2, b_2 \). Male \( D(a_2, b_2) \) mates with female \( C(c_2, d_2) \) and the resultant female offspring is \( E(c_3, d_3) \) and so on.
The probability arrays are

\[ c_2 = r_2 a_1 + r_1 b_1 \]
\[ d_2 = \frac{1}{2} c_1 + \frac{1}{2} d_1 \]
\[ a_2 = r_1 a_1 + r_2 b_1 \]
\[ b_2 = \frac{1}{2} c_1 + \frac{1}{2} d_1 \]
\[ = \frac{1}{2} r_2 a_1 + \frac{1}{2} r_1 b_1 + \frac{1}{4} c_1 + \frac{1}{4} d_1 \]

Now evaluating the probabilities, we have,

\[ P(a_2 = c_2) = r_1 r_2 + r_1^2 P(a_1 = b_1) + r_2^2 P(a_1 = b_1) + r_1 r_2 \]
\[ P(a_2 = d_2) = \frac{1}{2} r_1 P(a_1 = c_1) + \frac{1}{2} r_1 P(a_1 = d_1) + \frac{1}{2} r_2 P(b = c_1) + \frac{1}{2} r_2 P(b = d_1) \]
\[
\begin{align*}
\Pr(b_2 = c_2) &= \frac{1}{2} r_2^2 + \frac{1}{2} r_1 r_2 \Pr(a_1 = b_1) + \frac{1}{4} r_2^2 \Pr(a_1 = c_1) + \frac{1}{4} r_2^2 \Pr(a_1 = d_1) \\
&\quad + \frac{1}{2} r_1 r_2 \Pr(a_1 = b_1) + \frac{1}{2} r_2^2 + \frac{1}{4} r_1 \Pr(b_1 = c_1) + \frac{1}{4} r_1 \Pr(b_1 = d_1) \\
\Pr(b_2 = d_2) &= \frac{1}{4} r_2 \Pr(a_1 = c_1) + \frac{1}{4} r_1 \Pr(b_1 = c_1) + \frac{1}{8} \Pr(c_1 = d_1) \\
&\quad + \frac{1}{4} r_2 \Pr(a_1 = d_1) + \frac{1}{4} r_1 \Pr(b_1 = d_1) + \frac{1}{8} \Pr(c_1 = d_1) + \frac{1}{8} \\
\Pr(c_2 = d_2) &= \frac{1}{2} r_2 \Pr(a_1 = c_1) + \frac{1}{2} r_1 \Pr(b_1 = c_1) + \frac{1}{2} r_2 \Pr(a_1 = d_1) + \frac{1}{2} r_1 \Pr(b_1 = d_1) \\
\Pr(a_2 = b_2) &= \frac{1}{2} r_2^2 \Pr(a_1 = b_1) + \frac{1}{2} r_1 r_2 + \frac{1}{2} r_2 \Pr(a_1 = c_1) + \frac{1}{4} r_2 \Pr(a_1 = d_1) \\
&\quad + \frac{1}{2} r_2^2 \Pr(a_1 = b_1) + \frac{1}{2} r_1 r_2 + \frac{1}{2} r_2 \Pr(b_1 = c_1) + \frac{1}{4} r_2 \Pr(b_1 = d_1) \\
\end{align*}
\]

Now defining \( \theta \) as usual, we have,

\[
\begin{bmatrix}
1 \\
\Pr(a_2 = c_2) \\
\Pr(a_2 = d_2) \\
\Pr(b_2 = c_2) \\
\Pr(b_2 = d_2) \\
\Pr(a_2 = b_2) \\
\Pr(c_2 = d_2)
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 r_2 & 0 & 0 & 0 & 0 & r_2^2 + r_2^2 & 0 \\
0 & \frac{1}{2} r_1 & \frac{1}{2} r_1 & \frac{1}{2} r_2 & \frac{1}{2} r_2 & 0 & 0 \\
\frac{1}{2} (r_1^2 + r_2^2) & \frac{1}{4} r_2 & \frac{1}{4} r_2 & \frac{1}{4} r_1 & \frac{1}{4} r_1 & \frac{1}{2} r_1 & \frac{1}{2} r_2 \\
\frac{1}{4} & \frac{1}{4} r_2 & \frac{1}{4} r_2 & \frac{1}{4} r_1 & \frac{1}{4} r_1 & 0 & \frac{1}{4} \\
r_1 r_2 & \frac{1}{4} r_1 & \frac{1}{4} r_2 & \frac{1}{4} r_1 & \frac{1}{4} r_1 & \frac{1}{2} (r_1^2 + r_2^2) & 0 \\
0 & \frac{1}{2} r_2 & \frac{1}{2} r_2 & \frac{1}{2} r_1 & \frac{1}{2} r_1 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
1 \\
\Pr(a_1 = c_1) \\
\Pr(a_1 = d_1) \\
\Pr(b_1 = c_1) \\
\Pr(b_1 = d_1) \\
\Pr(a_1 = b_1) \\
\Pr(c_1 = d_1)
\end{bmatrix}
\]

Now expressing the recurrence relation in matrix notation we have

\[
\theta_2 = T \theta_1
\]

or, \( \theta_3 = T^2 \theta_2 \)

or, \( \theta_n = T^n \theta_1 \)
where the matrix $T$ is the transitional matrix. After we get the canonical roots and canonical vectors with particular values of $r$ ranging from 0 to $\frac{1}{2}$, we can solve for any components of the vector $\theta_n$. In this case, in fact, we will be interested in achieving solutions for $P(a_1 = b_1)$ and $P(c_1 = d_1)$ and by definition they are $F_{D_m}$ and $F_{C_f}$ respectively. We can reduce these difference equations to the form

$$F_{n+2} = \frac{1}{2}F_{n+1} + \frac{1}{4}F_n + \frac{1}{4}$$

when we have $r_1 = r_2 = \frac{1}{2}$.

**Covariance between relatives** We have indicated in the preceding chapter that the correlations between relatives under specified systems of inbreeding have been obtained by Kempthorne (1955) and Horner (1956) and the results are unpleasantly complex. No attempt will be made to proceed in the same direction to obtain more results in this study. The population we shall be concerned with is derived by inbreeding a random mating population to the extent measured by $F$. It will be of great practical interest to study the covariance between random members of such a population. This study will be made under the assumption of partially sexlinked inheritance and so there will be two inbreeding coefficients to be defined - one for the male and the other for the female. In the completely sexlinked case the heterogametic sex has $F = 0$. If $X(ab)$ is a female and $Y(cd)$ is a male, then we define,

$$P(a = b) = F_f$$

$$P(c = d) = F_m$$
where \( F_f \) indicates the coefficient of inbreeding for female and \( F_m \) indicates the coefficient of inbreeding for male. To keep the derivation of the covariances general we will use both the symbols and if it is assumed that

\[
F_f = F_m
\]

the algebraic simplification of the expressions will not be difficult.

Now we proceed to express the individuals in question in terms of additive and dominance effects and go through the usual procedure and arguments to derive the general structure of the covariances between (i) male and male, (ii) female and male and (iii) female and female. The covariances have the same structure as given in the preceding section, that is

\[
\text{Cov}(X_m', Y_m') = 2 \sigma_{XY}^2 \sigma_{\rho m}^2 + \sigma_{XY}^2 \sigma_{\rho m}^2
\]

\[
\text{Cov}(X_f, Y_m) = 2 \sigma_{XY}^2 C_{\rho} + \sigma_{XY}^2 C_{\rho}
\]

\[
\text{Cov}(X_f, Y_f) = 2 \sigma_{XY}^2 \sigma_{\rho f}^2 + \sigma_{XY}^2 \sigma_{\rho f}^2
\]

In the following we shall examine the application of above formulae as applied to the members of the population under consideration. We shall illustrate the situation by presenting the detailed derivations of the parent-offspring and full-sib covariances. Consider the following pedigree with \( A(ab) \) as the Dam, \( B(cd) \) as the Sire, \( C(ef) \) as the son and \( D(gh) \) as the daughter.
We have the following probability arrays,

\[
\begin{align*}
e &= \frac{1}{2} a + \frac{1}{2} b \\
f &= r_1 c + r_2 d \\
g &= \frac{1}{2} a + \frac{1}{2} b \\
h &= r_2 c + r_1 d \\
F_f &= P(a = b) \\
F_m &= P(c = d)
\end{align*}
\]

The genotypic covariance of Sire and son is equal to

\[
2v^t_{BC} \sigma^2_{A_{\rho m}} + u^t_{BC} \sigma^2_{D_{\rho m}}
\]

\[
= \frac{1}{2} \left[ P(c = e) + P(c = f) + P(d = e) + P(d = f) \right] \sigma^2_{A_{\rho m}}
\]

\[
+ \left[ P(c = e, d = f) + P(c = f, d = e) \right] \sigma^2_{D_{\rho m}}
\]

\[
= \frac{1}{2} \left[ r_1 + r_2 \right] + (r_1 + r_2) F_m \sigma^2_{A_{\rho m}}
\]

\[
= \frac{1}{2} \left( 1 + F_m \right) \sigma^2_{A_{\rho m}}
\]

The genotypic covariance of Sire and daughter is equal to
\[ 2^w_{DB} C_A^t + u^1_{DB} C_D^t \]

\[ = \frac{1}{2} \left[ P(c = g) + P(c = h) + P(d = g) + P(d = h) \right] C_A^t \]

\[ + \left[ P(c = g, d = h) + P(c = h, d = g) \right] C_D^t \]

\[ = \frac{1}{2} \left( 1 + F_m \right) C_A^t \]

The genotypic covariance of Dam and son is equal to

\[ 2^w_{AC} C_A^t + u^1_{AC} C_D^t \]

\[ = \frac{1}{2} \left[ P(a = e) + P(a = f) + P(b = e) + P(b = f) \right] C_A^t \]

\[ + \left[ P(a = e, b = f) + P(a = f, b = e) \right] C_D^t \]

\[ = \frac{1}{2} \left( 1 + F_f \right) C_A^t \]

The genotypic covariance of Dam and daughter is equal to

\[ 2^w_{AD} \sigma^2_{A_{pf}} + u^1_{AD} \sigma^2_{D_{pf}} \]

\[ = \frac{1}{2} \left[ P(a = g) + P(a = h) + P(b = g) + P(b = h) \right] \sigma^2_{A_{pf}} \]

\[ + \left[ P(a = g, b = h) + P(a = h, b = g) \right] \sigma^2_{D_{pf}} \]

\[ = \frac{1}{2} \left( 1 + F_f \right) \sigma^2_{A_{pf}} \]

We shall now proceed with the derivation of the covariances of full sibs.

Assume that C and D are two brothers in the pedigree given above, then the probability arrays are as follows,
\[ e = \frac{1}{2} a + \frac{1}{2} b \]
\[ f = r_1 c + r_2 d \]
\[ g = \frac{1}{2} a + \frac{1}{2} b \]
\[ h = r_1 c + r_2 d \]
\[ F_f = P(a = b) \]
\[ F_m = P(c = d) \]

The genotypic covariance of two brothers is equal to

\[ 2 \nu_{CD} \sigma_{A_{\rho m}}^2 + u_{\nu_{CD}} \sigma_{D_{\rho m}}^2 \]

\[ = \frac{1}{2} \left[ \sum P(e = g) + P(e = h) + P(f = g) + P(f = h) \right] \sigma_{A_{\rho m}}^2 \]

\[ + \sum P(e = g, f = h) + P(e = h, f = g) \sigma_{D_{\rho m}}^2 \]

\[ = \frac{1}{2} \left( 1 + F_f \right) + \left( r_1^2 + r_2^2 \right) + 2 r_1 r_2 F_m \sigma_{A_{\rho m}}^2 \]

\[ + \sum \left( 1 + F_f \right) \left( r_1^2 + r_2^2 + 2 r_1 r_2 F_m \right) \sigma_{D_{\rho m}}^2 \]

\[ = \frac{3}{4} + \frac{1}{4} F_f - r_1 r_2 \left( 1 - F_m \right) \sigma_{A_{\rho m}}^2 \]

\[ + \frac{3}{2} \left( 1 + F_f \right) - r_1 r_2 \left( 1 + F_f \right) \left( 1 - F_m \right) \sigma_{D_{\rho m}}^2 \]

If in the above pedigree we assume that C and D are two sisters, then we have the following probability arrays,

\[ e = \frac{1}{2} a + \frac{1}{2} b \]
\[ f = r_2 c + r_1 d \]
The genotypic covariance of two sisters has the same structure as that of the two brothers, that is,

\[ \frac{3}{4} + \frac{1}{4} F_f - r_1 r_2 (1 - F_m) \mathcal{J} \sigma^2_{A_{pm}} + \mathcal{J} \left( 1 + F_f \right) - r_1 r_2 (1 + F_f) (1 - F_m) \mathcal{J} \sigma^2_{D_{pm}} \]

In the pedigree above consider \( C \) and \( D \) as the brother and the sister, and the probability arrays are as follows,

\[ e = \frac{1}{2} a + \frac{1}{2} b \]
\[ f = r_1 c + r_2 d \]
\[ g = \frac{1}{2} a + \frac{1}{2} b \]
\[ h = r_2 c + r_1 d \]
\[ F_f = P(a = b) \]
\[ F_m = P(c = d) \]

The genotypic covariance of brother and sister is equal to

\[ 2^{w^I_{DC}} \sigma^2 \mathcal{J} C_{A_{\rho}} + 2^{w^J_{DC}} \sigma^2 \mathcal{J} C_{D_{\rho}} \]

\[ = \frac{1}{2} \mathcal{J} P(e = g) + P(e = h) + P(f = g) + P(f = h) \mathcal{J} C_{A_{\rho}} \]
\[ + \mathcal{J} P(e = g, f = h) + P(e = h, f = g) \mathcal{J} C_{D_{\rho}} \]
In the case of multiple loci theory, we can make a simultaneous consideration of autosomal, completely sexlinked and partially sexlinked genes and derive the general structures of the covariances of (i) male and male (ii) female and male and (iii) female and female. It would be seen that the results so obtained under the assumption of inbreeding followed by random mating conforms to that of the results obtained under random mating. So we shall not repeat the derivation here. Now if we consider the construction of the table of coefficients of the variance and covariance components under the assumption of no epistacy, we have twelve such coefficients. But if we assume that the effects of second order epistatic effects are negligible that is if we assume that the main effects and first order epistatic variances are important, we have 66 such components and hence 66 coefficients. But in the preceding chapter we have demonstrated in the table the principle involved in construction of such tables under the assumption of no linkage that is when independent conditions holds. So in the following we are presenting a table of covariances algebraically expressed under the assumption of inbreeding and partially sexlinked inheritance. These coefficients can be utilized while constructing tables of coefficients assuming the presence of any degree of epistacy. In the Table 10 the coefficients are expressed in terms of $r_1$, $r_2$, $F_f$ and $F_m$, where $r_1 + r_2 = 1$ and $F_f$ and $F_m$ have non-zero values.
Table 10. Genotypic covariance of some of the special cases of interest under the assumption of partially sexlinked inheritance and inbreeding

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Sire and son</td>
<td>( \frac{1}{2} (1 + F_m) \sigma^2_{A_{\rho m}} )</td>
</tr>
<tr>
<td>2. Sire and daughter</td>
<td>( \frac{1}{2} (1 + F_m) C_{A_{\rho}} )</td>
</tr>
<tr>
<td>3. Dam and son</td>
<td>( \frac{1}{2} (1 + F_f) C_{A_{\rho}} )</td>
</tr>
<tr>
<td>4. Dam and daughter</td>
<td>( \frac{1}{2} (1 + F_f) \sigma^2_{A_{\rho f}} )</td>
</tr>
<tr>
<td>5. Full brothers</td>
<td>( \frac{3}{4} + \frac{1}{4} - r_1 r_2 (1 - F_m) \sum \sigma^2_{A_{\rho m}} ) + ( \frac{1}{2} (1 + F_f) - r_1 r_2 (1 + F_f) (1 - F_m) \sum \sigma^2_{D_{\rho m}} )</td>
</tr>
<tr>
<td>6. Brother and sister</td>
<td>( \frac{3}{4} + \frac{1}{4} F_f - \frac{1}{2} (r_1^2 + r_2^2) (1 - F_m) \sum C_{A_{\rho}} ) + ( \frac{1}{2} (1 + F_f) - \frac{1}{2} (r_1^2 + r_2^2) (1 + F_f) (1 - F_m) \sum C_{D_{\rho}} )</td>
</tr>
<tr>
<td>7. Full sisters</td>
<td>( \frac{3}{4} + \frac{1}{4} F_f - r_1 r_2 (1 - F_m) \sum \sigma^2_{A_{\rho f}} ) + ( \frac{1}{2} (1 + F_f) - r_1 r_2 (1 + F_f) (1 - F_m) \sum \sigma^2_{D_{\rho f}} )</td>
</tr>
<tr>
<td>8. Paternal half brothers</td>
<td>( \frac{1}{2} - r_1 r_2 (1 - F_m) \sum \sigma^2_{A_{\rho m}} )</td>
</tr>
<tr>
<td>9. Paternal half brother - sister</td>
<td>( \frac{1}{2} \sum_1 - (r_1^2 + r_2^2) (1 - F_m) \sum C_{A_{\rho}} )</td>
</tr>
<tr>
<td>10. Paternal half sisters</td>
<td>( \frac{1}{2} - r_1 r_2 (1 - F_m) \sum \sigma^2_{A_{\rho f}} )</td>
</tr>
<tr>
<td>Relationship</td>
<td>Covariance</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>11. Maternal half brother</td>
<td>$\frac{1}{4} (1 + F_f) \sigma^2_{A_{\rho, m}}$</td>
</tr>
<tr>
<td>12. Maternal half sisters</td>
<td>$\frac{1}{4} (1 + F_f) \sigma^2_{A_{\rho, f}}$</td>
</tr>
<tr>
<td>13. Maternal half brother-sister</td>
<td>$\frac{1}{4} (1 + F_f) C_{A_{\rho}}$</td>
</tr>
<tr>
<td>14. Paternal uncle nephew</td>
<td>$\frac{1}{2} \frac{1}{2} r_1^1 (1+F_f) + r_2^2 (r_1^2 + r_2^2) + 2r_1 r_2^2 F_m \mathcal{J} \sigma^2_{A_{\rho, m}}$</td>
</tr>
<tr>
<td>15. Paternal uncle niece</td>
<td>$\frac{1}{2} \frac{1}{2} r_1^1 (1+F_f) + r_2^2 (r_1^2 + r_2^2) + 2r_1 r_2^2 F_m \mathcal{J} C_{A_{\rho}}$</td>
</tr>
<tr>
<td>16. Paternal aunt nephew</td>
<td>$\frac{1}{2} \frac{1}{2} r_1^1 (1+F_f) + 2r_1 r_2^2 + r_1^2 r_2^2 F_m \mathcal{J} C_{A_{\rho}}$</td>
</tr>
<tr>
<td>17. Paternal aunt niece</td>
<td>$\frac{1}{2} \frac{1}{2} r_1^1 (1+F_f) + 2r_1 r_2^2 + r_1^2 r_2^2 F_m \mathcal{J} \sigma^2_{A_{\rho, f}}$</td>
</tr>
<tr>
<td>18. Maternal uncle nephew</td>
<td>$\frac{1}{2} \frac{1}{4} (1+F_f) + r_1 r_2^2 + \frac{1}{2} (r_1^2 + r_2^2) F_m \mathcal{J} \sigma^2_{A_{\rho, m}}$</td>
</tr>
<tr>
<td>19. Maternal uncle niece</td>
<td>$\frac{1}{2} \frac{1}{4} (1+F_f) + r_1 r_2^2 + \frac{1}{2} (r_1^2 + r_2^2) F_m \mathcal{J} C_{A_{\rho}}$</td>
</tr>
<tr>
<td>20. Maternal aunt nephew</td>
<td>$\frac{1}{2} \frac{1}{4} (1+F_f) + \frac{1}{2} (r_1^2 + r_2^2) + r_1 r_2 F_m \mathcal{J} C_{A_{\rho}}$</td>
</tr>
<tr>
<td>21. Maternal aunt niece</td>
<td>$\frac{1}{8} \mathcal{L} (1+F_f) + 4(r_1^2 + r_2^2) + 8 r_1 r_2 F_m \mathcal{J} \sigma^2_{A_{\rho, f}}$</td>
</tr>
<tr>
<td>22. Sire and maternal grand daughter</td>
<td>$\frac{1}{4} (1 + F_m) C_{A_{\rho}}$</td>
</tr>
<tr>
<td>23. Dam and paternal grand daughter</td>
<td>$\frac{1}{2} \frac{1}{2} r_2^2 (1 + F_f) \mathcal{J} \sigma^2_{A_{\rho, f}}$</td>
</tr>
</tbody>
</table>
Table 10. (Continued)

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Covariance</th>
</tr>
</thead>
<tbody>
<tr>
<td>24. Sire and paternal grand daughter</td>
<td>( \frac{1}{2} \gamma_1 (1 + F_m) ) ( \gamma ) ( C_{A_p} )</td>
</tr>
</tbody>
</table>
CHAPTER V. GENE EFFECTS DEPENDING ON SEX

General Considerations

The mathematical model usually adopted to derive the partition of the genotypic variance and covariance between relatives under the assumption of autosomal inheritance is the following,

\[ y_{ij} = \mu + a_i + a_j + e_{ij} \]

where,

- \( y_{ij} \) is the genotypic response of individual \( A_iA_j \)
- \( \mu \) is the genotypic mean of the population
- \( a_i \) is the effect of gene \( A_i \)
- \( e_{ij} \) is the dominance deviation

Now by applying the least square procedure we obtain the value of \( a_k \) in terms of \( y_{ik} \) and so on, and have the following partition of the total genotypic variance,

\[ \sigma_G^2 = \sigma_A^2 + \sigma_D^2 \]

and the algebraic expression of the genotypic covariance of two random individuals irrespective of their sexes, is

\[ 2r_{XY} \sigma_A^2 + u_{XY} \sigma_D^2 \]

The cardinal assumption that has been made in this derivation is that the gene-effects are independent of sex. The physiological make-up of the sexes are such that one would expect that the same gene exerts different influence in two sexes and hence the gene-effects are different in different sexes. This can be true for autosomal genes, completely sexlinked
genes and partially sexlinked genes. In quantitative genetics, then, if an autosomal gene $A$ has an effect $z_m$ (measured in some units) in males and an effect $z_f$ in females, the general case would be to assume

$$z_m \neq z_f$$

rather than

$$z_m = z_f$$

In the following we shall present two biological examples based on qualitative traits.

**Biological Evidence**

In Ayrshire cattle the animals are spotted either red-white or mahogany-white. Crosses between homozygous red-white and homozygous mahogany-white animals result in peculiar and interesting ratios as shown below,

Parent: mahogany-white $\times$ red-white  
(male)  (female)

$F_1$:  
males : mahogany-white  
males : red-white

$F_2$:  
males : 3 mahogany-white and 1 red-white  
males : 3 red-white and 1 mahogany-white

The reciprocal crosses of red-white male with mahogany-white female give identical results. It may be seen that the character is in some way associated with sex, yet it is not sexlinked. If we consider only the males we see that mahogany-white acts as a dominant, since the $F_1$ males are mahogany-white and the $F_2$ males show a ratio of three
mahogany-white to one red-white. But when we consider only the females red-white appears to be dominant, as the $F_1$ females are all red-white, while the $F_2$ females give a ratio of three red-white to one mahogany-white. Of this pair of alleles one seems to be dominant in males and the other dominant in females.

In sheep the inheritance of horns is of the same nature. In the Dorset-breed both sexes are horned, whereas in the Suffolk-breed neither sex is horned. A cross between these breeds gives horned males and hornless females in the $F_1$ and a ratio of 3 horned:1 hornless among the $F_2$ males, while the $F_2$ females show 3 hornless:1 horned. Clearly the gene for horns is dominant in males and recessive in females, while its allele for hornless is dominant in females and recessive in males.

Mathematical Consequences

These illustrations are sufficient to establish the facts that the same gene differs in its action in different sexes. Now we shall proceed to examine the partition of the genotypic variances and covariances of relatives under the assumption that gene effects depend on sex. In the following we shall be dealing with autosomal genes under random mating and a single locus segregating.

Let a male $A_iA_j$ has the genotypic value $y_{ij}^m$ and a female $A_iA_j$ has the genotypic value $y_{ij}^f$ measured from the respective population means. Here it can be seen that even though the genic constitution of the sexes are same, they differ in their genotypic values. Now we shall express their genotypic values in terms of additive and dominance effects as
follows,

\[ y_{ij}^m = a_i + a_j + \pi_{ij} \]
\[ y_{ij}^f = a_i + a_j + d_{ij} \]

where,

the effect of \( A_i \) in male is \( a_i \)

the effect of \( A_i \) in female is \( a_i \)

the dominance deviation in male is \( \pi_{ij} \)

and the dominance deviation in female is \( d_{ij} \)

Since the genotypic values \( y_{ij}^m \) and \( y_{ij}^f \) are deviations from the mean, we have

\[ E(y_{ij}^m) = \sum_{g} \sum_{h} p_g p_h y_{gh}^m = 0 \]

if \( y_{ij}^m \) is the genotypic value of a random individual and similarly for females

\[ E(y_{ij}^f) = \sum_{g} \sum_{h} p_g p_h y_{gh}^f = 0 \]

The total genotypic variance in male is

\[ \sum_{g} \sum_{h} p_g p_h (y_{gh}^m)^2 \]

The total genotypic variance in female is

\[ \sum_{i} \sum_{j} p_i p_j (y_{ij}^f)^2 \]

where \( p_i \) and \( p_j \) are the gene frequencies of \( A_i \) and \( A_j \) respectively.

We now fit \( y_{ij}^f = a_i + a_j \) and \( y_{ij}^m = a_i + a_j \) by least squares and on solving the normal equations, we find
$$a_k = \sum_j \pi_{jk}^m$$

$$a_k = \sum_i \pi_{ik}^f$$

The additive genotypic variance is, therefore,

$$V_{A}^m = 2 \sum_g p_g a_g^2$$

$$V_{A}^f = 2 \sum_h p_h a_h^2$$

where $g$, $h$, are dummy (variables) subscripts.

The orthogonality of the components of the genotypic values can be proved by showing that, if we have random genes with effects $a_i$ and $a_j$ then,

$$E(a_i a_j) = 0$$

provided there is no association of the genes. Similarly,

$$E(a_k a_i) = 0$$

and

$$E (a_i + a_j)(\pi_{ij}) = 0$$

and

$$E (a_i + a_j)(d_{ij}) = 0$$

To achieve the partition we use the fact that the covariances of the components in the models adopted are zero, so that

$$\text{Var}(y_{ij}^m) = \text{Var}(a_i) + \text{Var}(a_j) + \text{Var}(\pi_{ij})$$

and

$$\text{Var}(y_{ij}^f) = \text{Var}(a_i) + \text{Var}(a_j) + \text{Var}(d_{ij})$$
Symbolically we have the partition,

\[ V^m_G = V^m_A + V^m_D \]

and

\[ V^f_G = V^f_A + V^f_D \]

where,

- \( V^m_G \) stands for total genotypic variance in male
- \( V^f_G \) stands for total genotypic variance in female
- \( V^m_A \) stands for additive genetic variance due to autosomal genes in male
- \( V^f_A \) stands for additive genetic variance due to autosomal genes in female
- \( V^m_D \) stands for variance due to dominance deviation in male
- \( V^f_D \) stands for variance due to dominance deviation in female

Now we proceed to derive the general structures of the covariances between relatives under the assumption of gene effects depending on sex.

At the moment we shall be concerned with a random mating population with a single locus segregating. Three types of covariances are encountered.

(a) male and male

(b) female and female

(c) female and male

Consider the case (a) first, Let \( X^m_m \) and \( Y^m_m \) be two males with genotypes \( A_i A_j \) and \( A_k A_l \) respectively. The genotypic values of \( X^m_m \)
and $Y_m$ can be expressed as follows,

$$\text{genotypic value of } X_m = y_{ij}^m$$

and

$$\text{genotypic value of } Y_m = y_{kl}^m$$

We have the following mathematical models adopted to study the covariances,

$$y_{ij}^m = a_i + a_j + \pi_{ij}$$

and

$$y_{kl}^m = a_k + a_l + \pi_{kl}$$

where,

- $a_i$ represents the additive effects of gene $A_i$,
- $\pi_{ij}$ represents the dominance deviation of genes $A_iA_j$,
- the subscript $s$ indicates the origin of the gene from the Sire side,
- and the subscript $d$ indicates the origin of the gene from the Dam side.

It can be seen above that each gene effect is expressed in terms of the $a$'s and $\pi$'s because the individuals under consideration emanate from the same population that is, the male population. Adopting the procedure described earlier for obtaining expectations, we find that

$$\operatorname{Cov}(X_m, Y_m) = 2 r_{XY} V_A^m + u_{XY} V_D^m$$

where

- $r_{XY}$ is Malecot's coefficient of parentage,

and

- $u_{XY}$ is the quantity defined by Kempthorne (1957).
We shall be concerned only with autosomal genes for the moment.

Next consider the case (b). Let \( X_f \) and \( Y_f \) be two females chosen from a random mating population and we shall derive the expression for the covariances of relatives under the assumption of autosomal inheritance. Let \( X_f \) and \( Y_f \) have the genotypes \( A_iA_j \) and \( A_kA_l \) respectively. The mathematical models used to derive the required covariances are

\[
\gamma_{ij}^f = a_i + a_j + d_{ij} + d_{isjd}
\]

and

\[
\gamma_{kl}^f = a_k + a_l + d_{kl} + d_{ksld}
\]

where,

- \( a_i \) is the additive gene effect of gene \( A_i \)
- \( d_{ij} \) is the dominance deviation of genes \( A_iA_j \)

and \( s \) and \( d \) are the subscripts indicating the origin of the genes from the respective parents.

Since the individuals in question emanate from the same female population, the gene effects are expressed in terms of \( a \) and \( d \). Now adopting the usual procedure of obtaining the expectations, we find

\[
\text{Cov}(X_f, Y_f) = 2r_{XY}V_A^f + u_{XY}V_D^f
\]

where \( r_{XY} \) and \( u_{XY} \) retain the same definitions as described earlier.

Finally we consider the case (c). Let \( X_m \) and \( Y_f \) have the genotypes \( A_iA_j \) and \( A_kA_l \) respectively. Then with the respective
genotypic values $y^m_{ij}$ and $y^f_{kl}$, we adopt the following models,

$$y^m_{ij} = a_i + a_j + \pi_{i} a_{i,j}$$
$$y^f_{ij} = a_k + a_l + d_{k,l}$$

Now define

$$E(a_i a_k) = \sum_g a_i a_k = C_A$$
$$E(\pi_{i,j} d_{i,j}) = \sum_{g,h} \pi_{i,j} d_{i,j} = C_D$$

We have then

$$\text{Cov}(y^m_{ij}, y^f_{kl}) = E(a_i a_k + \pi_{i,j} a_{i,j} + d_{i,j})$$

Using the definitions (1) and (2),

$$\text{Cov}(X^m, Y^f) = 2r_{XY} C_A + u_{XY} C_D$$

where,

$C_A$ stands for the covariance between additive effects of autosomal genes in male and female,

$C_D$ denotes the covariance between dominance deviation of autosomal genes in male and female.

In the case when the genotypes of the male and female are same, we have

$$\text{Cov}(X^m, Y^f) = C_A + C_D$$

The algebraic structures of the covariances between relatives under multiple loci theory are presented below. The details of the derivations are not given since they follow the same procedure as described in
previous chapters.

\[
\text{Cov}(X_m, Y_m) = 2r_{XY} V_A^m + u_{XY} V_D^m + \sum_{t_1+t_2>1} (2r_{XY})^{t_1} (u_{XY})^{t_2} V_{A1D2}^m
\]

\[
\text{Cov}(X_r, Y_r) = 2r_{XY} V_A^f + u_{XY} V_D^f + \sum_{t_1+t_2>1} (2r_{XY})^{t_1} (u_{XY})^{t_2} V_{A1D2}^f
\]

\[
\text{Cov}(X_m, Y_f) = 2r_{XY} C_A + u_{XY} C_D + \sum_{t_1+t_2>1} (2r_{XY})^{t_1} (u_{XY})^{t_2} C_{A1D2}
\]

We shall now make a simultaneous consideration of the autosomal genes, completely sexlinked genes and partially sexlinked genes. The procedure of decomposition of the genotypic values into orthogonal components and the partition of the total genotypic variance into respective components of variance will be the same as presented in previous chapters. The algebraic structure of the covariances between relative will be of the following form under the assumption of

(i) independent segregation

(ii) no inbreeding

\[
\text{Cov}(X_m, Y_m) = 2r_{XY} V_A^m + u_{XY} V_D^m + v_{XY} V_{A\rho m} + 2v_{XY}^f V_{A\rho m} + u_{XY}^f V_{D\rho m}
\]
\[ \text{Cov}(X'_f, Y'_f) \]
\[ = 2 r_{XY} V_f + u_{XY} V_f + 2 r_{XY} A_f V_f + u_{XY} A_f V_f + 2 r_{XY} \rho_f V_f + u_{XY} \rho_f V_f \]
\[ + \sum \frac{1}{t_1 t_2 t_4 t_5 t_7 t_8} (2r_{XY})^t_1 (u_{XY})^t_2 (v_{XY})^t_3 (v_{XY})^t_4 (u_{XY})^t_5 (u_{XY})^t_7 (u_{XY})^t_8 \]
\[ A_{1D} A_{3m} A_{7m} A_{8m} \rho_{mD} \rho_{mD} \]

\[ \text{Cov}(X_m', Y'_f) \]
\[ = 2 r_{XY} C_A + u_{XY} C_D + w_{XY} C_A + 2 w_{XY} C_A + u_{XY} C_D \]
\[ + \sum \frac{1}{t_1 t_2 t_6 t_7 t_8} (2r_{XY})^t_1 (u_{XY})^t_2 (w_{XY})^t_6 (w_{XY})^t_7 (u_{XY})^t_8 \]
\[ A_{1D} A_{6m} A_{7m} A_{8m} \rho_{mD} \rho_{mD} \]
CHAPTER VI. THE ESTIMATION OF GENETIC PARAMETERS

General Considerations

There are a number of different experimental plans by which components of genotypic variance and covariance can be estimated. The one to be discussed here is related to Experiment II of Comstock and Robinson (1952). The experiment consists of mating each of a random sample of $s$ Sires to each of a random sample of $d$ Dams and the quantitative character is measured on $r$ offsprings of each mating. A joint analysis of the measurements of the character in the Sires, Dams as well as in the offspring can be made, but here the consideration is only extended to the analysis of the measurements obtained from the progeny.

The above experiment leads to the analysis of variance given in Table 11, where $\sigma_p^2$ is the total phenotypic variance.

The analysis of variance is based on the structure of the observations in that Sire and Dams have a cross classification and offsprings are nested within Sire x Dam cells. The expectations of mean squares are based on the assumptions that if

$$y_{ijk}$$ is the $k^{th}$ observation in the $(ij)^{th}$ cell

then

$$\text{Var}(y_{ijk}) = \frac{\sigma_p^2}{s}$$

and

$$C_1 = \text{Cov}(y_{ijk}, y_{ijk'}), \quad k \neq k'$$

$$C_2 = \text{Cov}(y_{ijk}, y_{ij'k}), \quad j \neq j', \ k \neq k'$$

$$C_3 = \text{Cov}(y_{ijk}, y_{ij'k'}), \quad i \neq i', \ k \neq k'$$

$$C_4 = \text{Cov}(y_{ijk}, y_{ij'k'}), \quad i \neq i', \ j \neq j', \ k \neq k'$$
Table 11. Analysis of variance for estimating genetic parameters

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>d.f.</th>
<th>Mean Square</th>
<th>Estimated Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sire</td>
<td>(s-1)</td>
<td>MS(S)</td>
<td>((\sigma_p^2 - C_1)) + r(C_1 - C_2 - C_3 + C_4) + rd(C_2 - C_4)</td>
</tr>
<tr>
<td>Dams</td>
<td>(d-1)</td>
<td>MS(D)</td>
<td>((\sigma_p^2 - C_1)) + r(C_1 - C_2 - C_3 - C_4) + rs(C_3 - C_4)</td>
</tr>
<tr>
<td>Sire x Dams</td>
<td>(s-1)(d-1)</td>
<td>MS(SD)</td>
<td>((\sigma_p^2 - C_1)) + r(C_1 - C_2 - C_3 - C_4)</td>
</tr>
<tr>
<td>Residual</td>
<td>sd(r-1)</td>
<td>MS(R)</td>
<td>((\sigma_p^2 - C_1))</td>
</tr>
</tbody>
</table>

The biological interpretation of the covariance components, in general, is as follows,

\[C_1 = \text{Cov(full sibs)}\]
\[C_2 = \text{Cov(paternal half sibs)}\]
\[C_3 = \text{Cov(maternal half sibs)}\]
\[C_4 = \text{Cov(genetically unrelated individuals)}\]

If then \(y_{ijk}\) is written as

\[y_{ijk} = \mu + s_i + d_j + (sd)_{ij} + e_{ijk},\]

in which \(\mu\) is a constant, and all other terms are independent random variables with the \(s_i\) independently distributed around zero with variance \(\sigma^2_s\), the \(d_j\) independently distributed around zero with variance \(\sigma^2_d\), the \((sd)_{ij}\) independently distributed around zero with variance \(\sigma^2_{sd}\).
and \( e_{ijk} \) distributed independently around zero with variance \( \sigma^2_e \), then,

\[
\begin{align*}
\sigma^2_s &= \text{Cov(paternal half sibs)} \\
\sigma^2_d &= \text{Cov(maternal half sibs)} \\
\sigma^2_{sd} &= \text{Cov(full sibs)} - \text{Cov(paternal half sibs)} \\
&\quad - \text{Cov(maternal half sibs)}.
\end{align*}
\]

We shall take the covariance of unrelated individual to be zero.

In the following, we shall express the above components of variance in terms of additive, dominance and epistatic variance for autosomal, sexlinked and partially sexlinked gene effects for males and females separately under the following assumptions,

1. normal diploid segregation
2. autosomal, sexlinked and partially sexlinked transmission
3. no maternal effects
4. arbitrary epistacy and arbitrary number of alleles per locus
5. males heterogametic sex
6. independent segregation that is, no linkage and no selection
7. random mating population

Characteristics Expressed By Males

Here the measurements of the quantitative character are taken on the male progeny and an analysis of variance is conducted. The components of variance thus obtained have the following interpretation,

\[
\begin{align*}
\sigma^2_s &= \text{Cov(paternal half brothers)} \\
\sigma^2_d &= \text{Cov(maternal half brothers)}
\end{align*}
\]
\[ \sigma_{sd}^2 = \text{Cov(full brothers)} - \text{Cov(paternal half brothers)} - \text{Cov(maternal half brothers)} \]

Under the assumption that the additive effects, dominance effects and the first order epistatic effects are important, we have the following decomposition of the variance components noted above,

\[
\begin{align*}
\sigma_{s}^2 & \approx \frac{1}{4} \sigma_A^2 + \frac{1}{16} \sigma_{AA}^2 + \frac{1}{2} f^2 \sigma_{A}^2_{\rho_m} + \frac{1}{4} f^4 \sigma_A^2_{\rho_m} A_{\rho_m} + \frac{1}{8} f^2 \sigma_{AA}^2_{\rho_m} \\
\sigma_{d}^2 & \approx \frac{1}{4} \sigma_A^2 + \frac{1}{16} \sigma_{AA}^2 + \frac{1}{2} \sigma_{A_m}^2 + \frac{1}{4} \sigma_A^2_{A_m} A_{m} + \frac{1}{8} \sigma_{AA}^2_{A_m} + \frac{1}{4} \sigma_A^2_{A_m} \rho_m \\
\sigma_{sd}^2 & \approx \frac{1}{4} \sigma_D^2 + \frac{1}{8} \sigma_{AA}^2 + \frac{1}{16} \sigma_{DD}^2 + \frac{1}{8} \sigma_{AD}^2 + \frac{1}{8} \sigma_{AA}^2_{m} + \frac{1}{8} \sigma_{DA}^2_{m} \\
\sigma_{e}^2 & \approx \frac{1}{2} \sigma_A^2 + \frac{3}{4} \sigma_{D}^2 + \frac{3}{4} \sigma_{AA}^2 + \frac{15}{16} \sigma_{DD}^2 + \frac{7}{8} \sigma_{AD}^2 + \frac{1}{2} \sigma_{A_m}^2 \\
& \quad + \frac{3}{4} \sigma_{A_m}^2 A_{m} + \frac{3}{4} \sigma_{AA}^2_{m} + \frac{7}{8} \sigma_{DA}^2_{m} + \frac{3}{4} \sigma_{A_m}^2 (\frac{3}{2} f^2 - \frac{1}{2} f^2) \sigma_{A_{\rho_m}}^2 + \text{and so on.} 
\end{align*}
\]

where, \( f^2 = (r_1^2 + r_2^2) \)

and \( \approx \) means "approximately equal to" and "leads to the estimation of", depending on what use we make of the equations.
The following heritability estimates are obtained from the analysis:

1. \[ 4 \frac{\sigma_d^2}{\sigma_p^2} \]
   contains all additive (autosomal) and twice the additive sexlinked male variance. This is an overestimation of \( h^2 \) due to sexlinkage.

2. \[ 4 \frac{\sigma_s^2}{\sigma_p^2} \]
   contains all additive (autosomal) variance one fourth of additive x additive (autosomal) variance.

3. \[ 4 \frac{\sigma_{sd}^2}{\sigma_p^2} \]
   contains all dominance (autosomal), half of additive x additive (autosomal), half of additive x dominance (autosomal), half of additive (autosomal) x dominance (sexlinked male) and half of dominance (autosomal) x additive (sexlinked female) variances. This estimate is sensitive to dominance and epistacy.

4. \[ 2 \frac{(\sigma_D^2 - \sigma_s^2)}{\sigma_p^2} \]
   contains all additive sexlinked male and half of additive (sexlinked male) x additive (sexlinked male) variances.

Here it is assumed that \( r^2 \) and fractions less than \( \frac{1}{4} \) are negligible.
Characteristics Expressed By Females

For the purpose of the estimation of the genetic parameters associated with the analysis of female progeny, we conduct a similar experiment and the components of variance given by the analysis of variance have the following interpretation:

\[ \sigma_s^2 = \text{Cov(paternal half sisters)} \]

\[ \sigma_d^2 = \text{Cov(maternal half sisters)} \]

\[ \sigma_{sd}^2 = \text{Cov(full sister)} - \text{Cov(paternal half sisters)} - \text{Cov(maternal half sisters)} \]

Under the supposition that additive effects, dominance effects and first order epistatic effects are important, we have the following decomposition,

\[ \sigma_s^2 = \frac{1}{4} \sigma_A^2 + \frac{1}{16} \sigma_{AA}^2 + \frac{1}{2} \sigma_{A_f}^2 + \frac{1}{8} \sigma_{AA_f}^2 + \frac{1}{4} \sigma_{A_f A_f}^2 + \frac{1}{4} f^2 \sigma_{A_f}^2 \rho_f \]

\[ + \frac{1}{4} f^2 \sigma_{A_f}^2 \rho_f \]

\[ \sigma_d^2 = \frac{1}{4} \sigma_A^2 + \frac{1}{4} \sigma_{A_f}^2 + \frac{1}{16} \sigma_{AA}^2 + \frac{1}{16} \sigma_{A_f A_f}^2 + \frac{1}{16} \sigma_{AA_f}^2 + \frac{1}{4} \sigma_{A_f}^2 \rho_f \]

\[ + \frac{1}{4} \sigma_{A_f}^2 \rho_f + \frac{1}{16} \sigma_{A_f}^2 \rho_f \]

\[ + \frac{1}{16} \sigma_{AA_f}^2 + \frac{1}{16} \sigma_{A_f A_f}^2 \rho_f \]

\[ \sigma_{sd}^2 = \frac{1}{4} \sigma_D^2 + \frac{1}{2} \sigma_{D_f}^2 + \frac{1}{8} \sigma_{AA}^2 + \frac{1}{8} \sigma_{AD}^2 + \frac{1}{16} \sigma_{DD}^2 + \frac{3}{16} \sigma_{AA_f}^2 \]

\[ + \frac{1}{4} \sigma_{AD_f}^2 + \frac{3}{16} \sigma_{DA_f}^2 + \frac{1}{8} \sigma_{DD_f}^2 + \frac{1}{4} \sigma_{A_f A_f}^2 + \frac{1}{4} \sigma_{D_f D_f}^2 \]
\[
\sigma_e^2 \cong \frac{1}{2} \sigma_A^2 + \frac{3}{4} \sigma_D^2 + \frac{3}{4} \sigma_{AA}^2 + \frac{7}{8} \sigma_{AD}^2 + \frac{15}{16} \sigma_{DD}^2 \\
+ \frac{1}{4} \sigma_A^2 + \frac{3}{8} \sigma_A^2 D_f + \frac{1}{8} \sigma_A^2 D_f + \frac{5}{8} \sigma_D^2 D_f + \frac{3}{4} \sigma_D^2 D_f \\
+ \frac{5}{8} \sigma_{AA} D_f + \frac{13}{16} \sigma_{DA} D_f + \frac{3}{4} \sigma_{AD} D_f + \frac{7}{8} \sigma_{DD} D_f 
\]
and so on.

where \( f^2 = (r_1^2 + r_2^2) \)

and \( \cong \) stands for "approximately equal to" and "leads to the estimation of".

The following heritability estimates are obtained from the analysis,

1. \( 4 \sigma_s^2 / \sigma_p^2 \) contains all additive (autosomal) and twice additive (sexlinked) variances. This estimate is inflated by sexlinkage.

2. \( 4 \sigma_d^2 / \sigma_p^2 \) contains all additive (autosomal) and all additive (sexlinked) variances.

3. \( 4 \sigma_{sd}^2 / \sigma_p^2 \) contains all dominance (autosomal), half of additive x additive (autosomal), half of additive x dominance (autosomal), one
fourth of dominance x dominance (autosomal), twice dominance (sex-linked), all additive x additive (sex-linked) and all dominance x dominance (sexlinked) variances. This estimate is sensitive to dominance and epistacy.

\( 4 \frac{\sigma_s^2 - \sigma_D^2}{\sigma_p^2} \)

contains all additive (sexlinked), variances

Here it is assumed that \( f^2 \) and fraction less than \( \frac{1}{4} \) are negligible.

Characteristics Expressed By Both Sexes

The experiment to be considered consists of crossing a random sample of \( s \) Sires to a random sample of \( d \) Dams and the basic experiment is replicated \( r \) times. The measurements are made on the members of both sexes. The following biological relationships are established among the members in a cell,

(i) the members within a cell are either full brothers (FB) or full sisters (FS),

(ii) the members in different replicates resulting from a particular cross are either full brothers or full sisters,

(iii) the members in the same or different replicates resulting from a common Sire but different Dams are either paternal half brothers (PHB) or paternal half sisters (PHS),
(iv) the members in the same or different replicates resulting from a common Dam but different Sires are either maternal half brothers (MHB) or maternal half sisters (MHS).

The following linear mathematical model is adopted to estimate the various variances due to effects and due to interactions,

\[ y_{ijk\ell} = \mu + s_i + d_j + (sd)_{ij} + r_k + (sr)_{ik} + (dr)_{jk} + (sdr)_{ijk} + c_t + (sc)_{iti} + (dc)_{jt} + (sdc)_{ijt} + (r)_{kt} + (src)_{ikt} + (drc)_{jkt} + (sdc)_{ijkt} + (g + \epsilon)_{ijkl} + e_{ijkl} \]

where,

\[
\begin{align*}
  i & = 1, 2, \ldots, s \\
  j & = 1, 2, \ldots, d \\
  k & = 1, 2, \ldots, r \\
  t & = 1, 2 \\
  \ell & = 1, 2, \ldots, n
\end{align*}
\]

\( y_{ijk\ell} \) is the \( 1^{\text{st}} \) observation of the \( t^{\text{th}} \) sex (either male or female) of the \( k^{\text{th}} \) replicate resulting from the cross between \( i^{\text{th}} \) Sire and \( j^{\text{th}} \) Dam,

\( \mu \) is the constant general mean of the population under consideration,

\( s_i \) is the \( i^{\text{th}} \) Sire-effect (S) which is random,

\( d_j \) is the \( j^{\text{th}} \) Dam-effect (D) which is random,
(sd)_{ij} \text{ is the interaction-effect (SD) of } i^{th} \text{ Sire and } j^{th} \text{ Dam which is random,}

r_k \text{ is the } k^{th} \text{ replicate-effect (R) which is fixed,}

(sr)_{ik} \text{ is the interaction-effect (SK) of } i^{th} \text{ Sire and } k^{th} \text{ replicate,}

(dr)_{jk} \text{ is the interaction-effect (DR) of the } j^{th} \text{ Dam and } k^{th} \text{ replicate,}

(sdr)_{ijk} \text{ is the interaction-effect (SDR) of the } i^{th} \text{ Sire, } j^{th} \text{ Dam and } k^{th} \text{ replicate,}

c_t \text{ is the } t^{th} \text{ sex-effect (C) which is fixed,}

(sc)_{it} \text{ is the interaction-effect (SC) of the } i^{th} \text{ Sire and } t^{th} \text{ sex,}

(dc)_{jt} \text{ is the interaction-effect (DC) of the } j^{th} \text{ Dam and } t^{th} \text{ sex,}

(sdc)_{ijt} \text{ is the interaction-effect (SDC) of the } i^{th} \text{ Sire, } j^{th} \text{ Dam and } t^{th} \text{ sex,}

(rc)_{kt} \text{ is the interaction-effect (RC) of the } k^{th} \text{ replicate and } t^{th} \text{ sex,}

(src)_{ikt} \text{ is the interaction-effect (SRC) of the } i^{th} \text{ Sire, } k^{th} \text{ replicate and } t^{th} \text{ sex,}

(drc)_{jkt} \text{ is the interaction-effect (DRC) of the } j^{th} \text{ Dam, } k^{th} \text{ replicate and } t^{th} \text{ sex,}

(sdr)_{ijk} \text{ is the interaction-effect (SDRC) of the } i^{th} \text{ Sire, } j^{th} \text{ Dam, } k^{th} \text{ replicate and } t^{th} \text{ sex,}
\( \varepsilon_{ijktl} \) is a random error specific to each individual and uncorrelated from individual to individual with expectation zero and variance \( \sigma^2_\varepsilon \),

\( g_{ijktl} \) is a random genotypic deviation associated with each offspring with expectation zero and variance \( \sigma^2_{G_m} \) or \( \sigma^2_{G_f} \) depending upon whether the individual under consideration is a male or a female,

\( \varepsilon_{ijkt} \) is that part of the random error which is common to all individuals in a particular cell, uncorrelated between different cells with expectation zero and variance \( \sigma^2_\varepsilon \),

\[
E(s_i d_j) = E(s_i (sd)_{ij}) = \ldots = E(g t) = E(g e) = E(\varepsilon \varepsilon) = 0,
\]

that is, all terms in the model are uncorrelated,

\[
E(s_i) = 0, \quad E(s_i^2) = \sigma^2_s, \quad E(s_i s_j) = 0
\]

\[
E(d_j) = 0, \quad E(d_j^2) = \sigma^2_d, \quad E(d_j d_j) = 0
\]

\[
E(sd)_{ij} = 0, \quad E(sd)_{ij}^2 = \sigma^2_{sd}
\]

\[
\Sigma r_k = 0, \quad \Sigma r_k^2 = \sigma^2_r
\]

\[
\Sigma c_t = 0, \quad \Sigma c_t^2 = \sigma^2_c
\]

\[
E(sc)_{it} = 0, \quad E(sc)_{it} = 0 \quad E(sc)_{it}^2 = \sigma^2_{sc}
\]

\[
E(sr)_{ik} = 0, \quad \Sigma (sr)_{ik} = 0 \quad \Sigma (sr)_{ik}^2 / r-1 = \sigma^2_{sr}
\]

and so on,
\[ E(e_{ijkt}) = 0, \quad E(e_{ijkt})^2 = \sigma_e^2 \]

and

\[ E(\epsilon_{ijkl}) = 0, \quad E(\epsilon_{ijkl})^2 = \sigma_\epsilon^2 \]

No other distributional properties of the errors have been assumed since the chief interest here is the estimation problems.

For the purpose of estimation of genetic parameters our chief aim is the derivation of the expectation of mean squares and it will be of interest to study the structure of the expectations particularly when the \( g \)'s are correlated as in this study. Since we are involved with 16 sources of variation, it will be both time and space consuming to present the detail description of the derivation. We shall, however, demonstrate the behaviour of \( (g + \epsilon)_{ijkl} \) component in the derivation of the expectation of the sum of square due to Sire x Dam interaction as an illustration. During the course of the development of the expectation of the sum of square due to Sire x Dam interaction, we meet the following expression to be evaluated:

\[
E \sum_{ij} \sum_{rpn} \frac{1}{rpn} \sum_{kl} (g + \epsilon)_{ijklm} + \frac{1}{rpn} \sum_{kl} (g + \epsilon)_{ijklf} - \frac{1}{sdrpn} \sum_{ijkl} (g + \epsilon)_{ijklm} - \frac{1}{sdrpn} \sum_{ijkl} (g + \epsilon)_{ijklf}^2
\]

(where subscript \( m \) stands for male and subscript \( f \) stands for female and they should be regarded as fixed subscripts throughout this derivation.)
\[
\begin{align*}
E &= \frac{1}{p} \sum_{ij} \left[ (g + \epsilon)_{ijm} + (g + \epsilon)_{ijf} \right] - \frac{1}{sd} \sum_{ij} (g + \epsilon)_{ijm} \\
&\quad - \frac{1}{sd} \left( g + \epsilon \right)_{ijf} \\
&= \left( g + \epsilon \right)_{ijm} \left( g + \epsilon \right)_{ijf} - \frac{1}{sd} \sum_{ij} (g + \epsilon)_{ijm} \\
&\quad - \frac{1}{sd} \sum_{ij} \left( g + \epsilon \right)_{ijf} \\
&\quad \text{(here } \sum_{kl} (g + \epsilon)_{ijkl} = (g + \epsilon)_{ij}.f \text{ and we have dropped the dots)}
\end{align*}
\]
\[
+ \sum_{j \neq j'} \sum_{i} (g + \epsilon)_{ijf} (g + \epsilon)_{ij'f} + \sum_{i \neq i'} \sum_{j \neq j'} (g + \epsilon)_{ijf} (g + \epsilon)_{ij'f}
\]

\[
+ E \frac{2}{sd} \left\{ (g + \epsilon)_{ijm} (g + \epsilon)_{ij'm} + \sum_{i \neq i'} (g + \epsilon)_{ijm} (g + \epsilon)_{i'j'm} \right\}
\]

\[
= \frac{1}{p} \sum_{ij} \left[ \sigma^2_{G_m} + \sigma^2_{G_f} + \frac{1}{s^2d^2} \right] \text{sd} \sigma^2_{G_m} + \text{sd} \sigma^2_{G_f} + \text{sd} (s - 1) \text{Cov(MHB)} + \text{sd} (d - 1) \text{Cov(PHB)} + \frac{1}{s^2d^2} \text{sd} (s - 1) \text{Cov(MHS)} + \text{sd} (d - 1) \text{Cov(PHS)} \right\}
\[ + 2 \text{Cov}(FB, FS) - \frac{2}{s^2d^2} \left\{ \sigma^2_m + \sigma^2_\epsilon + (s - 1) \text{Cov}(MHB) \right\} \]

\[ + (d - 1) \text{Cov}(PHB) \right\} - \frac{2}{s^2d^2} \left\{ \text{Cov}(FB, FS) + (s - 1) \text{Cov}(MHB, HS) \right\} \]

\[ + (d - 1) \text{Cov}(PHB, HS) \right\} - \frac{2}{s^2d^2} \left\{ \text{Cov}(FB, FS) \right\} \]

\[ + (s - 1) \text{Cov}(MHB, HS) + (d - 1) \text{Cov}(PHB, HS) \right\} \]

\[ \left\{ \sigma^2_m + \sigma^2_\epsilon + (s - 1) \text{Cov}(MHS) + (d - 1) \text{Cov}(PHS) \right\} \]

\[ + \frac{2}{s^2d^2} \left\{ s\text{d Cov}(FB, FS) + s\text{d}(s - 1)\text{Cov}(MHB, HS) + s\text{d}(d - 1)\text{Cov}(PHB, HS) \right\} \]

We can pool similar terms together and after the algebraic simplification we will have the required result. In the following the expectation of mean square for each source of variation is presented. The first column denotes the source of variation in symbols, second column has the degrees of freedom and the third column is devoted to the expectation of mean square as presented in Table 12.

In this table Cov(X, Y) will be taken as \( \frac{1}{2} \) the covariance of X and Y and p is equal to two.
Table 12. Analysis of variance of characteristics expressed by both sexes

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>Estimated Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>(r-1)</td>
<td>( \left{ \sigma_{GM}^2 + \sigma_{\epsilon}^2 - \text{Cov}(FB) + \sigma_{GF}^2 + \sigma_{\epsilon}^2 - \text{Cov}(FS) \right} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( + n \sigma_{\epsilon}^2 + \alpha \sigma_{sdr}^2 + s \alpha \sigma_{dr}^2 + d \alpha \sigma_{sr}^2 + sdn \sigma_{r}^2 )</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>( \left{ \sigma_{GM}^2 + \sigma_{\epsilon}^2 - \text{Cov}(FB, FS) + \sigma_{GF}^2 + \sigma_{\epsilon}^2 - \text{Cov}(FB, FS) \right} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( + n \sigma_{\epsilon}^2 + \alpha \sigma_{sdc}^2 + s \alpha \sigma_{dc}^2 + s \alpha \sigma_{sc}^2 + sdn \sigma_{c}^2 )</td>
</tr>
<tr>
<td>RC</td>
<td>(r-1)</td>
<td>( \left{ \sigma_{GM}^2 + \sigma_{\epsilon}^2 - \text{Cov}(FB, FS) + \sigma_{GF}^2 + \sigma_{\epsilon}^2 - \text{Cov}(FB, FS) \right} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( + n \sigma_{\epsilon}^2 + \alpha \sigma_{scd}^2 + s \alpha \sigma_{drc}^2 + d \alpha \sigma_{src}^2 + sdn \sigma_{rc}^2 )</td>
</tr>
<tr>
<td>S</td>
<td>(s-1)</td>
<td>( \left{ \sigma_{GM}^2 + \sigma_{\epsilon}^2 - \text{Cov}(MHB) + \sigma_{GF}^2 + \sigma_{\epsilon}^2 - \text{Cov}(MHS) \right} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( + 2 \left{ \text{Cov}(FB, FS) - \text{Cov}(MHB, HS) \right} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( + n \sigma_{\epsilon}^2 )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( + rpn \left{ \text{Cov}(FB) - \text{Cov}(MHB) - \text{Cov}(PHB) + \text{Cov}(FS) \right} )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( - \text{Cov}(MHS) - \text{Cov}(PHS) + 2 \left{ \text{Cov}(FB, FS) - \text{Cov}(MHB, HS) \right} )</td>
</tr>
<tr>
<td>Source</td>
<td>d.f.</td>
<td>Estimated Mean Square</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$- \text{Cov}(\text{PHB, HS})$ + drn ${L \text{Cov}(\text{PHB}) + \text{Cov}(\text{PHS})$ + 2 Cov(\text{PHB, HS})$ + \text{Cov}(\text{MHB})$ + $\sigma_e^2 - \text{Cov}(\text{FB, FS})$ $+ n \sigma_e^2 + dpn \sigma_{sr}^2 + pn \sigma_{sdr}^2$</td>
</tr>
</tbody>
</table>
Table 12. (Continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>Estimated Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(- \text{Cov(MHB)} - \text{Cov(PHB)}) + (\text{Cov(FS)} - \text{Cov(MHS)} - \text{Cov(PHS)})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>+ 2 (\text{Cov(FB, FS)} - \text{Cov(MHB, HS)} - \text{Cov(PHB, HS)})</td>
</tr>
<tr>
<td>DR</td>
<td>((d-1)(r-1))</td>
<td>(+ \text{srn} \left{\text{Cov(MHB)} + \text{Cov(MHS)} + 2 \text{Cov(MHB, HS)}\right} )</td>
</tr>
<tr>
<td>DC</td>
<td>((d-1))</td>
<td>(+ \text{n} \sigma^2 + \text{pn} \sigma_{sdr}^2 + \text{spn} \sigma_{dr}^2)</td>
</tr>
<tr>
<td>DRC</td>
<td>((d-1)(r-1))</td>
<td>(+ \text{n} \sigma^2 + \text{rn} \sigma_{sdc}^2 + \text{srn} \sigma_{dc}^2)</td>
</tr>
<tr>
<td>SD</td>
<td>((s-1)(d-1))</td>
<td>(+ \text{n} \sigma_{e}^2 + \text{dn} \sigma_{sdrc}^2 + \text{sn} \sigma_{drc}^2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(+ \text{n} \sigma_{e}^2 + \text{pn} \sigma_{sdr}^2 + \text{spn} \sigma_{dr}^2)</td>
</tr>
</tbody>
</table>

\(\sigma^2_{m}\), \(\sigma^2_{f}\), \(\sigma^2_{e}\), \(\sigma^2_{sdc}\), \(\sigma^2_{sdrc}\), \(\sigma^2_{drc}\)
Table 12. (Continued)

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>Estimated Mean Square</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( + \sum \text{Cov}(FS) - \text{Cov}(MHS) - \text{Cov}(PHS) ) + 2 ( \sum \text{Cov}(FB, FS) )</td>
</tr>
<tr>
<td>SDR</td>
<td>((s-1)(d-1)(r-1))</td>
<td>( \left{ \frac{c^2}{G_m} + \sigma^2 + \text{Cov}(FB, FS) \right} )</td>
</tr>
<tr>
<td>SDC</td>
<td>((s-1)(d-1))</td>
<td>( \left{ \frac{c^2}{G_m} + \sigma^2 - \text{Cov}(FB, FS) \right} )</td>
</tr>
<tr>
<td>SDRC</td>
<td>((s-1)(d-1)(r-1))</td>
<td>( \left{ \frac{c^2}{G_m} + \sigma^2 - \text{Cov}(FB, FS) \right} )</td>
</tr>
<tr>
<td>Residual</td>
<td>(sdrp) ((n-1))</td>
<td>( \left{ \frac{c^2}{G_m} + \sigma^2 - \text{Cov}(FB) \right} )</td>
</tr>
</tbody>
</table>
Now for the purpose of estimation if we equate the mean squares to their respective expected mean squares, then we have 16 equations with 25 unknowns. These 25 unknowns can be reduced to lower number of unknowns by making some additional assumptions depending upon the experimenter's interest. Now if we are interested in the estimation of the components of genetic variances such as $\sigma^2_A$, $\sigma^2_D$ and so on and components due to the interaction of sex and genotype then we will not have difficulty in reducing these 25 to 16 unknowns as follows,

1. $\text{Cov}(FB)$,
2. $\text{Cov}(FS)$,
3. $\text{Cov}(FB, FS)$
4. $\text{Cov}(PHB)$,
5. $\text{Cov}(PHS)$,
6. $\text{Cov}(PHB, HS)$
7. $\text{Cov}(MHB)$,
8. $\text{Cov}(MHS)$,
9. $\text{Cov}(MHB, HS)$
10. $\sigma^2_{sc}$
11. $\sigma^2_{dc}$
12. $\sigma^2_{sd}$
13. $\sigma^2_r$
14. $\sigma^2_c$
15. $\sigma^2_G$
16. $\sigma^2_e$

The additional assumptions we make are that the variance components involving interactions with replicates are assumed to be a part of the plot-error and so they are all pooled into $\sigma^2_e$. The individual error $\sigma^2_e$ is also pooled to $\sigma^2_e$. The genotypic variances of both sex are combined into one, that is $\sigma^2_G$. In this situation then, we have an estimate of

(i) variance due to sex influence ($\sigma^2_c$)
(ii) variance due to environmental effect ($\sigma^2_r$)
(iii) variance due to genotype x sex interaction ($\sigma^2_{sc}$, $\sigma^2_{dc}$, $\sigma^2_{sd}$)
and (iv) all the covariances which estimate the genetic variance components in groups.
Now if we consider that the components of genetic variances such as $\sigma^2_A$, $\sigma^2_D$ etc. and genotype x environment variance components are important, then pooling the variance components due to genotype x sex interaction into $\sigma^2_e$ and assuming that second order interactions are negligible, we have the following 16 unknowns that can be uniquely estimated.

\[
\begin{align*}
(1) \quad & \sigma^2_{sr} \\
(2) \quad & \sigma^2_{dr} \\
(3) \quad & \sigma^2_{sdr} \\
(4) \quad & \sigma^2_{r} \\
(5) \quad & \sigma^2_c \\
(6) \quad & \sigma^2_G \\
(7) \quad & \sigma^2_e \\
(8) \quad & \text{Cov}(FB) \\
(9) \quad & \text{Cov}(FS) \\
(10) \quad & \text{Cov}(FB,FS) \\
(11) \quad & \text{Cov}(PHB) \\
(12) \quad & \text{Cov}(PHS) \\
(13) \quad & \text{Cov}(PHB,HS) \\
(14) \quad & \text{Cov}(MHB) \\
(15) \quad & \text{Cov}(MHS) \\
(16) \quad & \text{Cov}(MHB,HS)
\end{align*}
\]

In this situation then we have an estimate of

(i) variance due to genotype x environment interaction

$$\left( \sigma^2_{sr}, \sigma^2_{dr}, \sigma^2_{sdr} \right)$$

(ii) variance due to sex influence ($\sigma^2_c$)

(iii) variance due to replicates ($\sigma^2_r$)

and (iv) all the covariance components.

We could also absorb the covariance components to estimate all the variance components. The covariance components can be decomposed into additive, dominance and epistatic components as shown in previous tables. The meaningful groupings of the components of genetic variances will facilitate estimation of such groups and when the latter is expressed as fraction of $\sigma^2_p$, heritability estimates are obtained. The preceding two sections give the mechanics involved in detail and they
may be referred to in this connection.

Inbreeding Followed by Random Mating

Now consideration is extended to the case where the Sires and Dams arise by random mating of a population which has been inbred to an extent measured by the coefficient of inbreeding $F$. In case of analysis of characteristics expressed by males, we have the following decomposition:

$$
\sigma_s^2 \approx \frac{1}{4} (1+F) \sigma_A^2 + \frac{1}{16} (1+F)^2 \sigma_{AA}^2
$$

$$
\sigma_d^2 \approx \frac{1}{4} (1+F) \sigma_A^2 + \frac{1}{2} (1+F) \sigma_{A_m}^2 + \frac{1}{16} (1+F)^2 \sigma_{AA}^2 + \frac{1}{16} (1+F)^2 \sigma_{AA_m}^2
$$

$$
\sigma_{sd}^2 \approx \frac{1}{4} (1+F)^2 \sigma_D^2 + \frac{1}{8} (1+F)^2 \sigma_{AA}^2 + \frac{1}{16} (1+F) \sigma_{DD}^2 + \frac{1}{8} (1+F)^3 \sigma_{AD}^2
$$

$$
\sigma_e^2 \approx \frac{1}{2} (1+F) \sigma_A^2 + \frac{3}{4} (1+F)^2 \sigma_D^2 + \frac{3}{4} (1+F)^2 \sigma_{A_m}^2 + \frac{15}{16} (1+F) \sigma_{AD}^2
$$

In case of females we have the following decomposition,

$$
\sigma_s^2 \approx \frac{1}{4} (1+F) \sigma_A^2 + \frac{1}{16} (1+F)^2 \sigma_{AA}^2 + \frac{1}{2} \sigma_{Af}^2 + \frac{1}{8} (1+F) \sigma_{AA_f}^2 + \frac{1}{4} \sigma_{Af}^2 \sigma_{Af}
$$

$$
\sigma_d^2 \approx \frac{1}{4} (1+F) \sigma_A^2 + \frac{1}{4} (1+F) \sigma_{A_f}^2 + \frac{1}{16} (1+F)^2 \sigma_{AA}^2 + \frac{1}{16} (1+F)^2 \sigma_{A_f}^2 \sigma_{Af}
$$

$$
+ \frac{1}{16} (1+F)^2 \sigma_{AAf}^2
$$
\[ \sigma^2_{sd} \approx \frac{1}{4}(1+F)^2 \sigma^2_D + \frac{1}{2}(1+F) \sigma^2_{D_f} + \frac{1}{8}(1+F)^2 \sigma^2_{AA} + \frac{1}{8}(1+F)^3 \sigma^2_{AD} \\
+ \frac{1}{16}(1+F)^4 \sigma^2_{DD} + \frac{1}{16}(1+F)(3+F) \sigma^2_{AA_f} + \frac{1}{4}(1+F)^2 \sigma^2_{AD_f} \\
+ \frac{1}{16}(1+F)^2(3+F) \sigma^2_{DA_f} + \frac{1}{8}(1+F)^3 \sigma^2_{DD_f} + \frac{1}{16}(3+F)^2 \sigma^2_{A_f A_f} \\
+ \frac{1}{4}(1+F)^2 \sigma^2_{D_f D_f} \\
\sigma^2_e \approx \frac{1}{2}(1+F) \sigma^2_A + \frac{3}{4}(1+F)^2 \sigma^2_D + \frac{3}{4}(1+F)^2 \sigma^2_{AA} + \frac{7}{8}(1+F)^3 \sigma^2_{AD} \\
+ \frac{15}{16}(1+F)^4 \sigma^2_{DD} + \frac{1}{4}(3+F) \sigma^2_{A_f} + \text{and so on.} \]

The heritability estimates are constructed in the same way as has been described in the first two sections of this chapter. Since \( F \) is a known constant, there will be no difficulty in obtaining the numerical values of the estimates.

Non-orthogonal Situation

In actual experiments there will be unequal frequencies in the cells. The procedure to handle such cases are given in the literature (see Kemphthorne 1957 Chapter 20 for example). Instrinsically the basic idea is to perform an unweighted analysis of cell means and take account of the unequal frequencies in the expectation of mean squares.
CHAPTER VII. THE MONTE CARLO APPROACH

General Consideration

The field of mathematics which is based on the simulation of stochastic processes has achieved great prominence in recent years. This approach to the mathematical problems has come to be known as the Monte Carlo approach. Its prominence is directly attributable to the introduction of high speed digital computers and the resolution of a problem into thousands of elementary arithmetic steps which automatic computers can do at a fantastic speed. In this study the simulation of genetic systems is our prime interest and before we embark on the description of the details of the algebraic steps and the repetitive sequences leading to the simulation processes, we shall briefly describe the operation of a digital computer with special reference to the Cyclone.

It can be shown that any integer $N$ can be represented as a sum of powers of any whole number $r$ greater than one as,

$$N = a_n r^n + a_{n-1} r^{n-1} + \ldots + a_1 r + a_0$$

where $a_i$ is the integer in the range $0 \leq a_i < r$

and $r$ is the radix,

and for any fraction

$$M = b_1 r^{-1} + b_2 r^{-2} + \ldots + b_m r^{-m}$$

where $b_i's$ are the integers in the range $0 \leq b_i < r$.

A digital computer in representing numbers provides a place to hold each coefficient of the power sum representation for that number. If the
computer operates on the decimal scale, each such coefficient position would have to be able to assume one of the ten states namely, 0, 1, \ldots, 9. One can visualize that the lower the number of states maintained in the computers the more economical and easier would be the operations of the computer. If the binary number system is used, only two states have to be maintained for each of the two coefficient positions namely 0 and 1.

In a positional notation situation,

\[ .25_{10} = .01_{2} \]
\[ 502_{10} = 111110110_{2} \]

where the subscripts are the radices.

The Cyclone uses the binary system and each number is assumed to be a fraction and 40 binary bits (digits) long. Since a number 40 bits long represents the same information as one 12 digit decimal number at most, it is economical for the computers but it is not convenient for the programmer to use the system. So given a 40 bit number one reduces this to 10 digits by expressing 4 bits by a single digit and this is achieved by a system called the sexadecimal system which has 10 symbols namely 0, 1, \ldots, 9, K, S, N, J, F, L. The computer has 1024 memory locations and three registers namely A, Q and R. The two registers important to the programmer are the arithmetic register (A) and the quotient register (Q). The A and Q registers can be made to act like a 79 bit shifting register called AQ. A left shift of one place moves every bit one position to the left and generates a zero at the right-most bit of Q. A right shift of one place moves every bit one position to the right and propagates the left-most bit in A. These properties of AQ are taken
advantage of while forming the gametes in the simulation process. An order is an elementary operation, such as shifting or addition. A single location contains two orders called an order pair, and an order pair is called a word. The programmer writes these words on a tape sequentially and receives the output on a tape in the binary system which can be easily decoded to decimal system by the printer. The orders are of the following type,

(i) arithmetic
(ii) decision
(iii) logical
(iv) manipulative

All the above types of orders have been fully utilized in the program under consideration. A big problem is often broken into several semi-independent steps each of which can be programmed separately. The advantage of this is that the programmer is able to concentrate on one step at a time and test each part of the programme separately before incorporating it in the total program. Such programs are called subroutines. Collections of such subroutines are available in the library of routines associated with the computer. For a special problem as this, special subroutines had to be programmed and incorporated in the main body of the master routine.

Simulation Procedure

The binary representation of a genetic structure is achieved by assigning 1 and 0 to the two alternate alleles of a locus, where 1
implies the plus gene and 0 implies the not-plus gene. The paternal genotype and the maternal genotype of an individual are placed side by side so that the genes are in a linear arrangements. For example, suppose a genotype consists of 5 loci, then the usual expression of the same is

\[
\begin{array}{c}
1 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 1
\end{array}
\]

but in a linear arrangement situation it has the following structure

\[
1101010011
\]

where the 1st bit and the 6th bit are allelic, as also are the second and the seventh and so on. The advantage of such a representation is that the complete genotype can be stored in one memory location. Since every location has only 40 binary bits, the genotypes consisting of more than 20 loci can be stored in two or more locations.

The mechanism involved in evaluating a phenotype of a given genotype will be elucidated by an example. Suppose the dominance relationships of two alleles at a locus are as follows

\[
1/1 = a, \quad 0/1 \text{ or } 1/0 = \beta \quad \text{and} \quad 0/0 = \rho
\]

and the genotype under consideration has four loci with the following structure,

\[
10011101
\]

We then shift the genotypic structure next to the sign bit and for the purpose of the evaluation of the phenotype, we shift the two bits representing alleles one after the other for each locus to the sign bit position by shifting to the left. We store the phenotype thus recorded at an assigned
memory location. In this particular case, 1st and 5th, 2nd and 6th and so on are allelic and they are shifted in that order and the total phenotype thus recorded is

\[ 2a + \beta + e \]

which is brought about by converting the given dominance relationships in the machine to the following form,

\[-/- = a, +/- or -/+ = \beta and +/- = e\]

where a number is positive if the sign bit is 0 and a number is negative if the sign bit is 1.

For the case where epistacy is involved, bits that are in epistatic relationship are brought to the sign bit one after the other and a value for that order epistacy is added to the phenotype. In general then the value for the \( n^{th} \) order epistacy is incorporated into the cumulating phenotype by bringing the bits with epistatic relationship successively to the sign bit position and taking the appropriate decision after the \((n+1)^{th}\) bit arrives at the sign bit position.

The simulation of recombination and formation of gametes is handled as follows. Consider a genotype with three loci with the following structure

\[
\begin{array}{c}
T \\
B
\end{array} \quad \begin{array}{c}
T \\
B
\end{array} \quad \begin{array}{c}
T \\
B
\end{array}
\]

where T and B stand for top and bottom respectively and let the frequencies of the recombination for 1st and 2nd loci be \( r_1 \) and for 2nd and 3rd loci be \( r_2 \), that is the probability associated with a chosen gamete say, BTT is \( \frac{1}{2} r_1 (1 - r_2) \). In this situation then we have the
following probability relations in the gametic formation,

- probability of getting $T$ or $B$ at the first locus $= \frac{1}{2}$
- probability of getting $T$ in the second locus given that the first locus has a $B$ $= r_1$
- probability of getting $T$ in the second locus given that the first locus has a $T$ $= (1 - r_1)$
- probability of getting $T$ in the third locus given that the second locus has a $B$ $= r_2$
- and probability of getting $T$ in the third locus given that the second locus has a $T$ $= (1 - r_2)$

If we interchange $T$ for $B$ and $B$ for $T$ we have the same results. A gamete is produced from an individual by what may be called a random mask method. The actual mechanics involved are elaborate and difficult to understand without detailed knowledge of the operation of the machine, so we shall indicate briefly only the salient features of the method. For every individual, a random mask is constructed taking into account the probability relation appropriate to each locus. For example, a mask is of the following type for a genotype with five loci,

$$BBTBT$$

means that a random gamete from a genotype consists of the bottom gene at the first locus, bottom gene at the second locus, the top gene at the third locus, the bottom gene at the fourth locus and the top gene at the fifth locus. The consideration of the probability relation existant at each locus assures that the probability associated with the random mask and
the probability associated with the random gamete produced by the mask are equal. Now suppose the genotype has the following structure

\[
\begin{array}{cc}
11010 \\
10011
\end{array}
\]

then the random mask obtained above is applied on the linearly arranged genotype and by the extraction process, which is a command, we have the following random gamete,

\[
10010
\]

and the probability associated with the gamete is the same as the probability associated with the mask namely \( \frac{1}{2} (1-r_1) r_2 r_3 r_4 \) where \( r_1, r_2, r_3 \) and \( r_4 \) are the recombination frequencies between the 1st and 2nd, 2nd and 3rd, 3rd and 4th and 4th and 5th loci, respectively. So a mask is dependent upon the linkage relations under consideration and not the genotypic constitution of the individual. Most of the time consumed in constructing a mask is in generating random numbers because, for a genotype with \( n \) loci, we have to generate \( n \) random numbers. If there is enough storage space, it may be economical to store all the masks and then associate a random mask with a random individual. By trial and error, we have come to the conclusion that the use of random masks is one of the fastest methods for extracting a random gamete for the individual.

The following formula has been developed to simulate selection by truncation,

\[
I = -\frac{1}{4} \left| \frac{x - c}{k} \right|
\]

that is, the index \( I \) is equal to the negative absolute value of the absolute
value of \( (x - c) \) minus \( k \) and

where, \( x \) is the phenotype of the individual under consideration,
\( c \) is the numerical point around which the individuals are selected which we call the selection-point. The quantity \( k \) takes the discrete values 0, 1, 2... and so on in succession if the phenotypes are expressed in discrete values and takes continuous values if the phenotypes are expressed on a continuous scale.

For a given value of \( c \), the individual is selected when \( I \) becomes non-negative. The formula is rather general with regard to any level of truncation. We shall demonstrate the actual operation of the formula by an example. Suppose we are selecting 20 individuals out of 100 individuals whose phenotypes range from 0 to 30 and the selection point is at 30, that is, the upper extreme. Then we wish to have as many individuals as possible with 30 as their phenotypes, and if we cannot have all the 20 with this qualification, we shall pick individuals with phenotypes 29, 28 and so on until we collect 20 individuals. The formula takes care of this situation rather easily. The phenotypes are in random order in the memory locations and we do not have to arrange them in descending order of their phenotypes. We fix the value of \( c \) at 30, and bring the first individual whose phenotype is, say, 18 to the arithmetic register and find the value of \( |x - c| = |18 - 30| \). In the meantime we set the value of \( k = 0 \). We see that \( I \neq 0 \), and it is in fact equal to \(-12\), and so the individual is not to be selected. Suppose the phenotype of the individual were 30, then when \( c = 30 \) and \( k = 0 \), \( I \) is non-negative and the
individual is selected. So in this situation we fix the value of $c$ at 30 and set $k = 0$ and change the $x$ successively to $x_1, x_2', \ldots, x_{100}'$, the phenotypic values of the 100 individuals, and all the $x's$ whose phenotypes are 30 will be selected. When we set $k$ equal to 1, all the individuals whose phenotypes equal 29 are selected, and then when we set $k$ equal to 2, the individuals with phenotype 28 are selected and so on until the desired number of individuals is obtained. For a given selection point $t$ and for a given $k = n^t$, all the individuals with the phenotype $(t-n)$ are selected in an extreme truncation scheme. Now let us examine the case where an intermediate value is taken as the selection point. We shall consider the same example and let the value 8 be the selection point. Then when $k = 1$, the index $I$ will pick all the individuals with 9 as their phenotype, but the formula has been constructed in such a way that it will also select those individuals with 7 as their phenotype. So for extreme truncation, the index has a particular value for one phenotype, but for intermediate selection the index takes the same value for two phenotypes, symmetrically placed around the selection point. Since the phenotypes are considered in a random order for the purpose of selection, the values are selected on a first-come-first-served basis.

Fraser (1957) is the first one to introduce the binary representation of a genetic formula. In his method, maternal genotype and paternal genotypes of an individual are stored in two different locations, which would take more location space. His method of evaluating the phenotype of a genotype consists of obtaining the results from three logical
operations, viz, logical product, logical equivalent and logical not-sum of the two homologous genotypes, constructing three diagonal matrices and evaluating the phenotype by certain matrix multiplications. All these operations are time consuming and the evaluation of the phenotype by shifting process proposed in this study is much faster. Enumeration of all the gametes and their frequencies and then choosing one of the gametes at random is both time and space consuming and the random mask method proposed in this study gives faster results. So also the method of simulation of truncation selection proposed in this study is more general and much faster than the one proposed by Fraser (1957), since recording of the individual phenotypes in ascending or descending order would take both time and space.
CHAPTER VIII. THE PROBLEM AND THE PROCEDURE

The Problem

The genetic problem under investigation is the effect of linkage on the efficiency of selection. The following assumptions and parameters are involved in examining the problem:

1. There is one population of cross-breeding diploid organisms.
2. There is no over-lapping of generations.
3. There are two alleles per locus.
4. There are two autosomal chromosomes each with 5 loci and one sex chromosome with 5 loci.
5. The males are heterogametic and the sexlinked genes exhibit hemizygosis.
6. Two initial population compositions are considered
   (i) with maximum repulsion
   (ii) with maximum coupling.

The repulsion and coupling females and males have the following genetic constitution (expressed in binary form).

<table>
<thead>
<tr>
<th>Phase</th>
<th>Chromosome</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repulsion</td>
<td>I. autosome</td>
<td>01010</td>
<td>10101</td>
</tr>
<tr>
<td></td>
<td>II. autosome</td>
<td>10101</td>
<td>01010</td>
</tr>
<tr>
<td></td>
<td>III. sex chromosome</td>
<td>01010</td>
<td>10101</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10101</td>
<td>01010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10101</td>
<td>01010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10101</td>
<td>01010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10101</td>
<td>01010</td>
</tr>
<tr>
<td>Phase</td>
<td>Chromosome</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>--------</td>
<td>-------</td>
</tr>
<tr>
<td>Coupling</td>
<td>I. autosome</td>
<td>11111</td>
<td>11111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>00000</td>
<td>00000</td>
</tr>
<tr>
<td></td>
<td>II. autosome</td>
<td>11111</td>
<td>11111</td>
</tr>
<tr>
<td></td>
<td></td>
<td>00000</td>
<td>00000</td>
</tr>
<tr>
<td></td>
<td>III. sex chromosome</td>
<td>11111</td>
<td>01010</td>
</tr>
<tr>
<td></td>
<td></td>
<td>00000</td>
<td>01010</td>
</tr>
</tbody>
</table>

The sex chromosomes in the male have been expressed in a diploid form in which the two homologues have the same genes for manipulative purposes only.

(7) The following dominance relations were considered,

(i) no dominance
(ii) complete dominance
(iii) over dominance
(iv) mixed dominance

For all loci no dominance implies

\[
1/1 = 2, \quad 1/0 \ or \ 0/1 = 1 \quad \text{and} \quad 0/0 = 0
\]

For all loci complete dominance implies

\[
1/1 = 1, \quad 1/0 \ or \ 0/1 = 1 \quad \text{and} \quad 0/0 = 0
\]

For all loci over dominance implies

\[
1/1 = 1, \quad 1/0 \ or \ 0/1 = 2 \quad \text{and} \quad 0/0 = 0
\]

The case of mixed dominance was based on the following phenotype determination:
<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. autosome</td>
<td>1/1 = 2</td>
<td>1/1 = 2</td>
</tr>
<tr>
<td></td>
<td>0/1 or 1/0 = 2</td>
<td>1/0 or 0/1 = 1</td>
</tr>
<tr>
<td></td>
<td>0/0 = 0</td>
<td>0/0 = 0</td>
</tr>
<tr>
<td>II. autosome</td>
<td>1/1 = 1</td>
<td>1/1 = 1</td>
</tr>
<tr>
<td></td>
<td>1/0 or 0/1 = 2</td>
<td>1/0 or 0/1 = 2</td>
</tr>
<tr>
<td></td>
<td>0/0 = 0</td>
<td>0/0 = 0</td>
</tr>
<tr>
<td>III. sex chromosome</td>
<td>1/1 = 1</td>
<td>1/1 = 1</td>
</tr>
<tr>
<td></td>
<td>1/0 or 0/1 = 1</td>
<td>0/0 = 0</td>
</tr>
<tr>
<td></td>
<td>0/0 = 0</td>
<td>0/0 = 0</td>
</tr>
</tbody>
</table>

(8) The phenotypes are determined by the dominance relations.

(9) Three types of characteristics are considered,

(i) characters expressed by both sexes

(ii) characters expressed by females

(iii) characters expressed by males

(10) Selection is based on the individual's phenotype and is of the truncation type.

(11) Three types of truncation are under consideration,

(i) upper extreme

(ii) intermediate

(iii) lower extreme

(12) Two selection intensities are studied,

(i) 20 individuals out of 40 individuals for each sex

(ii) 5 individuals out of 40 individuals for each sex
(13) The values of the selection points for the different types of truncation and dominance relation are as follows,

<table>
<thead>
<tr>
<th>Dominance</th>
<th>Truncation</th>
<th>Selection point</th>
</tr>
</thead>
<tbody>
<tr>
<td>No dominance</td>
<td>upper</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>intermediate</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>lower</td>
<td>0</td>
</tr>
<tr>
<td>Complete dominance</td>
<td>upper</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>intermediate</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>lower</td>
<td>0</td>
</tr>
<tr>
<td>Over dominance</td>
<td>upper</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>intermediate</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>lower</td>
<td>0</td>
</tr>
<tr>
<td>Mixed dominance</td>
<td>upper</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>intermediate</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>lower</td>
<td>0</td>
</tr>
</tbody>
</table>

(14) The mean phenotype of the initial population under different dominance conditions are as follows,

<table>
<thead>
<tr>
<th>Dominance</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>No dominance</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Complete dominance</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Over dominance</td>
<td>30</td>
<td>23</td>
</tr>
<tr>
<td>Mixed dominance</td>
<td>20</td>
<td>18</td>
</tr>
</tbody>
</table>

(15) The selected parents are mated at random with replacement.

(16) Every individual has an equal chance of being chosen as a parent and each "mating" results in one offspring.
(17) Constant selection intensity is maintained in every generation.

(18) Selection is practiced separately in both sexes.

(19) If the characteristic is expressed by both sexes, the selected males and the selected females are mated at random to produce the progeny for the ensuing generation. This is true for,
   (i) selection for upper extreme
   (ii) selection for intermediate
   (iii) selection for lower extreme

(20) If the characteristic is expressed by either sex the selected males or females are mated with a randomly chosen female or male as the case may be. This situation also simulates the case where the selection is practiced only on one sex.

(21) The population under consideration is subjected to the following pressure,
   (i) selection
   (ii) linkage
   (iii) dominance
   (iv) random fluctuation

(22) The phenotypic values are unaffected by environmental forces.

(23) There is no epistacy.

(24) There is no mutation.

(25) There is no interference

(26) There is no differential viability

(27) The progressive changes in the mean and the variance of the selected and the unselected populations are the chief interest in the study.
The following linkage relations are considered:

(i) 0.5   (ii) 0.3   (iii) 0.1
(iv) 0.03  (v) 0.015 (vi) 0.007
(vii) 0.003 (viii) $t_f$ (ix) $t_m$

where a linkage relation of 0.5 or 0.3 and so on implies that all the adjacent loci have the same linkage relation irrespective of chromosome and sex but a linkage relation of $t_f$ and $t_m$ implies as follows:

<table>
<thead>
<tr>
<th>Cases</th>
<th>Sex</th>
<th>Locus and linkage relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_f$</td>
<td>female (I, II, III)</td>
<td>0.003 0.007 0.003 0.007</td>
</tr>
<tr>
<td></td>
<td>male (I, II)</td>
<td>0.5 0.3 0.5 0.3</td>
</tr>
<tr>
<td>$t_m$</td>
<td>female (I, II, III)</td>
<td>0.5 0.3 0.5 0.3</td>
</tr>
<tr>
<td></td>
<td>male (I, II)</td>
<td>0.003 0.007 0.003 0.007</td>
</tr>
</tbody>
</table>

The roman letters indicate chromosomes.

The Procedure

A program has been written to suit the problem at hand but it can also take care of many different situations arising out of both qualitative and quantitative expansion of the general situation by making appropriate inputs and modifications. We shall deal with such expansion in the latter part of this section. The program has the following constituent parts:

(i) control reproduction subroutine
(ii) main reproduction subroutine
(iii) phenotypic subroutine
(iv) selection subroutine
(v) master routine
(vi) random number generator (v 125)
(vii) input subroutine (N 2)
(viii) print subroutine (P 2)

The first four subroutines (i), (ii), (iii) and (iv) are special subroutines and they are specially programmed for the purpose. The last three subroutines are library subroutines available ready-made for use and their names are given in parentheses above. The master routine is the main program which controls the activities of the subroutines. Now we shall briefly describe the functions of the individual parts with particular reference to the problem at hand.

The control reproduction subroutine controls the main reproduction subroutine. Its chief function is to choose two random individuals (one male and one female), to accept offspring produced in the main reproduction subroutine and to store them in sequence in the memory locations. It repeats this process 40 times for each sex. At the end of this operation we have 40 male offspring and 40 female offspring. The control reproduction subroutine gives the sex switch for the main reproduction subroutine, so that after the desired number of female offspring have been produced by the main reproduction subroutine, the control reproduction subroutine allows the reproduction of male offspring only. The control reproduction subroutine chooses the random parents in the following way. First, it brings a random number, say \( r \), where \( 0 \leq r < 1 \), from the random number generator, then multiplies by 20 and takes the integral part and adds it to a constant. The resultant
number pertains to a memory location and the individual stored in that location is chosen as one of the parents. For example, suppose $r = 0.66$ and the individuals are stored in 200 to 239 memory location, then $0.66 \times 20 = 13.2$ and $200 + 13.0 = 213$, so then the individual located at 213 is chosen. Since the females are stored from 200 to 219 and males are stored from 220 to 239, 200 is added to the integral part if we want a female and 220 is added to the integral part if we want a male. One thing we note here is that we do not round the product while obtaining the integral part of the product.

The function of the main reproduction subroutine is to construct random masks for females and males to extract gametes from the female individual and male individual respectively. Then by associating the maternal gamete and the paternal gamete it produces one offspring per "mating." These operations are repeated 80 times and as already mentioned, by the end of the control reproduction subroutine and main reproduction subroutine operations, we have 40 female genotypes and 40 male genotypes stored in sequence.

We shall now describe the whole process in detail for the particular problem studied. First we produce a random number, say $r_1$, and then evaluate the following quantity,

$$\left[ r_1 - (1 - (lr)_1) \right]$$

where $r_1$ is the random number and $(lr)_1$ is the linkage relation of the 1st locus of the genotype under consideration. If after evaluating the quantity it is found to be negative then the location for mask is shifted (by bringing to the A register) one place to the right and in that way the
left mask bit is perpetuated. If the quantity is negative, then $-1$ is added to it and the sign is changed. This process is repeated a number of times with, $I \cdot r_2 - (1 - (1r)_2)$, $I \cdot r_3 - (1 - (1r)_3)$ and so on. In the process of evaluating the quantities, the mask location has collected 1's and 0's, depending upon the sign of the quantity $I \cdot r_1 - (1 - (1r)_1)$. If this is repeated 15 times there results a female mask and if it is repeated 10 times there results a male mask. The biological relevance of this process is that when the quantity becomes negative it indicates that there is no crossover between the two loci and so we stay either at the "bottom" or at the "top" and this is brought about by the perpetuation of the left-most bit. When the quantity becomes positive, we have a crossover at that point and so we go from top to bottom or bottom to top depending on our initial position and this is brought about by changing the sign bit from + to - or - to +. Now we evaluate the complement of the mask with regard to 0 and 1 and place it side by side with the mask in the same location. The female mask now has 30 bits and the male mask has 20 bits. Now we perform the logical multiplication of the 30 bit female genotype and 30 bit female mask. The first fifteen bits of the logical product is added to the last fifteen bits and the sum is the desired (15 bit) gamete from the female individual. The same procedure is followed for the male except that the paternal gamete has only 10 bits.

Let us illustrate the operations by a small example:

Let the genotype be

\[
\begin{array}{c}
1101 \\
1010
\end{array}
\]

and let the mask be

\[
TBBT
\]

then the gamete is

\[
1011
\]

(A)
Now let us examine if we can produce the same gamete by the above procedure:

The genotype is \(11011010\) \((B)\)

and the mask is \(1001\) if \(T = 1\) and \(B = 0\)

The mask and the complement of the mask placed side by side is \(10010110\) \((C)\)

The logical product of \((A)\)

and \((B)\) is \(10010010\)

Then add the first four bits to the last four bits and we have \(1011\) which is the same as \((A)\)

If the gametes are needed to produce a female then we extract the last 5 bits of the male individual and insert them into the last 5 positions of the paternal gamete. If the gametes are needed to produce a male, then we extract the last 5 bits of the female individual and insert them into the last 5 positions of the paternal gamete. The last bits of an individual are the sexlinked genes. At the end of the whole operation, we have 40 female genotypes and 40 male genotypes stored in sequence.

Next we evaluate the phenotypes of the genotypes stored in sequence. The chief function of the phenotypic subroutine is to evaluate the phenotypes of the individual by bringing the two bits in allelic relationship in succession to the sign bit position and then cumulating the numerical values associated with \(1/1\) or \(0/1\) or \(1/0\) or \(0/0\) \((-\,-, +-, -+, ++)\) at the location where the phenotype of the individual
is stored. The actual process has already been detailed in the last chapter. This subroutine also calculates the mean and variance of the unselected population after the whole operation is repeated 40 times for each sex and the 80 phenotypes and the number of bits for each genotype are stored in sequence. The other feature of this subroutine is that it counts the number of bits in each genotype and calculates the mean after collecting the information from 80 individuals.

Now we proceed with the selection of the phenotypes of the individuals. The actual selection mechanism and the procedure have already been described in the preceding chapter. The selection subroutine also calculates the mean and the variance of the selected population and prints them out by means of the print subroutine. If the female alone is exhibiting the characteristic, the females are subjected to the selection procedure and the males are ignored, and likewise for the case of males exhibiting the characteristic. The mean and the variance of the selected females or males are calculated and printed out. The numbers of bits associated with selected individuals are saved, and the mean is printed out.

Now we consider the chief functions of the main routine. It inputs the dominance relations for the 1/1, 0/1 or 1/0 and 0/0, phases. It inputs the linkage relations between all adjacent loci for each chromosome and sex. It transfers the members of the initial population from their permanent locations to the temporary locations. It prints the "case number" pertaining to a truncation, that is, it prints case 1 if there is selection for upper extreme, case 2 for intermediate selection, case 3 for lower extreme selection, case 4 for the upper extreme selection
when the females alone are showing the character and case 5 for the upper extreme selection when the males alone are exhibiting the characteristics. Now given a particular case number it prints the generation number and then shifts the control to the (i) control reproduction subroutine, (ii) main reproduction subroutine, (iii) phenotypic subroutine, (iv) selection subroutine in sequence. If sex is limiting the expression, it first picks random males or random females as the case may be, before shifting control to the sequence of subroutines mentioned above. Since the genotypes are produced in a random order, it picks the first 20 males or 20 females as the mates for the selected females and selected males respectively.

The main routine provides facilities for stopping at any generation while going through a particular case, and going to the next case if so desired. There is provision also for skipping a case or more and starting at the desired case.

The program offers a wide range of flexibility in the parameters involved. We could introduce different initial populations, different dominance relations, a different batch of linkage relations and different selection intensities by introducing modification tapes.

The number of words each part of the program contained is as follows; the control reproduction subroutine has 25 words, the main reproduction subroutine has 40 words, the phenotypic subroutine has 66 words, the selection subroutine has 57 words and the main routine has 85 words. The total length of the program is 350 words. Only 300 locations are used as working space. Each generation consumed
approximately 12 seconds to complete all the operations.

The functions of the routine and subroutines described above are related to one generation and to go to the next generation the whole cycle of events has to be repeated. Now briefly let us recapitulate the whole process. Two members (a male and a female) are picked at random by the control reproduction subroutine for the initial population under consideration. They are passed on to the main reproductive subroutine to extract their gametes. The offspring so resulted, is passed back to the control reproduction subroutine to be stored. When we have 80 genotypes (40 females and 40 males) stored in sequence, the phenotypic subroutine becomes operative. The latter evaluates the phenotypes of all the 80 genotypes, given the dominance relation. Given a selection point, the selection subroutine selects 20 males and 20 females and their genotypes occupy the locations where the members of the initial populations were stored at the beginning of the whole operation, and this heralds the termination of the first generation and the beginning of the next one. If the selection intensity is 5 out of 40, then the number 20 is replaced by 5 in the subroutines by certain modifications.

This program can also be used for,

(i) more loci per individual
(ii) more offspring per "mating"
(iii) any type of partial dominance
(iv) inclusion of epistatic effects
(v) more number of chromosomes
(vi) different dominance relations in different chromosomes and in two sexes
(vii) different linkage relations in different chromosomes and in two sexes.

(viii) other selection intensities and different selection intensities at different times

(ix) any level of truncation

(x) selection by intervals rather than selection aimed at one point.
CHAPTER IX. RESULTS AND DISCUSSION

General Considerations

The results of this type of study are best visualized by means of graphs, which are incorporated in the appendix for reference. Each graph contains a plotting of means and variances of the selected and the unselected populations against time, that is, against generation number. A selection of the cases considered is presented in Appendix Figures 1 to 70. On each graph the four attributes are indicated as follows:

- — mean of the parent (selected) population
- — mean of the progeny (unselected) population
- - variance of the parent (selected) population
- - variance of the progeny (unselected) population

The title of a graph has the following information,

(i) level of truncation
(ii) sex
(iii) dominance relation
(iv) linkage relation
(v) selection intensity
(vi) phase of the population under consideration.

When we refer to a situation in the text, we have to specify the above noted six parameters for identification and it will use both time and space to present the information without some sort of symbolical representation. We, therefore, propose to symbolize the specifications as follows:

$T_1$ represents selection for upper extreme
195

\[ T_2 \] represents selection for intermediate

\[ T_3 \] represents selection for lower extreme

\[ T_4 \] represents selection for upper extreme in females when male
does not exhibit the characteristic

\[ T_5 \] represents selection for upper extremes in male when female
does not exhibit the characteristic

ND represents no dominance

CD represents complete dominance

OD represents over-dominance

MD represents mixed dominance

20/40 represents the selection of 20 individuals out of 40
individuals for each sex.

5/40 represents the selection of 5 individuals out of 40
individuals for each sex.

R represents the repulsion phase of the initial population

C represents the coupling phase of the initial population

If, then, a figure has the specifications

\[(T_3)(.3)(OD)(5/40)(R)\]

they denote the case when there is selection of lower extreme, the
population has .3 as the linkage relation, over-dominance as the
dominance relation between two alleles, 5/40, that is, 5 individuals out
of 40 individuals, as the selection intensity, and repulsion as the initial
phase of the population. This type of specification will be followed
throughout the text.
Results

The following cases have been examined and only the salient features of each case will be mentioned. Graphs are given in the appendix for the cases which seemed of special interest. In counting plus genes in the males those on the sex chromosome are counted twice.

1. \( (T_1)^{1.5}(ND)(20/40)(R) \) The selection point for this case is 30. The opening mean and variance were 15.47 and 7.07 respectively for females and 15.72 and 13.54 respectively for males in the progeny population. In this situation there is no linkage and the selection force is predominant. Random fluctuation plays a part in the beginning but as the selection machinery moves the population closer to the selection point homozygosis is rapidly achieved and random fluctuation becomes negligible. The parent population and the progeny population attained the selection point at the 12\(^{th}\) and 13\(^{th}\) generation respectively. The greater variability with males is a result of the sexlinkage. This case is presented in Figures 1 and 2.

2. \( (T_2)^{1.5}(ND)(20/40)(R) \) The selection point in this case is 15. The opening mean and variance were 15.22 and 6.90 respectively for females and 15.90 and 10.40 respectively for males in the progeny population. The linkage force is rather inactive and whenever the selection is trying to move the population towards the selection point, random fluctuation is moving it back to another point which is not too far from the selection point. Since the population is at repulsion phase, this movement back and forth is almost perpetually maintained. Fixation cannot be attained because of unrestricted
recombinations. So the population forms a fluctuating genetic plateau. The number of generations studied was 15 and the mean varied between 14.6 and 15.9 during those 15 generations. Again the greater variability of males is the effect of sexlinkage. This case has been presented in Figures 3 and 4.

3. \((T^3)(0.5)(ND)(20/40)(R)\) The selection point in this case is 0. The opening mean and the variance were 15.88 and 5.77 respectively for females and 14.72 and 13.95 respectively for males in the progeny population. The interpretation of this case is almost identical with case 1 described above as far as the operation of the various forces is concerned. The parent population and the progeny population attained the selection point at the 13\(^{th}\) and 14\(^{th}\) generation respectively. But in case 1 the populations achieved the selection point one generation sooner. This can be explained by the fact that the mean phenotype of males in the initial population is 16 which offers a slight advantage to the upper extreme selection. This case has been presented in Figures 5 and 6.

4. \((T^4)(0.5)(ND)(20/40)(R)\) The selection point in this case is 30 and only the females are exhibiting the characteristic. The opening mean and variance were 15.35 and 6.18 respectively for females in the progeny population. In this case selection is weak. As the selection is trying to keep favorable alleles into the population, the unselected males are continuously introducing some unfavorable alleles. Random fluctuations and free recombinations became complementary to each other and the population reaches a fluctuating genetic plateau.
The number of generations studied was 15 and the means were between 15.5 to 19.9 during those 15 generations. This case has been presented in Figure 7.

5. \((T_g)^{(5)}(ND)^{(20/40)}(R)\) The selection point in this case is at 30 and the males only are exhibiting the characteristic. The opening mean and variance were 15.35 and 4.233 respectively for males in the progeny population. The situation is almost identical with that of case 4. In this situation the influence of sexlinkage particularly on the variances of the progeny population is strongly evidenced by the wide fluctuations in the graph. The population reaches a fluctuating genetic plateau. This case has been presented in Figure 8.

6. \((T_1)^{(3)}(ND)^{(20/40)}(R)\) The selection point in this case is at 30. The opening mean and variance were 15.37 and 7.01 respectively for females and 15.42 and 8.56 respectively for males in the progeny population. Linkage is a weak force here and so selection still remains the dominant force, and it almost reproduces the results of case 1. This is evidenced by the fact that the progeny population and the parent population attained the selection point at the 13th and 12th generations respectively. The selection differential between generations are almost the same as in case 1. There is evidence of a sexlinkage effect. This case has been presented in Figures 9 and 10.

7. \((T_3)^{(3)}(ND)^{(20/40)}(R)\) The selection point in this case is 0. The opening mean and variance were 15.10 and 4.913 respectively for females and 13.65 and 9.053 respectively for males in the progeny
population. The effect of linkage is very mild and selection is the
dominant force. The situation is almost identical with that of case 6.
The parent population and the progeny population attained the selection
point at the 13\textsuperscript{th} and 14\textsuperscript{th} generations respectively. The population in
case 6 attained the point one generation sooner. The reason for this
has been given in case 3. This case is presented in Figures 11 and 12.

8. \((T_1)(.1)(ND)(20/40)(R)\) The selection point in this case is 30. The
opening mean and the variance were 15.37 and 1.163 respectively
for females and 14.85 and 3.310 respectively for males in the
progeny population. The parent population and the progeny population
attained the selection point at the 16\textsuperscript{th} and 17\textsuperscript{th} generation respectively.
This is a clear evidence of effect of linkage in the population and the
progress has been slowed down by 4 generations as compared to case
1. The selection force is still powerful enough to move the population
up to the selection point. The variances of the selected individuals
were mostly less than 1. This case is presented in Figures 13 and 14.

9. \((T_3)(.1)(ND)(20/40)(R)\) The selection point in this case is 0. The
opening mean and variance were 15.67 and 1.25 respectively for
females and 15.05 and 1.74 for males in the progeny population. The
parent population and the progeny population attained the selection
point at the 17\textsuperscript{th} and 18\textsuperscript{th} generation respectively. The effect of
linkage is evident. The reason for the population to attain the selec­tion point one generation later than that of case 8 has been given in
case 3. Most of the variances of the selected remained less than 1.
This case is presented in Figures 15 and 16.

10. \( T_1 \{ .03 \} (ND)(20/40)(R) \) The selection point in this case is 30. The opening mean and the variance were 15.67 and 1.456 respectively for the female and 15.20 and 1.959 respectively for the male in the progeny population. This was an interesting case in that the population was under both linkage pressure and selection pressure but the force of selection was able to make the population advance for 22 generations. The restricted recombination favored fixation and 14 loci became fixed with favorable genes and one locus with unfavorable genes by the 22\(^{nd}\) generation. So selection became completely powerless after the 22\(^{nd}\) generation. The parent population and the progeny population reached a stationary genetic plateau at the 21\(^{st}\) and 22\(^{nd}\) generation, respectively, with a mean phenotype of 28 for both sexes. The population could not reach the selection point of 30. The variances for the selected and unselected were below 1 after first generation of selection. This case is presented in Figures 17 and 18.

11. \( T_3 \{ .03 \} (ND)(20/40)(R) \) The selection point in this case is 0. The opening mean and the variance were 15.125 and 1.292 respectively for females and 14.975 and 2.281 respectively for males in the progeny population. The situation here is almost identical to the previous case. The parent population and the progeny population reached a stationary genetic valley at the 22\(^{nd}\) and 23\(^{rd}\) generations respectively with a mean phenotype of 2 for both sexes. The population could not reach the selection point at 0. Most of the variances
after the first few generations were below 1. This case is presented in Figures 19 and 20.

12. (T₄)(.03)(ND)(20/40)(R) The selection point in this case is 30. The opening mean and variance were 15.75 and 1.217 respectively for females in the progeny population. The number of generations studied was 16, during which the parent population and the progeny population reached a fluctuating genetic plateau of which the range of the means was from 15.5 to 17.4. All the variances were below 1 except in the first two generations. This case is presented in Figure 21.

13. (T₃)(.03)(ND)(20/40)(R) The selection point in this case is 30. The opening mean and variance were 15.05 and 1.125 respectively for males in the progeny population. The number of generations studied was 15, during which the parent population and the progeny population reached a fluctuating genetic plateau with the range of means from 15 to 17.2. This case is given in Figure 22.

14. (T₂)(.015)(ND)(20/40)(R) The selection point here is 30. The opening mean and variance were 15.55 and 1.587 respectively for females and 14.725 and 1.845 respectively for male in the progeny population. This population behaved very interestingly in that the linkage force and the selection force tried to pull the population in opposite directions, but the selection force advanced the population for 15 generations and then when the variability of the population went down considerably, selection became inoperative. At this point the parent population reached a temporary plateau for a period
of two or three generations. Then random fluctuation became a powerful force and the fate of the population was almost entirely dependent upon this force. It so happened that the restricted recombination at this point had not yet completed fixation. Random fluctuation enabled selection to advance the population by four more units, and the parent population and the progeny population reached a final stationary genetic plateau in the 27th and 28th generations, respectively, with a mean phenotype of 28 for both sexes. Except in the first two generations, the variances were below 1. The population could not reach the selection point at 30. This case is presented in Figures 23 and 24.

15. \( (T_1)(.007)(ND)(20/40)(R) \) The selection point in this case is 30.

The opening mean and the variance were 15.25 and 1.78 respectively for females and 14.50 and 2.10 respectively for males in the progeny population. A short temporary plateau was observed at the 5th and 6th generations with mean phenotype of 18. Then selection and random fluctuation enabled the population to advance by two more units, and the population achieved another temporary plateau for 5 generations beginning at the 13th with a mean phenotype of 20. Again selection and random fluctuations enabled the population to advance by two more units, and the parent population and the progeny population reached a stationary genetic plateau at the 23rd and 24th generation, respectively, with a mean phenotype of 22 for both sexes. The variances were below 1 for all generations except the first two. The plateau finally reached was 8 units short of the selection point. This case is presented in Figures 25 and 26.
16. \((T_3)(.007)(ND)(20/40)(R)\)  The selection point in this case is 0.  The population did not form distinct temporary valleys except that there was an indication of a temporary valley in the 5\(^{th}\) generation, with a mean phenotype of 12 which is the complement of the mean phenotype of 18 reached in the corresponding upper extreme selection, and of another valley in the 12\(^{th}\) generation with a mean phenotype of 10 which corresponds to the mean phenotype of 20 in the previous case. The parent population and the progeny population attained a stationary valley at the 20\(^{th}\) and 21\(^{st}\) generation respectively with a mean phenotype of 8 for both sexes. This is 3 generations sooner than for the corresponding upper selection, which was probably due to the absence of distinct temporary valleys. This case is presented in Figures 27 and 28.

17. \((T_1)(.003)(ND)(20/40)(R)\)  The selection point here is 30. The linkage force was extremely dominant and selection effect small. The selection force and random fluctuation advanced the population for 5 generations, and the parent population and the progeny population reached a stationary genetic plateau in the 4\(^{th}\) and 5\(^{th}\) generations, respectively, with a mean phenotype of 18 for both sexes. This case is given in Figures 29 and 30.

18. \((T_3)(.003)(ND)(20/40)(R)\)  The selection point here is 0, and the results are almost identical in nature to those of the previous case. The parent population and the progeny population reached stationary genetic valleys at the 5\(^{th}\) and 6\(^{th}\) generations, respectively, with a phenotype of 12 for both sexes. This case is given in Figures 31 and 32.
19. \( (T_4)(.003)(ND)(20/40)(R) \) The selection point here is 30. The population formed a fluctuating genetic plateau. Due to restricted recombination the fluctuations were not too wide. The number of generations examined was 12 and the selected female population maintained a uniform variance throughout. The males showed wide fluctuations due to the effect of sex linkage. This case is presented in Figure 33.

20. \( (T_5)(.003)(ND)(20/40)(R) \) The selection point here is 30. The situation is almost identical with the previous one. The number of generations studied was 13 and the population reached a fluctuating genetic plateau. This case is given in Figure 34.

21. \( (T_1)(t_1)(ND)(20/40)(R) \) The selection point in this case is 30. Here the female loci had tight linkage and the male loci had free recombination. So only 10 loci out of the 30 loci possessed by females and males were segregating and the population became fixed when these 10 loci had achieved fixation. The parent population and the progeny population reached a stationary genetic plateau at the 11\(^{th}\) and 12\(^{th}\) generations respectively with a mean phenotype of 26 for both sexes. There was no indication of temporary plateau formation. This case has been presented in Figures 35 and 36.

22. \( (T_3)(t_1)(ND)(20/40)(R) \) The selection point here is zero. The situation is almost identical with the preceding case. The parent population and progeny population attained a stationary genetic valley at the 13\(^{th}\) generation with a mean phenotype of 4 for both sexes.
23. \(T_{4}\)(\(t_{f}\))(ND)(20/40)(R)\) The number of generations studied was 18 and the population reached a fluctuating plateau. The phenotypic means varied between 15.5 and 18.8.

24. \(T_{5}\)(\(t_{f}\))(ND)(20/40)(R)\) The number of generations studied was 12 and the population reached a fluctuating plateau. The phenotypic means varied between 15.5 and 17.

25. \(T_{1}\)(\(t_{m}\))(ND)(20/40)(R)\) The selection point in this case is 30. Here the male loci had tight linkage and females had free recombination. This meant that out of 30 loci between male and female, 15 loci had segregating genes and we would expect the behaviour of this case to be different than case 21 with \(t_{f}\). The parent population and the progeny population attained the selection point at the 13\(^{th}\) and 14\(^{th}\) generations respectively.

26. \(T_{3}\)(\(t_{m}\))(ND)(20/40)(R)\) The selection point in this case is 0. The results were almost identical in nature with those of the previous case. The parent population and the progeny population attained the selection point at 14\(^{th}\) and 15\(^{th}\) generations respectively.

27. \(T_{4}\)(\(t_{m}\))(ND)(20/40)(R)\) This population formed a fluctuating genetic plateau. The number of generations studied was 16 and the range of means were between 15.6 to 18.5.

28. \(T_{5}\)(\(t_{m}\))(ND)(20/40)(R)\) This population reached a fluctuating plateau. The number of generations studied was 15 and the range of means were 15.4 to 18.4.
29. \((T_1)(.5)(ND)(20/40)(C)\) The selection point is 30 and the original population was in the coupling phase. The parent population and the progeny population attained the selection point at the 11\(^{th}\) and 12\(^{th}\) generations respectively. This case gave results similar to case 1. This case is presented in Figure 37.

30. \((T_2)(.5)(ND)(20/40)(C)\) The selection point is at 15 and the population reached a fluctuating plateau. The range of means was 15 to 15.5. This case is presented in Figure 38.

31. \((T_3)(.5)(ND)(20/40)(C)\) This case gave results almost identical to case 29. The parent population and the progeny population attained the selection point at the 11\(^{th}\) and 12\(^{th}\) generations respectively. This case is presented in Figure 37.

32. \((T_4)(.5)(ND)(20/40)(C)\) The selected and unselected populations formed a fluctuating plateau. The difference between this case and the repulsion case was that both mean and variance curves in this case had wider fluctuations. This case is presented in Figure 41.

33. \((T_5)(.5)(ND)(20/40)(C)\) The results are almost identical to those of the previous case and are presented in Figure 42.

34. \((T_1)(.1)(ND)(20/40)(C)\) This is rather an interesting case since the parent population and the progeny population attained the selection point at the 8\(^{th}\) and 9\(^{th}\) generations respectively which was 8 generations sooner than the corresponding repulsion case and 3 generations sooner than case 29. This shows that the linkage force and the selection force were moving the population in the same direction.
So with a population in the coupling phase, restriction of recombinations favors upward selection by keeping the plus genes in a block. This case is presented in Figure 43.

35. \((T_2)(1)(ND)(20/40)(C)\) In this case of an initially coupled population the linkage force and the selection force combined to give a plateau which fluctuated only narrowly. The range of means was 14.7 to 15.2 except for the first generation, so the successive progeny means were almost identical. Random fluctuation produces the small variation in means which occurred. This case is given in Figure 44.

36. \((T_3)(1)(ND)(20/40)(C)\) This case is almost identical with case 34. The parent population and the progeny population attained the selection point at the 9\(^{th}\) and 10\(^{th}\) generations respectively.

37. \((T_4)(1)(ND)(20/40)(C)\) This case leads to a fluctuating genetic plateau. The variation introduced by the lack of selection in the males is reflected in the wide fluctuations around the plateau. This case is given in Figure 45.

38. \((T_5)(1)(ND)(20/40)(C)\) This case gave a plateau with wide fluctuations. This case is given in Figure 46.

39. \((T_1)(0.003)(ND)(20/40)(C)\) The parent population and the progeny population attained the selection point (30) at the 5\(^{th}\) and 6\(^{th}\) generations respectively. The corresponding repulsion case gave a plateau 12 units short of the selection point in that it led to a stationary plateau with a mean phenotype of 18 in 4 or 5 generations. This clearly shows that the genetic behaviour of repulsion
and coupling population may be distinctly different. In the case of a coupling population the tighter the linkage relation, the faster the progress, whereas in the case of a repulsion population tight linkage impedes the progress with the formation of a plateau at a point short of the selection point. This case is given in Figure 47.

40. \( (T_2)(0.003)(ND)(20/40)(C) \) This case leads to a fluctuating plateau with means lying in a very close range, and is given in Figure 48.

41. \( (T^H-0.003)(ND)(20/40)(C) \) This is almost identical with that of case 39 except the selection is in the reverse direction. The parent population and the progeny population attained the selection point at the 5th and 6th generations respectively.

42. \( (T_1)(0.5)(CD)(20/40)(R) \) The selection point here is 15. In a complete dominance situation both the homozygous dominant and the heterozygotes are favored when there is upper extreme truncation. So it will take longer to achieve complete fixation. In this particular case the mean phenotype of the initial population was 15 which was the same as the selection point, and because of free recombination selection and random fluctuation are both operative. The parent population achieved the selection point at the 10th generation with 27.3 and 27.4 as the mean number of plus genes in female and male respectively. The progeny population achieved the selection point at the 22nd generation with 29.25 and 29.52 as the mean number of plus genes in females and males respectively. At this point the parent population had averages of 29.3 and 29.65 plus genes in females and males respectively. From then on random
fluctuation carried the population to complete fixation by 42 generations, at which time there were 30 plus genes in the females. To increase the mean number of plus genes from 29.4 to 30 took 20 generations. This case is presented in Figure 49.

43. \((T_2)(.5)(CD)(20/40)(R)\) The selection point in this case is 8. The opening mean and variance were 11.42 and 2.55 respectively for females and 9.3 and 4.11 respectively for males. The population reached a fluctuating genetic valley. The number of generations studied was 29 and the mean numbers of plus genes were 13.45 and 12.9 for female and male respectively.

44. \((T_3)(.5)(CD)(20/40)(R)\) The selection point in this case is 0. The parent population and progeny population attained the selection point at the 14\(^{th}\) and 15\(^{th}\) generations respectively. In lower extreme truncation with complete dominance attainment of selection point and attainment of homozygosis are simultaneous. This case is given in Figures 50 and 51.

45. \((T_1)(.3)(CD)(20/40)(R)\) The selection point here is 15. There was a very mild effect of linkage and this population behaved almost like case 42. The selected population attained the selection point at the 11\(^{th}\) generation with 27.2 as the mean number of plus genes in both sexes. The number of generations studied was 19.

46. \((T_2)(.3)(CD)(20/40)(R)\) This case led to a fluctuating equilibrium similar to case 43.
47. \((T_3)(.3)(CD)(20/40)(R)\) This gave results similar to case 44. The parent and progeny populations reached zero at the 15\(^{th}\) and 16\(^{th}\) generations respectively, as indicated in Figures 52 and 53.

48. \((T_1)(.1)(CD)(20/40)(R)\) The selected population attained the selection point at the 16th generation with 27.1 and 28.35 as the mean number of plus genes for female and male respectively. There was an indication of an effect of linkage because 6 generations more were needed than the corresponding case with .5 as the linkage relation. The number of generations studied was 20 and one might predict that by 30 generations the parental population would attain the selection point but it would be difficult to predict the number of generations required to attain complete fixation.

49. \((T_2)(.1)(CD)(20/40)(R)\) The number of generations studied was 15 and a fluctuating plateau was formed.

50. \((T_3)(.1)(CD)(20/40)(R)\) The number of generations studied was 26, and it was seen that there was a steady decline of number of plus genes (number of 1's), so that the population would probably attain the selection point eventually. There was an indication of a temporary valley at 19, 20 and 21 generations with 3 as the mean phenotype. This case is given in Figure 54.

51. \((T_1)(.03)(CD)(20/40)(R)\) The number of generations studied was 41. The effect of linkage is evidenced by the fact that 37 generations were required for the parental population to attain the selection point, compared with 10 generations when there was free
recombination. The progeny population attained the selection point at the 40th generation with 29.1 and 28.95 as the mean number of plus genes for females and males respectively.

52. \((T_2) (.03) (CD)(20/40) (R)\) The number of generations studied was 15 and the population formed a fluctuating plateau during that time. This case was repeated and it was observed that the selected population attained the selection point at the 18th generation and in the 25th generation the progeny population did likewise, but the population had not achieved complete fixation. After 53 generations the population became completely homozygous with 16 plus genes in both sexes.

53. \((T_2) (.03) (CD)(20/40) (R)\) The population formed a stationary valley at the 24th generation with 5 as the mean phenotype and with 10 as the mean number of plus genes for both sexes. There was an indication of a temporary valley in generations 13 through 18. This case is given in Figure 55.

54. \((T_1) (.007) (CD)(20/40) (R)\) The number of generations studied was 22 and there was evidence of effects of linkage and selection. It would be difficult to predict the time at which the population might attain fixation without attaining the selection point.

55. \((T_2) (.007) (CD)(20/40) (R)\) The selected population and the progeny population attained the selection point at the 12th and 13th generations respectively. This case is given in Figures 56 and 57.
56. \( T_2 \{.007\} \) (CD) (20/40) (R) The population reached a stationary genetic valley at the 5\(^{th}\) generation with 7 as the mean phenotype and 14 as the mean number of plus genes for both sexes.

57. \( T_1 \{.003\} \) (CD) (20/40) (R) The number of generations studied was 20 and the mean number of plus genes was 16.8 or so during that time. The population might reach a genetic plateau before reaching the selection point.

58. \( T_2 \{.003\} \) (CD) (20/40) (R) The selected population and the progeny population attained the selection point at the 15\(^{th}\) and 23\(^{rd}\) generations respectively. Genetic fixation would take more than 15 generations.

59. \( T_2 \{.003\} \) (CD) (20/40) (R) The population attained a stationary genetic valley at the 18\(^{th}\) generation with 6 as the mean phenotype and 12 as the mean number of plus genes for both sexes. This case is given in Figure 58.

60. \( T_1 \{.5\} \) (OD) (20/40) (R) The selection point here is 30 and selection favors the heterozygotes. The population is initially in the repulsion phase and a fluctuating plateau is formed since fixation will not be possible in this situation. The number of generations studied was 19. and the results are given in Figure 59.

61. \( T_2 \{.5\} \) (OD) (20/40) (R) The mean phenotypes of the females and males in the initial generation are 30 and 23 respectively. Here the selection point is 15, and so the population reached a fluctuating plateau. The number of generations studied was 19.
The population reached a stationary genetic valley at the 22nd generation with 1 as the mean phenotype and 2 as the mean number of plus genes. The population attained fixation when the phase was 1/1 at one locus. See Figure 60.

The population reached a fluctuating plateau.

The population reached a fluctuating valley.

This case is interesting in that the population reached a stationary valley at the 26th generation with 3 as the mean phenotype and 6 as the mean number of plus genes. The results are given in Figure 61.

The population reached a fluctuating plateau.

The population reached a fluctuating valley.

The population attained a stationary genetic valley at the 15th generation with 7 as the mean phenotype and 14 as the mean number of plus genes. See Figure 62.

The number of generations studied was 20 and a fluctuating plateau was attained.

The population reached a fluctuating valley.

The population reached a stationary genetic valley at the 12th generation with 8 as the mean phenotype and 16 as the mean number of plus genes. See figures 63 and 64.
72. \((T_1)(.007)(OD)(20/40)(R)\) The number of generations studied was 15 and a fluctuating genetic plateau was reached.

73. \((T_2)(.007)(OD)(20/40)(R)\) The number of generations studied were 16 and the population reached a fluctuating valley.

74. \((T_3)(.007)(OD)(20/40)(R)\) The population reached a stationary valley at the 8th generation with 6 as the mean phenotype and 12 as the number of plus genes.

75. \((T_4)(.003)(OD)(20/40)(R)\) The number of generations studied was 17 and the population reached a fluctuating plateau.

76. \((T_5)(.003)(OD)(20/40)(R)\) The number of generations studied was 15 and the population reached a fluctuating valley.

77. \((T_6)(.003)(OD)(20/40)(R)\) The population reached a stationary valley at the 7th generation with 7 as the mean phenotype and 14 as the mean number of plus genes.

78. \((T_7)(.5)(MD)(20/40)(R)\) The selection point is 25. Here we have a peculiar situation. The loci of the first chromosome of the female have no dominance and favor progress for upper extreme truncation, but the loci of the 2nd and 3rd chromosome have over-dominance and complete dominance relations respectively and they impede complete fixation. In case of the males the absence of dominance in the 1st chromosome favors progress but the loci of the 2nd chromosome have over-dominant relations and this impedes fixation. Even though the recombination is free, selection cannot lead the population to fixation because of intra-chromosomal conflict, so in this
situation the population reached a fluctuating plateau. The number of generations studied was 34 and the mean number of plus genes in the selected population at the 24th generation was 26.2 for both sexes. See Figure 65.

79. \( (T_2)(.5)(MD)(20/40)(R) \) The number of generations studied was 10 and a fluctuating plateau was reached.

80. \( (T_3)(.5)(MD)(20/40)(R) \) The selected population and the progeny population attained the selection point at the 16th and 17th generations respectively. See Figure 66.

81. \( (T_1)(.3)(MD)(20/40)(R) \) The number of generations studied was 25 and the population reached a fluctuating plateau. No effect of linkage seems to be evident.

82. \( (T_2)(.3)(MD)(20/40)(R) \) The population reached a fluctuating plateau.

83. \( (T_3)(.3)(MD)(20/40)(R) \) The population reached a stationary genetic valley at the 16th generation with 1 as the mean phenotype and 2 as the mean number of plus genes. See Figure 67.

84. \( (T_1)(.1)(MD)(20/40)(R) \) The number of generations studied was 40 and the population reached a fluctuating plateau. No effect of the linkage was evident.

85. \( (T_2)(.1)(MD)(20/40)(R) \) The population reached a fluctuating plateau.

86. \( (T_3)(.1)(MD)(20/40)(R) \) The population reached a stationary valley
at the 38th generation with 3 as the mean phenotype of the population and 6 as the mean number of plus genes. There was a distinct temporary valley from the 21st to the 35th generation.

87. \( (T_1)(.03)(MD)(20/40)(R) \) The number of generations studied was 29 and the population reached a fluctuating plateau. There was some evidence of effect of linkage.

88. \( (T_2)(.03)(MD)(20/40)(R) \) The population reached a fluctuating plateau.

89. \( (T_3)(.03)(MD)(20/40)(R) \) The population reached a stationary valley at the 24th generation with 4 as the mean phenotype and 8 as the number of plus genes. There was only a slight indication of a temporary valley. See Figure 68.

90. \( (T_1)(.007)(MD)(20/40)(R) \) The number of generations studied was 20 and the population reached a fluctuating plateau. There was some indication of effect of linkage.

91. \( (T_2)(.007)(MD)(20/40)(R) \) The population reached a fluctuating plateau.

92. \( (T_3)(.007)(MD)(20/40)(R) \) The population reached a stationary valley at the 12th generation with 6 as the mean phenotype and interestingly enough 10 as the mean number of plus genes which was not (6 x). This was due to peculiar combination of dominance relations. See Figure 69.

93. \( (T_1)(.003)(MD)(20/40)(R) \) The number of generations studied was 30 and the population reached a fluctuating plateau. There was some
evidence of effect of linkage.

94. (T_2)(.003)(MD)(20/40)(R)  The population reached a fluctuating plateau.

95. (T_3)(.003)(MD)(20/40)(R)  The population reached a stationary valley at the 16^{th} generation with 9 as the mean phenotype and 14 as the mean number of plus genes in the population. See Figure 70.

96. (T_1)(.5)(ND)(5/40)(R)  This population was subjected to more intense selection and with free recombinations the progress was very rapid. Selection was the dominant force. The parent and the progeny populations attained the selection point at the 7^{th} and 8^{th} generations respectively. This was almost half as much time as in the corresponding case with 20/40.

97. (T_2)(.5)(ND)(5/40)(R)  This parent population attained the selection point right at the first generation and was almost stable. But the progeny population formed a fluctuating plateau.

98. (T_3)(.5)(ND)(5/40)(R)  The parent and the progeny populations attained the selection point at the 6^{th} and 7^{th} generations respectively. This was almost half the time taken as compared to the corresponding case with 20/40.

99. (T_1)(.3)(ND)(5/40)(R)  The parent and the progeny populations attained the selection point at the 7^{th} and 8^{th} generations respectively. This case was almost identical with case 96. There was no evidence of effect of linkage.

100. (T_3)(.3)(ND)(5/40)(R)  The parent and the progeny populations
attained the selection point at the 8\textsuperscript{th} and 9\textsuperscript{th} generations respectively.

101. \(\text{T}_1(.1)(\text{ND})(5/40)(\text{R})\) The population reached a stationary plateau at the 11\textsuperscript{th} generation with 28 as the mean phenotype. There was indication of effect of linkage.

102. \(\text{T}_3(.1)(\text{ND})(5/40)(\text{R})\) The parent and the progeny populations attained the selection point at the 10\textsuperscript{th} and 11\textsuperscript{th} generations respectively. This population did not form a valley. This was due to the fact that when the selection was intense, there was a rapid fixation process and so the random fluctuation became a major force to guide the destiny of the population.

103. \(\text{T}_1(.03)(\text{ND})(5/40)(\text{R})\) The population reached a stationary plateau at the 10\textsuperscript{th} generation with 24 as the mean phenotype. There was definite effect of linkage here.

104. \(\text{T}_3(.03)(\text{ND})(5/40)(\text{R})\) The population reached a stationary valley at the 15\textsuperscript{th} generation with 4 as the mean phenotype. There appears to be definite effect of linkage, by comparison with case 102.

105. \(\text{T}_1(.007)(\text{ND})(5/40)(\text{R})\) The population reached a stationary plateau at the 7\textsuperscript{th} generation with 20 as the mean phenotype. There was evidence of effect of linkage.

106. \(\text{T}_3(.007)(\text{ND})(5/40)(\text{R})\) The population reached a stationary valley at the 9\textsuperscript{th} generation with 10 as the mean phenotype. Both linkage and random fluctuation were strong forces in this small population with intense selection.
107. \( T_1 \{ (0.003) \text{ND} (5/40) \} (R) \) The population reached a stationary plateau at the 12\(^{th}\) generation with 18 as the mean phenotype.

108. \( T_3 \{ (0.003) \text{ND} (5/40) \} (R) \) The population reached a stationary valley at the 12\(^{th}\) generation with 10 as the mean phenotype.

109. \( T_1 \{ (0.5) \text{CD} (5/40) \} (R) \) Due to intense selection the parent population attained the selection point at the 3\(^{rd}\) generation with 22 as the mean number of plus genes. The progeny population attained the same at the 22\(^{nd}\) generation with 29.6 as the mean number of plus genes. Fixation was not achieved even at the 26\(^{th}\) generation.

110. \( T_2 \{ (0.5) \text{CD} (5/40) \} (R) \) The parent population attained the selection point at the 3\(^{rd}\) generation with 9.2 and 11 as the mean number of plus genes for female and male respectively. The progeny population reached a fluctuating plateau.

111. \( T_3 \{ (0.5) \text{CD} (5/40) \} (R) \) The parent and the progeny populations attained the selection point at the 8\(^{th}\) and 9\(^{th}\) generations respectively.

112. \( T_1 \{ (0.3) \text{CD} (5/40) \} (R) \) The parent population attained the selection point at the 5\(^{th}\) generation with 24.8 and 25 as the mean number of plus genes for female and male respectively. The progeny population attained the same at the 25\(^{th}\) generation with 29.8 as the mean number of plus genes. At the 30\(^{th}\) generation the population attained complete homozygosis.

113. \( T_2 \{ (0.3) \text{CD} (5/40) \} (R) \) The parent population attained the selection point at the 3\(^{rd}\) generation with 11.4 as the mean
number of plus genes. The progeny population reached a fluctuating plateau.

114. \((T^3)(.3)(CD)(5/40)(R)\) The population reached a stationary valley at the 9th generation with 3 as the mean phenotype.

115. \((T^1)(.1)(CD)(5/40)(R)\) The parent population attained the selection point at the 9th generation with 27.4 and 27.8 as the mean number of plus genes for female and male respectively. The study was discontinued before the progeny population attained the selection point. The mean number of plus genes was 28.5.

116. \((T^2)(.1)(CD)(5/40)(R)\) The parent population attained the selection point at the 4th generation with 13.6 and 14.6 as the mean number of plus genes for males and females respectively. The progeny population reached a fluctuating plateau.

117. \((T^3)(.1)(CD)(5/40)(R)\) The population reached a stationary genetic valley at the 8th generation with 5 as the mean phenotype.

118. \((T^4)(.03)(CD)(5/40)(R)\) The number of generations studied for this case was 50. This case showed strong effect of linkage. The parent population attained the selection point at 22nd generation with 22.4 and 21.8 as the mean number of plus genes for female and male respectively. The fixation of the progeny population was still incomplete at the 50th generation. The steady increase of plus genes might lead to the conclusion that the former might attain the selection point sooner.
119. \( T_2 \cdot 0.03 \cdot CD \cdot 5/40 \cdot R \) The parent population attained the selection point at the 3\(^{rd}\) generation with 13.6 and 13.2 as the mean number of plus genes for females and males respectively. The progeny population reached a fluctuating plateau.

120. \( T_3 \cdot 0.03 \cdot CD \cdot 5/40 \cdot R \) The population reached a stationary valley at the 12\(^{th}\) generation with 5 as the mean phenotype.

121. \( T_1 \cdot 0.007 \cdot CD \cdot 5/40 \cdot R \) The number of generations studied was 25 and the parent population did not attain the selection point. There was only 17 plus genes in the population. This was a strong evidence of linkage which was impeding the progress.

122. \( T_2 \cdot 0.007 \cdot CD \cdot 5/40 \cdot R \) The parent population attained the selection point at the 4\(^{th}\) generation and the progeny population attained the selection point at the 15\(^{th}\) generation. Complete fixation was attained at the 19\(^{th}\) generation.

123. \( T_3 \cdot 0.007 \cdot CD \cdot 5/40 \cdot R \) The population reached a stationary genetic plateau at the 4\(^{th}\) generation with 7 as the mean number of plus genes.

124. \( T_1 \cdot 0.003 \cdot CD \cdot 5/40 \cdot R \) There is strong effect of linkage here. The parent and the progeny population did not attain the selection point even after 16 generations of selection. The progress was impeded to a great extent.

125. \( T_2 \cdot 0.003 \cdot CD \cdot 5/40 \cdot R \) The parent and the progeny populations attained the selection point at the 2\(^{nd}\) and 4\(^{th}\) generations respectively. At the 4\(^{th}\) generation the population attained com-
plete fixation (homozygosis).

126. \((T_3)(0.003)(CD)(5/40)(R)\) The population reached a stationary genetic valley at the 4th generation with 8 as the mean phenotype.

127. \((T_1)(0.5)(OD)(5/40)(R)\) The number of generations studied was 15 and the population reached a fluctuating plateau.

128. \((T_2)(0.5)(OD)(5/40)(R)\) The parent population attained the selection point at the 12th generation but the progeny population reached a fluctuating plateau.

129. \((T_3)(0.5)(OD)(5/40)(R)\) The population reached a stationary genetic valley at the 7th generation with 3 as the mean phenotype.

130. \((T_1)(0.3)(OD)(5/40)(R)\) The number of generations studied was 17 and the population reached a fluctuating plateau.

131. \((T_2)(0.3)(OD)(5/40)(R)\) The parent population attained the selection point at the 9th generation with 17.4 and 18 as the mean number of plus genes. The progeny population reached a fluctuating plateau.

132. \((T_3)(0.3)(OD)(5/40)(R)\) The population reached a stationary valley at the 8th generation with 3 as the mean phenotype.

133. \((T_1)(0.1)(OD)(5/40)(R)\) The population reached a fluctuating plateau.

134. \((T_2)(0.1)(OD)(5/40)(R)\) The parent population attained the selection point at the 8th generation with 15 as the mean number of plus genes. The progeny population reached a fluctuating plateau.
135. \((T_3)(.1)(OD)(5/40)(R)\) The population reached a stationary valley at the 5\(^{th}\) generation with 8 as the mean phenotype.

136. \((T_1)(.03)(OD)(5/40)(R)\) There is some indication of the effect of linkage in this case. The population reached a fluctuating plateau.

137. \((T_2)(.03)(OD)(5/40)(R)\) The parent population attained the selection point at the 2\(^{nd}\) generation and the progeny population reached a fluctuating plateau.

138. \((T_3)(.03)(OD)(5/40)(R)\) The population reached a stationary valley at the 4\(^{th}\) generation with 7 as the mean phenotype.

139. \((T_1)(.007)(OD)(5/40)(R)\) There is some indication of effect of linkage. The population reached a fluctuating plateau.

140. \((T_2)(.007)(OD)(5/40)(R)\) The parent population attained the selection point at the 2\(^{nd}\) generation and the progeny population reached a fluctuating plateau.

141. \((T_3)(.007)(OD)(5/40)(R)\) The population reached a stationary plateau at the 4\(^{th}\) generation with 8 as the phenotypic mean.

142. \((T_1)(.003)(OD)(5/40)(R)\) There was strong indication of effect of linkage. The population reached a fluctuating plateau.

143. \((T_2)(.003)(OD)(5/40)(R)\) The parent population attained the selection point at the 1\(^{st}\) generation itself with 15 as the mean number of plus genes. The selected population maintained the same number of plus genes (15) all throughout. The progeny population reached a fluctuating plateau.
144. \(T_3\{.003\}(OD)(5/40)(R)\) The population reached a stationary valley at the 4\textsuperscript{th} generation with 7 as the phenotypic mean.

145. \(T_1\{.5\}(MD)(5/40)(R)\) The number of generations studied was 20. The population reached a fluctuating plateau.

146. \(T_2\{.5\}(MD)(5/40)(R)\) The parent population attained the selection point (13) at the 7\textsuperscript{th} generation with 13.8 and 13 as the mean number of plus genes for female and male respectively. The progeny population reached a fluctuating plateau.

147. \(T_3\{.5\}(MD)(5/40)(R)\) The population reached a stationary valley at the 11\textsuperscript{th} generation with 1 as the mean phenotype of the population.

148. \(T_1\{.3\}(MD)(5/40)(R)\) The number of generations studied was 17 and the population reached a fluctuating plateau.

149. \(T_2\{.3\}(MD)(5/40)(R)\) The parent population attained the selection point at the 3\textsuperscript{rd} generation and the progeny population reached a fluctuating plateau.

150. \(T_3\{.3\}(MD)(5/40)(R)\) The population reached a stationary valley at the 11\textsuperscript{th} generation with 2 as the phenotypic mean.

151. \(T_1\{.1\}(MD)(5/40)(R)\) A peculiar result was obtained. The parent population reached a stationary plateau at the 19\textsuperscript{th} generation with 24 as the mean phenotype but the progeny population reached a fluctuating plateau. This was a strong indication of effect of linkage. The number of plus genes for the parent population was 23 at the plateau.
152. \( (T_2)^{(1.1)(MD)(5/40)(R)} \) The parent population attained the selection point at the 3\textsuperscript{rd} generation with 14.2 and 15.8 as the mean number of plus genes for female and male respectively.

153. \( (T_3)(1)(MD)(5/40)(R) \) The population reached a stationary valley at the 10\textsuperscript{th} generation with 4 as the mean phenotype.

154. \( (T_4)(.03)(MD)(5/40)(R) \) The progeny population reached a fluctuating plateau but the parent population behaved rather peculiarly due to random fluctuations. At the 2\textsuperscript{nd} generation the females reached a temporary plateau with 21 as the mean phenotype and 16 as the mean number of plus genes. The males at the same time reached a temporary plateau with 19 as the mean phenotype and 17 as the mean number of plus genes. The parent population maintained almost the same phenotypes for all further generations.

155. \( (T_5)(.03)(MD)(5/40)(R) \) The parent population attained the selection point at the 3\textsuperscript{rd} generation and the progeny population reached a fluctuating plateau.

156. \( (T_6)(.03)(MD)(5/40)(R) \) The population reached a stationary valley at the 15\textsuperscript{th} generation with a mean phenotype of 4.

157. \( (T_7)(.007)(MD)(5/40)(R) \) At the 3\textsuperscript{rd} generation the parent population reached a stationary plateau with 21 as the mean phenotype and 16 plus genes for females and with 19 as the mean phenotype and 17 plus genes for males. This is the first case that male and female in the same selected population achieved
different plateaus.

158. \((T_2)(.007)(MD)(5/40)(R)\) The parent population attained the selection point at 13 and the progeny population attained the same at generation 19. It is interesting that the population actually attained the selection point with 20 plus genes.

159. \((T_3)(.007)(MD)(5/40)(R)\) The population reached a stationary valley at the 4th generation with 9 as the mean phenotype and 14 as the mean number of plus genes.

160. \((T_1)(.003)(MD)(5/40)(R)\) The progeny population reached a fluctuating plateau. The parent population reached a stationary plateau; the females had 21 as the mean phenotype and 16 as the mean number of plus genes and the males had 19 as the mean phenotype and 17 as the mean number of plus genes.

161. \((T_2)(.003)(MD)(5/40)(R)\) The number of generations studied was 32. The progeny population reached a fluctuating plateau. But the parent population reached a stationary plateau with 13 and 18 as the mean phenotype and mean number of plus genes respectively for females and 12 and 18 as the mean phenotype and mean number of plus genes respectively for males. This plateau is reached at the 23rd generation.

162. \((T_3)(+003)(MD)(5/40)(R)\) The population reached a stationary genetic valley at the 3rd generation with 9 as the mean phenotype and 14 as the mean number of plus genes in the population at the
point of plateau formation.

Discussion

The complexity of the venetic situation incorporated into the program is modest compared to that associated with actual biological material. More complex situations can be studied in the same program but the knowledge of the behaviour of relatively simple ones is a necessary prerequisite to comprehension of more elaborate situations. The choices of the values of the parameters investigated is somewhat arbitrary and the results of some of the preliminary runs influenced the later choices.

As a general conclusion from the results of the cases studied, the behaviour of a population in maximum repulsion is opposite to that of a population in maximum coupling. When there was unrestricted recombination both populations attained the selection point approximately in the same duration, as was to be expected. When there are tighter linkage relations in a repulsion population the selection force and linkage force exert pressure in opposite directions and the progress of the population is impeded. For example, with no dominance and .003 as the linkage value, selection is almost powerless and linkage and random fluctuation lead the population to a stationary plateau which is far short of the selection point under consideration. So, for a repulsion population, the tighter the linkage, the less progress selection makes. When there is tight linkage in a coupling population the forces of selection and linkage behave in a complementary way in that each produces pressure in the same direction.
and the progress of the population is accelerated. For example, with no dominance and .003 as the linkage value the population attained the selection point after 6 generations of upper extreme selection but under the same dominance relation the selection point was achieved after 9 and 12 generations with .1 and .5 as the linkage values respectively and in this particular case the population never formed a plateau. So, for a coupling population, the tighter the linkage, the faster is the progress.

When a population is initially in maximum repulsion and there is no dominance among the two alleles at a locus, the progress of the population under selection depends very much upon the relative strength of two opposing forces, linkage and selection. If the selection force is powerful, the population achieves the selection point under consideration and if the linkage force is powerful, it results in the formation of a plateau. In this case, selection carried the population up to the selection point only when the linkage values were .5, .3 and .1, but when the linkage values were .03 and .003 the population reached a plateau. Linkage values intermediate between .1 and .03 resulted in interesting behaviour of the population. In a non-stationary population with intermediate linkage values the two opposing forces establish a partially balanced relationship in that sometimes one and sometimes the other determined the progress. It is interesting to note however that as the population progressed a complete balance is achieved between the two forces for a short time and as a consequence, the population passes into a stationary state. When a population has reached such a state, it is said to have formed a temporary plateau. At this point random fluctuation becomes the powerful force and leads the population to fixation. In
the case when the linkage was .015, the population formed a temporary plateau for a brief period and then the random fluctuation favored the selection force which moved the population very close to the selection point. But in the case when the linkage was .007, the population reached a distinct temporary plateau and the random fluctuation favored the linkage force and hence the population reached a stationary plateau at a short distance from the temporary one. In case of intense selection, a slight restriction on the formation of recombinants favors a plateau formation. The linkage force becomes powerful and leads the population to rapid fixation. The plateaus for the different linkage values are lower in heights with the increase of the intensity of the selection. Unrestricted recombination and intense selection favors fast progress.

In the case when there is no dominance between the two alleles at each locus there exists a symmetric relation between upper extreme truncation and lower extreme truncation for less intense selection. For example, with .003 as the linkage relation and 20/40 as the intensity of the selection, a stationary plateau was formed with 18 as the mean phenotype for upper extreme truncation and a stationary valley was formed with 12 as the phenotype for lower extreme truncation. The symmetric relation is completely lost when the intensity of selection and tightness of linkage is increased. There is ample evidence of this principle situation in this study.

In the complete dominance situation where selection favors the heterozygotes and the plus homozygotes equally, the variability of the population is maintained longer than for the case of no dominance. Selection is sensitive over a longer period of time. Random fluctuations
consume more time in leading the population to complete fixation. The effect of the restricted recombination is best visualized by examining the relative duration of the attainment of selection point by the parent population with different linkage values. The tighter the linkage, the longer is the duration. In lower extreme truncation the selection point is attained after a large number of generations of selection as compared to the corresponding cases when there is no dominance and the formation of valleys is favored when the linkage falls below .03. The symmetrical relation between the upper and lower truncations is lost even in the case of less intense selection. It is interesting to note that the selection point is attained when there is selection for intermediates particularly under the condition of tight linkage in contrast to the case of no dominance, where a fluctuating plateau is invariably favored. In the case of intense upper extreme selection the whole process, that is, reaching the selection point or a plateau, is accelerated, selection is more powerful, and random fluctuations consume less time to achieve fixation. In the lower extreme selection, the progress is faster under the condition of unrestricted recombination but a slight restriction on recombination favors valley formation. For example, when the linkage was .3 and the intensity of selection was 5/40 a stationary valley was formed with 3 as the mean phenotypic value.

When the initial population has maximum repulsion and there is over-dominance, the progress of the population depends largely on random fluctuation. The heterozygotes are favored at the cost of both homozygotes and hence the variability of the population is perpetually maintained. Both the opposing forces, selection and linkage, became weak and the random fluctuation leads the population to a fluctuating plateau when there
is selection for upper extremes. The progeny population and the parent population form two distinct plateaus. But the interesting feature here is that, as the linkage relations became tighter, the selection force and the linkage force exert pressure on these two plateaus in such a way that the two plateaus are pulled away from each other further depending on the degree of restriction on recombination, thus forming two non-overlapping distinct fluctuating plateaus. For example, in the case where the linkage was .007 the mean distance between two plateaus was about 6 units. When there was intense selection the mean distance between the two plateaus was 10 units for the case where .03 was the linkage relation and 5/40 was the selection intensity. In the case of lower extreme truncation the linkage force and random fluctuation favor rapid fixation and even when there is no restriction on recombination the process leads to valley formation. As the linkage becomes tighter, the valleys get more and more shallow. This is even more pronounced in case of intense selection. All these lead to a complete asymmetric relation between the two truncations. The selection force in an intermediate truncation scheme becomes a victim of the random fluctuation and a fluctuating plateau results when there is less restriction on recombinations. It is interesting to note that as the linkage becomes tighter the selection point is achieved, establishing the fact that selection and random fluctuations become complementary to each other in this situation.

In the case of mixed dominance, the dominance relation between two alleles at a locus on one chromosome is different from that on another chromosome in the same individual. The first chromosome and the second chromosome have no dominance and over-dominance respectively
in both sexes and the third chromosome in the female had complete
dominance. When the initial population has maximum repulsion and has
the mixed dominance and is subjected to a selection pressure, the
existence of an asymmetrical dominance relation among the chromo-
somes induces an intense interchromosomal competition, the first
chromosome of both sexes favoring rapid fixation and the other chromo-
somes of both sexes favoring the perpetuation of the variability in the
population. So when there is upper extreme trunction the population
reaches a fluctuating plateau. The progeny population always attains a
fluctuating plateau irrespective of linkage values and intensities of
selection. It is interesting to note that the parent population reaches a
stationary plateau when there is intense selection and tighter linkage
relation as opposed to the corresponding cases in less intense selection.
It is more interesting to note that under mixed dominance conditions,
intense selection and restricted recombinations, the female parent and the
male parent population reach two distinct non-overlapping plateaus
contrary to the results obtained in relation to other corresponding
situations. For example, when there was selection for the upper extreme
with .003 as the linkage values and 5/40 as the selection intensity, the
female parent population reached a plateau with 21 and 16 as the mean
phenotype and mean number of plus genes respectively and the male
parent population reached a plateau with 19 and 17 as the mean phenotype
and mean number of plus genes respectively after 3 generations of
selection. The number of generations studied was 15 and the results
were consistent over the generations. In the case of lower extreme
truncation with less intense selection, the selection point is achieved but as the linkage becomes tighter the formation of valleys is favored. Distinct temporary valleys are formed for the case with linkage relations less than .1 and the number of plus genes in the individuals in the valley is not half of the mean phenotype of the population as happened in the corresponding cases of other types of dominance. For example, a valley was reached at the 12th generation for the case when .003 was the linkage value with 9 as the phenotypic mean and 14 as the number of plus genes in the population. Lower extreme truncation and intense selection favors valley formation even when there is no restriction on recombination. These factors establish an asymmetric relation between the two extreme truncations. When there is selection for intermediates the population reaches a fluctuating plateau, but the selection point is achieved with intense selection.

If the selection pressure is exerted in a population in which the loci in the females have very tight linkage, the loci in the male have no restriction on recombination and there is no dominance for both male and female locus, the population reaches a stationary plateau or a stationary valley depending upon the level of truncation. This is because the non-segregating loci out-number the segregating ones and lead the population to rapid fixation. If on the contrary the male has tight linkage and the female has loose linkage the population attains the selection point. That is to say that if the heterogametic sex has tight linkage and the homogametic sex has unrestricted recombination of loci, the population achieves the selection point. The latter situation partially resembles the case of Drosophila where there is complete restriction on recombination.
in the heterogametic sex.

If the selection pressure is exerted on a population in which either sex is exhibiting the characteristic and selection is only operative on the sex exhibiting the characteristic, the population reaches a fluctuating plateau under any linkage relation. The favorable genes introduced by the selection of upper extremes are nullified by the unfavorable genes introduced by the sex not exhibiting the characteristic and hence the fluctuating plateau results.

Fraser (1958) and Martin and Cockerham (1960) have attacked a similar problem but their studies were restricted to the case where selection for upper extremes was practiced in a repulsion population with no dominance and complete dominance. Their conclusions for these cases are similar to the present findings in that tight linkage slows down the progress from selection in populations which are in maximum repulsion. In this study, as already seen, consideration has been extended to two distinct phases of a population, four types of dominance relations, three types of characteristics and 9 different linkage relationships.

Extrapolation of the results at this stage is considered rather dangerous but the fact that can be clearly visualized is that in general intense selection in small populations does not necessarily result in progress because the forces in operation may lead the population to an undesired plateau.
CHAPTER X. SUMMARY

The chief objective of this study was to examine the role of sexlinked (complete and partial) genes in quantitative inheritance. Two methods of approach were followed, the mathematical approach and the Monte Carlo approach. Numerical solutions were obtained by the Monte Carlo method to the problems which did not yield to algebraic solution. The partition of the genotypic variance into their respective constituents and the derivation of the covariance between relatives were obtained by mathematical approach. A Monte Carlo investigation was made to study the effect of linkage on the efficiency of selection.

In Chapter III we expressed the genotypic values of male and female by a mathematical model under the assumption of single locus, multiple alleles and complete sexlinked transmission. The gene effects were defined by the application of least squares. After proving the orthogonality of each term in the model, the total genotypic variance of the homogametic sex yielded to the following decomposition,

\[ \sigma^2_{G_f} = \sigma^2_A_f + \sigma^2_D_f \]

where \( A \) and \( D \) stand for additive and dominance. Three types of covariances between relatives were recognized, male and male, female and female, and male and female. Corresponding to each type of combination three probability quantities were introduced, \( w_{XY} \), \( r_{XY} \) and \( s_{XY} \) and the following covariances were obtained in terms of these quantities:
\[
\text{Cov}(X^m, Y^m) = v_{XY} \sigma^2_{Am}
\]

\[
\text{Cov}(X^f, Y^f) = \frac{2 r^*_{XY} \sigma^2_{Af}}{\sigma^2_{XY}} + u^2_{XY} \sigma^2_{Df}
\]

\[
\text{Cov}(X^m, Y^f) = w_{XY} \frac{1}{2} C_{As}
\]

The multiple loci theory was approached under the joint consideration of autosomal genes and sexlinked genes. The total genotypic variances for male and female were partitioned into their respective constituents under the assumptions of arbitrary number of alleles, arbitrary number of loci, arbitrary epistacy and independent segregation. The genotypic values of male and female were expressed in terms of a population identity, each term was assigned its appropriate genetic interpretation, the orthogonality of each term was established and the total variances were decomposed into appropriate variance components. The general structure of the covariance between relatives was constructed under the same assumptions. The important feature of the derivation is the inclusion of the epistatic deviations arising from autosomal x sexlinked and sexlinked x sexlinked gene interactions.

Simple formulae were derived for establishing the probabilities of relationships between genes for the case of complete sexlinkage. These were used to study the progress of the panmictic index with regular and irregular systems of inbreeding under the assumption of sexlinked transmission. The recurrence relations over generations for some of the special cases of interest were tabulated.

The general structure of the covariances between individuals belonging to a population which is generated by random mating of the members
of a population which, in turn, resulted by inbreeding a random mating population to the extent measured by $F$, the inbreeding coefficient of homogametic sex, was constructed for both single locus and multiple loci cases. Tables of coefficients were constructed under the assumption that the main effects and the first order epistatic effects were important.

In Chapter IV, we examined the role of partially sexlinked genes in quantitative inheritance. The partition of the total genotypic variance and the construction of the general structure of the covariances between relatives was achieved by the same procedure adopted in the case of completely sexlinked genes. The peculiar situation involved in this case was that there was symmetrical frequency relationship between two genes located in an autosomal locus but an asymmetric probability relation existed between two genes at a partially sexlinked locus and the situation was taken care of by evaluating the probability quantities $v_{XY}^I$, $r_{XY}^I$, and $w_{XY}^I$ in terms of $r$, the recombination frequency.

The fundamental structure of the covariance between relatives was derived under the simultaneous consideration of autosomal, completely sexlinked, and partially sexlinked inheritance. The important feature in the derivation was the inclusion of the epistatic deviations arising from autosomal x sexlinked, autosomal x partially sexlinked, sexlinked x partially sexlinked, sexlinked x sexlinked and partially sexlinked x partially sexlinked gene interactions.

Due to the asymmetrical frequency relations between the genes at a locus, difficulty was encountered in achieving simple relationships among the quantities $r_{XY}^I$, $v_{XY}^I$, and $w_{XY}^I$. The recurrence relation
over generations of the panmictic index was established by expressing the vector associated with the $n^{th}$ generation in terms of the vector associated with the first generation through a transitional matrix. It was pointed out that the solution of the canonical roots and the canonical vectors would give the required result. The transitional matrices of some of the special cases were presented.

The general structure of the covariance between relatives was derived under the assumption of inbreeding followed by random mating, in terms of $F_m$ and $F_f$, the inbreeding coefficients of male and female respectively, to achieve some degree of generality. Special cases of interest were tabulated.

In Chapter V we examined the mathematical consequences of the assumption of gene effects being dependent on sex on the structure of the total genotypic variances and covariances between relatives. Under autosomal inheritance, the total genotypic variances of male and female yielded to the following decomposition,

$$V_G^m = V_A^m + V_D^m$$

and

$$V_G^f = V_A^f + V_D^f$$

Multiple loci theory was developed under the simultaneous consideration of autosomal, sexlinked and partially sexlinked gene transmission.

In Chapter VI we applied statistical procedure to estimate the various components of variances and covariances. The components of variances obtained from the analysis of variance for the characteristic
expressed by either sex were expressed in terms of components of variances due to additive, dominance, additive x additive effects and so on, and the heritability estimates under the assumption of sex-linked inheritance were tabulated. The analysis of variance constructed for the characteristic expressed by both male and female yielded the estimation of the components of variances due to sex influence, genotype sex interaction, genotype x environment interaction and all the genetic components of variances such as additive, dominance and so on. The estimation procedure was also studied for inbreeding followed by random mating.

Chapters VII and VIII were devoted to the description of the program for the Cyclone to study the effect of linkage on the efficiency of selection. The enumeration of results and discussion of the Monte Carlo investigation were presented in Chapter IX.

The biological parameters involved in this study were as follows: two types of population, repulsion and coupling; four types of dominance, no dominance, complete dominance, over dominance and mixed dominance; three types of characters, characters expressed by male, female and both sexes; three types of truncation, upper extreme, intermediate and lower extreme; two types of selection intensities, 20/40 and 5/40, and nine types of linkage relations, .5, .3, .1, .03, .015, .007, .003, t_f and t_m where t_f stands for tight linkage in female and t_m stands for tight linkage in male.

A short resume of the interpretation of the results was as follows. When the linkage force was predominant, the progress of the population under selection was impeded and the population had a tendency to reach a plateau or a valley far short of the selection point.
selection force was predominant, the population attained the selection point. When both the forces were balances, the population tended to reach a temporary plateau, and when random fluctuation was powerful a fluctuating plateau was formed. When there was selection for intermediates, tight linkage, and all other dominances except the no-dominance relation, the populations attained the respective selection points in most of the cases. No extrapolation of the results was made but the fact that was clearly visualized was that intense selection in small populations did not necessarily result in progress.
CHAPTER XI. LITERATURE CITED


_______ 1939. Cytological and genetic basis of sex in sex and internal secretions. 2nd ed. Williams and Wilkins, Baltimore, Maryland.


1917. The numerical results of diverse systems of breeding with respect to two pairs of characters, linked or independent, with special relation to the effects of linkage. Genetics, 2: 97-154.


1955. The correlation between relatives in inbred populations. Genetics, 40: 681-691.


Lush, J. L. 1948. The genetics of populations. (Mimeo.) Animal Husbandry Dept., Iowa State University of Science and Technology, Ames, Iowa.


CHAPTER XII. ACKNOWLEDGEMENTS

The author wishes to express his deepest gratitude to Professor Oscar Kempthorne for introducing the author to the vast fields of genetic statistics and experimental design and suggesting the problem. Without Professor Kempthorne's encouragement and invaluable suggestions during the preparation of the manuscript, the work could not have taken its final form.

The author wishes to thank Dr. Earl Schweppe and Mr. Paul Sampson for helping in programming for the Cyclone computer.

The author wishes to thank Dr. J. L. Lush and Dr. T. A. Bancroft for encouragement during graduate work, and Iowa State University for granting free time on the Cyclone.

Heartfelt thanks are due to Mrs. Audrey Graham for her patience, kindness and intelligence in deciphering the author's manuscript and her skillful preparation of the typewritten manuscript.

Finally, the author wishes to thank Mrs. Margaret Willey for the beautiful job of inking.
Figure 1. The graph relates to females with the following specifications: 
\( T_1(5)(ND)(20/40)(R) \)

- - - - - mean of the parent (selected) population
- - - - - mean of the progeny (unselected) population
- - - - - variance of the parent (selected) population
- - - - - variance of the progeny (unselected) population
Figure 2. The graph relates to males with the following specifications:
$(T_1)(.5)(ND)(20/40)(R)$
VARIANCES

MEANS

GENERATIONS
Figure 3. This graph relates to females with the following specifications: 
\( (T_2)(.5)(\text{ND})(20/40)(R) \)
Figure 4. This graph relates to males with the following specifications:
(T₂(.5)(ND)(20/40)(R)}
Figure 5. This graph relates to females with the following specifications: 
$(T_3)(0.5)(ND)(20/40)(R)$
Figure 6. This graph relates to males with the following specifications:
\( (T_3) .5 (\text{ND}) (20/40)(R) \)
Figure 7. This graph relates to females with the following specifications: $(T_4)(.5)(ND)(20/40)(R)$
Figure 8. This graph relates to males with the following specifications: 
\(T_5(.5)(ND)(20/40)(R)\)
Figure 9. This graph relates to females with the following specifications: $\{T_1\}(.3)(ND)(20/40)(R)$
Figure 10. This graph relates to males with the following specifications: 
\( T_1 \) (.3) (ND) (20/40) (R)
Figure 11. This graph relates to females with the following specifications:
\((T_3)(3)(ND)(20/40)(R)\)
Figure 12. This graph relates to males with the following specifications:
\( \{T_3\}(.3)(ND)(20/40)(R) \)
Figure 13. This graph relates to females with the following specifications:
\((T_1)(N)(D)(20/40)(R)\)
Figure 14. This graph relates to males with the following specifications: \( T_1 \cdot 1 \cdot \text{ND} \cdot \frac{20}{40} \cdot \text{R} \)
Figure 15. This graph relates to females with the following specifications: $(T_3)(1.1)(ND)(20/40)(R)$
Figure 16. This graph relates to males with the following specifications:
(T3)(.1)(ND)(20/40)(R)
Figure 17. This graph relates to females with the following specifications:
\((T_1)^{(0.03)}(ND)(20/40)(R)\)
Figure 18. This graph relates to males with the following specifications: \( T_1 \) (0.03) (ND) (20/40) (R)
Figure 19. This graph relates to females with the following specifications:
\{T_3\}(.03)(ND)(20/40)(R)
Figure 20. This graph relates to males with the following specifications:
\( T_3(0.03)(ND)(20/40)(R) \)
Figure 21. This graph relates to females with the following specifications:

\( T_4 \{ 0.03 \} \{ ND \} \{ 20/40 \} \{ R \} \)
Figure 22. This graph relates to males with the following specifications: 
($T_5$)(.03)(ND)(20/40)(R)
Figure 23. This graph relates to females with the following specifications: $(T_1)(.015)(ND)(20/40)(R)$
Figure 24. This graph relates to males with the following specifications:
$(T_1)(.015)(ND)(20/40)(R)$
Figure 25. This graph relates to females with the following specifications: $\{T_1\}_{<.007}$ (ND) (20/40) (R)
Figure 26. This graph relates to males with the following specifications: 
(T1)(.007)(ND)(20/40)(R)
Figure 27. This graph relates to females with the following specifications:
$(T_3)(.007)(ND)(20/40)(R)$
Figure 28. This graph relates to males with the following specifications:
\[ T_3(0.007)(ND)(20/40)(R) \]
Figure 29. This graph relates to females with the following specifications:
\[(T_1)(.003)(ND)(20/40)(R)\]

Figure 30. This graph relates to males with the following specifications:
\[(T_1)(.003)(ND)(20/40)(R)\]
Figure 31. This graph relates to females with the following specifications: $T_3(.003)(\text{ND})(20/40)(R)$

Figure 32. This graph relates to males with the following specifications: $T_3(.003)(\text{ND})(20/40)(R)$
Figure 33. This graph relates to females with the following specifications: $T_4(0.003)(ND)(20/40)(R)$

Figure 34. This graph relates to males with the following specifications: $T_3(0.003)(ND)(20/40)(R)$
VARIANCES

MEANS

GENERATIONS

90
Figure 35. This graph relates to females with the following specifications:
$$\{T_1, t_1, \text{ND}, (20/40), \text{R}\}$$

Figure 36. This graph relates to males with the following specifications:
$$\{T_1, t_1, \text{ND}, (20/40), \text{R}\}$$
Figure 37. This graph relates to females with the following specifications: $(T_1 \cdot 5)(ND)(20/40)(C)$

Figure 38. This graph relates to females with the following specifications: $(T_2 \cdot 5)(ND)(20/40)(C)$
Figure 39. This graph relates to females with the following specifications: 
\( (T_3)(.5)(ND)(20/40)(C) \)

Figure 40. This graph relates to females with the following specifications: 
\( (T_3)(.1)(ND)(20/40)(C) \)
Figure 41. This graph relates to females with the following specifications: 
$T_4(0.5)(ND)(20/40)(C)$
Figure 42. This graph relates to females with the following specifications: $(T_B, 0.5, ND, 20/40, C)$
Figure 43. This graph relates to females with the following specifications: \( T_1 \times 1 \times \text{ND} \times 20/40 \times \text{C} \)

Figure 44. This graph relates to females with the following specifications: \( T_2 \times 1 \times \text{ND} \times 20/40 \times \text{C} \)
Figure 45. This graph relates to females with the following specifications:
$(T_4)(.1)(ND)(20/40)(C)$
VARIANCES
MEANS
Figure 46. This graph relates to females with the following specifications: 
\((T_5)(.1)(ND)(20/40)(C)\)
Figure 47. This graph relates to females with the following specifications: 
(T₁)(.003)(ND)(20/40)(C)

Figure 48. This graph relates to females with the following specifications: 
(T₂)(.003)(ND)(20/40)(C)
Figure 49. This graph relates to females with the following specifications:
\[(T_1)(.5)(CD)(20/40)(R)\]
Figure 50. This graph relates to females with the following specifications:
\( \{T_3(0.5)CD(20/40)R\} \)

Figure 51. This graph relates to males with the following specifications:
\( \{T_3(0.5)CD(20/40)R\} \)
Figure 52. This graph relates to females with the following specifications:
$T_3(3)(CD)(20/40)(R)$

Figure 53. This graph relates to males with the following specifications:
$T_3(3)(CD)(20/40)(R)$
Figure 54. This graph relates to females with the following specifications:
(T₃)(0.1)(CD)(20/40)(R)
Figure 55. This graph relates to females with the following specifications: $(T_3)(.03)(CD)(20/40)(R)$
Figure 56. This graph relates to females with the following specifications:
\[(T_3)(0.007)(CD)(20/40)(R)\]

Figure 57. This graph relates to males with the following specifications:
\[(T_3)(0.007)(CD)(20/40)(R)\]
Figure 58. This graph relates to females with the following specifications: 
\((T_3)(.003)(CD)(20/40)(R)\)
Figure 59. This graph relates to females with the following specifications: 
($T_1$)(.5)(OD)(20/40)(R)
Figure 60. This graph relates to females with the following specifications: \((T_3)(.5)(OD)(20/40)(R)\)
Figure 61. This graph relates to females with the following specifications:
\((T_3)(.3)(OD)(20/40)(R)\)
Figure 62. This graph relates to females with the following specifications:
$(T_3)(.1)(OD)(20/40)(R)$
Figure 63. This graph relates to females with the following specifications: \( T_3 (.03)(OD)(20/40)(R) \)

Figure 64. This graph relates to males with the following specifications: \( T_3 (.03)(OD)(20/40)(R) \)
Figure 65. This graph relates to females with the following specifications:
\( T_1(0.5) \text{MD}(26/40) \text{R} \)
Figure 66. This graph relates to females with the following specifications: \( T_3 \times 5 \times (MD) \times (20/40) \times (R) \)
Figure 67. This graph relates to females with the following specifications:
\[ T_3(3)(MD)(20/40)(R) \]
Figure 68. This figure relates to females with the following specifications: $(T_3)(.03)(MD)(20/40)(R)$
Figure 69. This graph relates to females with the following specifications: (T₃)(.007)(MD)(20/40)(R)

Figure 70. This graph relates to females with the following specifications: (T₃)(.003)(MD)(20/40)(R)
VARIANCES
MEANS

0 10 20

0 10 20

0 5 10

GENERATIONS