CONSEQUENCES OF SELECTION
IN FINITE POPULATIONS
by

Peter Michael Burrows

Institute of Statistics
Mimeograph Series No. 758
Raleigh - July 1971
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>vii</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.1 Selection Models</td>
</tr>
<tr>
<td></td>
<td>1.2 Literature Review</td>
</tr>
<tr>
<td></td>
<td>1.3 Notation</td>
</tr>
<tr>
<td>2. SELECTION DIFFERENTIALS</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2.1 Selection Differentials Defined</td>
</tr>
<tr>
<td></td>
<td>2.2 Directional Selection</td>
</tr>
<tr>
<td></td>
<td>2.3 Approximation k for Directional Selection</td>
</tr>
<tr>
<td></td>
<td>2.4 Performance of Approximation k</td>
</tr>
<tr>
<td></td>
<td>2.5 An Expression for $v$ under Directional Selection</td>
</tr>
<tr>
<td></td>
<td>2.6 The Nature of $v$ under Directional Selection</td>
</tr>
<tr>
<td></td>
<td>2.7 Performance of Approximation $v$</td>
</tr>
<tr>
<td></td>
<td>2.8 Symmetry Considerations</td>
</tr>
<tr>
<td>3. LINEAR RESPONSES TO SELECTION</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>3.1 Bivariate Notations</td>
</tr>
<tr>
<td></td>
<td>3.2 The Nature of Response to Selection</td>
</tr>
<tr>
<td></td>
<td>3.3 Pre- and Post-test Projections of Response</td>
</tr>
<tr>
<td></td>
<td>3.4 Expectations of Individual and Average Responses</td>
</tr>
<tr>
<td></td>
<td>3.5 Variances of Individual and Average Responses</td>
</tr>
<tr>
<td></td>
<td>3.6 Covariances of Individual and Average Responses</td>
</tr>
<tr>
<td></td>
<td>3.7 Distributions of Responses</td>
</tr>
<tr>
<td></td>
<td>3.8 Approximation $\hat{\beta}(y_0,n)$ for $\beta(y_0,n)$ under Directional Selection</td>
</tr>
<tr>
<td>4. THE BIVARIATE NORMAL MODEL</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>4.1 Expectations and Variances of Response</td>
</tr>
<tr>
<td></td>
<td>4.2 Distributions of Response</td>
</tr>
<tr>
<td></td>
<td>4.3 Approximation $\hat{\beta}(y_0,n)$</td>
</tr>
<tr>
<td></td>
<td>4.4 Symmetry Considerations</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

5. THE SINGLE POLYGENE MODEL .................................. 69
   5.1 Reference Population $P_0$ .................................. 69
   5.2 Selective Value ............................................. 75
   5.3 Gene Frequency Response .................................... 77
   5.4 Variance Effective Number $n_V$ ............................ 80
   5.5 Trait Mean Responses ....................................... 81
   5.6 Consequences of No Selection ............................... 83
   5.7 Consequences of Directional Selection ..................... 84
   5.8 Consequences of Divergent Selection ...................... 86
   5.9 Additional Consequences for the Additive Model .......... 91
   5.10 Effects of Allelic Disequilibrium ......................... 92
   5.11 Consequences of Selection in a Dioecious System ........ 95
   5.12 Consequences of Disruptive Selection ................... 99

6. SINGLE POLYGENES AND RECURRENT SELECTION .................... 102
   6.1 Review ..................................................... 102
   6.2 Selection Limits in a Monoecious System ................. 105
   6.3 Selection Limits in a Dioecious System ................. 115
   6.4 Rate Effective Number $n_{net}$ under Recurrent Mass Selection 125
   6.5 Trait Mean Advance and Variation among Replicates .... 130
   6.6 Life of Polymorphism ...................................... 136

7. THE DOUBLE POLYGENE MODEL ................................... 148
   7.1 Reference Population $P_0$ .................................. 148
   7.2 Frequencies in Selected Zygotes .......................... 154
   7.3 Reference Population $P_r$ .................................. 158
   7.4 Trait Mean Response ....................................... 159
   7.5 Consequences of No Selection ................................ 160
   7.6 Consequences of Directional Selection .................... 161

8. DISCUSSION AND SUMMARY .................................... 165
   8.1 Selection Practices ......................................... 165
   8.2 Parameters of Selection Differentials ..................... 167
   8.3 Linearity of Response ...................................... 171
   8.4 Nature of Response ......................................... 172
   8.5 Pre- and Post-test Projections ............................. 173
   8.6 Prediction or Projection? .................................. 175
   8.7 Selection Designs and Strategies ........................... 177
   8.8 The Single Polygene Model .................................. 180
   8.9 Selection Limits in a Monoecious System .................. 182
   8.10 Selection Limits in a Dioecious System ................... 183
   8.11 Rate Effective Number $n_{net}$ ............................ 188
   8.12 Trait Mean Advance and Variation among Replicates .... 189
   8.13 Life of Polymorphism ..................................... 189
   8.14 The Double Polygene Model ................................ 190
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.</td>
<td>LIST OF REFERENCES</td>
<td>192</td>
</tr>
<tr>
<td>10.</td>
<td>APPENDICES</td>
<td>199</td>
</tr>
<tr>
<td>10.1</td>
<td>Variances and Covariances of Normal Order Statistics</td>
<td>199</td>
</tr>
<tr>
<td>10.2</td>
<td>Some Properties of $v$ under Directional Selection and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normally Distributed $X$</td>
<td>201</td>
</tr>
<tr>
<td>10.3</td>
<td>Generalization of Approximations $k$ and $\hat{\beta}(y_0,n)$ in</td>
<td></td>
</tr>
<tr>
<td></td>
<td>the Case of Directional Selection</td>
<td>205</td>
</tr>
<tr>
<td>10.4</td>
<td>Numerical Values of Parameters Used in the Diffusion Equation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Approximations for Selection Limits</td>
<td>208</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparison of $k_0$, $\hat{k}$ and $\hat{k}$ in the case of</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>directional selection and Normally distributed $X$</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>Comparison of $k_0$, $\hat{k}$ and $k$ in the case of</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>directional selection and Exponentially distributed $X$</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>Comparison of $v_0$, $\hat{v}$ and $v$ in the case of</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>directional selection and Normally distributed $X$</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>Comparison of $k_0$, $\hat{k}$ and $k$ in the case of</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>directional selection and Exponentially distributed $X$; $n$ bottom-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ranking test units selected</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Pre-test: post-test information ratio for</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>average response in the case of</td>
<td></td>
</tr>
<tr>
<td></td>
<td>directional selection and bivariate Normal distribution for $(X,Y)$</td>
<td></td>
</tr>
<tr>
<td>4.2</td>
<td>Pre-test projections of average response</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>in the case of directional selection and bivariate Normal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>distribution for $(X,Y)$</td>
<td></td>
</tr>
<tr>
<td>4.3</td>
<td>$\Pr[\hat{\delta}_n &gt; 0</td>
<td>\Delta_n = k]$ in the case of</td>
</tr>
<tr>
<td></td>
<td>directional selection and bivariate Normal distribution for $(X,Y)$</td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>Lower bounds to pre-test $\Pr[\hat{\delta}_n &gt; 0]$ in the case of</td>
<td>63</td>
</tr>
<tr>
<td></td>
<td>directional selection and bivariate Normal distribution for $(X,Y)$</td>
<td></td>
</tr>
<tr>
<td>4.5</td>
<td>Values of $\hat{\delta}(0,n)$ for $N = 20$ in the case of</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>directional selection and bivariate Normal distribution for $(X,Y)$</td>
<td></td>
</tr>
<tr>
<td>4.6</td>
<td>Values of $\hat{\alpha}(0,i)$ and $\hat{\bar{\alpha}}(0,i)$ for</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>$N = 20$ and bivariate Normal distribution for $(X,Y)$</td>
<td></td>
</tr>
</tbody>
</table>
### LIST OF TABLES (continued)

**Page**

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Values of pre-test $\Pr[(++)</td>
<td>r]$ in a test sample of eight zygotes</td>
</tr>
<tr>
<td>6.1</td>
<td>Comparison of selection limits $\hat{u}(0.5)$ and $v(2n;\infty)$ obtained by diffusion equation approximation and transition matrix methods for recurrent mass selection in a monoecious system</td>
<td>113</td>
</tr>
<tr>
<td>6.2</td>
<td>Comparison of selection limits $\hat{u}(0.5)$ and $v(2n_m,2n_f;\infty)$ obtained by diffusion equation approximation and transition matrix methods for recurrent mass selection in a dioecious system; $N_m = N_f = 8$, $n_m = n_f$</td>
<td>121</td>
</tr>
<tr>
<td>6.3</td>
<td>Comparison of selection limits $\hat{u}(0.5)$ and $v(2n_m,2n_f;\infty)$ obtained by diffusion equation approximation and transition matrix methods for recurrent mass selection in a dioecious system; $N_m = N_f = 8$, $n_m \neq n_f$</td>
<td>121</td>
</tr>
<tr>
<td>6.4</td>
<td>Asymptotic rate effective numbers for recurrent mass selection in a monoecious system; $N = 16$ and $32$</td>
<td>128</td>
</tr>
<tr>
<td>6.5</td>
<td>Asymptotic rate effective numbers for recurrent mass selection in a monoecious system; $8n = N$, $(a/\sigma_x) = 0.05$ and $0.10$</td>
<td>129</td>
</tr>
<tr>
<td>6.6</td>
<td>Asymptotic rate effective numbers for recurrent mass selection in a dioecious system; $N_m = N_f = 8$, $n_m = n_f$</td>
<td>130</td>
</tr>
<tr>
<td>6.7</td>
<td>Asymptotic rate effective numbers for recurrent mass selection in a dioecious system; $N_m = N_f = n_f = 8$</td>
<td>130</td>
</tr>
<tr>
<td>6.8</td>
<td>Expectation and standard deviation of trait mean ($x \times 100$) in a subdivided monoecious system under recurrent mass selection; $rN = 240$, $8n = N$, $(a/\sigma_x) = 0.1$, $d = 0$</td>
<td>132</td>
</tr>
<tr>
<td>6.9</td>
<td>Expectation and standard deviation of trait mean ($x \times 100$) in a subdivided monoecious system under recurrent mass selection; $rN = 240$, $8n = N$, $(a/\sigma_x) = 0.1$, $d = a$</td>
<td>132</td>
</tr>
<tr>
<td>6.10</td>
<td>Expectation and standard deviation of trait mean ($x \times 100$) in a monoecious system under recurrent mass selection; $N = 16$, $(a/\sigma_x) = 0.1$, $p_0 = 0.25$ and $0.50$, $d = 0$</td>
<td>134</td>
</tr>
<tr>
<td>6.11</td>
<td>Expectation and standard deviation of trait mean ($x \times 100$) in a dioecious system under recurrent mass selection; $N_m = N_f = 8$, $(a/\sigma_x) = 0.1$, $p_0 = 0.5$, $d = 0$</td>
<td>137</td>
</tr>
</tbody>
</table>
LIST OF TABLES (continued)

6.12 Median life of polymorphism and $t_{95}$ for recurrent mass selection in a monoecious system; $p_0 = 0.5$ ............... 142

6.13 Median life of polymorphism and $t_{95}$ for recurrent mass selection in a monoecious system; $p_0 = 0.25$ ............... 142

6.14 Median life of polymorphism and $t_{95}$ for recurrent mass selection in a monoecious system; $(a/\sigma_x) = 0.05$ and 0.10 .... 143

6.15 Median life of polymorphism and $t_{95}$ for recurrent mass selection in a dioecious system; $N_m = N_f = 8$, $n_m = n_f$ .... 144

6.16 Median life of polymorphism and $t_{95}$ for recurrent mass selection in a dioecious system; $N_m = N_f = 8$, $n_m \neq n_f$ .... 144

6.17 Values of $\sigma_t^2$ for symmetric disruptive selection and for neutral genes; $(a/\sigma_x) = 0.2$, $p_0 = 0.5$, $d = 0$ .... 146

6.18 Distribution function $Pr[T < t]$ for symmetric disruptive selection and for neutral genes; $(a/\sigma_x) = 0.2$, $p_0 = 0.5$, $d = 0$ .... 147

7.1 Trait $X$ means per genotype and indicator variables for the two loci model .... 149

8.1 Comparison of selection limits for monoecious and dioecious systems in cases where the diffusion equation approximation provides the same selection limit for both; $(a/\sigma_x) = 0.1$, $(k; n, N = 16)$ for monoecious, $(k = k, n_m = n_f = \frac{1}{2} n)$ for dioecious, $p_0 = 0.5$, $d = (2h-1)a$ .... 186

8.2 Comparison of monoecious and dioecious systems in cases where the diffusion equation approximation provides the same selection limit for both; $(a/\sigma_x) = 0.1$, $d = 0$, $(k; n, N = 16)$ for monoecious, $(k = k, n_m = n_f = \frac{1}{2} n)$ for dioecious, $p_0 = 0.5$ .... 187
LIST OF FIGURES

5.1 Ratio \( n_v/n \) as a function of \( p \) for four degrees of dominance ........................................ 82

6.1 Progression of rate effective number under recurrent mass selection in a monoecious system .............. 126

6.2 Progression of rate effective number under recurrent mass selection in a dioecious system ................. 127
1. INTRODUCTION

Selection is both a natural phenomenon and an ubiquitous feature of directed activity in human society. In the latter case selection and its consequences are often collectively formulated and discussed in language specific to context and discipline. A representative sample of operational synonyms includes "discrimination," "restriction," "screening," "streaming," "culling," "grading," "upgrading" and even "improvement." Additionally, the term selection may be preceded by one or more of an assortment of adjectives describing style of selection, particularly in the context of plant and animal breeding plans.

Whatever the terminology used, intentional selection and its consequences in relation to its objectives are amenable to quantitative analysis. By providing prior projections of consequences of selection such analysis is valuable in the planning and design of selection practices. After selection such analysis is useful in evaluating selection achievements relative to prior objectives and projections. It is with such analysis that this thesis is concerned. Specifically, attention is focused initially on a general model characterized by stochastic consequences of selection; later developments are restricted to some special genetic models.

1.1 Selection Models

In models of selection practices three primary elements are universally encountered: operational units, a selection basis (denoted here by $X$) and a measure of response to selection (denoted here by $Y$).

That there exists a population of operational units is fundamental; selection is a defined prescription for manipulation of these units.
In plant and animal breeding plans for example, operational units are typically individuals or families. Selection basis X is observable on all units and provides a means of assigning units to selected status according to the selection criterion defined as a function of X. Response Y is potentially observable on all units but is realized only on selected units. It cannot be over-emphasized that response Y is defined for a unit as a function of utilization of that unit if selected.

It is both axiomatic and intuitive that in all nontrivial cases basis X and response Y vary from unit to unit and that they are related in some manner. Without variation in X there is no opportunity for selective discrimination among units. Without covariation between X and Y selection basis X is not relevant to response Y. Without variation in Y there is no covariation between X and Y.

Further structure of selection models is provided by organization of populations of operational units. First there is the collection of units upon which selection is practiced. This will be referred to as the "test population" or "test material" in accordance with common usage of the term test in many applications, including progeny testing in animal breeding, test planting in plant breeding, aptitude tests for personnel selection in civil service and commercial employment, quality testing for batched improvement of cropped or manufactured produce, scholastic achievement tests for entrance to academic institutions and competence tests for entrance to professional associations. In all of these contexts the collection of tested units forms the test material and the test score per unit becomes selection basis X.
Next there is the collection of units selected from test material according to prescribed functions of basis X. Objectives and style of selection are embodied in these selection criteria.

Finally there is the response population defined in terms of utilization of selected units. Particularly in the context of plant and animal breeding, units of response populations may not be the selected units themselves but rather some reproduction of them or their properties. As stressed previously this distinction does not require special treatment because definition of response Y accommodates such variations in utilization of selected units.

The sources and sizes of these three populations of units determine the mathematical and statistical methods required for analysis of selection and its consequences. It is convenient to consider their sizes first. In some cases the test population and the population of selected units may be so large as to be effectively infinite and treated as such for analytical purposes. This is a commonly adopted feature for analysis of selection practices in sociological and educational contexts; all of the ten examples drawn from literature in a variety of fields by Owen (1959) fall into this category. In other cases the test population may again be considered infinite but the population of selected units is treated as being of finite size \( n \) with the implication of some sampling feature being involved. Alternatively, test material itself may be treated as being of finite size \( N \); then the number of selected units is necessarily finite (\( n \leq N \)) and sampling features are implicated for both test material and selected units.

Whenever sampling features are implicated in a selection model, response is stochastic and may be characterized by parameters of
reference distributions defined by appropriately chosen reference populations. These latter are conveniently introduced when considering sources of test material as is done next.

When \( N \) is considered infinite the test population is itself a reference distribution specifying various properties of the dispersion of \( X \) and \( Y \) over test units as well as the relationship between \( X \) and \( Y \). When \( N \) is considered finite, test material may be considered as a well-defined "fixed" population or as a "random sample" from some reference population defining a reference distribution of \( X \) and \( Y \) specified by known parameters.

Gathering together these notions and introducing formulations of selection criteria suggests four formal models of selection; \( \Pi_0 \) denotes the reference distribution for test material:

**Selection model I:** \( N \) infinite, \( n \) infinite. The population of selected units is defined by some style of truncation applied to the \( X \)-dimension of \( \Pi_0 \). Simple examples are directional selection (single truncation to an \( X \)-tail of \( \Pi_0 \)) and modal selection (double truncation to a central \( X \)-portion of \( \Pi_0 \)).

**Selection model II:** \( N \) infinite, \( n \) finite. Selected units are defined as a random sample, size \( n \), from that portion of \( \Pi_0 \) remaining after truncation in the \( X \)-dimension according to selection style.

**Selection model III:** \( N \) finite, \( n \leq N \). Test material is a random sample, size \( N \), from \( \Pi_0 \). Selected units are a subset of the test material, a subset that can always be defined by the ranks (on basis \( X \)) of test units it contains. Thus,
selection is a censorship, rather than a truncation, operation in the X-dimension.

**Selection model IV:** N finite, n ≤ N. Test material is a "fixed" finite population. Selected units are a subset of the test material.

It is useful at this point to recall the distinction between truncation and censorship (Kendall and Buckland, 1957) as it applies to these selection models. Truncation as a selection criterion is applied with reference to some predetermined fixed points in the X-dimension; censorship is applied with reference to rank positions in the X-dimension. Whereas censorship is entirely a sample concept, the sample being the test material, truncation may be applied to both the parent distribution in \( \Pi_0 \) and the sampled test material. Thus, a possible selection criterion not specifically mentioned in selection models III and IV would be truncation of sampled test material, in which case \( n \) itself would become a random variable and could even take the value zero if none of the N test units fall in the truncated range of the X-dimension. Application of such a selection criterion would produce a set of selected units the X-ranks of which would be known; consequences of selection could then be analyzed as for selection model III.

Selection model IV is included in the above listing for the sake of completeness and will not be considered analytically here. It has received considerable attention in a variety of forms. For example, test units may be a set of N populations with unknown means (Y) for which there exist estimators (X). The selection problem is choice of that population (tested unit) with the largest mean, or a subset of populations containing that with the largest mean with a specified
probability (Sobel and Huyett, 1957; Dunnett, 1960). That \( X \) is a measurement of \( Y \) with error is a common feature of selection model IV (Finney, 1956).

Selection model III is the model of major interest in this thesis; there are few applications for which \( N \) and \( n \) can be considered as anything but finite and the assumption of test material randomly sampled from some reference population is a satisfactory approximation in many applications.

Selection model II may have some applications, especially in certain screening procedures, but is introduced here to provide an intermediate step in the derivation of some theoretical properties of selection model III. For some properties of response to selection, selection model I will be shown to be a particular asymptotic form of model III; namely, as \( N \to \infty \) with \( n/N \) held constant. It should be noted that selection model I poses no statistical problems for it contains no stochastic elements; analysis requires only calculus of variation and covariation. In contrast, analysis of selection model II requires sampling distributions for samples from truncated distributions, and selection model III introduces the additional complexities of distributional properties of order statistics.

Thus far, little has been said concerning response populations about which it is more difficult to generalize. That response populations may, and usually will, be finite is clear. Thus, another stochastic element of selection models is introduced, this time in the utilization of selected test units, and a reference population \( \Pi_r \) for sampling of response units is required. This additional stochastic feature is of little direct consequence since projections of response
now refer to the parameterization of $\Pi_r$. Indeed the analysis of selection and its consequences, under selection model III, may be summarized as

"the problem of relating parameters of $\Pi_r$ to those of $\Pi_0$ with a proper accounting of stochastic elements involved in the transition $\Pi_0 \rightarrow \Pi_r"$

which is taken to be the definitive theme of this thesis.

1.2 Literature Review

Literature is drawn from a variety of fields in statistics and population genetics. A comprehensive review of each such field is obviously not required. Collecting together all the literature used into one section here, or perhaps a separate chapter, results in a rambling unconnected miscellany unless organized in the same manner as the presented thesis material is developed, thus creating an unnecessary and repetitive duplication of citations and descriptions of their content. Instead, appropriate literature is introduced and reviewed where relevant and required, particularly at the beginning of Chapters 4 and 6.

1.3 Notation

Notations of all variables, parameters and functions are defined as they are introduced. Expectation, variance and covariance operators are denoted by $E$, $V$ and $C$ respectively with their arguments in {}; for example $E\{X\}$ and $C\{X,Y\}$. Conditioning is indicated by the common practice of separation by the "$|"$ symbol; for example $E\{Y|X\}$ denotes the conditional expectation of $Y$ given $X$. In some places these moment operators are used iteratively in an obvious notation; thus $E\{V\{Y|X\}\}$
denotes expectation, with respect to the distribution of $X$, of the conditional variance of $Y$ given $X$. 
2. SELECTION DIFFERENTIALS

In the context of Model III described in Chapter 1, let $X_1, X_2, \ldots, X_N$ denote the observed X-values of the $N$ test units and let $X_1 \leq X_2 \leq \ldots \leq X_N$ denote the corresponding order statistics arranged in increasing order of magnitude; $X_{(r)}$ is the $r$th smallest observed value in the test material and test unit corresponding to $X_{(r)}$ is allocated rank $r$. Any form of selecting $n$ units from the available test material may be specified by defining the set of ranks corresponding to selected units. In general such a set containing $n$ specified ranks will be written

$$R_n = \{r_1, r_2, \ldots, r_n \}, \ r_i \in \{1, 2, \ldots, N\}, \ n \leq N.$$

2.1 Selection Differentials Defined

Let $\overline{X}_n$ denote the average X-value of selected units defined by set $R_n$. Algebraically

$$\overline{X}_n = \frac{1}{n} \sum_{R_n} X_{(r)}, \quad (2.1.1)$$

where $\sum_{R_n}$ denotes summation over ranks $r \in R_n$.

Let $\mu_x$ and $\sigma^2_x$ denote the mean and variance respectively of the distribution of $X$ in $\Pi_0$, the reference population for sampling of test material. Let $D_n$, a random variable, denote the selection differential defined as the difference $(\overline{X}_n - \mu_x)$ and let $k$, a parameter, denote the standardized selection differential defined as the expectation of $D_n$ expressed as a multiple of $\sigma_x$:

$$E(D_n) = E(\overline{X}_n - \mu_x) = k\sigma_x, \quad (2.1.2)$$
If \( \overline{X}_N \) denotes the average \( X \)-value of all \( N \) test units, an alternative definition of selection differential would be \( D_n^- = (\overline{X}_n - \overline{X}_N) \), sometimes referred to as "reach" (e.g., Lush, 1945). The expectations of \( D_n \) and \( D^- \) are the same because \( E\{\overline{X}_n\} = \mu_x \), but their variances differ (see Section 10.2).

Let \( \nu \), a parameter, denote the standardized variance of the selection differential defined as

\[
\nu \{ D_n \} = \nu \{ \overline{X}_n \} = \frac{\nu \sigma^2}{n} .
\]  

(2.1.3)

When the selection style specifies that two groups of test units are selected from the same test material, defined by rank sets \( R_n \) and \( R_m \), let \( \omega \), a parameter, denote the standardized covariance between selection differentials \( D_n \) and \( D_m \) defined as

\[
C \{ D_n, D_m \} = C \{ \overline{X}_n, \overline{X}_m \} = \frac{\omega \sigma^2}{\nu nm} .
\]  

(2.1.4)

Whatever the distribution of \( X \) in \( \Pi_0 \), it is standardized to zero mean and unit variance by the transformation

\[
x_i = \frac{(X_i - \mu_x)}{\sigma_x} , \quad x(r) = \frac{(X(r) - \mu_x)}{\sigma_x}
\]  

(2.1.5)

\[
\Delta_n = \frac{D_n}{\sigma_x} = \frac{1}{n} \sum_{R_n} x(r) .
\]

Then, denoting by \( \mu_{x|r|N} \) the expectation of the \( r \)th order statistic in a random sample of size \( N \) from the standardized distribution, combining (2.1.2) and (2.1.5) provides
\[ k = \mathbb{E}(\Delta_n) = \frac{1}{n} \sum_{r \in R_n} \mu_{r|N}. \]  

(2.1.6)

Similarly, denoting by \( \sigma_{r,s|N} \) the covariance of the \( r \)th and \( s \)th order statistics in a random sample of size \( N \) from the standardized distribution, combining (2.1.5) with (2.1.3) and (2.1.4) in turn provides

\[ v = nV(\Delta_n) = \frac{1}{n} \sum_{r \in R_n} \sum_{s \in R_n} \sigma_{r,s|N}. \]  

(2.1.7)

and

\[ w = \sqrt{nm} C(\Delta_n, \Delta_m) = \frac{1}{\sqrt{nm}} \sum_{r \in R_n} \sum_{s \in R_m} \sigma_{r,s|N}. \]  

(2.1.8)

As usual, the variance parameter \( v \) is a special case \((R_n = R_m)\) of the covariance parameter \( w \). Because of the adopted notation \( \sigma_{i,i|N} \) for the variance of the \( i \)th order statistic, the formulation of \( w \) is general in the sense that it is not required that \( R_n \cap R_m \) is null, and as the notation indicates, sets \( R_n \) and \( R_m \) may have different cardinal numbers \( n \) and \( m \).

### 2.2 Directional Selection

A common style of selection, having as its objective some directional change, involves selection of the group of top-ranking, or bottom-ranking, test units. Because of the rank symmetry of this defined operation, there is no loss of generality in assuming that the \( n \) top-ranking units are selected from test material. Then selection is specified by

\[ R_n \equiv \{r+1, r+2, \ldots, N\}, \quad r = N-n \]

with standardized selection differential and standardized variance of selection differential given by (2.1.6) and (2.1.7) as
\[ k = \frac{1}{n} \sum_{i=1}^{n} \mu_{r+i|N} \] (2.2.1)

and

\[ \nu = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{r+i, r+j|N}. \] (2.2.2)

Unless the \( X \)-distribution in \( H_0 \) is very simple (e.g., the Uniform density), suitable algebraic expressions for \( \mu_{i|N} \) and \( \sigma_{i,j|N} \) as functions of \( i, j \) and \( N \) are not available. Even for the well-studied Normal distribution it is necessary to resort to numerical values for low moments of order statistics. Thus, Harter (1960) presents a tabulation of \( \mu_{i|N} \) given to five decimal places for \( N = 2(1)100 \) and through other values of \( N \) up to 400; Becker (1968) presents a tabulation of \( k \) itself but for a much less comprehensive range in \( N \). Availability of tables for \( \sigma_{i,j|N} \) in the Normal case is severely restricted to low values of \( N \). Ruben (1954) provides a tabulation of low moments (including \( \sigma_{1,1|N} \)) of just the extreme order statistics, \( x_{(1)} \) or \( x_{(N)} \), in samples of size \( N = 2(1)50 \).

Sarhan and Greenberg (1956) provide a tabulation of all \( \sigma_{i,j|N} \) in samples of size \( N = 2(1)20 \). For other distributions such tabulations do not exist. In any case, for purposes of theoretical manipulation of \( n \) and \( N \) in evaluation of selection response, simple algebraic expressions for \( k \) and \( \nu \) are required in preference to numerical values. An alternative route to \( k \) and \( \nu \) would be the distribution of \( \Delta_n \) itself (or its moment generating function). This is soon found to be an even more intractable approach than that via the moments \( \mu_{i|N} \) and \( \sigma_{i,j|N} \).

Comments in the last paragraph also apply to selection styles other than directional selection, the only exception being the trivial case.
where $R_n$ contains $n$ ranks sampled at random from integers \{1,2,...,N\}, for which $k = 0$ and $v = 1$ irrespective of the $X$-distribution in $\Pi_0$.

The remainder of this chapter is concerned with the derivation and performance of simple expressions serving as approximations for $k$ and $v$ under directional selection. The strategy adopted is to first find the conditional expectation and variance of $\Delta_n$, given $x(r)$, and then to remove this conditioning. Consideration of this procedure was prompted by the results obtained by Kojima (1961) for $k$ alone.

The density function for the $X$-distribution in $\Pi_0$ is denoted by $f(x)$ throughout. Continuity, differentiability and other analytic properties of $f(x)$ are assumed without later specific mention; in particular it is assumed that the integral

$$m_p(x) = \int_0^\infty t^p f(t) dt, \quad p = 0,1,2$$

(2.2.3)

is finite and that its derivatives with respect to $x$ exist. The corresponding distribution function is denoted by $F(x) = 1 - G(x)$.

Without loss of generality $f(x)$ is taken to be standardized to $\mu_x = 0$ and $\sigma^2_x = 1$ so that $X_n = x_n$, $D_n = \Delta_n$ and $k$ and $v$ are found directly.

2.3 Approximation $\hat{k}$ for Directional Selection

Kojima (1961) derived the following standardized selection differential as appropriate for selection of $n$ top-ranking test units in the case of a special $X$-distribution that was a compound of three Normal distributions with the same variance but slightly differing means:

$$k^* = \frac{N}{n} \left[ \phi(0) + \frac{1}{2!} u_2 \phi(2) + \frac{1}{3!} u_3 \phi(3) + \ldots \right] ,$$

(2.3.1)

where $u_i$ is the $i$th central moment of the distribution of the $(N-n)$th order statistic in a sample of size $(N-1)$ from the unit Normal distribution.
and \( \phi^{(1)} \) is the \( i \)th derivative of the unit Normal density function evaluated at \( \mu_{N-n|N-1} \).

On summing this series, Pike (1969) showed that \( k^* \) was identically \( k \) given at (2.2.1) with \( \mu_{r+1|N} \) defined for the unit Normal density. Thus, a suitable curtailment of the series in (2.3.1) would provide an approximation for \( k \) in the case of Normally distributed \( X \). Kojima himself curtailed the series after the first two terms. Such an approximation suffers from two disadvantages. First, in the case of no selection (\( n = N \)) the value \( k^* = 0 \) is not obtained. Second, a tabulation of \( \mu_{N-n|N-1} \) and \( \mu_2 \) is required although the latter may be replaced by its asymptotic value. Avoidance of these disadvantages is a guiding feature in the approximation process described below.

When test material may be considered an infinite population and the \( n \) selected units treated as a random sample from the upper tail of \( f(x) \) truncated at \( x_0 \) (that is Model II described in Chapter 1), the mean of the truncated distribution, and hence of the corresponding \( \Delta_n \), is given by (2.2.3) as

\[
k = \frac{m_1(x_0)}{G(x_0)}.
\]  

(2.3.2)

In the case where \( f(x) \) is the unit Normal density, \( m_1(x_0) = f(x_0) \) and \( k_0 \) becomes the well-known and widely used expression for the standardized selection differential:

\[
k_0 = \frac{f(x_0)}{G(x_0)}.
\]

This particular \( k_0 \) is also of historical importance being the reciprocal of Mills' Ratio evaluated at \( x_0 \) (see e.g., Kendall and Stuart, 1958). It is also well-known that in the Normal case, choosing \( x_0 \) to satisfy
$G(x_0) = n/N$ produces a value $k_0$ greater than $k$ although $k + k_0$ as $N \to \infty$ with $n/N$ held constant (see Box 11 of Lerner, 1958, and Table 2.1 of this chapter).

At the outset therefore, an approximation $\hat{k}$ of the following form is sought:

$$k \approx \hat{k} = k_0 - C(n,N), \quad k_0 \text{ defined at (2.3.2)}$$  \hspace{1cm} (2.3.3)

where $C(n,N)$ denotes some function of $n$ and $N$ having the property that $C(n,N) \to 0$ as $N \to \infty$ with $n/N$ held constant. Such a form would indicate the correction (most probably a reduction) applied to $k_0$ due to finiteness of sampled test material and thus contrast selection Models I and III.

Following Kojima (1961), $x_{(r)}$ is considered a fixed point initially with $x_{(r+1)}, x_{(r+2)}, \ldots, x_{(N)}$ treated conditionally as a random sample of size $n$ from the upper tail of $f(x)$ truncated at $x_{(r)}$. Using (2.3.2) the conditional expectation of $\Delta_n$, given $x_{(r)}$ is

$$E(\Delta_n | x_{(r)}) = \frac{m_1(x_{(r)})}{G(x_{(r)})}, \quad r = N-n$$  \hspace{1cm} (2.3.4)

and it remains to remove the conditioning by taking expectation over the distribution of $x_{(r)}$ which has density function

$$z(x_{(r)}; N, r) = \frac{N!}{(r-1)! (N-r)!} [F(x_{(r)})]^{r-1} [G(x_{(r)})]^{N-r} f(x_{(r)}).$$  \hspace{1cm} (2.3.5)

Thus, $k = E(\Delta_n) = E[E(\Delta_n | x_{(r)})]$ is given by

$$k = \int_{-\infty}^{\infty} \frac{m_1(t)}{G(t)} z(t; N, r) \, dt = \frac{N}{(N-r)} \int_{-\infty}^{\infty} m_1(t) z(t; N-1, r) \, dt.$$  \hspace{1cm} (2.3.6)
Next, expanding $m_1(t)$ in a Taylor series about the point $t = a$ provides

$$m_1(t) = m_1(a) - (t-a) af(a) - \frac{1}{2} (t-a)^2 \{af'(a) + f(a)\} - \ldots.$$  

Then taking the integration at (2.3.6) term by term, $k$ can be written as

$$k = \frac{m_1(a)}{n/N}$$

$$- \frac{N}{n} \{ (\mu_{r|N-1} - a) af(a) + \frac{1}{2} \{ (\mu_{r|N-1} - a)^2 + \sigma_{r,r|N-1} \} \}$$

$$\cdot \{af'(a) + f(a)\}$$

$$+ \frac{1}{6} \{ (\mu_{r|N-1} - a)^3 + 3(\mu_{r|N-1} - a) \sigma_{r,r|N-1} + \kappa_r^{(3)} \}$$

$$\cdot \{af''(a) + 2f'(a)\} + \ldots \} \right\} , \quad (2.3.7)$$

where $\kappa_r^{(3)}$ is the third central moment of the $r$th order statistic in a sample of size $(N-1)$ from distribution with density function $f(x)$.  

At this stage Kojima (1961) chooses $a = \mu_{r|N-1}$ with the consequence that all terms $(\mu_{r|N-1} - a)^p$, $p \geq 1$ are zero, but leaves terms involving $m_1(\mu_{r|N-1})$, $f(\mu_{r|N-1})$, $f'(\mu_{r|N-1})$, etc., together with the central moments $\sigma_{r,r|N-1}$, $k_r^{(3)}$, $\kappa_r^{(3)}$, etc.  When $f(x)$ is the unit Normal density $k$ at (2.3.7) readily reduces to $k^*$ at (2.3.1); this development can be compared with that of Pike (1969) who begins with $k^*$ and shows that it is equal to $k$.  

As an alternative consider the choice $a = x_0$ satisfying $G(x_0) = n/N$.  This choice has several advantages over that just considered.  In the first place $x_0$ is more readily available, both analytically and numerically, than is $\mu_{r|N-1}$, for general $f(x)$.  Second, the first term of $k$ at (2.3.7) becomes
\[
\frac{m_1(x_0)}{G(x_0)} = k_0 \text{ defined at (2.3.2)}
\]

and so the required form at (2.3.3) for approximation \( \hat{k} \) is yielded.

Additionally, it is shown next that \( \mu_{r|N-1 - x_0}, \sigma_{r|N-1}, \kappa_{r|N-1}^{(3)} \) and higher order cumulants may be replaced by simple approximations.

David and Johnson (1954) provide the following general results for cumulants of order statistics. If \( y_{(i)} \) is the \( i \)th order statistic in a random sample of size \( M \) from a population with distribution function \( F(y) \), then subject to differentiability conditions on \( F(y) \),

\[
E\{y_{(i)}\} = Y_1 + \frac{p_i q_i}{2(M+2)} Y_1^2 + (M + 2)^{-2} \{ \text{further terms} \},
\]

\[
V\{y_{(i)}\} = \frac{p_i q_i}{(M+2)} (Y_1)^2 + (M + 2)^{-2} \{ \text{further terms} \},
\]

\[
\kappa_{(3)}\{y_{(i)}\} = 0 + (M + 2)^{-2} \{ \text{further terms} \},
\]

\[
\kappa_{(4)}\{y_{(i)}\} = 0 + (M + 2)^{-3} \{ \text{further terms} \},
\]

etc.,

where \( Y_1 \) satisfies \( F(Y_1) = 1/(M+1) = p_1 = 1 - q_1 \), and

\[
Y_1' = \left[ \frac{dy}{dF} \right]_{y=Y_1}, \quad Y_1'' = \left[ \frac{d^2y}{dF^2} \right]_{y=Y_1}.
\]

Saw (1960) derives bounds for the remainder after the first two terms of \( E\{y_{(i)}\} \) above. The inverse series expansions, that are curtailed to provide these cumulant expressions, may be slow to converge and even not converge if \( i/M \) is too close to 0 or 1.

In the current context required cumulants are those of the \( r \)th order statistic \( (r = N-n) \) in a sample of size \( M = N-1 \). Thus, \( Y_r \) is given
by

\[ p_r = 1 - q_r = \frac{r}{N} = 1 - \frac{n}{N} = F(Y_r) \]

and, serendipitously, the required \( Y_r \) happens to be \( x_0 \) the choice adopted above for Taylor series expansion of \( m_1(t) \). Elaboration of the required \( Y_i \) and \( Y_i' \) yields:

\[ Y'_r = \left[ \frac{dx}{dF} \right]_{x=x_0} = \frac{1}{f(x_0)} \]

and

\[ Y''_r = \left[ \frac{d^2 x}{dF^2} \right]_{x=x_0} = \frac{f'(x_0)}{\{f(x_0)\}^3}. \]

Then compiling the cumulants provides

\[ (\mu_r|N-1 - x_0) \approx -\frac{r}{N} \frac{(1 - \frac{x}{N})}{2(N+1)} \frac{f'(x_0)}{\{f(x_0)\}^3}, \quad (2.3.8) \]

\[ \sigma_{r,r|N-1} \approx \frac{r}{N} \frac{(1 - \frac{x}{N})}{(N+1)} \frac{1}{\{f(x_0)\}^2}, \quad (2.3.9) \]

and

\[ \kappa^{(p)}_{r|N-1} \approx 0, \; p \geq 3. \quad (2.3.10) \]

Since \( r = (N-n) \) and \( (1 - \frac{x}{N}) = G(x_0) \), it is convenient to combine (2.3.8) and (2.3.9) into

\[ (\mu_{r|N-1 - x_0}) \approx -\frac{1}{2} \frac{f'(x_0)}{f(x_0)} \sigma_{r,r|N-1} \]

\[ \sigma_{r,r|N-1} \approx \frac{2G(x_0)}{f(x_0)} \frac{C(n,N)}{f(x_0)^3}, \quad (2.3.11) \]
where \( C(n,N) \), a parameter of the \( X \)-distribution, is given by

\[
C(n,N) = \frac{(N-n)}{2n(N+1)} \frac{G(x_0)}{f(x_0)} = \frac{1}{2(N+1)} \frac{F(x_0)}{f(x_0)}. \tag{2.3.12}
\]

Finally, choosing \( a = x_0 \) in (2.3.7), and substituting these cumulant approximations,

\[
k \approx k_0 - \frac{1}{G(x_0)} \left[ (\mu_{r|N-1} - x_0) f(x_0) + \frac{1}{2} \frac{f'(x_0)}{f(x_0)} \right] \\
\text{and}
\]

\[
k \approx k_0 - \frac{x_0 f(x_0)}{G(x_0)} \left[ (\mu_{r|N-1} - x_0) + \frac{1}{2} \frac{f'(x_0)}{f(x_0)} \right] \\
- \frac{1}{2} \frac{f(x_0)}{G(x_0)}
\]

and using (2.3.11) yields the proposed approximation \( \hat{k} \):

\[
k = \hat{k} + (N+1)^{-2} \text{ (further terms)} \tag{2.3.13}
\]

\[
\hat{k} = k_0 - C(n,N),
\]

where \( k_0 \) was defined at (2.3.2) and \( C(n,N) \) at (2.3.12).

2.4 Performance of Approximation \( \hat{k} \)

Recalling the definitions of \( k_0 \) and \( C(n,N) \),

\[
k_0 = \frac{m_1(x_0)}{G(x_0)}, \quad G(x_0) = \frac{n}{N}, \quad C(n,N) = \frac{1}{2(N+1)} \frac{F(x_0)}{f(x_0)} \geq 0,
\]

the performance of \( \hat{k} \) as an approximation for \( k \) can be evaluated for two special circumstances with general \( f(x) \). First, in the case of no selection, \( n = N, G(x_0) = 1, F(x_0) = 0, x_0 = -\infty \) and \( m_1(-\infty) = \mu_x = 0 \)
because \( f(x) \) is assumed standardized to zero mean. Thus, \( k_0 = 0 \), \( C(N,N) = 0 \) and \( \hat{k} = 0 \) as required. Second, in the case where \( N \to \infty \) with \( G(x_0) = n/N \) held constant, \( C(n,N) \to 0 \) from above and \( \hat{k} \to k_0 \) the mean of the upper tail of the original \( X \)-distribution truncated at \( x_0 \). Thus, as far as \( k \) is concerned, selection Model I is an asymptotic form of Model III.

For less trivial cases, the performance of \( \hat{k} \) must be judged in relation to exact (algebraic or numerical) values of \( k \) for distributions where the latter is available. Three such distributions are considered next.

The Uniform density \( f(x) = 1/\theta \) for \( |x| \leq \frac{1}{2} \theta \) has zero mean and unit variance when \( \theta^2 = 12 \); expectations of order statistics in a random sample of size \( N \) are well-known:

\[
\mu_j|N = \frac{j \theta}{(N+1)} - \frac{\theta}{2} .
\]

Then by definition (2.2.1), \( k \) is given exactly as

\[
k = \frac{1}{n} \sum_{i=1}^{n} \mu_{r+1}|N = \frac{\theta (N-n)}{2 (N+1)} .
\]  

(2.4.1)

For approximation \( \hat{k} \), applying the required definitions yields

\[
m_1(x) = \frac{\theta}{2} \left( \frac{1}{2} - \frac{x}{\theta} \right) \left( \frac{1}{2} + \frac{x}{\theta} \right) , \quad \frac{n}{N} = G(x_0) = \left( \frac{1}{2} - \frac{x_0}{\theta} \right) , \quad F(x_0) = \left( \frac{1}{2} + \frac{x_0}{\theta} \right)
\]

and

\[
m_1(x_0) = \frac{\theta}{2} \left( \frac{n}{N} \right) \left( 1 - \frac{n}{N} \right) , \quad C(n,N) = \frac{\theta}{2N} \left( \frac{N-n}{N+1} \right) .
\]

Thus,

\[
\hat{k} = k_0 - C(n,N) = \frac{\theta}{2} \left( \frac{N-n}{N+1} \right) ,
\]

and so in this case \( \hat{k} \) yields the exact result \( k \) at (2.4.1).
Initially this is a somewhat surprising result; it certainly was not an intentional feature of the procedure by which \( k \) was constructed. Closer analysis shows that those terms of the Taylor series expansion that are neglected after integration in (2.3.7) are in fact zero. Moreover, David and Johnson (1954) indicate that the approximations for \( \mu_{r|N-1} \) and \( \sigma_{r,r|N-1} \) employed at (2.3.8) and (2.3.9) were organized by them to provide exact results in the case of the uniform distribution.

In the case of the unit Normal density, \( m_1(x) = f(x) \) and so

\[
k_0 = \frac{f(x_0)}{G(x_0)} , \quad C(n|N) = \frac{(N-n)}{2n(N+1)} \frac{1}{k_0},
\]

where \( x_0 \) can be obtained from standard tabulations of the unit Normal distribution function \( F(x_0) = (N-n)/N \). Then for the Normal case, approximation \( \hat{k} \) takes the simple form

\[
\hat{k} = k_0 \left[ 1 - \frac{(N-n)}{2n(N+1)} \frac{1}{k_0} \right], \tag{2.4.2}
\]

with performance shown in Table 2.1 for \( n/N = 0.5, 0.2 \) and \( 0.1 \) at each of \( N = 10(10) 50 \) and \( 100 \). Values of \( k_0 \) and \( \hat{k} \) in this table were computed with the aid of Tables 4 and 5 of Hartley and Pearson (1958); values of \( k \) were computed from the tabulation of \( \mu_{1|N} \) presented by Harter (1960). It is clear that \( \hat{k} \) is a satisfactory approximation for almost all applications in the case of Normally distributed \( X \) and directional selection.

The exponential density \( f(x) = \exp\{-\lambda x\} \), \( x \geq 0 \), has zero mean and variance unity. The expectation of the \( j \)th order statistic in a random sample of size \( N \) is given by Lehman (1959):

\[
\mu_{j|N} = -1 + \frac{\sum_{\ell=1}^{j} (N - \ell + 1)^{-1}}{N}, \tag{2.4.3}
\]
Table 2.1. Comparison of $k_0$, $\hat{k}$ and $k$ in the case of directional selection and Normally distributed $X$

<table>
<thead>
<tr>
<th>$n/N$</th>
<th>N</th>
<th>$k_0$</th>
<th>$\hat{k}$</th>
<th>$k$</th>
<th>100 $k_0/k$</th>
<th>100 $\hat{k}/k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>0.7979</td>
<td>0.7409</td>
<td>0.7389</td>
<td>107.98</td>
<td>100.27</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>0.7979</td>
<td>0.7680</td>
<td>0.7675</td>
<td>103.96</td>
<td>100.07</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>0.7979</td>
<td>0.7777</td>
<td>0.7774</td>
<td>102.64</td>
<td>100.04</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td>0.7979</td>
<td>0.7826</td>
<td>0.7825</td>
<td>101.97</td>
<td>100.01</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>0.7979</td>
<td>0.7865</td>
<td>0.7855</td>
<td>101.58</td>
<td>100.01</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>0.7979</td>
<td>0.7917</td>
<td>0.7917</td>
<td>100.78</td>
<td>100.00</td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1.3998</td>
<td>1.2699</td>
<td>1.2701</td>
<td>110.21</td>
<td>99.98</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>1.3998</td>
<td>1.3318</td>
<td>1.3318</td>
<td>105.11</td>
<td>100.00</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>1.3998</td>
<td>1.3537</td>
<td>1.3537</td>
<td>103.41</td>
<td>100.00</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td>1.3998</td>
<td>1.3649</td>
<td>1.3649</td>
<td>102.56</td>
<td>100.00</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>1.3998</td>
<td>1.3718</td>
<td>1.3718</td>
<td>102.04</td>
<td>100.00</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>1.3998</td>
<td>1.3857</td>
<td>1.3857</td>
<td>101.02</td>
<td>100.00</td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>1.7550</td>
<td>1.5219</td>
<td>1.5388</td>
<td>114.05</td>
<td>98.90</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>1.7550</td>
<td>1.6329</td>
<td>1.6375</td>
<td>107.18</td>
<td>99.72</td>
</tr>
<tr>
<td>30</td>
<td></td>
<td>1.7550</td>
<td>1.6723</td>
<td>1.6744</td>
<td>104.81</td>
<td>99.87</td>
</tr>
<tr>
<td>40</td>
<td></td>
<td>1.7550</td>
<td>1.6925</td>
<td>1.6937</td>
<td>103.62</td>
<td>99.93</td>
</tr>
<tr>
<td>50</td>
<td></td>
<td>1.7550</td>
<td>1.7047</td>
<td>1.7055</td>
<td>102.90</td>
<td>99.95</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td>1.7550</td>
<td>1.7296</td>
<td>1.7298</td>
<td>101.46</td>
<td>99.99</td>
</tr>
</tbody>
</table>

so that $k$ is given exactly by

$$k = \frac{1}{n} \sum_{i=1}^{N-n} \mu_{r+i/N} = \frac{N-n}{N} \sum_{i=1}^{N} (n+i)^{-1}, \ n < N.$$  

(2.4.4)

Applying the definitions to this density where $f(x) = G(x)$,

$$m_1(x) = (1 + x) e^{-(1+x)}, \quad \frac{n}{N} = G(x_0) = e^{-(x_0+1)}, \quad (x_0 + 1) = \ln \cdot \frac{N}{n}$$

and

$$k_0 = \frac{m_1(x_0)}{G(x_0)} = (1 + x_0) = \ln \cdot \frac{N}{n}, \quad C(n,N) = \frac{(N-n)}{2n(N+1)}$$

and so

$$\hat{k} = \ln \cdot \frac{N}{n} - \frac{(N-n)}{2n(N+1)}$$  

(2.4.5)
with performance shown in Table 2.2. Values of k in this table were computed from (2.4.4) with the aid of Table 51 of Hartley and Pearson (1958). Again there is excellent agreement between \( \hat{k} \) and k.

<table>
<thead>
<tr>
<th>( \frac{n}{N} )</th>
<th>N</th>
<th>( k_0 )</th>
<th>( \hat{k} )</th>
<th>k</th>
<th>100 ( k_0/k )</th>
<th>100 ( \hat{k}/k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>10</td>
<td>0.6932</td>
<td>0.6477</td>
<td>0.6456</td>
<td>107.37</td>
<td>100.33</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.6932</td>
<td>0.6693</td>
<td>0.6688</td>
<td>103.65</td>
<td>100.07</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.6932</td>
<td>0.6770</td>
<td>0.6768</td>
<td>102.42</td>
<td>100.03</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.6932</td>
<td>0.6810</td>
<td>0.6808</td>
<td>101.82</td>
<td>100.02</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.6932</td>
<td>0.6834</td>
<td>0.6833</td>
<td>101.45</td>
<td>100.01</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.6932</td>
<td>0.6882</td>
<td>0.6882</td>
<td>100.73</td>
<td>100.00</td>
</tr>
</tbody>
</table>

| 0.2             | 10  | 1.6094  | 1.4276 | 1.4290 | 112.62         | 99.90          |
|                 | 20  | 1.6094  | 1.5142 | 1.5144 | 106.27         | 99.99          |
|                 | 30  | 1.6094  | 1.5449 | 1.5450 | 104.17         | 100.00         |
|                 | 40  | 1.6094  | 1.5607 | 1.5607 | 103.12         | 100.00         |
|                 | 50  | 1.6094  | 1.5702 | 1.5702 | 102.50         | 100.00         |
|                 | 100 | 1.6094  | 1.5896 | 1.5896 | 101.25         | 100.00         |

| 0.1             | 10  | 2.3026  | 1.8935 | 1.9290 | 119.37         | 98.16          |
|                 | 20  | 2.3026  | 2.0883 | 2.0977 | 109.77         | 99.55          |
|                 | 30  | 2.3026  | 2.1574 | 2.1617 | 106.52         | 99.80          |
|                 | 40  | 2.3026  | 2.1928 | 2.1952 | 104.89         | 99.89          |
|                 | 50  | 2.3026  | 2.2144 | 2.2159 | 103.91         | 99.93          |
|                 | 100 | 2.3026  | 2.2580 | 2.2584 | 101.96         | 99.98          |

### 2.5 An Expression for \( \nu \) under Directional Selection

An approximation for the standardized variance of the selection differential is provided by a procedure similar to that used in deriving \( \hat{k} \) as an approximation for k. When test material may be considered an infinite population and the n selected units treated as a random sample from the upper tail of \( f(x) \) truncated at \( x_0 \) (that is Model II described in Chapter 1), the variance of the truncated distribution is given by (2.2.3) as
\[ v_0 = \frac{m_2(x_0)}{G(x_0)} - k_0^2, \quad (2.5.1) \]

where \( k_0 \) was defined at (2.3.2) as \( m_1(x_0)/G(x_0) \). Then the variance of the corresponding \( \Delta_n \) is

\[ \Psi(\Delta_n) = \frac{v_0}{n}. \quad (2.5.2) \]

In the case where \( f(x) \) is the unit Normal density, \( m_2(x_0) = x_0 m_1(x_0) + G(x_0) \), and \( v_0 \) reduces to

\[ v_0 = 1 - k_0(k_0 - x_0) \quad (2.5.3) \]

as given by Finney (1956). Thus, in the Normal case parameters \( k_0 \) and \( v_0 \) are simply related: \( (1 - v_0) \) is the derivative of \( k_0 \) with respect to \( x_0 \).

Under selection Model III, and general standardized \( f(x) \), let \( x(r) \) be considered as a fixed point initially, with \( x_{r+1}, x_{r+2}, \ldots, x_{(N)} \) treated conditionally as a random sample of size \( n \) from the upper tail of \( f(x) \) truncated at \( x(r) \). Using (2.5.1) and (2.5.2) conditional variance of \( \Delta_n \), given \( x(r) \), is

\[ \Psi(\Delta_n | x(r)) = \frac{1}{n} \left[ \frac{m_2(x(r))}{G(x(r))} - \left( \frac{m_1(x(r))}{G(x(r))} \right)^2 \right]. \quad (2.5.4) \]

The expectation, over the distribution of \( x(r) \), of the first term on the right-hand side of (2.5.4) is

\[ E[\frac{m_2(x(r))}{G(x(r))}] = \int_{-\infty}^{\infty} \frac{m_2(t)}{G(t)} z(t; N, r) dt \]

\[ = \frac{N}{(N-r)} \int_{-\infty}^{\infty} m_2(t) z(t; N-1, r) dt. \]

Expanding \( m_2(t) \) in a Taylor series about the point \( t = a \) provides
\[ m_2(t) = m_2(a) - (t-a) a^2 f(a) - \frac{1}{2} (t-a)^2 a \{ a f'(a) + 2 f(a) \} - \ldots , \]

and taking the integration term by term,

\[ E \left[ \frac{m_2(x(r))}{G(x_r)} \right] = \frac{m_2(a)}{n/N} \]

\[ - \frac{N}{n} [(\mu_{r|N-1} - a) a^2 f(a) + \frac{1}{2} \{ (\mu_{r|N-1} - a)^2 + \sigma_{r,r|N-1} \} \]

\[ \cdot \{ a^2 f'(a) + 2 f(a) \} + \ldots ] . \]

Choosing \( a = x_0 \) satisfying \( G(x_0) = n/N \), and using the moment expressions at (2.3.11),

\[ E \left[ \frac{m_2(x(r))}{G(x_r)} \right] = \frac{m_2(x_0)}{G(x_0)} - x_0^2 \frac{f(x_0)}{G(x_0)} [(\mu_{r|N-1} - x_0) + \frac{1}{2} \sigma_{r,r|N-1} \frac{f'(x_0)}{f(x_0)}] \]

\[ - \frac{x_0 f(x_0)}{G(x_0)} \sigma_{r,r|N-1} \]

\[ = (v_0 + k_0^2) - 2x_0 \sigma_{r,r|N-1} \]

(2.5.5)

where \( v_0 \) was defined at (2.5.1).

The expectation, over the distribution of \( x(r) \), of the second term on the right-hand side of (2.5.4) is

\[ E \left[ \frac{m_1(x(r))^2}{G(x(r))} \right] = \int_{-\infty}^{\infty} \frac{m_1(t)}{G(t)} z(t;N,r) dt \]

\[ = \frac{N}{N-r} \int_{-\infty}^{\infty} \frac{m_1(t)^2}{G(t)} z(t;N-1,r) dt . \]

Expanding \( m_1(t)^2/G(t) \) in a Taylor series about the point \( t = a \) provides

\[ \frac{m_1(a)^2}{G(a)} + \{(t-a) + \frac{1}{2} (t-a)^2 \frac{f'(a)}{f(a)} \} \left[ \frac{m_1(a)}{G(a)} - a^2 - a^2 \right] f(a) \]

\[ + (t-a)^2 \frac{f(a)}{G(a)} \left[ \frac{m_1(a)}{G(a)} - a^2 \right] f(a) - m_1(a) \] + \ldots ,
and taking the integration term by term,

\[
E[\{\frac{1}{G(x_r)}\}^2] = \frac{m_1(a)}{G(a)} \cdot \frac{m_1(a)}{G(a)} - a^2
\]

\[
+ \frac{f(a)}{G(a)} [(\mu_{r|N-1} - a) + \frac{1}{2} (\mu_{r|N-1} - a)^2 + \sigma_{r,r|N-1} f'(a)]
\]

\[
\cdot \left[\frac{m_1(a)}{G(a)} - a^2 - a^2\right]
\]

\[
+ \{\mu_{r|N-1} - a\}^2 + \sigma_{r,r|N-1}\] \left[\frac{m_1(a)}{G(a)} - a\right]^2 \frac{f(a)}{G(a)} - \frac{m_1(a)}{G(a)} \frac{f(a)}{G(a)} + \ldots
\]

Choosing \( a = x_0 \) satisfying \( G(x_0) = n/N \), and using the moment expressions at (2.3.11),

\[
E[\{\frac{1}{G(x_r)}\}^2] \cong \frac{m_1(x_0)}{G(x_0)} \cdot \frac{m_1(x_0)}{G(x_0)} - x_0^2
\]

\[
+ \left[\frac{m_1(x_0)}{G(x_0)} - x_0\right]^2 \frac{f(x_0)}{G(x_0)} - \frac{m_1(x_0)}{G(x_0)} \frac{f(x_0)}{G(x_0)} + \sigma_{r,r|N-1}
\]

\[
= k_0^2 + 2(k_0 - x_0) \left(1 - (k_0 - x_0) \frac{f(x_0)}{G(x_0)} - k_0\right) G(n,N)
\]

(2.5.6)

where \( k_0 \) was defined at (2.3.2).

Taking expectation, over the distribution of \( x_r \), through (2.5.4) and substituting from (2.5.5) and (2.5.6) for the terms on the right-hand side,

\[
E[\Delta_n | x_r] \cong \frac{1}{n} \left[v_0 + 2(k_0 - x_0) \left(1 - (k_0 - x_0) \frac{f(x_0)}{G(x_0)} - k_0\right) G(n,N)\right]
\]

(2.5.7)

In the case where \( f(x) \) is the unit Normal density function,

\( k_0 = f(x_0)/G(x_0) \) and \( v_0 = 1 - k_0(k_0 - x_0) \), (2.5.7) reduces to
\[ E[V(\Delta_n | x_{(r)})] \approx \frac{v_0}{n} \{1 + 2(k_0 - x_0) \text{C}(n,N)\} . \quad (2.5.8) \]

Consider next the variance, over the distribution of \( x_{(r)} \), of the conditional expectation of \( \Delta_n \) given \( x_{(r)} \):

\[
V[E(\Delta_n | x_{(r)})] = \frac{\frac{m_1(x_{(r)})}{G(x_{(r)})}}{G(x_{(r)})} = E[\left(\frac{\frac{m_1(x_{(r)})}{G(x_{(r)})}}{G(x_{(r)})}\right)^2] - E^2[\frac{m_1(x_{(r)})}{G(x_{(r)})}] ,
\]

the first term of which has already been evaluated at (2.5.6) and the second term is simply \( k^2 \) (see 2.3.5) which can be written, using (2.3.13), as

\[
k^2 = (\hat{k})^2
= k^2 - 2k_0 \text{C}(n,N) .
\]

Thus,

\[
V[E(\Delta_n | x_{(r)})] = 2(k_0 - x_0)^2 \frac{f(x_0)}{G(x_0)} \text{C}(n,N) . \quad (2.5.9)
\]

The unconditional variance of \( \Delta_n \) can now be compiled from (2.5.7) and (2.5.9):

\[
V(\Delta_n) = E[V(\Delta_n | x_{(r)})] + V[E(\Delta_n | x_{(r)})]
\]

\[
= \frac{1}{n} \left[ v_0 + 2(k_0 - x_0) \text{C}(n,N) \right] + (1 - \frac{1}{n}) 2(k_0 - x_0)^2 \frac{f(x_0)}{G(x_0)} \text{C}(n,N) . \quad (2.5.10)
\]

\[
= \frac{1}{n} \left[ v_0 + 2(k_0 - x_0)(1 + (n-1)(k_0 - x_0) \frac{f(x_0)}{G(x_0)}) \text{C}(n,N) \right] .
\]
Finally, \( v \) is defined at (2.2.2) as \( nV(\Delta_n) \) so that an approximation for \( v \) is provided by (2.5.10):

\[
v = v_1 + n(N+1)^{-2} \{ \text{further terms} \}
\]

(2.5.11)

\[
v_1 = v_0 + 2(k_0 - x_0)(1 + (n-1)(k_0 - x_0) \frac{f(x_0)}{g(x_0)} C(n,N).}
\]

2.6 The Nature of \( v \) under Directional Selection

As discussed in Section 2.4, as far as \( k \) is concerned, selection Model I is an asymptotic form \( (N \to \infty \) with \( n/N \) constant) of Model III. This is not the case for \( v \). As \( N \to \infty \) with \( n/N = G(x_0) \) constant, \( V(\Delta_n) \) under Model III is \( v/n \to v_0/n \) the \( V(\Delta_n) \) under Model II, simply because both \( v/n \) and \( v_0/n \to \text{zero} \); but \( v \neq v_0 \) for in (2.5.11), letting \( N \to \infty \) with \( n/N = G(x_0) \) constant,

\[
v \to \hat{v} = v_0 + \frac{(N-n)}{N} (k_0 - x_0)^2
\]

(2.6.1)

because \( C(n,N) \to \text{zero} \), being of order \( (N+1)^{-1} \), but

\[
2(n-1) f(x_0) C(n,N)/G(x_0) + F(x_0).
\]

Thus, unlike \( k_0 \) for \( k \), if \( n \neq N \), \( v_0 \) cannot be a satisfactory approximation for \( v \), not even asymptotically.

The reason of course is that \( v_0 \) fails to account for the \( n(n-1) \) covariances \( \sigma_{i,j|N} \) between selected test units. In the course of preparing a tabulation of \( v \) for the case of Normally distributed \( X \) (see next section) it was found that when \( N = 20 \) and \( n = 10 \), 83 percent of \( v \) is contributed by the 90 covariances included, and when \( N = 50 \), \( n = 25 \), 92 percent of \( v \) is contributed by the 600 covariances included. Thus, the term

\[
\hat{v} - v_0 = (k_0 - x_0)^2 F(x_0) \geq 0
\]

(2.6.2)
in (2.6.1) is the asymptotically required correction that accounts for these covariances in the case of general \( f(x) \).

Another empirical finding (see next section) is that for given \( n/N \), \( v \) is a relatively stable function of \( N \), ranging close to its asymptotic \( (N \to \infty) \) value \( \hat{v} \). The simplicity of \( \hat{v} \) leads to the proposed approximation for \( v \) as \( \hat{v} \) defined at (2.6.1) with performance discussed in the next section.

2.7 Performance of Approximation \( \hat{v} \)

In the case of general \( f(x) \) and no selection, \( n = N, G(x_0) = 1, F(x_0) = 0, x_0 = -\infty, m_1(-\infty) = \mu_x = 0 \) and \( m_2(-\infty) = \sigma_x^2 = 1 \) because \( f(x) \) is assumed standardized to zero mean and unit variance. Thus, \( k_0 = 0, v_0 = 1 \) and \( \hat{v} = 1 \) as is required in this case.

In the case of the unit Normal density, \( v_0 = 1 - k_0(k_0 - x_0) \) as at (2.5.3) and \( \hat{v} \) reduces to

\[
\hat{v} = 1 - (k_0 - x_0) \{f(x_0) + x_0 F(x_0)\}
\]

with performance shown in Table 2.3. The term \( \{f(x_0) + x_0 F(x_0)\} \) in \( \hat{v} \) is the unit Normal linear loss integral \( l(u) = f(u) - uG(u) \) defined by Raiffa and Schlaifer (1961), evaluated at \( u = -x_0 \). Values of \( v_0 \) and \( \hat{v} \) in this table were computed with the aid of Tables 4 and 5 of Hartley and Pearson (1958). For \( N = 10 \) and 20, values of \( v \) were computed from the tabulation of \( \sigma_{i,j|N} \) given by Sarhan and Greenberg (1956); for \( N = 30, 40, 50 \) and 100, values of \( v \) were calculated by first preparing a tabulation of \( \sigma_{i,j|N} \) using the method described in Section 10.1; as discussed there they are of doubtful accuracy beyond the third decimal place. Although not generally as good as \( \hat{k} \) for \( k \), \( v \) is a satisfactory approximation to \( \hat{v} \) for most purposes.
Table 2.3. Comparison of $v_0$, $\hat{v}$ and $v$ in the case of directional selection and Normally distributed $X$

<table>
<thead>
<tr>
<th>$n/N$</th>
<th>$v_0$</th>
<th>$\hat{v}$</th>
<th>$N$</th>
<th>$v$</th>
<th>$100 \hat{v}/v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.3634</td>
<td>0.6817</td>
<td>10</td>
<td>0.6740</td>
<td>101.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>0.6771</td>
<td>100.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>0.6791</td>
<td>100.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td>0.6800</td>
<td>100.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>0.6806</td>
<td>100.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>0.6815</td>
<td>100.03</td>
</tr>
<tr>
<td>0.2</td>
<td>0.2186</td>
<td>0.4679</td>
<td>10</td>
<td>0.4507</td>
<td>103.82</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>0.4570</td>
<td>102.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>0.4613</td>
<td>101.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td>0.4634</td>
<td>100.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>0.4647</td>
<td>100.69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>0.4674</td>
<td>100.11</td>
</tr>
<tr>
<td>0.1</td>
<td>0.1692</td>
<td>0.3709</td>
<td>10</td>
<td>0.3443</td>
<td>107.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td>0.3521</td>
<td>105.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td>0.3587</td>
<td>103.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td>0.3624</td>
<td>102.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>0.3648</td>
<td>101.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>0.3698</td>
<td>100.30</td>
</tr>
</tbody>
</table>

Continuing with the case of $f(x)$ as the unit Normal density, examination of the numerical values of $v$ in Table 2.3 (and others not shown) together with the previously mentioned special case $v = 1$ when $n = N$, leads to the assertion that for directional selection

$$0 < v < 1, \text{ all } n \text{ satisfying } 0 < n < N,$$

the upper bound to $v$ being reached only when $n = N$.

That $v < 0$ is trivial since variances of nondegenerate random variables are positive and $\Delta_n$ is nondegenerate for $n > 0$. That $v = 1$ when $n = N$ has already been shown. It remains, therefore, to show that

$$v < 1, \text{ all } n < N.$$

(2.7.2)

This is done in Section 10.2; in fact a more restrictive condition than
(2.7.1) is established at (10.2.9); namely,
\[
\frac{n}{N} < v < 1, \ 0 < n < N. \tag{2.7.3}
\]

That \( \hat{v} \) satisfies
\[
\frac{n}{N} < \hat{v} < \frac{n}{N} + (1 - \frac{n}{N})(1 - \frac{2}{\pi}), \ 0 < n < N, \tag{2.7.4}
\]
and, therefore, satisfies (2.7.3), is shown as follows. Consider the quantity
\[
\frac{(1 - \hat{v})}{F(x_0)} = (k_0 - x_0) \left\{ \frac{f(-x_0)}{G(-x_0)} - (-x_0) \right\}, \ 0 < n = NG(x_0) < N.
\]
This is a symmetric function of \( x_0 \) with a minimum of \( 2/\pi \) at \( x_0 = 0 \) and maxima of unity as \( x_0 \to \pm \infty \). Thus,
\[
\frac{2}{\pi} \left(1 - \frac{n}{N}\right) < (1 - \hat{v}) < \left(1 - \frac{n}{N}\right), \ 0 < n < N
\]
as required for (2.7.4).

2.8 Symmetry Considerations

Thus far in the treatment of directional selection it has been assumed that \( n \) top-ranking units are selected from test material. If, instead, the \( n \) bottom-ranking test units are selected, results derived above may be applied with appropriate amendments based on symmetry considerations.

If the distribution of \( X \) in \( \Pi_0 \) has a symmetric standardized density, \( f(x) = f(-x) \), approximations \( \hat{k} \) and \( \hat{v} \) at (2.3.13) and (2.6.1) respectively apply directly since only the sign of \( k \) changes and \( v \) remains unchanged. If the distribution of \( X \) has standardized density \( w(x) \) that is not symmetric, \( w(x) \neq w(-x) \), and the \( n \) bottom-ranking test units are selected, definitions of \( \hat{k} \) and \( \hat{v} \) may be applied by the simple device of first
writing \( f(x) = w(-x) \) and then proceeding as before; when this is done it is necessary to change the sign of the derived \( \hat{k} \).

An example of this latter procedure is provided by the exponential density considered in Section 2.4. Suppose the distribution of \( X \) in \( \Pi_0 \) has density function \( w(x) = \exp(-1 + x) \), \( x \geq -1 \) with zero mean and variance unity. When \( n \) bottom-ranking units are selected from \( N \) test units randomly sampled from this \( X \) distribution, the required \( \hat{k} \) is obtained as follows. Define \( f(x) = w(-x) = \exp(x-1) \), \( x \leq 1 \). Then applying the definition of Sections 2.2 and 2.3 to this \( f(x) \),

\[
m_1(x) = (1-x)e^{x-1}, \quad F(x) = f(x), \quad G(x_0) = 1 - e^{x_0-1} = \frac{n}{N}
\]

and so

\[
k_0 = \left(\frac{N}{n} - 1\right) \ln\left(\frac{N}{N-n}\right), \quad C(n,N) = \frac{1}{2(N+1)}
\]

yielding the applicable \( \hat{k} \) as

\[
\hat{k} = -\{k_0 - C(n,N)\} = -\left(\frac{N-n}{n}\right) \ln\left(\frac{N}{N-n}\right) + \frac{1}{2(N+1)}.
\]

Using the exact moment expression at (2.4.3) for \( \mu_j|N \), \( k \) is given explicitly by

\[
k = \frac{1}{n} \sum_{j=1}^{n} \mu_j|N = -1 + \frac{1}{n} \sum_{j=1}^{n} j(N - n + j)^{-1}.
\]

Finally, the performance of approximation \( \hat{k} \) in this particular case is shown in Table 2.4 for the same \( n,N \) combinations as in Table 2.2. Comparison of these two tables illustrates the difference between selecting top-ranking or bottom-ranking test units in a case where \( f(x) \) is not a symmetric density function.

While the procedure just described treats the case of \( n \) bottom-ranking test units directly, an indirect method is also available. The
<table>
<thead>
<tr>
<th>( n/N )</th>
<th>N</th>
<th>(-k_0)</th>
<th>( \hat{k})</th>
<th>k</th>
<th>100 ( k_0/k )</th>
<th>100 ( \hat{k}/k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.6932</td>
<td>-0.6477</td>
<td>-0.6456</td>
<td>107.37</td>
<td>100.32</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>-0.6932</td>
<td>-0.6693</td>
<td>-0.6688</td>
<td>103.64</td>
<td>100.09</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>-0.6932</td>
<td>-0.6770</td>
<td>-0.6768</td>
<td>102.42</td>
<td>100.04</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>-0.6932</td>
<td>-0.6810</td>
<td>-0.6808</td>
<td>101.82</td>
<td>100.02</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>-0.6932</td>
<td>-0.6834</td>
<td>-0.6833</td>
<td>101.45</td>
<td>100.01</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-0.6932</td>
<td>-0.6882</td>
<td>-0.6882</td>
<td>100.75</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.8926</td>
<td>-0.8471</td>
<td>-0.8444</td>
<td>105.71</td>
<td>100.32</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>-0.8926</td>
<td>-0.8688</td>
<td>-0.8680</td>
<td>102.83</td>
<td>100.08</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>-0.8926</td>
<td>-0.8764</td>
<td>-0.8761</td>
<td>101.88</td>
<td>100.04</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>-0.8926</td>
<td>-0.8804</td>
<td>-0.8802</td>
<td>101.41</td>
<td>100.02</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>-0.8926</td>
<td>-0.8828</td>
<td>-0.8827</td>
<td>101.12</td>
<td>100.01</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-0.8926</td>
<td>-0.8876</td>
<td>-0.8876</td>
<td>100.56</td>
<td>100.00</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-0.9483</td>
<td>-0.9028</td>
<td>-0.9000</td>
<td>105.37</td>
<td>100.31</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>-0.9483</td>
<td>-0.9244</td>
<td>-0.9237</td>
<td>102.66</td>
<td>100.08</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>-0.9483</td>
<td>-0.9321</td>
<td>-0.9318</td>
<td>101.77</td>
<td>100.04</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>-0.9483</td>
<td>-0.9361</td>
<td>-0.9359</td>
<td>101.32</td>
<td>100.02</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>-0.9483</td>
<td>-0.9385</td>
<td>-0.9383</td>
<td>101.07</td>
<td>100.01</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-0.9483</td>
<td>-0.9433</td>
<td>-0.9433</td>
<td>100.53</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

n bottom-ranking test units may be regarded as residual from the selection of \((N-n)\) top-ranking test units. Consider then the sum of standardized \(X\)-values over all \(N\) test units:

\[
\sum_{i=1}^{N} x(i) = \sum_{i=1}^{N} x(i) + \sum_{i=n+1}^{N} x(i) = n \Delta_n + (N-n) \Delta_{N-n}.
\]

But \(E\{\sum_{i=1}^{N} x(i)\} = 0\), being the expectations of the whole sample sum, so that

\[
E\{\Delta_n\} = -\frac{(N-n)}{n} E\{\Delta_{N-n}\} \tag{2.8.2}
\]

which relates the required standardized selection differential for \(n\)
bottom-ranking units, $E\{\Delta \} \mid_n$, to that for (N-n) top-ranking units, $E\{\Delta \} \mid_{N-n}$, for which approximation $\hat{k}$ is formulated directly. In the case of Exponentially distributed X, (2.4.5) provides the following expression for $E\{\Delta \} \mid_{N-n}$ on substituting (N-n) for n:

$$E\{\Delta \} \mid_{N-n} \approx \ln \left( \frac{N}{N-n} \right) - \frac{n}{2(N-n)(N+1)}.$$ 

Multiplying by $-(N-n)/n$, as indicated at (2.8.2), yields

$$E\{\Delta \} \mid_n \approx -\frac{(N-n)}{n} \ln \left( \frac{N}{N-n} \right) + \frac{1}{2(N+1)}$$

which is exactly the result obtained at (2.8.1) for n bottom-ranking test units.

Property (2.8.2) applies for general $f(x)$ and in the case of symmetric $f(x)$ implies another symmetry property, this time symmetry with respect to selection of n and (N-n) top-ranking units. As noted above when $f(x)$ is symmetric, only the sign of $k$ changes when directional selection changes from n top-ranking to n bottom-ranking units. Thus, $E\{\Delta \} \mid_n$ in (2.8.2) referring to n bottom-ranking units is $-E\{\Delta \} \mid_n$ referring to n top-ranking units, and (2.8.2) provides

$$k \mid_n = \frac{(N-n)}{n} k \mid_{N-n},$$

(2.8.3)

where $k \mid_m$ refers to selection of m top-ranking units. In the case of symmetric $f(x)$, therefore, function $m_k \mid_{(m)}$ is symmetric about $m = (1/2)N$ when N even a result discussed by Hill (1969a). Approximation $\hat{k}$ appropriately satisfies identity (2.8.3) because for symmetric $f(x)$, $G(x_0) = F(-x_0)$, $f(x_0) = f(-x_0)$, $m \mid_{(x_0)} = m \mid_{(-x_0)}$ and in (2.3.13),
\[
\frac{n}{(N-n)} \hat{k}_n = \frac{n}{(N-n)} \left\{ k_0 - C(n,N) \right\} \\
= \frac{n}{(N-n)} \left\{ \frac{m_1(x_0)}{G(x_0)} - \frac{1}{2(N+1)} \frac{F(x_0)}{f(x_0)} \right\}, \ G(x_0) = \frac{n}{N} \\
= \frac{m_1(-x_0)}{(N-n) N} - \frac{1}{2(N+1)} \frac{F(-x_0)}{f(-x_0)} \\
= \frac{m_1(-x_0)}{G(-x_0)} - \frac{1}{2(N+1)} \frac{F(-x_0)}{f(-x_0)}, \ G(-x_0) = \frac{(N-n)}{N} \\
= \hat{k}_{(N-n)}
\]

as required.

Writing \(v_n\) for \(v\), the symmetry property analogous to (2.8.3) in the case of Normal samples is

\[
n\{1 - v_n\} = (N-n)\{1 - v_{(N-n)}\} \quad (2.8.4)
\]

established at (10.2.6) in Section 10.2 where it is also shown that \(\hat{v}\) satisfies (2.8.4):

\[
n\{1 - \hat{v}_n\} = (N-n)\{1 - \hat{v}_{(N-n)}\}. \quad (2.8.5)
\]
3. LINEAR RESPONSES TO SELECTION

Analysis of response to selection requires some specification of the relationship between selection basis X and response variable Y defined as a function of post-selection utilization of selected test units. A complete specification of the bivariate distribution of X and Y in the reference population $\Pi_0$ is not a prerequisite for characterization of some features of selection response, such as low moments, particularly when linearity of response is assumed. Generally however, the more questions that are asked about selection response, the more definitive must be the specification of the bivariate distribution of X and Y.

3.1 Bivariate Notations

As in Chapter 2, the marginal distribution of X in $\Pi_0$ has mean $\mu_x$ and variance $\sigma_x^2$; the standardized marginal density is $f(x), x = (X-\mu_x)/\sigma_x$, with distribution function $F(x) = 1 - G(x)$. Similarly, the marginal distribution of Y in $\Pi_0$ has mean $\mu_y$ and variance $\sigma_y^2$, with standardized density $h(y), y = (Y-\mu_y)/\sigma_y$, and distribution function $H(y)$.

A common assumption underlying projections of the consequences of selection is linearity of response, often formulated by representation as a linear regression model:

$$E(Y|X) = \mu + \rho(X-\mu_x)(\frac{\sigma_y}{\sigma_x}); E(y|X) = \rho x$$  \hspace{1cm} (3.1.1)

from which predictions of mean response may be readily derived (see e.g., Robertson, 1955).

Projections of variance of response require further specification of the relationship between X and Y in $\Pi_0$. The possibilities that conditional variance of Y, given X, is a constant function of X or a
linear function of $X$, are accommodated in the following assumption:

$$
V(y|x) = \left[ b_0 + b_1 \left( \frac{x - \mu_x}{\sigma_x} \right) \right] \sigma_y^2, \quad V(y|x) = b_0 + b_1 x.
$$

(3.1.2)

The standardized bivariate density in $\Pi_0$ is denoted by $\ell(x,y;\rho)$ with associated conditional densities $\ell_x(x|y;\rho)$ and $\ell_y(y|x;\rho)$ satisfying

$$
f(x) \cdot \ell_y(y|x;\rho) = \ell(x,y;\rho) = h(y) \cdot \ell_x(x|y;\rho).
$$

(3.1.3)

The corresponding bivariate distribution function and conditional distribution functions are formulated as

$$
L(x_0,y_0;\rho) = \int_0^\infty \int_0^\infty \ell(x,y;\rho) \, dy \, dx,
$$

(3.1.4)

$$
L_x(x_0|y_0;\rho) = \int_0^\infty \ell_x(x|y;\rho) \, dx,
$$

and

$$
L_y(y_0|x_0;\rho) = \int_{y_0}^\infty \ell_y(y|x;\rho) \, dy.
$$

Continuity, integrability and other properties of these density and distribution functions are assumed as they are required.

As described in Chapter 2, test material is assumed to be a random sample, size $N$, from $\Pi_0$ and selection of $n$ test units is defined by a set $R_n$ of $n$ ranks referring to ranking on selection basis $X$. Formally, and in standardized variables, the test material may be written

$$
(x^{(1)}_1, y^{(1)}_1), \ (x^{(2)}_2, y^{(2)}_2), \ \ldots, \ (x^{(N)}_N, y^{(N)}_N),
$$

where $x^{(i)}_i$ is the $i$th order statistic, arranged in increasing order of magnitude as before, and $y^{(i)}_1$ is the concomitant standardized $Y$ value for
test unit with rank i. Selected test units are then defined by the set
\[(x_{(r)}, y_r); r \in \mathbb{N}_n\].

3.2 The Nature of Response to Selection

Implicit in the definition of response variable Y, as a function of post-selection utilization of selected test units, is the notion that all relevant consequences of selection are properties of the set
\[S_n(y) = \{y_r; r \in \mathbb{N}_n\}. \quad (3.2.1)\]

Ideally therefore, a full specification of the joint distribution of the n random variables \(y_r\) in \(S_n(y)\) represents a complete characterization of consequences of selection and emphasizes the stochastic nature of those consequences; this is "response" to selection in the broadest possible sense. In contrast, the historically more usual treatment of selection response is in the narrow sense of average response \(\delta_n\):

\[\delta_n = \frac{1}{n} \sum_{r \in \mathbb{N}_n} y_r, \quad \frac{1}{n} \sum_{r \in \mathbb{N}_n} y_r = \mu_y + \delta_n \sigma_y \quad (3.2.2)\]

particularly in the case where \(n\) defines directional selection as in Section 2.2. The expectation of \(\delta_n\) is referred to as "expected response."

Some properties of the distribution of \(\delta_n\), that are of interest in characterizing consequences of selection, are derived in this chapter, as are some properties of marginal distributions of individual responses \(y_r\) in \(S_n(y)\).

3.3 Pre- and Post-test Projections of Response

Let \(S_n(x)\) denote the set \(\{x_{(r)}; r \in \mathbb{N}_n\}.\) Once the test material has been sampled from \(\Pi_0\) the realized \(S_n(x)\) is available for any choice of rank set \(\mathbb{N}_n;\) for example, the realized selection differential \(\Delta_n \sigma_x\) is
available. Projections of consequences of selection may use this information. Analytically, probability statements or low moments of the response distributions are expressed conditional on the realization of \( S_n(x) \) and will be referred to as "post-test" projections. In contrast, "pre-test" projections are made prior to sampling of test material and, therefore, in the absence of perfect information on \( S_n(x) \).

The information content of pre-test projection relative to that of post-test projection will be measured by the corresponding statistical efficiency defined as

\[
\text{Eff}\{y_r; N\} = \frac{\mathcal{V}\{y_r \mid x(r)\}}{\mathcal{V}\{y_r\}}
\]  

(3.3.1)

for individual response \( y_r, r \in R_n \), and as

\[
\text{Eff}\{\delta_n; N\} = \frac{\mathcal{V}\{\delta_n \mid S_n(x)\}}{\mathcal{V}\{\delta_n\}}
\]

(3.3.2)

for average response \( \delta_n \), the numerator in each case being the post-test variance and the denominator in each case being the pre-test variance.

3.4 Expectations of Individual and Average Responses

Post-test projections of expected individual and average responses are provided immediately by the linearity assumption at (3.1.1):

\[
E\{y_r \mid x(r)\} = \rho \cdot x(r)
\]

(3.4.1)

and

\[
E\{\delta_n \mid S_n(x)\} = \frac{1}{n} \sum_{R_n} E\{y_r \mid x(r)\} = \rho \cdot \frac{1}{n} \sum_{R_n} x(r)
\]

or, in a notation indicating the sufficiency of \( \Delta_n = \frac{1}{n} \sum_{R_n} x(r) \),

\[
E\{\delta_n \sigma_y \mid \Delta_n\} = \Delta_n \rho \cdot \sigma_y
\]

(3.4.2)
Pre-test projections of expected individual and average responses are obtained by removing the conditioning in (3.4.1) and (3.4.2):

$$E(y_r) = \rho \mu_r | N$$

and

$$E(\delta_n \sigma_y) = k \rho \sigma_y, \quad (3.4.3)$$

where $k = E(\Delta_n)$ is the standardized selection differential corresponding to $R_n$ as defined at (2.1.6). In the context of directional selection, expected responses at (3.4.2) and (3.4.3) are of a familiar form (see e.g., Lush, 1945) involving the product of factors selection pressure ($k$), precision of selection ($\rho$) and vulnerable variance ($\sigma_y^2$).

### 3.5 Variances of Individual and Average Responses

Post-test projections of variances of individual and average responses are provided by assumption (3.1.2):

$$V(y_r | x(r)) = b_0 + b_1 x(r) \quad (3.5.1)$$

and

$$V(\delta_n | s_n(x)) = \frac{1}{n^2} \sum_{R_n} V(y_r | x(r))$$

because for $r \neq s$ random variables $y_r$ and $y_s$ are conditionally independent, given $x(r)$ and $x(s)$, since test units are independently sampled from $\Pi_0$; hence

$$C(y_r, y_s | x(r), x(s)) = 0, \quad r \neq s. \quad (3.5.2)$$

Thus,

$$V(\delta_n | \Delta_n) = \frac{1}{n} (b_0 + b_1 \Delta_n) \quad (3.5.3)$$
Pre-test projections of variances of individual and average responses are obtained by removing the conditioning in (3.5.1) and (3.5.3). For individual response, using (3.5.1) and (3.4.1),

\[ V(y_r) = E[V(y_r | x_r^{(r)})] + V[E(y_r | x_r^{(r)})] \]
\[ = E[b_0 + b_1 x_r^{(r)}] + V[\rho x_r^{(r)}] \]
\[ = b_0 + b_1 \mu_r | N + \rho^2 \sigma_r, r | N. \]

For average response, using (3.5.3) and (3.4.2),

\[ V(\delta_n) = E[V(\delta_n | \Delta_n)] + V[E(\delta_n | \Delta_n)] \]
\[ = E[\frac{1}{n} (b_0 + b_1 \Delta_n)] + V[\rho \Delta_n] \]
\[ = \frac{1}{n} (b_0 + b_1 k + \rho^2 v), \]

where \( v = n V(\Delta_n) \) is the standardized variance of the selection differential corresponding to \( R_n \) as defined at (2.1.7).

Pre-test:post-test information ratios defined at (3.3.1) and (3.3.2) may now be compiled as

\[ \text{Eff}(y_r; N) = \frac{b_0 + b_1 x_r^{(r)}}{b_0 + b_1 \mu_r | N + \rho^2 \sigma_r, r | N} \]

with pre-test expectation equal to

\[ 1 - \frac{\rho^2 \sigma_r, r | N}{b_0 + b_1 \mu_r | N + \rho^2 \sigma_r, r | N} \]

and

\[ \text{Eff}(\delta_n; N) = \frac{b_0 + b_1 \Delta_n}{b_0 + b_1 k + \rho^2 v} \]
with pre-test expectation equal to

\[ 1 - \frac{\rho^2 \nu}{(b_0 + b_1 k + \rho^2 \nu)} \]

3.6 Covariances of Individual and Average Responses

Post-test covariance of individual response \( y_r \) and \( y_s \), \( r \neq s \), is zero as discussed at (3.5.2). Removing the conditioning yields the pre-test covariance as

\[
C(y_r, y_s) = E[C(y_r, y_s | x(r), x(s))] + C[E(y_r | x(r)), E(y_s | x(s))]
\]

(3.6.1)

\[ = \text{zero} + C[\rho x(r), \rho x(s)] \]

\[ = \rho^2 \sigma_{r,s} | N \ , \ r \neq s . \]

Thus far styles of selection defined by just one rank set \( R_n \) have been considered; when there are two defining sets, such as \( R_n \) and \( R_m \) as in Section 2.1, the post-test projection of covariance between average responses \( \delta_n \) and \( \delta_m \) is

\[
C(\delta_n, \delta_m | S_n(x), S_m(x)) = \frac{1}{nm} \sum_{r \in R_n} \sum_{s \in R_m} C(y_r, y_s | x(r), x(s))
\]

(3.6.2)

\[ = \frac{1}{nm} \sum_{r \in R_a} \nu[y_r | x(r)] \]

\[ = \frac{1}{nm} \sum_{r \in R_a} (b_0 + b_1 x(r)) , \]

where \( R_a \) is \( R_n \cap R_m \) the set of ranks common to \( R_n \) and \( R_m \). If \( R_n \cap R_m \) is null, \( a = 0 \) and \( \sum_{r \in R_a} (b_0 + b_1 x(r)) = 0 \).
Removing the conditioning yields the pre-test covariance of average responses:

$$C(\delta_n, \delta_m) = E[C(\delta_n, \delta_m | S_n(x), S_m(x))] + C[E\{\delta_n | S_n(x)\}, E(\delta_m | S_m(x))]$$

(3.6.3)

$$= \frac{1}{nm} \sum_{R_a} (b_0 + b_1 x(r))} + C[\rho \Delta_n, \rho \Delta_m]$$

$$= \frac{1}{\sqrt{nm}} \{A + \rho^2 w\},$$

where \(w\) is the standardized covariance between selection differentials as defined at (2.1.8), and

$$A = \frac{1}{\sqrt{nm}} \sum_{R_a} (b_0 + b_1 \mu_r | N).$$

If \(R_n \cap R_m\) is null, \(A\) is zero and \(C(\delta_n, \delta_m) = \rho^2 w/\sqrt{nm}\). If \(R_n \equiv R_m\), then \(R_a = R_n\), \(A = (b_0 + b_1 k)\) and \(w = v\) the standardized variance of \(\Delta_n\); thus, this case

$$C(\delta_n, \delta_n) = \frac{1}{n} (b_0 + b_1 k + \rho^2 v) = V(\delta_n)$$

as required at (3.5.5).

3.7 Distributions of Responses

Further characterization of individual and average responses is provided by probability statements concerning \(y_r\) and \(\delta_n\). In the context of directional selection there is particular interest in the probabilities of positive response, \(Pr[y_r > 0]\) for \(r \in R_n\) and \(Pr[\delta_n > 0]\); these of course are just special cases of the corresponding distribution functions \(Pr[y_r \leq y_0]\) and \(Pr[\delta_n \leq \delta_0]\). Determination of these distribution functions requires a complete specification of the bivariate distribution of \(X\) and \(Y\) in the reference population \(\Pi_0\); even then they may be analytically
intractable because of the complexities introduced by order statistics in the X-dimension.

Let y denote "response;" \( \hat{y} = y_r \) for individual response, \( \hat{y} = \delta_n \) for average response. If all that can be specified about \( \Pi_0 \) is that information required to obtain the pre-test or post-test expectations and variances described in Sections 3.4 and 3.5, then all that can be specified about the distribution of \( \hat{y} \) is its mean \( \mu \) and variance \( \sigma^2 \), whether pre-test or post-test projections are to be made. Then the most that can be achieved in the way of probability statements about response \( \hat{y} \) is provided by the Bienayme-Tchebycheff inequality:

\[
\Pr[|\hat{y} - \mu| \leq \frac{\sigma}{\lambda^2}] \leq \lambda^2.
\]  

(3.7.1)

When probabilities of positive response are of interest, a lower bound for \( \Pr[\hat{y} > 0] \) is provided by (3.7.1) as follows:

\[
\Pr[\hat{y} \leq \mu - \frac{\sigma}{\lambda}] \leq \Pr[\hat{y} \geq \mu + \frac{\sigma}{\lambda} \text{ or } \hat{y} \leq \mu - \frac{\sigma}{\lambda}]
\]

\[
= \Pr[|\hat{y} - \mu| \geq \frac{\sigma}{\lambda}]
\]

\[
\leq \lambda^2
\]

and setting \( \lambda = \sigma/\mu = C \), the coefficient of variation of response \( \hat{y} \),

\[
\Pr[\hat{y} \leq 0] \leq C^2
\]

or, equivalently,

\[
\Pr[\hat{y} > 0] \geq 1 - C^2.
\]

(3.7.2)

It has to be emphasized that although exact and rigorous, (3.7.2) is a very weak inequality using only the existence, and appropriate pre-test or post-test values, of \( \mu \) and \( \sigma^2 \). As discussed by Mallows (1956) almost any further piece of information about the distribution of a random variable may be used to sharpen the Bienayme-Tchebycheff inequality.
If it can be assumed that the density function for the distribution of \( \hat{y} \) is continuous and has a single maximum, use can be made of Selberg's inequality (see Godwin, 1955):

\[
\Pr[|\hat{y} - \mu| \geq \frac{\sigma}{\lambda}] \leq \theta \lambda^2,
\]

(3.7.3)

where \( \theta = 0.565376... \), is a root of \( \theta^3 - 9\theta^2 + 3\theta + 1 = 0 \). In comparison with (3.7.1) this sharpens the Bienayme-Tchebycheff inequality by a factor of \( \theta \), and using the same argument as before

\[
\Pr[\hat{y} > 0] \geq 1 - \theta \epsilon^2.
\]

(3.7.4)

Inequalities (3.7.2) and (3.7.4) stress the importance of a low coefficient of variation of response \( \hat{y} \) in order to attain a high probability of positive response.

Moving to the other extreme, when the bivariate distribution of \( X \) and \( Y \) in \( \Pi_0 \) is completely specified by density function \( \lambda(x, y; \rho) \) at (3.1.3), explicit expressions for the distribution function of \( \hat{y} \), and hence for \( \Pr[\hat{y} > 0] \), can be derived in some cases.

The post-test distribution of \( y_r \) follows directly from (3.1.4):

\[
\Pr[y_r > y_0 | x_r] = L_y (y_0 | x_r; \rho)
\]

(3.7.5)

and since the \( n \) random variables \( y_r \), \( r \in R_n \), are conditionally independent given \( S_n(x) \), their joint post-test density function is

\[
\prod_{r \in R_n} \lambda_y (y_r | x_r; \rho)
\]

and so the post-test density function for average response \( \delta_n \) is available provided the required convolution is tractable.

Pre-test projections are more complex. For individual response \( y_r \), pre-test projection of \( \Pr[y_r > y_0] \) is obtained by removing the conditioning
in (3.7.5):

\[ a(y_0, r) = \Pr[y_r > y_0] = \int_{-\infty}^{\infty} L(y_0 | x(r); \rho) z(x(r); N, r) \, dx(r), \]

(3.7.6)

where \( z(x(r); N, r) \) is the density function for the distribution of \( x(r) \) defined at (2.3.5). Explicit evaluation of this pre-test projection is dependent on tractability of the integration in (3.7.6); numerical values could be obtained by routine quadrature of course. The same comments apply to average response \( \delta_n \) but with the additional complication that not only must the convolution for the conditional distribution of \( \delta_n \), given \( S_n(x) \), be available, but also removal of the conditioning requires the \( n \)-fold integration with respect to the joint density of \( n \) order-statistics in \( S_n(x) \). This situation may be alleviated in those special cases where the density function for the conditional distribution of \( \delta_n \), given \( S_n(x) \), is a function of the \( n \) random variables in \( S_n(x) \) only through \( \Delta_n \), corresponding to sufficiency of \( \Delta_n \) in the context of estimation theory; then removal of the conditioning requires integration over the distribution of \( \Delta_n \) rather than over the multivariate distribution of \( S_n(x) \). However, the distribution of \( \Delta_n \) would be required in such cases. The alternative approach of first removing the conditioning upon \( S_n(x) \) to yield the pre-test multivariate distribution of \( S_n(y) \) and then attempting a reduction to the pre-test distribution of the average, \( \delta_n \), is equally formidable.

Instead of the pre-test "probability of positive average response," \( \Pr[\delta_n > 0] \), consider the "average probability of positive response," \( \beta(0,n) \) defined as

\[ \beta(0,n) = \frac{1}{n} \sum_{R_n} \Pr[y_R > 0] = \Pr[y_R > 0], \]
where \( y^* \) is one of the \( n \) random variables in \( S_n(y) \) sampled at random.

In the case of directional selection a simple approximation for

\[
\beta(y_0, n) = \Pr[y^* > y_0] = \frac{1}{n} \sum_{i=1}^{n} \Pr[y_{N-n+i} > y_0]
\]

(3.7.7)

and hence for the pre-test distribution function of \( y^* \), is derived in the next section by a method similar to that used in Section 2.3 in deriving approximation \( \hat{k} \) for \( k \).

Using definition (3.7.7), this approximation to \( \beta(y_0, n) \) automatically provides approximations for the pre-test distribution functions of individual responses defined at (3.7.6):

\[
\beta(y_0, s) = \frac{1}{s} \sum_{j=j^-}^{N} \alpha(y_0, j), \quad j^- = N-s+1,
\]

\[
= \frac{1}{s} \left[ \alpha(y_0, N-s+1) + \sum_{j=j^-+1}^{N} \alpha(y_0, j) \right]
\]

or,

\[
s \cdot \beta(y_0, s) = \alpha(y_0, N-s+1) + (s-1) \beta(y_0, s-1)
\]

and so

\[
\Pr[y_i > y_0] = \alpha(y_0, i) = (N-i+1) \beta(y_0, N-i+1) - (N-i) \beta(y_0, N-i-1)
\]

(3.7.8)

for \( i=1,2,\ldots,N \). Thus, if the integration required for \( \alpha(y_0, i) \) at (3.7.6) proves intractable, a simple approximation is provided by (3.7.8) with \( \beta(y_0, N-i+1) \) and \( \beta(y_0, N-i) \) replaced by their approximate values derived in the next section.

3.8 **Approximation \( \hat{\beta}(y_0, n) \) for \( \beta(y_0, n) \) under Directional Selection**

Consider first the case where test material is assumed to form an infinite population with \( n \) selected test units as a random sample from
the upper tail of the standardized X-distribution truncated at \( x_0 \) (i.e., selection Model II of Chapter 1). Using the bivariate notations of Section 3.1,

\[
\Pr[y_* > y_0] = \Pr[y > y_0 | x > x_0] = \frac{L(x_0', y_0'; \rho)}{G(x_0)}.
\]  

(3.8.1)

Next, for selection Model III, assume \( x(r) \) is a fixed point initially \( (r = N-n) \), and treat \( x_{(r+1)}, x_{(r+2)}, \ldots, x_{(N)} \) as a random sample, size \( n \), from the upper tail of the standardized X-distribution truncated at \( x_{(r)} \). Then the conditional probability that \( y_* > y_0 \), given \( x_{(r)} \), is provided by (3.8.1) as

\[
\Pr[y_* > y_0 | x_{(r)}] = \frac{L(x_{(r)}', y_0'; \rho)}{G(x_{(r)})}.
\]

As in Chapter 2 the conditioning is removed by taking expectation over the distribution of \( x_{(r)} \) with density function \( z(x_{(r)}; N, r) \) at (2.3.5):

\[
\Pr[y_* > y_0] = \int_{-\infty}^{\infty} \frac{L(t, y_0'; \rho)}{G(t)} z(t; N, r) \, dt
\]

\[
= \frac{N}{(N-r)} \int_{-\infty}^{\infty} L(t, y_0'; \rho) z(t, N-1, r) \, dt.
\]

(3.8.2)

Now \( L(t, y_0'; \rho) \) can be written as

\[
\int_{-\infty}^{\infty} h(y) \, L_x(t | y; \rho) \, dy
\]

\( y_0 \)

so that

\[
\frac{d}{dt} \cdot L(t, y_0'; \rho) = -\int_{-\infty}^{\infty} \ell(t, y; \rho) \, dy = -f(t) \cdot L_y(y_0 | t; \rho)
\]

\( y_0 \)
and $L(t,y_0;\rho)$ may be expanded in a Taylor series about the point $t = a$ as

$$L(a,y_0;\rho) - (t-a) f(a)L_y(y_0|a;\rho)$$

$$- \frac{1}{2} (t-a)^2 f(a) \left[\frac{f'(a)}{f(a)} L_y(y_0|a;\rho) + L^*_y(y_0|a;\rho)\right] - \ldots ,$$

where

$$L^*_y(y_0|a;\rho) = \left[\frac{d}{dt} \cdot L_y(y_0|t;\rho)\right]_{t=a} \quad \text{(3.8.3)}.$$

Taking the integration at (3.8.2) term by term,

$$\frac{n}{N} \Pr[y_0 > y_0] = L(a,y_0;\rho)$$

$$- f(a) L_y(y_0|a;\rho)\{(\mu_r|N-1 - a) + \frac{1}{2} \sigma_{r,r|N-1} \frac{f'(a)}{f(a)}\}$$

$$- \frac{1}{2} \sigma_{r,r|N-1} (\mu_r|N-1 - a)^2 \{f'(a) L_y(y_0|a;\rho) + f(a) L^*_y(y_0|a;\rho)\}$$

$$- \frac{1}{2} \sigma_{r,r|N-1} f(a) L^*_y(y_0|a;\rho) - \ldots .$$

Choosing $a = x_0$ satisfying $G(x_0) = n/N$ and using the moment approximations (2.3.11) and (2.3.12),

$$\frac{n}{N} \Pr[y_0 > y_0] = L(x_0,y_0;\rho) - \frac{1}{2} \sigma_{r,r|N-1} f(x_0) L^*_y(y_0|x_0;\rho)$$

$$+ (N+1)^{-2} \{\text{further terms}\}$$

$$= L(x_0,y_0;\rho) - G(x_0) L^*_y(y_0|x_0;\rho) C(n,N)$$

$$+ (N+1)^{-2} \{\text{further terms}\}$$

and so the proposed approximation is $\hat{\beta}(y_0,n)$:

$$\beta(y_0,n) = \hat{\beta}(y_0,n) + (N+1)^{-2} \{\text{further terms}\} \quad \text{(3.8.4)}$$

and
\[ \hat{\beta}(y_0,n) = \frac{L(x_0,y_0;\rho)}{G(x_0)} - L_y^*(y_0|x_0;\rho) \cdot C(n,N) \]

with \( L_y^*(y_0|x_0;\rho) \) defined at (3.8.3) and \( C(n,N) \) is the parameter of the X-distribution in \( \Pi_0 \) defined at (2.3.12).

As was the case with approximations \( \hat{k} \) for \( k \) and \( \hat{v} \) for \( v \), evaluation of the performance of approximation \( \hat{\beta}(y_0,n) \) requires comparison with exact values of \( \beta(y_0,n) \) for some \( \Pi_0 \) where the latter are available. The same applies to approximation

\[ \hat{\alpha}(y_0,1) = (N-1+1) \cdot \hat{\beta}(y_0,N-1+1) - (N-1) \cdot \hat{\beta}(y_0,N-1) \quad (3.8.5) \]

suggested at (3.7.8). Approximation \( \hat{\beta}(y_0,n) \) does yield appropriate values in two special cases with general \( \ell(x,y;\rho) \). In the case of no selection, \( n = N, \ C(n,N) = 0, \ G(x_0) = 1, \ \hat{\beta}(y_0,N) = L(-\infty,y_0;\rho) = 1 - H(y_0) \) as required. In the case where \( N \to \infty \) with \( n/N = G(x_0) \) held constant, \( C(n,N) \to 0 \) and \( \hat{\beta}(y_0,n) \to L(x_0,y_0;\rho)/G(x_0) \) as required at (3.8.1) for selection Model I.

Consider next the special case where \( \Pi_0 \) has bivariate density function

\[ \ell(x,y;\rho) = \frac{(1 + \rho \ xy)}{\theta^2} \quad (3.8.6) \]

\[ |x| < \frac{1}{2} \theta , \quad |y| < \frac{1}{2} \theta , \quad |\rho| < \frac{1}{3} , \quad \theta^2 = 12 \]

which is standardized to zero means and unit variances, has uniform marginal densities \( f(x) = h(y) = 1/\theta \), and satisfies the linearity assumption (3.1.1):

\[ \ell_y(y|x;\rho) = \frac{(1 + \rho \ xy)}{\theta} , \quad L_y(y|x;\rho) = \left( \frac{1}{2} - \frac{Y}{\theta} \right) \left[ 1 + \frac{1}{2} \rho \ \theta \ x \left( \frac{1}{2} + \frac{Y}{\theta} \right) \right] \quad (3.8.7) \]
$E(y|x) = \rho \cdot x$.

The distribution of individual response $y_r$ has pre-test $\Pr[y_r > y_0]$ evaluated via (3.7.6) as

$$\alpha(y_0, r) = \int L_y(y_0|x(r); \rho) \cdot z(x(r); N, r) \, dx(r) \tag{3.8.8}$$

$$= L_y(y_0|\mu_r|N; \rho),$$

where $\mu_r|N = (r\theta)/(N+1) - \theta/2$ as in Section 2.4.

Thus, $\beta(y_0, n)$ is given explicitly as

$$\beta(y_0, n) = \frac{1}{n} \sum_{i=1}^{n} \alpha(y_0, N-n+1) \tag{3.8.9}$$

$$= L_y(y_0|k; \rho),$$

where $k = \theta(N-n)/2(N+1)$ as at (2.4.1).

The special case $y_0 = 0$ yields the corresponding probabilities of positive response:

$$\alpha(0, r) = \frac{1}{2} \left[ 1 + \frac{3}{2} \rho \left( \frac{2r-N-1}{N+1} \right) \right]$$

and

$$\beta(0, n) = \frac{1}{2} \left[ 1 + \frac{3}{2} \rho \left( \frac{N-n}{N+1} \right) \right].$$

Evaluation of approximation $\hat{\beta}(y_0, n)$ requires

$$L(x_0, y_0; \rho) = \left( \frac{1}{2} - \frac{x_0}{\theta} \right) \left( \frac{1}{2} - \frac{y_0}{\theta} \right) \left[ 1 + 3\rho \left( \frac{1}{2} + \frac{y_0}{\theta} \right) \left( \frac{1}{2} + \frac{x_0}{\theta} \right) \right];$$

$$\frac{n}{N} = \left( \frac{1}{2} - \frac{x_0}{\theta} \right)$$

and

$$L^*(y_0|x_0; \rho) = \frac{1}{2} \rho \theta \left( \frac{1}{2} - \frac{y_0}{\theta} \right) \left( \frac{1}{2} + \frac{y_0}{\theta} \right); \quad C(n, N) = \frac{\theta}{2N} \left( \frac{N-n}{N+1} \right).$$
Then in definition (3.8.4),

\[
\hat{\beta}(y_0, n) = \left( \frac{1}{2} - \frac{y_0}{\theta} \right) \left[ 1 + 3\rho \left( \frac{1}{2} + \frac{y_0}{\theta} \right) \left( \frac{N-n}{N} \right) \right] \\
- \left( \frac{1}{2} - \frac{y_0}{\theta} \right) \cdot 3\rho \left( \frac{1}{2} + \frac{y_0}{\theta} \right) \left( \frac{N-n}{N} \right) \frac{1}{N+1} \\
= \left( \frac{1}{2} - \frac{y_0}{\theta} \right) \left[ 1 + 3\rho \left( \frac{1}{2} + \frac{y_0}{\theta} \right) \left( \frac{N-n}{N+1} \right) \right]
\]

= \beta(y_0, n) at (3.8.9) on substituting for k.

Thus, for this particular bivariate distribution with Uniform marginal distributions and linear regressions, \( E(y|x) \) and \( E(x|y) \), approximation \( \hat{\beta}(y_0, n) \) provides the exact result \( \beta(y_0, n) \); approximation \( \hat{\alpha}(y_0, i) \) at (3.8.5), formed from \( \hat{\beta}(y_0, N-i+1) \) and \( \hat{\beta}(y_0, N-i) \), necessarily provides the exact result \( \alpha(y_0, i) \) at (3.8.8). The reason why \( \hat{\beta}(y_0, n) \) provides the exact result in this case is precisely that discussed in Section 2.4 in relation to approximation \( \hat{k} \) providing exact result \( k \) in the case of Uniformly distributed \( X \).

Further special cases indicating the adequacy of \( \hat{\beta}(y_0, n) \) as an approximation for \( \beta(y_0, n) \) are encountered in the next chapter.
4. THE BIVARIATE NORMAL MODEL

When the joint distribution of X and Y in reference population Π₀
is assumed to be bivariate Normal with standardized density function

\[ f(x,y;\rho) = \frac{1}{2\pi} \frac{1}{\sqrt{1-\rho^2}} \exp \left( -\frac{1}{2} \frac{x^2 - 2\rho xy + y^2}{1 - \rho^2} \right) \]

\[ |x| < \infty, \quad |y| < \infty, \quad |\rho| < 1 \]

some consequences of selection in finite populations are derived in this chapter.

As discussed in the previous chapter, such a distributional
assumption is not necessary for a determination of low moments, means
and variances, of response. However, the bivariate Normal distribution
is a reasonable assumption (or approximation) in many situations.

Additionally, the bivariate Normal distribution and the more general
multivariate Normal distribution form the basis for studies of truncation
selection in infinite test populations by Aitken (1934), Campbell¹ (1945),
Cochran (1950), Birnbaum and Meyer (1953), Cohen (1955,1957), Prout (1958),
and Finney (1962). Properties other than expectations, and sometimes
variances, of individual and average responses have received little
theoretical attention even in these studies of infinite test populations.

4.1 Expectations and Variances of Response

With \( f(x,y;\rho) \) as defined in this chapter, linearity condition
(3.1.1) is satisfied, as is the conditional variance expression at

¹F. L. Campbell. 1945. A doctoral dissertation from the University
of Michigan entitled "A study of truncated bivariate Normal distributions"
and unseen by me.
(3.1.2) with \( b_0 = (1 - \rho^2) \) and \( b_1 = 0 \):

\[
E[y|x] = \rho x, \quad V[y|x] = (1 - \rho^2).
\]

(4.1.1)

The overall summary of pre-test and post-test expectations and variances, compiled from results in Sections 3.4 and 3.5 together with (4.1.1), is listed next.

For individual responses:

\[
\begin{align*}
E[y_i] &= \rho \mu_{i|N} \\
E[y_i|x(i)] &= \rho x(i)
\end{align*}
\]

(4.1.2)

\[
\begin{align*}
V[y_i] &= 1 - \rho^2(1 - \sigma_{i,i|N}) \\
V[y_i|x(i)] &= (1 - \rho^2)
\end{align*}
\]

with pre-test:post-test information ratio

\[
\text{Eff} \{y_i;N\} = 1 - \frac{\rho^2 \sigma_{i,i|N}}{\{1 - \rho^2(1 - \sigma_{i,i|N})\}}.
\]

(4.1.3)

For average response:

\[
\begin{align*}
E[\delta_n] &= kp \\
E[\delta_n|\Delta_n] &= \rho \Delta_n
\end{align*}
\]

(4.1.4)

\[
\begin{align*}
V[\delta_n] &= \frac{1}{n} \{1 - \rho^2(1 - \nu)\} \\
V[\delta_n|\Delta_n] &= \frac{1}{n} (1 - \rho^2)
\end{align*}
\]

with pre-test:post-test information ratio

\[
\text{Eff} \{\delta_n;N\} = 1 - \frac{\rho^2 \nu}{\{1 - \rho^2(1-\nu)\}}.
\]

(4.1.5)

Since the marginal distribution of \( X \) is Normal, \( k \) and \( \nu \) are determined by expectations, \( \mu_{i|N} \), and covariances, \( \sigma_{i,j|N} \), of unit Normal order statistics. Thus, the above compilation of projections can
be evaluated using either exact or approximate values of \( k \) and \( v \), for any rank set \( R^* \) defining selected units, as described in Chapter 2.

Prout (1962) obtained the conditional variance of average response, given the observed selection differential, in the context of a bivariate Normal distribution of phenotypic values \((X)\) and genotypic values \((Y)\) with \( \rho^2 = \sigma^2_y / \sigma^2_x \). This was sufficient for his purpose of deriving the variance of the ratio of average response to observed selection differential, given the latter. In the current notation, and neglecting the error variance in measurement of response, his result may be obtained directly from the post-test variance of \( \delta \) at (4.1.4):

\[
V(\delta_n^{\Delta Y} | \Delta_n^x) = \frac{1}{\Delta_n^2} \frac{\rho^2 (1 - \rho^2)}{n}^2
\]

In the context of directional selection, pre-test expectation of average response at (4.1.4) with \( k \) appropriately calculated from expectations of Normal order statistics, is usually quoted in standard texts (see e.g., Lerner, 1958 and Falconer, 1964) although its pre-test nature is not emphasized.

For given \( N \), \( |\mu_1|_N = |\mu_{N-1}|_N \) increases as \( |i - \frac{1}{2} (N+1)| \) increases, that is as rank \( i \) becomes more extreme, whereas for given \( i \), \( \mu_i|_N \) decreases as \( N \) increases. These trends translate directly to pre-test expectation of individual response at (4.1.2) depending on the sign of \( \rho \).

For given \( N \), \( \sigma_{i,i}^N = \sigma_{N-i+1,N-i+1}^N \) increases as rank \( i \) becomes more extreme, whereas for given \( i \), \( \sigma_{i,i}^N \) decreases as \( N \) increases. Thus, from the form of \( \text{Eff}(y_1|N) \) at (4.1.3), the information content of pre-test projection of individual response \( y_1 \), relative to that of
post-test projection, decreases as rank \( i \) becomes more extreme given \( N \) and \( \rho^2 \) but increases as \( N \) increases given \( i \) and \( \rho^2 \). Typical examples of these trends are listed next for \( \rho = 0.6 \):

\[
\begin{align*}
\text{Eff}(y_{10};10) &= 0.8377 & \text{Eff}(y_4;4) &= 0.7833 \\
\text{Eff}(y_9;10) &= 0.8923 & \text{Eff}(y_4;8) &= 0.9047 \\
\text{Eff}(y_8;10) &= 0.9104 & \text{Eff}(y_4;12) &= 0.9271 \\
\text{Eff}(y_7;10) &= 0.9184 & \text{Eff}(y_4;16) &= 0.9378 \\
\text{Eff}(y_6;10) &= 0.9217 & \text{Eff}(y_4;20) &= 0.9444
\end{align*}
\]

For given \( N \) and \( i \) but unknown \( \rho \), a more practical situation, \( \text{Eff}(y_i;N) \) decreases as \( \rho^2 \) increases.

Similar considerations apply to \( \text{Eff}(\delta_n;N) \) at (4.1.5) under directional selection defined in Section 2.2 by

\[ R_n \equiv \{N-n+1,N-n+2,\ldots,N\} \, . \]

For this case values of \( \text{Eff}(\delta_n;N) \) are shown in Table 4.1 based on values of \( v \) in Table 2.3 and for the same \((N,n)\) combinations as listed there.

For given \( n/N \) and \( \rho \) the stability of \( \text{Eff}(\delta_n;N) \) over increasing \( N \) follows from the similar stability of \( v \) discussed in Section 2.6. The main feature of Table 3.1 is the decrease in \( \text{Eff}(\delta_n;N) \) as precision of selection \( (\rho^2) \) increases.

Continuing with the case of directional selection, pre-test projections of average response are shown in Table 4.2 based on values of \( k \) and \( v \) in Tables 2.1 and 2.3; tabulated are the expected response as a multiple of \( \sigma_y \) (columns headed \( kp \)), and standard deviation of response (columns headed \( s \)) also as a multiple of \( \sigma_y \) .
Table 4.1. Pre-test:post-test information ratio for average response in the case of directional selection and bivariate Normal distribution for (X,Y)

<table>
<thead>
<tr>
<th>( \frac{n}{N} )</th>
<th>N</th>
<th>( \rho = 0.2 )</th>
<th>( \rho = 0.4 )</th>
<th>( \rho = 0.6 )</th>
<th>( \rho = 0.8 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>10</td>
<td>0.9727</td>
<td>0.8862</td>
<td>0.7251</td>
<td>0.4549</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.9726</td>
<td>0.8858</td>
<td>0.7242</td>
<td>0.4538</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.9725</td>
<td>0.8855</td>
<td>0.7236</td>
<td>0.4530</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.9724</td>
<td>0.8853</td>
<td>0.7233</td>
<td>0.4527</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.9724</td>
<td>0.8852</td>
<td>0.7231</td>
<td>0.4525</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9724</td>
<td>0.8851</td>
<td>0.7229</td>
<td>0.4522</td>
</tr>
<tr>
<td>0.2</td>
<td>10</td>
<td>0.9816</td>
<td>0.9209</td>
<td>0.7978</td>
<td>0.5552</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.9813</td>
<td>0.9199</td>
<td>0.7955</td>
<td>0.5517</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.9811</td>
<td>0.9192</td>
<td>0.7940</td>
<td>0.5494</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.9811</td>
<td>0.9189</td>
<td>0.7932</td>
<td>0.5483</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.9810</td>
<td>0.9187</td>
<td>0.7928</td>
<td>0.5476</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9809</td>
<td>0.9182</td>
<td>0.7918</td>
<td>0.5462</td>
</tr>
<tr>
<td>0.1</td>
<td>10</td>
<td>0.9859</td>
<td>0.9384</td>
<td>0.8378</td>
<td>0.6203</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.9855</td>
<td>0.9371</td>
<td>0.8347</td>
<td>0.6150</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.9853</td>
<td>0.9360</td>
<td>0.8321</td>
<td>0.6106</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.9851</td>
<td>0.9354</td>
<td>0.8307</td>
<td>0.6082</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.9850</td>
<td>0.9350</td>
<td>0.8297</td>
<td>0.6066</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.9848</td>
<td>0.9342</td>
<td>0.8278</td>
<td>0.6033</td>
</tr>
</tbody>
</table>

\[ s^2 = \frac{1}{n} \{1 - \rho^2(1 - v)\} = V[\delta_n] \]  \hspace{1cm} (4.1.6)

as at (4.1.4).

The effect of increasing precision of selection (\( \rho^2 \)) is clear, both from the moments at (4.1.4) and from Table 4.2. For a given (N,n) combination, increasing \( \rho \) increases expected response and decreases the variance of average response.

Recalling the properties of k and v discussed in Chapter 2, for given N and \( \rho \), increased selection pressure (n decreasing, k increasing) naturally increases \( E[\delta_n] \) but invariably increases \( V[\delta_n] \) because although \( (1-v) \) increases and this has a decreasing effect on \( V[\delta_n] \), \( s^2 \) is
Table 4.2. Pre-test projections of average response in the case of directional selection and bivariate Normal distribution for (X,Y)

<table>
<thead>
<tr>
<th>N</th>
<th>N</th>
<th>ρ = 0.2</th>
<th></th>
<th>ρ = 0.4</th>
<th></th>
<th>ρ = 0.6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>kp</td>
<td>s</td>
<td>kp</td>
<td>s</td>
<td>kp</td>
<td>s</td>
</tr>
<tr>
<td>10</td>
<td>0.1478</td>
<td>0.4443</td>
<td>0.2956</td>
<td>0.4354</td>
<td>0.4433</td>
<td>0.4202</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.1535</td>
<td>0.3142</td>
<td>0.3070</td>
<td>0.3080</td>
<td>0.4605</td>
<td>0.2973</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.1555</td>
<td>0.2565</td>
<td>0.3110</td>
<td>0.2515</td>
<td>0.4664</td>
<td>0.2428</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.1565</td>
<td>0.2222</td>
<td>0.3130</td>
<td>0.2178</td>
<td>0.4695</td>
<td>0.2103</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.1571</td>
<td>0.1987</td>
<td>0.3142</td>
<td>0.1948</td>
<td>0.4713</td>
<td>0.1882</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.1583</td>
<td>0.1405</td>
<td>0.3167</td>
<td>0.1378</td>
<td>0.4750</td>
<td>0.1331</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>N</th>
<th>ρ = 0.2</th>
<th></th>
<th>ρ = 0.4</th>
<th></th>
<th>ρ = 0.6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.2540</td>
<td>0.6993</td>
<td>0.5080</td>
<td>0.6753</td>
<td>0.7621</td>
<td>0.6333</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.2664</td>
<td>0.4945</td>
<td>0.5327</td>
<td>0.4778</td>
<td>0.7991</td>
<td>0.4485</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.2707</td>
<td>0.4038</td>
<td>0.5415</td>
<td>0.3903</td>
<td>0.8122</td>
<td>0.3665</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.2730</td>
<td>0.3497</td>
<td>0.5460</td>
<td>0.3380</td>
<td>0.8189</td>
<td>0.3176</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.2744</td>
<td>0.3128</td>
<td>0.5487</td>
<td>0.3024</td>
<td>0.8231</td>
<td>0.2841</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.2771</td>
<td>0.2212</td>
<td>0.5543</td>
<td>0.2139</td>
<td>0.8314</td>
<td>0.2010</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N</th>
<th>N</th>
<th>ρ = 0.2</th>
<th></th>
<th>ρ = 0.4</th>
<th></th>
<th>ρ = 0.6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.3078</td>
<td>0.9870</td>
<td>0.6155</td>
<td>0.9461</td>
<td>0.9233</td>
<td>0.8740</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.3275</td>
<td>0.6979</td>
<td>0.6550</td>
<td>0.6695</td>
<td>0.9825</td>
<td>0.6192</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.3349</td>
<td>0.5699</td>
<td>0.6698</td>
<td>0.5469</td>
<td>1.0046</td>
<td>0.5063</td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>0.3387</td>
<td>0.4936</td>
<td>0.6775</td>
<td>0.4738</td>
<td>1.0162</td>
<td>0.4389</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>0.3411</td>
<td>0.4415</td>
<td>0.6822</td>
<td>0.4239</td>
<td>1.0233</td>
<td>0.3928</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0.3460</td>
<td>0.3122</td>
<td>0.6919</td>
<td>0.2999</td>
<td>1.0379</td>
<td>0.2781</td>
<td></td>
</tr>
</tbody>
</table>

-dominated by $n^{-1}$. This is illustrated in Table 4.2 by the magnitude of $s$ relative to $E(\delta_n)$; $s$ is frequently of the same order as $kp$, exceeds $kp$ when selection precision is low and $n$ is small, but is less than $kp$ when selection precision is intermediate to high and $n$ is large. The relative magnitude of $s$ to $kp$ defines the coefficient of variation of average response on a pre-test projection basis; as discussed in Section 3.7 a small coefficient of variation is necessary to attain a high lower bound to the pre-test probability of positive average response.
4.2 Distributions of Response

For bivariate Normal density function \( \lambda(x,y; \rho) \) the general bivariate notations of Section 3.1, for reference population \( \Pi_0 \), specialize to:

\[
f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} x^2}, \quad F(x) = \int_{-\infty}^{\infty} f(t) \, dt = 1 - G(x)
\]

\[
h(y) = f(y), \quad H(y) = F(y)
\]

\[
\lambda(x|y; \rho) = \frac{1}{\sqrt{1-\rho^2}} f\left(\frac{x-\rho y}{\sqrt{1-\rho^2}}\right), \quad \lambda(y|x; \rho) = \frac{1}{\sqrt{1-\rho^2}} f\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right), \quad (4.2.1)
\]

\[
L_x(x|y; \rho) = G\left(\frac{x-\rho y}{\sqrt{1-\rho^2}}\right), \quad L_y(y|x; \rho) = G\left(\frac{y-\rho x}{\sqrt{1-\rho^2}}\right) \quad (4.2.2)
\]

and

\[
L(x_0, y_0; \rho) = \int_{x_0}^{\infty} \int_{y_0}^{\infty} \lambda(x, y; \rho) \, dy \, dx
\]

\[
= \frac{1}{2\pi} \int_{0}^{\pi} \exp\left(-\frac{1}{2} \left(\frac{x_0^2 + y_0^2 - 2x_0 y_0 \cos \theta}{\sin^2 \theta}\right)\right) \, d\theta, \quad \cos \theta = \rho \quad (4.2.3)
\]

A nomogram yielding numerical values of \( L(x_0, y_0; \rho) \) is given by Zelen and Severo (1964) and an extensive tabulation of \( L(x_0, y_0; \rho) \) is provided by Anonymous (1959).

Using (3.7.5) and (4.2.2) the post-test distribution of individual response \( y_i \) is defined by

\[
Pr[y_i > y_0 | x(i)] = G\left(\frac{y_0 - \rho x(i)}{\sqrt{1-\rho^2}}\right)
\]

so that post-test probability statements about individual responses are readily available from standard tabulations of the unit Normal distribution function.
Using (4.2.1) and given $S_n(x)$, the sequence of $n$ random variables $y_x$, $x \in R_n$, is a set of conditionally independent Normal random variables with means $\rho \Delta_n$ and common variance $(1-\rho^2)$ as at (4.1.2); thus average response, $\delta_n$, has post-test distribution as Normal with mean $\rho \Delta_n$ and variance $(1-\rho^2)/n$ as at (4.1.4). In this case, therefore, the post-test convolution discussed in Section 3.7 is simply obtained and post-test probability statements about average response follow directly:

$$
Pr[\delta_n \leq \delta | S_n(x)] = F\left(\frac{\delta - \rho \Delta_n}{\sqrt{(1-\rho^2)}/\sqrt{n}}\right).
$$

In particular, the post-test probability of positive average response is

$$
Pr[\delta_n > 0 | S_n(x)] = G\left(-\Delta_n \left(\frac{\rho \Delta_n^2}{1-\rho^2}\right)^{1/2}\right), \quad \rho > 0,
$$

and since $G(x)$ is a decreasing function of $x$, this probability is maximized by choosing that rank set $R_n$ for which $\sqrt{n} \Delta_n$ is a maximum, implying directional selection of $n$ top-ranking test units but with $n$ varying from one sample of $N$ test units to another, all of this being independent of $\rho > 0$.

For the case of directional selection ($k > 0$, $\rho > 0$), post-test probability of positive average response, conditional on $\Delta_n$ observed at its expected value $k$, is shown in Table 4.3.

Apart from the obvious effect that increases in selection precision increase the probability of positive average response, Table 4.3 illustrates the above mentioned notion of a maximum value of

$$
Pr[\delta_n > 0 | S_n(x)]
$$

with respect to choice of $n$ for given $N$. For example, when $N = 30$ and $\rho = 0.4$, $Pr[\delta_n > 0 | \Delta_n = k] = 0.897$ when $n = 3$ top-ranking test units are selected, increases to 0.926 when $n = 6$ and then decreases.
Table 4.3. $\Pr[\delta_n > 0 | \Delta_n = k]$ in the case of directional selection and bivariate Normal distribution for $(X,Y)$

<table>
<thead>
<tr>
<th>$n$</th>
<th>N</th>
<th>$\rho = 0.2$</th>
<th>$\rho = 0.4$</th>
<th>$\rho = 0.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>0.6320</td>
<td>0.7646</td>
<td>0.8924</td>
</tr>
<tr>
<td>0.5</td>
<td>20</td>
<td>0.6898</td>
<td>0.8553</td>
<td>0.9656</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.7306</td>
<td>0.9056</td>
<td>0.9880</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.7625</td>
<td>0.9367</td>
<td>0.9957</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.7886</td>
<td>0.9567</td>
<td>0.9984</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.8734</td>
<td>0.9927</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.6431</td>
<td>0.7835</td>
<td>0.9110</td>
</tr>
<tr>
<td>0.2</td>
<td>20</td>
<td>0.7067</td>
<td>0.8775</td>
<td>0.9771</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.7507</td>
<td>0.9261</td>
<td>0.9936</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.7847</td>
<td>0.9540</td>
<td>0.9981</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.8121</td>
<td>0.9708</td>
<td>0.9994</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.8971</td>
<td>0.9966</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.6233</td>
<td>0.7491</td>
<td>0.8758</td>
</tr>
<tr>
<td>0.1</td>
<td>20</td>
<td>0.6818</td>
<td>0.8439</td>
<td>0.9588</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.7231</td>
<td>0.8972</td>
<td>0.9852</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.7554</td>
<td>0.9303</td>
<td>0.9945</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.7818</td>
<td>0.9520</td>
<td>0.9979</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.8679</td>
<td>0.9915</td>
<td></td>
</tr>
</tbody>
</table>

*Indicates $\Pr[\delta_n > 0 | \Delta_n = k] > 0.99995$.

to 0.906 when $n = 15$. Using (4.2.4), $\Pr[\delta_n > 0 | \Delta_n = k]$ is maximized for given $N$ when $\sqrt{n} k_0$ is at a maximum; since $n = NG(x_0)$, differentiating $\sqrt{n} k_0$ with respect to $x_0$ yields

$$\frac{1}{2} \sqrt{n} k_0 (k_0 - 2x_0)$$

which is zero for a maximum of $\sqrt{n} k_0$ at $k_0 = 2x_0$ with solution

$$n = (0.27)N.$$  \hspace{1cm} (4.2.5)

Pre-test distributions of individual response, $y_r$, and average response, $\delta_n$, are obtained by removal of the conditioning on $S_n(x)$ from post-test distributions as discussed in Section 3.7. For individual
response, (3.7.6) and (4.2.2) provide

\[ \alpha(y_0, r) = \Pr[y > y_0] \]

\[ = \int_{-\infty}^{\infty} G(y_0 - \rho t) \frac{N!}{(r-1)! (N-r)!} \left\{ F(t) \right\}^{r-1} \left\{ G(t) \right\}^{N-r} f(t) \, dt \]

which does not appear to be analytically tractable, and for average response, using (4.2.4),

\[ \Pr[\delta_n \leq \delta] = \int_{-\infty}^{\infty} F\left( \frac{\delta - \rho \Delta_n}{\sqrt{1-\rho^2}} \right) w(\Delta_n; N) \, d\Delta_n , \]

where \( w(\Delta_n; N) \) is the density function for the distribution of \( \Delta_n \) in samples of size \( N \) from \( f(x) \) and is not available.

Thus, it is necessary to resort to approximation \( \hat{\alpha}(y_0, r) \) for \( \alpha(y_0, r) \) as developed in the next section, and to lower bounds for \( \Pr[\delta_n > 0] \) as developed at (3.7.2) and (3.7.4). The pre-test coefficient of variation, \( C \), for \( \delta_n \) is obtained from (4.1.4) as

\[ C^2 = \frac{1 - \rho^2}{n k^2 \rho^2} = \left( \frac{\sigma}{kp} \right)^2 \]

and can be evaluated using Table 4.2. For the case of directional selection \( (k > 0, \rho > 0) \) values of the lower bounds to \( \Pr[\delta_n > 0] \) provided by Selberg's inequality at (3.7.4) are listed in Table 4.4.

In some cases the lower bound is practically worthless (as indicated by "_" in Table 4.4). However, in those cases characterized by intermediate to high precision of selection and larger values of \( N \), the existence of a maximum with respect to choice of \( n \) given \( N \), discussed previously in relation to post-test projections, is also indicated for pre-test projections.
Table 4.4. Lower bounds to pre-test $\Pr[\delta_n > 0]$ in the case of
directional selection and bivariate Normal distribution for
$(X,Y)$

<table>
<thead>
<tr>
<th>$n/N$</th>
<th>$N$</th>
<th>$\rho = 0.2$</th>
<th>$\rho = 0.4$</th>
<th>$\rho = 0.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>a</td>
<td>a</td>
<td>0.4920</td>
</tr>
<tr>
<td>0.5</td>
<td>20</td>
<td>a</td>
<td>0.4309</td>
<td>0.7644</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>a</td>
<td>0.6303</td>
<td>0.8468</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>a</td>
<td>0.7262</td>
<td>0.8866</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.0956</td>
<td>0.7827</td>
<td>0.9099</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.5546</td>
<td>0.8930</td>
<td>0.9556</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>a</td>
<td>0.0009</td>
<td>0.6096</td>
</tr>
<tr>
<td>0.2</td>
<td>20</td>
<td>a</td>
<td>0.5452</td>
<td>0.8219</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>a</td>
<td>0.7063</td>
<td>0.8849</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.0723</td>
<td>0.7833</td>
<td>0.9150</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.2653</td>
<td>0.8283</td>
<td>0.9326</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.6397</td>
<td>0.9158</td>
<td>0.9670</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>a</td>
<td>a</td>
<td>0.4934</td>
</tr>
<tr>
<td>0.1</td>
<td>20</td>
<td>a</td>
<td>0.4093</td>
<td>0.7754</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>a</td>
<td>0.6231</td>
<td>0.8564</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>a</td>
<td>0.7235</td>
<td>0.8945</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.0528</td>
<td>0.7817</td>
<td>0.9167</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>0.5397</td>
<td>0.8938</td>
<td>0.9594</td>
</tr>
</tbody>
</table>

* indicates a negative lower bound.

4.3 **Approximation $\hat{\beta}(y_0,n)$**

Recalling the definitions of $L^*(y_0|x_0;\rho)$ at (3.8.3) and $C(n,N)$ at
(2.3.12), the bivariate Normal notations at (4.2.1) and (4.2.2) provide

$$L^*_y(y_0|x_0;\rho) = \left[ \frac{d}{dt} \cdot \frac{y_0 - \rho t}{\sqrt{1-\rho^2}} \right]_{t=x_0}$$

$$= \frac{\rho}{\sqrt{1-\rho^2}} \frac{y_0 - \rho x_0}{\sqrt{1-\rho^2}}$$

and $C(n,N) = F(x_0)/2(N+1) f(x_0)$. Thus, in (3.8.4), the approximation
for $\Pr[y_* > y_0]$, as a function of $x_0$ satisfying $G(x_0) = n/N$, is given by
\[ \beta(y_0, n) = \hat{\beta}(y_0, n) + (N+1)^{-2} \{ \text{further terms} \} \]  

\[ \hat{\beta}(y_0, n) = \frac{L(x_0, y_0; \rho)}{G(x_0)} - \frac{(N-n)}{2n(N+1)k_0} \frac{\rho}{\sqrt{1-\rho^2}} \frac{y_0 - \rho x_0}{\sqrt{1-\rho^2}} f(\frac{\rho x_0}{\sqrt{1-\rho^2}}) \]  

and the approximate average probability of positive response under directional selection is

\[ \hat{\beta}(0, n) = \frac{L(x_0, 0; \rho)}{G(x_0)} - \frac{(N-n)}{2n(N+1)k_0} \frac{\rho}{\sqrt{1-\rho^2}} f(\frac{\rho x_0}{\sqrt{1-\rho^2}}). \]  

In the special case that \( n = \frac{1}{2} N \), \( x_0 = 0 \) and

\[ \hat{\beta}(0, \frac{1}{2} N) = \frac{1}{2} + \frac{\arcsin \rho}{\pi} - \frac{1}{4(N+1)} \frac{\rho}{\sqrt{1-\rho^2}}. \]

Calculation of \( \hat{\beta}(y_0, n) \) requires \( L(x_0, y_0; \rho) \) which can be obtained by quadrature on (4.2.3) or by interpolation in the tabulation of Anonymous (1959). An example of the results of such calculations is provided in Table 4.5; in calculating the values of \( \hat{\beta}(0, n) \) shown there, \( L(x_0, y_0; \rho) \) was obtained by 32-point Gaussian closed quadrature in double precision on an IBM 360/75 computer, and also by simple linear interpolation in Table I of Anonymous (1959). The interpolated values agreed with the quadrature results to within 0.001 always, and better in the cases of \( n \) closer to \( (1/2)N \).

The corresponding approximations for \( \alpha(0, i) = \Pr[y_i > 0] \), namely

\[ \hat{\alpha}(0, i) = (N-i+1) \hat{\beta}(0, N-i+1) - (N-i) \hat{\beta}(0, N-i) \]

as suggested at (3.8.5), are shown in Table 4.6 together with the post-test probability that \( y_i > 0 \), conditional on the observed \( x(i) = \nu_i \mid N \), derived from Section 4.2 as
Table 4.5. Values of $\hat{\beta}(0, n)$ for $N = 20$ in the case of directional selection and bivariate Normal distribution for $(X, Y)$

<table>
<thead>
<tr>
<th>n</th>
<th>$\rho = 0.2$</th>
<th>$\rho = 0.4$</th>
<th>$\rho = 0.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.6458</td>
<td>0.7834</td>
<td>0.9020</td>
</tr>
<tr>
<td>2</td>
<td>0.6298</td>
<td>0.7562</td>
<td>0.8736</td>
</tr>
<tr>
<td>3</td>
<td>0.6171</td>
<td>0.7335</td>
<td>0.8472</td>
</tr>
<tr>
<td>4</td>
<td>0.6064</td>
<td>0.7139</td>
<td>0.8227</td>
</tr>
<tr>
<td>5</td>
<td>0.5973</td>
<td>0.6964</td>
<td>0.7997</td>
</tr>
<tr>
<td>6</td>
<td>0.5890</td>
<td>0.6806</td>
<td>0.7776</td>
</tr>
<tr>
<td>7</td>
<td>0.5815</td>
<td>0.6658</td>
<td>0.7564</td>
</tr>
<tr>
<td>8</td>
<td>0.5746</td>
<td>0.6519</td>
<td>0.7357</td>
</tr>
<tr>
<td>9</td>
<td>0.5679</td>
<td>0.6385</td>
<td>0.7156</td>
</tr>
<tr>
<td>10</td>
<td>0.5617</td>
<td>0.6258</td>
<td>0.6958</td>
</tr>
</tbody>
</table>

Table 4.6. Values of $\hat{\alpha}(0, i)$ and $\bar{\alpha}(0, i)$ for $N = 20$ and bivariate Normal distribution for $(X, Y)$

<table>
<thead>
<tr>
<th>Rank</th>
<th>$\rho = 0.2$</th>
<th>$\rho = 0.4$</th>
<th>$\rho = 0.6$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{\alpha}(0, i)$</td>
<td>$\bar{\alpha}(0, i)$</td>
<td>$\hat{\alpha}(0, i)$</td>
</tr>
<tr>
<td>20</td>
<td>0.6458</td>
<td>0.6485</td>
<td>0.7834</td>
</tr>
<tr>
<td>19</td>
<td>0.6138</td>
<td>0.6141</td>
<td>0.7290</td>
</tr>
<tr>
<td>18</td>
<td>0.5913</td>
<td>0.5917</td>
<td>0.6881</td>
</tr>
<tr>
<td>17</td>
<td>0.5743</td>
<td>0.5746</td>
<td>0.6551</td>
</tr>
<tr>
<td>16</td>
<td>0.5609</td>
<td>0.5605</td>
<td>0.6264</td>
</tr>
<tr>
<td>15</td>
<td>0.5475</td>
<td>0.5480</td>
<td>0.6016</td>
</tr>
<tr>
<td>14</td>
<td>0.5365</td>
<td>0.5365</td>
<td>0.5770</td>
</tr>
<tr>
<td>13</td>
<td>0.5263</td>
<td>0.5256</td>
<td>0.5546</td>
</tr>
<tr>
<td>12</td>
<td>0.5143</td>
<td>0.5152</td>
<td>0.5313</td>
</tr>
<tr>
<td>11</td>
<td>0.5059</td>
<td>0.5051</td>
<td>0.5115</td>
</tr>
</tbody>
</table>

\[ \hat{\alpha}(0, i) = \Pr[y_i > 0 | x_{(i)}] = \mu | N = G(\frac{-\rho \mu_i}{\sqrt{1-\rho^2}}) \]

and included for comparison purposes.

The results in Table 4.6 suggest that, at least for $N = 20$, $\hat{\alpha}(0, i)$ is very similar to $\bar{\alpha}(0, i)$; since $\hat{\alpha}(0, i) \approx \alpha(0, i)$, this suggests that

\[ \alpha(0, i) = \Pr[y_i > 0] = E[G(\frac{-\rho x_{(i)}}{\sqrt{1-\rho^2}})] \]

\[ \approx G(\frac{-\rho \mu_i}{\sqrt{1-\rho^2}}) = \Pr[y_i > 0 | x_{(i)} = E[x_{(i)}]] = \bar{\alpha}(i, 0). \]
In respect to this suggestion it should be noted that \( G(-x) \) is a concave function of \( x \) when \( x > 0 \), so that for \( \rho > 0 \),

\[
G\left(\frac{-\rho x}{\sqrt{1-\rho^2}}\right) \text{ is concave when } x > 0
\]

and by Jensen's Inequality,

\[
E\left[G\left(\frac{-\rho x}{\sqrt{1-\rho^2}}(i)\right)\right] \leq G\left(\frac{-\rho i}{\sqrt{1-\rho^2}}\right)
\]

(4.3.3)

or \( \alpha(0,i) \leq \bar{\alpha}(0,i) \), \( i > \frac{1}{2} (N+1) \)

when \( Pr[x(i) > 0] \to 1 \) as is the case for \( i \to N \) and large \( N \). Examination of Table 4.6, where \( N = 20 \) throughout, reveals that \( \hat{\alpha}(0,i) \leq \bar{\alpha}(0,i) \) almost always, the only exceptions being when \( i \) close to \( \frac{1}{2} (N+1) \) and for low selection precision (\( \rho = 0.2 \)). This represents further confirmation of the adequacy of \( \hat{\alpha}(0,i) \) as an approximation for \( \alpha(0,i) \).

Further interpretation of Tables 4.3 and 4.6 involves some of the symmetry relationships developed in the next section.

4.4 Symmetry Considerations

Symmetry and continuity properties of the bivariate Normal distribution lead directly to some simple assertions regarding probabilities of positive response. The case \( \rho = 0 \) is trivial since then \( X \) does not qualify as a selection basis for response \( Y \). There is no loss of generality in assuming \( \rho > 0 \); if \( \rho < 0 \), simply change the sign of \( Y \).

Since \( x(i) \geq x(j) \) for \( i > j \), and \( G(x) \) is a monotone decreasing function of \( x \),

\[
G\left(\frac{-\rho x(0)}{\sqrt{1-\rho^2}}(i)\right) \geq G\left(\frac{-\rho x(0)}{\sqrt{1-\rho^2}}(j)\right) , i > j
\]

or
\[ \Pr[y_i > y_0|x(i)] \geq \Pr[y_j > y_0|x(j)], \ i > j \]

is an obvious property of post-test distributions of individual responses. But more than this, since \( \mu_i|N = -\mu_{N-i+1}|N \) and \( G(x) = 1 - G(-x) \),

\[ \Pr[y_i > 0|x(i)] = \mu_i|N + \Pr[y_{N-i+1}|N < 0|x_{(N-i+1)} = -\mu_i|N] \]

and

\[ \Pr[y_i > 0|x(i)] = \mu_i|N - \frac{1}{2} = \frac{1}{2} - \Pr[y_{N-i+1} > 0|x_{(N-i+1)} = -\mu_i|N] \]

with \( \Pr[y_i > 0|x(i)] = 0 ; i = \frac{1}{2} (N+1) = \frac{1}{2} \) when \( N \) is odd.

The same properties carry over to pre-test distributions of individual responses:

\[ \Pr[y_i > 0] > \Pr[y_j > 0], \ i > j \]

and

\[ \Pr[y_i > 0] - \frac{1}{2} = \frac{1}{2} - \Pr[y_{N-i+1} > 0] \]

so that in Table 4.6 where only \( i \geq 11 \) is considered

\( \hat{\alpha}(0,N-i+1) = 1 - \hat{\alpha}(0,i), \ \alpha(0,N-i+1) = 1 - \alpha(0,i). \)

In all cases where directional selection has been considered in this chapter, it has been assumed that the \( n \) top-ranking test units have been selected and that positive average response \( (\delta_n > 0) \) is of interest. For the pre-test distribution of average response it can be asserted that

\[ \Pr[\delta_n > 0] > \frac{1}{2} \text{ for } n < N \]

because applying the symmetry properties to the case \( n = N \) (no selection)

\[ \Pr[\delta_N > 0] = \Pr[\delta_N < 0] = \frac{1}{2} \]

and directional selection of \( n < N \) test units cannot lead to a lower probability of positive average response that retaining all \( N \) test units.
Thus, in Table 4.4, those cells where the lower bound to $\Pr[\delta_n > 0]$, obtained by Selberg's inequality, does not exceed 0.5 can have the value 0.5 inserted on the basis of symmetry considerations alone.

For the post-test distribution of $\delta_n$, it is clear that when $n = N$

$$\Pr[\delta_n > 0|\Delta_n] > \frac{1}{2} \text{ when } \Delta_n > 0 \text{ and } < \frac{1}{2} \text{ when } \Delta_n < 0$$

so that unlike the pre-test probability of positive average response, \( \Pr[\delta_n > 0|\Delta_n] = 1/2 \) only when $\Delta_n = 0$. Similar to the post-test distributions of individual responses, there is an ordering of probabilities in the post-test distribution of average response;

$$\Pr[\delta_n > \delta|\Delta_n] > \Pr[\delta > \delta|\Delta_n] \text{, if } \Delta_n > \Delta_n'$$

where $\Delta_n$ and $\Delta'_n$ refer to directional selection of the same number, $n$, of test units from different samples of $N$ test units.

Finally, when directional selection is specified by selecting the $n$ bottom-ranking units from test material, the above results apply with the simple replacement of "negative response" for "positive response" and reversal of appropriate inequalities; for example, $G(x)$ is a convex function of $x$ when $x > 0$ so that inequality (4.3.3) is reversed to read

$$\tilde{\alpha}(0,i) \leq \alpha(0,i) \text{, } i < \frac{1}{2} (N+1)$$

when $\Pr[x(i) < 0] \to 1$ as is the case for $i \to 1$ and large $N$. 
5. THE SINGLE POLYGENE MODEL

Restricting attention to the effects of a single polygenic locus in a diploid organism, some consequences of a single cycle of phenotypic selection in finite populations are derived in this chapter. The basis is derived for a treatment of repeated cycles of selection in Chapter 6.

5.1 Reference Population $\Pi_0$

At the polygenic locus of interest three diploid genotypes are generated by gene $a^+$ and its alleles $a^-$. The model for $\Pi_0$ adopted as a reference population for the $N$ randomly sampled test zygotes, assumes random mating without selection among parents of test material; $p$ denotes the frequency of gene $a^+$ in this parental group and $q = 1 - p$. Irrespective of the genotypic composition of the parental group, or its size, genotypic composition of $\Pi_0$ is defined initially to be Hardy-Weinberg equilibrium proportions with gene frequency $p$. Effects of allelic disequilibrium are accommodated in a later section. Sampled test material cannot be said to be in Hardy-Weinberg proportions except in the expectational sense.

Effects at the locus are assumed to influence trait $X$ the basis for phenotypic ranking in test material. Marginal trait $X$ means per genotype are shown in the following table.

<table>
<thead>
<tr>
<th>genotypes</th>
<th>frequencies</th>
<th>trait $X$ means</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(++)$</td>
<td>$p^2$</td>
<td>$+a$</td>
</tr>
<tr>
<td>$(+-)$</td>
<td>$2pq$</td>
<td>$+d$</td>
</tr>
<tr>
<td>$(--)$</td>
<td>$q^2$</td>
<td>$-a$</td>
</tr>
</tbody>
</table>

Marginal genotypic mean for the locus is

$$\mu_X = (p - q) a + 2pqd$$  \hspace{1cm} (5.1.1)
and the well-known orthogonal partition of marginal genotypic variance for the locus, into its additive and dominance components, is

\[
\text{additive component: } \sigma^2_\alpha = 2pqA^2 \\
\text{dominance component: } \sigma^2_\delta = (2pqd)^2
\]

where \( A = [a - (p-q)d] \) is the allelic substitution effect (see e.g., Falconer, 1964). Without loss of generality it may be assumed that \( a \) is non-negative identifying "favorable" allele \( a^+ \) with frequency \( p \). However, dominance contrast \( d \) may be positive, zero or negative; in writing the genotypic standard deviation \( \sigma_\delta \), the convention is adopted that \( \sigma_\delta \) is positive or negative according as \( d \) is positive or negative.

The phenotypically observed \( X \)-value for any zygote sampled from \( \Pi_0 \) is assumed to be the sum of the genotypic mean corresponding to its genotype at the locus and a random Normal deviate with zero mean and variance \( \sigma^2 \), independently and identically distributed for each genotype. Thus the distribution of \( X \) in \( \Pi_0 \) has the compounded density function

\[
\frac{1}{\sigma} \left\{ p^2 f\left(\frac{X-a}{\sigma}\right) + 2pqf\left(\frac{X-d}{\sigma}\right) + q^2 f\left(\frac{X+a}{\sigma}\right) \right\}, \quad (5.1.3)
\]

where \( f(\cdot) \) denotes the unit Normal density function. It is convenient to denote genotype mean deviations by

\[
g_2 = \frac{(a-\mu_x)}{\sigma}, \quad g_1 = \frac{(d-\mu_x)}{\sigma}, \quad g_0 = \frac{(-a-\mu_x)}{\sigma}
\]

so that

\[
p^2 g_2 + 2pqg_1 + q^2 g_0 = 0 \quad (5.1.4)
\]

and
\[ p^2 g_2^2 + 2pq g_1^2 + q^2 g_0^2 = \frac{(\sigma^2 + \sigma^2_\delta)}{\sigma^2} \] \hspace{1cm} (5.1.5)

The phenotypic mean of \( X \) in \( \Pi_0 \) is the same as the genetic mean, \( \mu_X \), and phenotypic variance is compounded as \( \sigma_X^2 = (\sigma^2 + \sigma^2_\delta) + \sigma^2 \).

This model, incorporating a common phenotypic distribution per genotype, is the same as that adopted by Kojima (1961), Curnow and Baker (1968), Hill (1969a, 1969b) and Pike (1969). Curnow and Baker (1968) use results obtained by Kojima (1961) for the polygenic model developed below. Hill (1969a) treats problems arising with order statistics in random samples from the distribution (5.1.3) numerically using quadrature; not all of his results are restricted to the polygenic case.

Each of the three densities compounded in (5.1.3) may be expanded in a Taylor series to provide

\[
f\left(\frac{X-a}{\sigma}\right) = f\left(\frac{X-\mu}{\sigma}\right) \left[ 1 + g_2\left(\frac{X-\mu}{\sigma}\right) \right] + \text{terms of order } (g_2)^2, \hspace{1cm} (5.1.6)
\]

\[
f\left(\frac{X-d}{\sigma}\right) = f\left(\frac{X-\mu}{\sigma}\right) \left[ 1 + g_1\left(\frac{X-\mu}{\sigma}\right) \right] + \text{terms of order } (g_1)^2,
\]

and

\[
f\left(\frac{X+a}{\sigma}\right) = f\left(\frac{X-\mu}{\sigma}\right) \left[ 1 + g_0\left(\frac{X-\mu}{\sigma}\right) \right] + \text{terms of order } (g_0)^2.
\]

Henceforth, the phrases "terms of order \((g_2)^2\), or \((g_1)^2\), or \((g_0)^2\)" will be collectively abbreviated to "\(0(g)^2\)." Substitution of expansions (5.1.6) into (5.1.3) yields

\[
\frac{1}{\sigma} f\left(\frac{X-\mu}{\sigma}\right) \left[ 1 + (p^2 g_2 + 2pq g_1 + q^2 g_0) \left(\frac{X-\mu}{\sigma}\right) \right] + 0(g)^2
\]

\[
= \frac{1}{\sigma} f\left(\frac{X-\mu}{\sigma}\right) + 0(g)^2, \hspace{1cm} \text{by virtue of } (5.1.4) \hspace{1cm} (5.1.7)
\]

as the density function of \( X \) in \( \Pi_0 \).
It is now assumed that relative to the phenotypic standard deviation per genotype (σ), effects a and d at the locus are sufficiently small to permit neglect of terms 0(g)^2; accordingly terms 0(g)^2 are generally omitted in the following developments. This assumption may be regarded as an operational definition of "polygenic" effects. The assumption is not uncommon; in particular it is the basis of formulating selection values for genotypes as proportional to their marginal means in analysis of selection in infinite test populations (Haldane, 1931; Griffing, 1960; Latter, 1965a). Adoption of this definition implies \( \sigma^2_x \approx \sigma^2 \) and that the distribution of \( X \) in \( \Pi_0 \) has standardized density function at (5.1.7) as

\[
f\left(\frac{X - \mu_x}{\sigma_x}\right) = f(x), \text{ the unit Normal density,}
\]

where \( x = (X - \mu_x) / \sigma_x \) is the standardization introduced in Section 2.1. Under these conditions, properties of selection differentials for selection in finite material randomly sampled from \( \Pi_0 \), have already been described in Chapter 2; required moments \( \mu_i|N \) and \( \sigma_{i,j}|N \) of order statistics are those appropriate for \( f(x) \), and in the case of directional selection approximations \( \hat{k} \) and \( \hat{v} \) at (2.4.2) and (2.6.1) respectively are available.

Applying the polygenic assumption, expansions at (5.1.6) may be curtailed to provide density functions of the three genotypic distributions as

\[
f\left(\frac{X - a}{\sigma}\right) = f(x) \left[1 + xg_2\right] \text{ for } (++) \text{ genotypes ,}
\]

\[
f\left(\frac{X - d}{\sigma}\right) = f(x) \left[1 + xg_1\right] \text{ for } (+-) \text{ genotypes ,}
\]

and
\[ f\left(\frac{x + a}{b}\right) = f(x) \left[1 + xg_0\right] \] for (--) genotypes.

Thus far nothing has been said about response variable \( Y \) in \( \Pi_0 \).

Consequences of selection are completely characterized by specifying \( Y \) as the (1 x 3) vector of genotypic indicator random variables \((Y_2, Y_1, Y_0)\) defined as

\[
Y_2 = 1 \text{ if the random zygote from } \Pi_0 \text{ has genotype } (++) ,
\]

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

\[
Y_1 = 1 \text{ if the random zygote from } \Pi_0 \text{ has genotype } (+-)
\]

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

\[
Y_0 = 1 \text{ if the random zygote from } \Pi_0 \text{ has genotype } (--) ,
\]

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

\[
1 = Y_2 + Y_1 + Y_0 .
\]

Different aspects of response to selection may then be defined as functions of \( Y \); for example, gene \( a^+ \) frequency response is characterized by

\[
Z = Y_2 + \frac{1}{2} Y_1 .
\]

As stressed previously in Chapters 3 and 4, a complete definition of response must include specification of post-selection utilization of selected test units which are zygotes in the current context. For gene frequency and trait mean response functions of \((Y_2, Y_1, Y_0)\) a useful specification of this type will be the large Hardy-Weinberg population obtained by random mating selected zygotes and, therefore, parameterized analogously with \( \Pi_0 \) but most likely with a different gene frequency than \( p \) and in any case with a gene frequency that is a random variable. Thus, it is convenient to write \( \Pi_0 = \Pi_0(p) \) as the reference population for
sampling of test material and \( \Pi_r = \Pi_0(p + \delta p) \) as the reference population for some response functions of \((Y_2, Y_1, Y_0)\). The genotypic mean of \( \Pi_r \) is given by (5.1.1) as

\[
(\mu_x + \delta \mu_x) = ((p + \delta p) - (q - \delta p)) a + 2(p + \delta p)(q - \delta p) d
\]

or

\[
\delta \mu_x = 2A(\delta p) - 2d(\delta p)^2 . \tag{5.1.9}
\]

Consider a random zygote from \( \Pi_0(p) \). Random variable \((Y_2, Y_1, Y_0)\) has the unit multinomial distribution:

\[
\Pr[Y_2, Y_1, Y_0] = (p^2)^2 (2pq)^1 (q^2)^0 ; (Y_2 + Y_1 + Y_0) = 1
\]

so that gene frequency indicator \( Z \) has expectation and variance as

\[
\mu_z = E(Z) = p , \quad \sigma_z^2 = V(Z) = \frac{1}{2} pq . \tag{5.1.10}
\]

The joint density function of \( X \) and \( Y_j \) in \( \Pi_0(p) \) is obtained using (5.1.8) and conditional densities for \( Y_j \), given \( X \), \( j=0,1,2 \), are then

\[
\lambda_j(Y_j = 1|X) = \theta_j[1 + x g_j] , \quad (Y_2 + Y_1 + Y_0) = 1 , \tag{5.1.11}
\]

where \( \theta_2 = p^2 , \theta_1 = 2pq , \theta_0 = q^2 \). Despite the neglect of terms \( O(g)^2 \), (5.1.11) defines a proper three-point density function inasmuch as

\[
\lambda_2(Y_2 = 1|X) + \lambda_1(Y_1 = 1|X) + \lambda_0(Y_0 = 1|X) = 1 .
\]

The joint density of \( Z \) and \( X \) in \( \Pi_0(p) \) is contained in three planes defined by (5.1.8) and provides the conditional density of \( Z \), given \( X \), as

\[
\pi(Z = \frac{1}{2}, j|X) = \lambda_j(Y_j = 1|X) , \quad j=2,1,0 . \tag{5.1.12}
\]

In analysis of response, using the results of Chapters 2 and 3, conditional expectations and variances of response functions of
(Y_2, Y_1, Y_0), given X, are required for post-test projections. These will be developed as needed.

5.2 Selective Value

The consequences of any particular style of selection, in relation to its objective, depend on the relationship between genotype of a test zygote and its phenotypic rank in test material. To each rank position there corresponds a selective value with respect to the objective of selection and such selective value must be some function of the genotypic indicator random variables (Y_2, Y_1, Y_0).

Post-test projections of genotype for the zygote ranked \( r \)th on the basis \( X_{(r)} \) follow directly from (5.1.11):

\[
\begin{align*}
\Pr[(++)|X_{(r)}] &= \frac{1}{2} (Y_2 = 1|X_{(r)}) = p^2[1 + x_{(r)} g_2] \quad \text{(5.2.1)} \\
\Pr[(+-)|X_{(r)}] &= \frac{1}{2} (Y_1 = 1|X_{(r)}) = 2pq[1 + x_{(r)} g_1] \\
\text{and} \\
\Pr[(-+)|X_{(r)}] &= \frac{1}{2} (Y_0 = 1|X_{(r)}) = q^2[1 + x_{(r)} g_0].
\end{align*}
\]

Pre-test projections of genotype for the \( r \)th ranked test zygote are obtained by removing the conditioning in (5.2.1) as discussed in Chapter 3; this is achieved by simply replacing \( x_{(r)} \) in (5.2.1) by \( \mu_r|N \). Selective value of the \( r \)th ranked test zygote may, therefore, be discussed generally in terms of the probability statements at (5.2.1) depending on the pre- or post-test context. For example, in the case of directional selection with the objective of increasing the frequency of allele \( a^+ \) conferring "favorable" additive effects, \( \Pr[(++)|r] \) is of interest if \( r \in \mathbb{R} \), the rank set defining selected zygotes. Values of...
Pr[(++)\mid r] for each rank position in a test sample of size N = 8 are shown in Table 5.1 for p = 0.5, \((a/\sigma_x) = 0.1\) and four degrees of dominance.

Table 5.1. Values of pre-test \(\text{Pr}[(++)\mid r]\) in a test sample of eight zygotes

<table>
<thead>
<tr>
<th>Rank</th>
<th>(d = -a) (recessivity)</th>
<th>(d = 0) (additivity)</th>
<th>(d = a) (dominance)</th>
<th>(d = 3a) (overdominance)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.197</td>
<td>0.214</td>
<td>0.232</td>
<td>0.268</td>
</tr>
<tr>
<td>2</td>
<td>0.218</td>
<td>0.229</td>
<td>0.239</td>
<td>0.261</td>
</tr>
<tr>
<td>3</td>
<td>0.232</td>
<td>0.238</td>
<td>0.244</td>
<td>0.256</td>
</tr>
<tr>
<td>4</td>
<td>0.244</td>
<td>0.246</td>
<td>0.248</td>
<td>0.252</td>
</tr>
<tr>
<td>5</td>
<td>0.256</td>
<td>0.254</td>
<td>0.252</td>
<td>0.248</td>
</tr>
<tr>
<td>6</td>
<td>0.268</td>
<td>0.262</td>
<td>0.256</td>
<td>0.244</td>
</tr>
<tr>
<td>7</td>
<td>0.282</td>
<td>0.271</td>
<td>0.261</td>
<td>0.239</td>
</tr>
<tr>
<td>8</td>
<td>0.303</td>
<td>0.286</td>
<td>0.268</td>
<td>0.232</td>
</tr>
</tbody>
</table>

The high degree of overdominance (\(d = 3a\)) included in this table was chosen because at p = 0.5 it has the effect of completely reversing listed probabilities in the adjacent column corresponding to complete dominance (\(d = a\)). The effect of positive dominance deviations (\(d > 0\)) in interfering with identification of \(+\) alleles, such as (++) combinations, by phenotypic selection is well known. In studies of repeated cycles of directional phenotypic selection with a constant number of parental zygotes per cycle, this effect is responsible for a measurable retardation in the rate of approach to fixation of \(+\) alleles (Robertson, 1962; Carr and Nassar, 1970b). As this type of effect will be encountered again in the analysis of trait mean response to selection, and again in relation to epistatic deviations, it is worthy of a convenient label; the phrase chosen here is "identification interference effect."
5.3 Gene Frequency Response

Conditional expectation of gene $a^+$ frequency indicator $Z$, given $X$, under the $\Pi_0(p)$ model specification, is found from (5.1.12) and (5.1.11) as

$$E[Z|X] = \pi(Z = 1|X) + \frac{1}{2} \pi(Z = \frac{1}{2}|X),$$

$$= p^2[1 + xg_2] + pq[1 + xg_1],$$

$$= p + xpq \left( \frac{A}{\sigma_x} \right). \tag{5.3.1}$$

Applying the standardization $z = (Z - \mu_z)/\sigma_z$, with $\mu_z = p$ and $\sigma_z^2 = \frac{1}{2} pq$ as at (5.1.10), yields

$$E[z|X] = \sqrt{2pq} \left( \frac{A}{\sigma_x} \right) x = \left( \frac{\sigma_A}{\sigma_x} \right) x. \tag{5.3.2}$$

Comparing this result with the formulation of linear response to selection at (3.1.1), (5.3.2) may be rewritten as

$$E[z|X] = \rho x, \quad \rho^2 = \frac{(2pqA^2)}{\sigma_x^2} = \frac{\sigma_A^2}{\sigma_x^2}. \tag{5.3.3}$$

for $\rho$ as used throughout Chapter 3, and $\sigma^2$ the component of genotypic variance attributable to additive effects, defined at (5.1.2). In the wider context of total genetic variance, rather than that of a single locus as here, the ratio of additive component to phenotypic variance is referred to as narrow sense heritability (Lush, 1945). Thus, $\rho^2$ as defined here might be referred to as narrow sense heritability contribution per locus; $\rho^2$ is negligible being $O(g)^2$. 


Conditional variance of $Z$, given $X$, is found from (5.1.12) and (5.1.11) as follows:

$$E[Z^2 | X] = \pi(Z = 1 | X) + \frac{1}{4} \pi(Z = \frac{1}{2} | X)$$

$$= E[Z | X] - \frac{1}{4} \pi(Z = \frac{1}{2} | X)$$

so that

$$V[Z | X] = E[Z^2 | X] - E^2[Z | X]$$

$$= \frac{1}{2} pq - \frac{1}{2} pq[(p-q)(\frac{A}{\sigma_x}) + 2pq(\frac{d}{\sigma_x})] x . \quad (5.3.4)$$

Again standardizing $Z$ to $z$, and identifying gentic variance components,

$$V[z | X] = 1 - \left[ (\frac{p-q}{\sqrt{2pq}})(\frac{\alpha}{\sigma_x}) + (\frac{\Delta}{\sigma_x}) \right] x \quad (5.3.5)$$

and in comparison with (3.1.2) this indicates that gene frequency response satisfies the conditional variance assumption of Chapter 3.

Thus, as far as gene frequency response is concerned all the results of Chapter 3 dealing with linear response to selection apply directly. In particular, for any general set $R_n$ of ranks defining a group of selected zygotes, gene frequency in the selected group is "average response" in the terminology of Chapter 3:

$$(p + \delta p) = \frac{1}{n} \Sigma_{R_n} Z_r , \quad (5.3.6)$$

where $Z_r$ is the value of gene frequency indicator $Z$ assigned to test zygote ranked $r$th on the basis of phenotype $X_{(r)}$; $\delta p$ of course is a random variable. For $\overline{x}_n$, $\Delta_n$, $\delta_n$, $k$ and $v$ defined for set $R_n$ as in Chapter 2, post- and pre-test projections of

$$\delta p = \sqrt{\frac{1}{2}pq} \cdot \frac{1}{n} \Sigma_{R_n} z_r = \sqrt{\frac{1}{2}pq} \cdot \delta_n$$
are provided by moments at (3.4.2), (3.4.3), (3.5.3) and (3.5.5), the complete summary being:

\[ E\{\delta p \mid \bar{X}_n \} = \rho A_n \sqrt{\frac{1}{2}pq} = pq\left(\frac{1}{\sigma_x^2}\right) \Delta_n \]  

(5.3.7)

and

\[ V\{\delta p \mid \bar{X}_n \} = \frac{pq}{2n} \left[ 1 - \Delta_n \left( (p-q)\left(\frac{A}{\sigma_x^2}\right) + 2pq\left(\frac{d}{\sigma_x^2}\right) \right) \right] \]

for post-test moments, and

\[ E\{\delta p\} = \rho k \sqrt{\frac{1}{2}pq} = kpq\left(\frac{A}{\sigma_x^2}\right), \]  

(5.3.8)

\[ V\{\delta p\} = \frac{pq}{2n} \left[ 1 - k\left( (p-q)\left(\frac{A}{\sigma_x^2}\right) + 2pq\left(\frac{d}{\sigma_x^2}\right) \right) + \rho^2 v \right] \]

\[ = \frac{pq}{2n} \left[ 1 - k\left( \frac{(p-q)}{\sqrt{2pq}}\rho + \frac{\sigma^2}{\sigma_x^2} \right) \right], \text{since } \rho \text{ is } O(g)^2, \]

for pre-test moments. Gene frequency \((p + \delta p)\) at (5.3.6) is that in selected zygotes and also that in reference population \(\Pi_r = \Pi_0 (p + \delta p)\) of Section 5.1; thus, these pre- and post-test projections apply to \(\Pi_r\) also.

Results (5.3.8) are algebraically equivalent to those given by Kojima (1961) for directional phenotypic selection; his \(q\) is \(p\), his \(\alpha\) is substitution effect \(A\), his \(\beta\) is the dominance contrast equivalent to \(-2d\) and he assumes \(\sigma_x^2 = 1\). Although algebraically identical, the results differ in derivation, and more importantly, in generality. As derived here they apply to any set \(R_n\) defining a group of selected zygotes including as a special case the set of \(n\) top-ranking zygotes as considered by Kojima.
5.4 Variance Effective Number \( n_v \)

In addition to response to selection other consequences associated with the combination of selection and sampling effects, referred to as "drift" by Wright (1931), may be parameterized in terms of \( \Pi_r \). One such parameter will be considered in detail—variance effective number \( n_v \). For the case of directional phenotypic selection Kojima (1961) described a region defined by \( p \) and \( d/a \) where variance of gene frequency change was less than in the case of no selection. Here this aspect is considered for general \( R_n \) in terms of a variance effective number \( n_v \) defined from the pre-test variance of \( \delta p \):

\[
V(\delta p) = \frac{pq}{2n_v}.
\]  

(5.4.1)

Then in general, the ratio of census number to variance effective number of parents of \( \Pi_r \) is given by (5.3.8) as

\[
\frac{n}{n_v} = 1 - k[(p-q)\left(\frac{A}{x}\right) + 2pq\left(\frac{d}{x}\right)].
\]  

(5.4.2)

Rewriting (5.4.2) as

\[
\frac{n}{n_v} = 1 - \left(\frac{ka}{x}\right) [(p-q) + (6pq - 1) \frac{d}{a}]
\]

yields the two values of gene frequency in \( \Pi_0(p) \) for which \( n_v/n \) is not a function of the degree of dominance, as solutions to \( 6pq = 1 \); namely,

\[
p_1 = \frac{1}{2}(1 - \frac{1}{\sqrt{3}}) \approx 0.211 \quad \text{and} \quad p_2 = \frac{1}{2}(1 + \frac{1}{\sqrt{3}}) \approx 0.789
\]

with

\[
\frac{n}{n_v} = \{1 + (0.57735)(\frac{ka}{x})\}, \quad \frac{n}{n_v} = \{1 - (0.57735)(\frac{ka}{x})\}
\]
respectively. Figure 5.1 shows ratio $n_v/n$ as a function of $p$ for $(ka/\sigma_x) = 0.1$ and four degrees of dominance defined as $d/a$. When $d = 0$ (additivity), $n_v/n$ is almost linear in $p$ and $n_v = n$ only at $p = 0.5$. The general effect of positive dominance deviations ($d > 0$) is to produce a variance effective number greater than the census number of parents over the intermediate range of gene frequency (approximately $0.3 < p < 0.8$).

5.5 Trait Mean Responses

Mean response for trait $X$, assessed in reference population

$\Pi_r = \Pi_0(p + \delta p)$, is defined at (5.1.9). Taking expectation, either pre- or post-test in nature, yields expected mean response as

$$E(\delta_{\epsilon x}) = 2\Pi E(\delta p) - 2d[E^2(\delta p) + V(\delta p)] \quad (5.5.1)$$

which can be evaluated using the appropriate pre- or post-test moments of $\delta p$ at (5.3.7) and (5.3.8).

Another specification of response variable $Y$ in $\Pi_0$ arises when there are pleiotropic effects at the locus; $Y$ becomes the trait subject to "correlated response" when selection is practiced on trait $X$. Analogous to the formulation of trait $X$ genotypic components in Section 5.1, marginal trait $Y$ means per genotype are written as follows:

<table>
<thead>
<tr>
<th>genotypes</th>
<th>trait $Y$ means</th>
</tr>
</thead>
<tbody>
<tr>
<td>(+•)</td>
<td>+c</td>
</tr>
<tr>
<td>(••)</td>
<td>+b</td>
</tr>
<tr>
<td>(−−)</td>
<td>−c</td>
</tr>
</tbody>
</table>

with genotypic mean for the locus analogous to (5.1.1):

$$\nu_y = (p-q)c + 2pqb$$
Figure 5.1. Ratio \( \left( \frac{n_y}{n} \right) \) as a function of \( p \) for four degrees of dominance.
and orthogonal partition of genotypic variance analogous to (5.1.2):

additive component: \( \sigma^2_Y = 2pqC^2 \)

dominance component: \( \sigma^2_\beta = (2pqb)^2 \)

where \( C = [c - (p-q)b] \) is the allelic substitution effect. Unlike parameter \( a \) for trait \( X \), parameter \( c \) for trait \( Y \) may be negative or positive since gene \( a^+ \) does not necessarily confer "favorable" additive effects on trait \( Y \). Again the convention is adopted that genotypic standard deviations \( \sigma_Y \) and \( \sigma_\beta \) have the sign of their corresponding effects. Marginal genotypic covariation between traits \( X \) and \( Y \), entirely attributable to pleiotropic effects at the locus, may also be orthogonally partitioned into additive and dominance components:

\[
\sigma_{\alpha Y} = 2pqAC = \sigma_\alpha \sigma_Y
\]

and

\[
\sigma_{\delta \beta} = (2pqd)(2pqb) = \sigma_\delta \sigma_\beta .
\]

Mean correlated response is analogous to (5.1.9)

\[
\delta u_Y = 2C(\delta p) - 2b(\delta p)^2
\]  
(5.5.3)

with pre- or post-test expectation analogous to (5.5.1):

\[
E\{\delta u_Y\} = 2CE\{\delta p\} - 2b[E^2\{\delta p\} + \sigma^2]\ .
\]  
(5.5.4)

5.6 Consequences of No Selection

For the case of no selection defined by \( R_n = \{1, 2, \ldots, N\} \), \( n = N \) and \( k = 0 \). Substitution in (5.3.7) yields post-test expectation

\[
E\{\delta p|\bar{X}_N\} = pq(\frac{A}{\sigma_X}) \Delta_N
\]
which is not zero unless \( \Delta_N = 0 \), whereas substitution in (5.3.8) yields the pre-test moments

\[
E(\delta p) = 0
\]

(5.6.1)

and

\[
V(\delta p) = \frac{pq}{2n}.
\]

This Binomial variance is the basis for defining variance effective number \( n_v \) as at (5.4.1).

Hardy-Weinberg proportions of reference population \( \Pi_r = \Pi_0(p + \delta p) \) are

\[
(p + \delta p)^2: 2(p + \delta p)(q - \delta p): (q - \delta p)^2
\]

with pre-test expectations given by (5.6.1) as

\[
p + \frac{pq}{2n}; 2pq(1 - \frac{1}{2n}); q^2 + \frac{pq}{2n}.
\]

\( \Pi_r \) is thus inbred with inbreeding coefficient \( F = 1/2n \). Substitution in (5.5.1) and (5.5.4) yields expected trait mean responses

\[
E(\delta u_x) = -\frac{1}{2n} 2pqd = -F \sigma_\delta
\]

(5.6.2)

and

\[
E(\delta u_y) = -\frac{1}{2n} 2pqb = -F \sigma_\beta
\]

for pre-test projections. When \( d \) and \( b \) are positive these negative expected changes in trait means are referred to as "inbreeding depressions."

5.7 Consequences of Directional Selection

When the \( n \) top-ranking zygotes are selected, \( R_n = \{N-n+1,N-n+2,...,N\} \), n < N, with standardized selection differential \( k \) at (2.2.1) and
approximation $k$ at (2.4.2). Post- and pre-test moments of gene frequency response are given by direct substitution in (5.3.7) and (5.3.8).

Trait mean responses in relation to \( \Pi_r = \Pi_0(p + \delta p) \) are obtained by substitution in (5.5.1) and (5.5.4). Those for pre-test projections are

\[
E(\delta_{\mu x}) = k \, 2pq \, \frac{A^2}{\sigma_x} - k \, 2pq \left( \frac{A}{\sigma_x} \right)^2 \, pqd - 2pqd \, \frac{1}{2n_v}
\]

\[
= \left\{ k \left( \frac{\sigma}{\sigma_x} \right)^2 - k \left( \frac{\sigma}{\sigma_x} \right) \frac{1}{2} \, k \left( \frac{\sigma}{\sigma_x} \right) \frac{1}{2n_v} \right\} \, \sigma_x
\]

(5.7.1)

and

\[
E(\delta_{\mu y}) = k \, 2pq \, \frac{AC}{\sigma_x} - k \, 2pq \left( \frac{A}{\sigma_x} \right)^2 \, pqb - 2pqb \, \frac{1}{2n_v}
\]

\[
= \left\{ k \left( \frac{\sigma}{\sigma_x} \right)^2 - k \left( \frac{\sigma}{\sigma_x} \right) \frac{1}{2} \, k \left( \frac{\sigma}{\sigma_y} \right) \frac{1}{2n_v} \right\} \, \sigma_y.
\]

(5.7.2)

Trait X mean response at (5.7.1) was given by Kojima (1961) in a slightly different representation—he did not introduce the variance of effective number \( n_v \). The first term, \( k \sigma^2 \sigma_x = k \rho \sigma_x \), is the usual projection of response to directional phenotypic selection in relation to offspring population \( \Pi_r \). The second term of (5.7.1) is a correction for identification interference effects of dominance as discussed in Section 5.2. The first two terms together may be written

\[
k \rho \sigma_x \left[ 1 - \frac{1}{2} \, k \left( \frac{\sigma}{\sigma_x} \right) \right]
\]

which is a function of \( n \) and \( N \) only through \( k \). The third term of (5.7.1) is a further correction to the usual projected response accounting for drift effects. In comparison with the inbreeding depression, \(- (\sigma_\delta / 2n)\) at (5.6.2), the representation \(- (\sigma_\delta / 2n_v)\) for this third term conveniently identifies its nature.
Correlated trait mean response at (5.7.2) has a similar form, the first term being the usual projection of correlated responses to directional phenotypic selection in relation to offspring population \( \Pi_r \), often written (see e.g., Falconer, 1964)

\[
kp_r \sigma_{\alpha \gamma} = k \frac{\sigma_{\alpha \gamma}}{\sigma_x \sigma_y} \sigma_y,
\]

where \( r_A \) is the correlation between additive gene effects on traits \( X \) and \( Y \). In the current context \( r_A = \pm 1 \) depending on the sign of \( c \neq 0 \), which is why covariance \( \sigma_{\alpha \gamma} \) factors into \( \sigma_{\alpha} \sigma_{\gamma} \) at (5.5.2). The second term of (5.7.2) is the correction for dominance effects, this time indirectly through perfect correlation between dominance effects on the two traits.

5.8 Consequences of Divergent Selection

When the \( n_h \) top-ranking zygotes, and separately the \( n_\ell \) bottom-ranking zygotes, are selected from the same sample of \( N \) test zygotes with \((n_h + n_\ell) \leq N\), selection is formally defined by

\[
R_{n_\ell} \equiv \{1, 2, \ldots, n_\ell\}
\]

\[
R_{n_h} \equiv \{N-n_h+1, N-n_h+2, \ldots, N\}
\]

\[
R_{n_\ell} \cap R_{n_h} \text{ is null},
\]

with corresponding standardized selection differentials parameterized by \( k_\ell, k_h \) and \( v_\ell, v_h \) and \( w \) all defined in Section 2.1. Reference populations for response are defined for "high" (signified by \( h \)) and "low" (signified by \( \ell \)) groups separately and identically as \( \Pi_h = \Pi_0 (p + \delta p(h)) \) and \( \Pi_\ell = \Pi_0(p + \delta p(\ell)) \) respectively. Expectations and variances of gene
frequency changes for high and low groups are those given at (5.3.7) and (5.3.8) with \( k \) replaced by \( k_h \) and \( k_\ell \) respectively. Similar replacements are required for variance effective numbers \( n_{vh} \) and \( n_{v\ell} \) respectively at (5.4.2), and for expected trait mean responses at (5.5.1) and (5.5.4).

Additionally, linearity of gene frequency response together with \( R \cap R \cap R \) null provides covariances between gene frequency changes via \( n_{h} \) and \( n_{\ell} \) (5.6.2) and (3.6.3):

\[
C[\delta p(h) , \delta p(\ell)| n_h \bar{X}_h \bar{X}_\ell] = 0
\]

(5.8.1)

\[
C[\delta p(h) , \delta p(\ell)] = \frac{2pqP^2}{\sqrt{n_h n_\ell}} \approx 0
\]

since \( P^2 \) is \( O(g)^2 \).

Divergence of trait mean responses as assessed in \( \Pi_h \) and \( \Pi_\ell \) is defined as

\[
D(\mu) = E\{\delta \mu(h)\} - E\{\delta \mu(\ell)\},
\]

(5.8.2)

where \( \mu = \mu_X \) for trait \( X \), \( \mu = \mu_Y \) for correled response trait \( Y \) and expectations are pre- or post-test as required. In the same notation, asymmetry of mean responses as assessed in \( \Pi_h \) and \( \Pi_\ell \) is defined as

\[
S(\mu) = E\{\delta \mu(h)\} + E\{\delta \mu(\ell)\}.
\]

(5.8.3)

Thus \( S(\mu) \) is zero if trait mean response is symmetric,

\[
E\{\delta \mu(\ell)\} = -E\{\delta \mu(h)\}; S(\mu) \) is negative if \( E\{\delta \mu(h)\} < -E\{\delta \mu(\ell)\} \) and positive if \( E\{\delta \mu(h)\} > -E\{\delta \mu(\ell)\} \).

Pre-test projection of \( D(\mu_X) \) is provided by (5.7.1) with \( k \) replaced first by \( k_h \) and then by \( k_\ell \):
\[ D(\mu_x) = (k_h - k_x) \rho \sigma_x \left[ 1 - \frac{1}{2} (k_h + k_x) \frac{\sigma_\delta}{\sigma_x} \right] - \frac{1}{2} \sigma_\delta \left( \frac{1}{n_{vh}} - \frac{1}{n_{v\lambda}} \right) \]  

(5.8.4)

the first term of which appropriately compounds expected mean responses with corrections for identification interference effects of dominance and the second term accounts for drift effects on divergence. Using (5.4.2) the factor \{ \} in (5.8.4) is evaluated as

\[ \left\{ \frac{1}{n_{vh}} - \frac{1}{n_{v\lambda}} \right\} = \left\{ \frac{1}{n_h} - \frac{1}{n_{v\lambda}} \right\} - \left\{ \frac{k_h}{n_h} - \frac{k_x}{n_{v\lambda}} \right\} \left( \frac{\rho (p-q)}{\sqrt{2pq}} + \frac{\sigma_\delta}{\sigma_x} \right). \]  

(5.8.5)

If \( n_h = n_{v\lambda} = n \), as is usually the case in planned selection experiments involving high and low lines, then as noted in Section 2.8, \( k_{v\lambda} = -k_h \) (= \(-k \) say) and

\[ D(\mu_x) = 2k \rho \sigma_x + \sigma_x \left( \frac{1}{n} - \frac{1}{n_v} \right), \]  

(5.8.6)

where \( n_v \) was defined at (5.4.2). Thus in this special case of symmetric high and low selection, divergence is not affected by interference identification effects of dominance but may be enhanced by drift effects at intermediate gene frequencies if \( d > 0 \), for then \( n_v > n \) as shown in Figure 5.1.

Pre-test projection of \( S(\mu_x) \) is similarly evaluated:

\[ S(\mu_x) = (k_h + k_x) \rho \sigma_x \left[ 1 - \frac{1}{2} (k_h + k_x) \frac{\sigma_\delta}{\sigma_x} \right] + k_h k_{v\lambda} \rho \sigma_x \frac{\sigma_\delta}{\sigma_x} \]

\[ - \frac{1}{2} \sigma_\delta \left( \frac{1}{n_{vh}} + \frac{1}{n_{v\lambda}} \right) \]  

(5.8.7)

and using (5.4.2),

\[ \left\{ \frac{1}{n_{vh}} - \frac{1}{n_{v\lambda}} \right\} = \left\{ \frac{1}{n_h} - \frac{1}{n_{v\lambda}} \right\} - \left\{ \frac{k_h}{n_h} - \frac{k_x}{n_{v\lambda}} \right\} \left( \frac{\rho (p-q)}{\sqrt{2pq}} + \frac{\sigma_\delta}{\sigma_x} \right). \]
The first term of $S(\mu_x)$ is the contribution to asymmetry of trait X mean response attributable to asymmetry of selection ($k_h \neq -k_l$) with correction for identification interference effects of dominance. The latter effects also contribute to asymmetry of mean response irrespective of symmetry of selection, as indicated by the second term of $S(\mu_x)$. The third term is a contribution from drift effects. In the special case of symmetric divergent selection, $n_h = n_l = n$ and $k_h = -k_l = -k$, (5.8.7) reduces to

$$S(\mu_x) = -k^2 \rho_\alpha \frac{\sigma_\delta}{\sigma_x} - \frac{\sigma_\delta}{n}.$$  

(5.8.8)

For a given number of test zygotes, N, asymmetry of trait mean response increases with increasing selection pressure because both $k$ and $1/n$ increase. Only in the case of no dominance, $d = 0$, is mean response expected to be symmetrical.

Similar results are obtained for pre-test projection of correlated trait mean response; only those for symmetric selection are given here:

$$D(\mu_y) = 2k\rho_\sigma_\gamma + \sigma_\beta \left( \frac{1}{n} - \frac{1}{n_y} \right)$$

(5.8.9)

$$S(\mu_y) = -k^2 \rho_\alpha \left( \frac{\sigma_\beta}{\sigma_x} - \frac{\sigma_\beta}{n} \right).$$

Results (5.8.4) through (5.8.9) are applicable to selection of high and low groups from the same test material as might be the case at the initiation of a planned selection experiment. In subsequent recurrent cycles of divergent selection high and low lines would then be maintained in reproductive isolation. The consequences of one such subsequent cycle of selection follow directly from results established in the previous section for directional phenotypic selection. Suppose that in some
advanced generation, parental groups of high and low lines have attained
gene frequencies $p_h$ and $p_\ell$ respectively. Then for the next cycle of
selection the reference populations are $\Pi_0(p_h)$ and $\Pi_0(p_\ell)$ respectively
for test material and $\Pi_r = \Pi_0(p_h + \delta p_h)$ and $\Pi_r = \Pi_0(p_\ell + \delta p_\ell)$ respectively
for evaluation of responses.

Existing divergence in trait $X$ mean response, between $\Pi_0(p_h)$ and
$\Pi_0(p_\ell)$ is given by (5.1.1) as

$$(p_h - p_\ell)(A_h + A_\ell),$$

where $A_h = [a - (p_h - q_h)d]$ and $A_\ell = [a - (p_\ell - q_\ell)d]$. Then the increment
in divergence, due to this cycle of selection, is given by repeated use of
(5.5.1):

$$E\{\delta u_X^{(h)} \} - E\{\delta u_X^{(\ell)} \}$$

$$= 2[A_h E\{\delta p_h \} - A_\ell E\{\delta p_\ell \} ]$$

$$- 2d[E^2\{\delta p_h \} - E^2\{\delta p_\ell \} + V\{\delta p_h \} - V\{\delta p_\ell \}].$$

Similarly, the increment to asymmetry is

$$E\{\delta u_X^{(h)} \} + E\{\delta u_X^{(\ell)} \}.$$  \hspace{1cm} (5.8.10)

The feature to be stressed here is that asymmetry of trait $X$ mean
response does not require the presence of dominance deviations. For
example, in the case of symmetric selection ($k_h = -k_\ell$) and $d = 0$, (5.8.10)
reduces to

$$[k 2p_h q_h \left( \frac{a}{\sigma_X} \right)^2 - k 2p_\ell q_\ell \left( \frac{a}{\sigma_X} \right)^2] \sigma_X$$

$$= -2k \sigma \left( \frac{a}{\sigma_X} \right)^2 (p_h - p_\ell)[(p_h - \frac{1}{2}) - (\frac{1}{2} - p_\ell)].$$
and the increment to asymmetry is nonzero unless \( p_h = p_l \) or \( p_h \) and \( p_l \) are symmetrically displaced from 1/2.

5.9 Additional Consequences for the Additive Model

In the case of no dominance \((d = 0)\) marginal genotypic variance for trait \( X \) in \( \Pi_0(p) \) reduces to

\[
\sigma^2 = 2pq a^2 = \rho^2 \sigma^2_x
\]

and in \( \Pi_r = \Pi_0(p + \delta p) \), to

\[
(\sigma^2 + \delta \sigma^2) = 2(p + \delta p)(q - \delta p) a^2
\]

\[
= \sigma^2 - 2\sigma^2 [((p-q)(\delta p) + (\delta p)^2].
\]

Thus for any general rank set \( R_n \), pre-test projection of changed genetic variance, in relation to \( \Pi_r \), is

\[
E(\sigma^2 + \delta \sigma^2) = \sigma^2 \left[ (\frac{n - 1}{n_v}) - \frac{1}{2} k^2 \rho^2 \right]
\]

on substituting for \( E(\delta p) \) and \( V(\delta p) \) from (5.3.8) and with \( n_v \) defined at (5.4.2) reducing to

\[
n_v = n \left[ 1 - k(p-q)(\frac{\sigma}{\sigma_x})^{-1} \right]
\]

with general \( k \) for \( R_n \). In the case of no selection, as in Section 5.6, \( k = 0 \) and \( n_v = n \), so that

\[
E(\sigma^2 + \delta \sigma^2) = \sigma^2 (1 - \frac{1}{2n})
\]

as would be predicted by drift theory alone. In the case of directional selection the expectation in (5.9.1) can be compared with the result obtained by Nei (1963) for truncation selection in large populations.
Identifying his (δ/σ) with k, his result may be written in the current notation as

$$E\{\delta \sigma^2_\alpha\} = -k(p-q)(\frac{a}{x}) \sigma^2_\alpha = \left(\frac{n-n_v}{n_v}\right) \sigma^2_\alpha,$$

whereas (5.9.1) yields

$$E\{\delta \sigma^2_\alpha\} = \left(\frac{n-n_v-\frac{1}{2}}{n_v}\right) \sigma^2_\alpha.$$

The difference, namely $-\sigma^2_\alpha/2n_v$, is the required adjustment for effects of drift in finite populations.

Other consequences that can be projected in the case of no dominance effects are variances of trait mean responses defined at (5.1.9) and (5.5.3) as

$$\delta \mu_x = 2a\delta p, \quad \delta \mu_y = 2c\delta p$$

with pre-test expectations $k\sigma_\alpha$ and $k\sigma_\gamma$ respectively and pre-test variances

$$V\{\delta \mu_x\} = \frac{\sigma^2_\alpha}{n_v}, \quad V\{\delta \mu_y\} = \frac{\sigma^2_\gamma}{n_v}$$

given by (5.4.1).

5.10 **Effects of Allelic Disequilibrium**

Returning to the case at the beginning of Section 5.8 where two separate groups of zygotes are selected from the same test sample, consider the large population of offspring obtained by random matings across selected groups. The genotypic composition of such an offspring population would not be expected to follow the Hardy-Weinberg equilibrium pattern because of different gene frequencies in parental groups. The
same situation arises in the case of a sexually dimorphic system when random matings are made between males of one group and females of another. Even if the gene frequencies of male and female groups of parents are expected to be the same, they may differ as random variables. This phenomenon necessitated the corrections described by Curnow and Baker (1969).

Under these circumstances it is necessary to redefine reference populations for test zygotes and response assessment. For $\Pi_0$ the genotypic composition is written in general as

$$p^2 + Upq, 2pa(1-U), q^2 + Upq$$  \hspace{1cm} (5.10.1)

for genotypes (++) , (+−), (−−) respectively; $U$ is the correlation between united gametes as defined by Wright (1922) and the quantity $Upq$ is referred to as allelic disequilibrium. $\Pi_0$ is now written as $\Pi_0(p,U)$.

When the cause of allelic disequilibrium is a difference between $p^m$ and $p^f$ the gene $a^+$ frequencies in male and female parental groups in a random mating system, the genotypic composition of $\Pi_0(p,U)$ is

$$p^mp^f, p^mq^f, q^mp^f, q^mq^f,$$

where $q^m = 1 - p^m$ and $q^f = 1 - p^f$. Then by identifying terms $\Pi_0(p,U)$ is specified by

$$p = \frac{1}{2} (p^m + p^f) = 1 - q, \hspace{0.5cm} U = -\frac{(p^m - p^f)^2}{4pq}.$$  \hspace{1cm} (5.10.2)

Mean $X$ and $Y$ trait values per genotype are parameterized by $(a,d)$ and $(c,b)$ as described in Section 5.1. Marginal genotypic means for the locus change to

$$\mu_x = (p-q) a + 2pq(1-U) d$$

$$\mu_y = (p-q) c + 2pq(1-U) b$$  \hspace{1cm} (5.10.3)
and marginal genotypic variances may be orthogonally partitioned into additive and dominance components using the formulation of Cockerham (1959):

\[
\begin{array}{ccc}
\text{variance component} & \text{trait X} & \text{trait Y} \\
\text{additive effects} & \sigma_a^2 = 2pq(1 + U) \ A^2 & \sigma_Y^2 = 2pq(1 + U) \ C^2 \\
\text{dominance effects} & \sigma_d^2 = (2pq \ U_d)^2 & \sigma_B^2 = (2pq \ U_b)^2 \\
\end{array}
\]

(5.10.4)

where

\[
A = a - \frac{(1-U)}{(1+U)} \ (p-q)d , \quad C = c - \frac{(1-U)}{(1+U)} \ (p-q)b
\]

are the substitution effects, and

\[
U_d^2 = \frac{(1-U)}{(1+U)} \ [(1-U)^2 + \frac{U}{pq}].
\]

Similarly the marginal genotypic covariance, attributable to pleiotropic effects, may be partitioned into corresponding additive and dominance components:

\[
\sigma_{\alpha \gamma} = 2pq(1 + U) \ AC = \sigma_a \sigma_Y
\]

and

\[
\sigma_{\delta \beta} = (2pq \ U_d)(2pq \ U_b) = \sigma_d \sigma_B.
\]

The assumption regarding nature of polygenic effects in Section 5.1 is unchanged; phenotypic distribution of trait X in \(\Pi_0(p,U)\) is considered Normal with mean \(\sigma_X\) and variance \(\sigma_X^2\). Derivations of gene frequency changes, and trait mean changes, follow the pattern of those in previous sections. For the sake of brevity complete algebraic details are omitted and attention is restricted to gene frequency changes.
For a sample of \( N \) test zygotes from \( \Pi_0(p,U) \), and a general rank set \( R_n \) defining \( n \) selected zygotes with standardized selection differential \( k \), gene frequency in the selected group is random variable \((p + \delta p)\) with

\[
E(\delta p) = kpq(1+U) \frac{A}{\sigma_x} = \rho k \sqrt{\frac{1}{2} pq(1+U)}
\]

\[
V(\delta p) = \frac{pq}{2n} \{ (1+U) - k[(1+3U)(p-q)\frac{\sigma^2}{\sigma_x} + 2pqU^2\frac{d}{\sigma_x}] + \rho^2 \}
\]

\[
= \frac{pq}{2n} \{ (1+U) - k[(1+3U) \frac{p-q}{\sqrt{2pq(1+U)}} + U \frac{\sigma^2}{\sigma_x}] \}
\]

where \( \rho^2 = (\sigma_\alpha / \sigma_x)^2 \) is \( 0(g)^2 \) and \( \frac{1}{2} pq(1+U) \) is \( \sigma_Z^2 \) the variance of gene frequency indicator \( Z \) in \( \Pi_0(p,U) \). When there is no allelic disequilibrium, \( U = 0 \), \( U_x = 1 \) and moments at (5.10.5) naturally reduce to those at (5.3.8).

### 5.11 Consequences of Selection in a Dioecious System

It is assumed that \( \Pi_0(p,U) \), with \( p \) and \( U \) as defined at (5.10.2), serves as the reference population for independent test samples of \( N_m \) males and \( N_f \) females. Then \( n_m \) male zygotes defined by rank set \( R_n \) and \( n_f \) female zygotes defined by rank set \( R_n \) are selected from those test samples with operative standardized selection differentials \( k_m \) and \( k_f \) respectively. Gene frequencies in the selected groups are denoted by \((p + \delta p_m)\) and \((p + \delta p_f)\) respectively with pre-test expectations and variances given by (5.10.5) on appropriate substitution of \( k_m \) or \( k_f \) for \( k \).

As a reference population for responses it is assumed that the \( n_m \) males and \( n_f \) females are random mated to form a large offspring population which will therefore have genotypic composition defined by gene frequency
(p + δp) = \frac{1}{2} (p + δp_m + p + δp_f) = p + \frac{1}{2} (δp_m + δp_f) \quad (5.11.1)

and allelic disequilibrium

(U + δU)(p + δp)(q - δp) = -\frac{1}{4} \{(p + δp_m) - (p + δp_f)\}^2

= -\frac{1}{4} (δp_m - δp_f)^2 \quad (5.11.2)

derived from (5.10.2). Thus, reference population Π_r is equivalent to Π_0(p + δp, U + δU). Independence of sampled test material and selection practices ensures that C{δp_m, δp_f} is zero.

Gene frequency response, as assessed in Π_r, is δp with pre-test moments compounded from (5.10.5):

\begin{equation}
\mathcal{E}(δp) = \overline{kvq(1 + U)}(\frac{A}{\sigma_x}) = \rho k \sqrt{\frac{1}{2} pq(1+U)}
\end{equation}

\begin{equation}
\mathcal{V}(δp) = \frac{pq}{2n_d} \{(1+U) - \overline{k}(1+3U) \frac{(p-q)}{\sqrt{2pq(1+U)}} \rho + U \frac{\overline{\sigma}}{\sigma_x}\}
\end{equation}

where

\begin{equation}
\overline{k} = \frac{1}{2} (k_m + k_f), \quad \overline{k} = \frac{(n_k k_m + n_f k_f)}{(n_m + n_f)}
\end{equation}

and \(n_d = n_m n_f / (n_m + n_f)\) is the variance effective number for a dioecious mixture of two Binomial samplings given by Wright (1931) corresponding to the case of no selection and zero allelic disequilibrium in Π_0, as can be readily verified by setting \(k_m = k_f = 0\) and \(U = 0\).

Defining variance effective number \(n_v\) from the pre-test \(\mathcal{V}(δp)\) as at (5.4.1) yields

\begin{equation}
\frac{n_d}{n_v} = (1 + U) - \overline{k}(1+3U) \frac{(p-q)}{\sqrt{2pq(1+U)}} \rho + U \frac{\overline{\sigma}}{\sigma_x}\) \quad (5.11.4)
\end{equation}

Allelic disequilibrium in Π_r, defined at (5.11.2), has pre-test expectation
\[
E\left(-\frac{1}{4} (\delta p_m - \delta p_f)^2\right) = -\frac{1}{4} \left[ E(\delta p_m) - E(\delta p_f) \right]^2 - V(\delta p)
\]

the first term of which is attributable to different selection differentials and is zero if \( k_m = k_f \), and is negligible if \( k_m \neq k_f \) being \( \rho^2 (k_m - k_f)^2 \frac{1}{2} pq(1+U) \) and therefore \( O(g)^2 \). The second term is attributable to a combination of selection and drift effects and can be written

\[
E\left(-\frac{1}{4} (\delta p_m - \delta p_f)^2\right) = -\frac{pq}{2n_v} .
\] (5.11.5)

Some special cases are of interest. When \( \Pi_0 \) is in Hardy-Weinberg equilibrium, setting \( U = 0 \) yields the following expectations for \( \Pi_r \):

\[
E(\delta p) = \rho k \sqrt{\frac{1}{2}} \frac{1}{pq}
\]

and

\[
V(\delta p) = \frac{pq}{2n_d} \left[ 1 - \tilde{k} \left( \frac{(p-q)}{\sqrt{2pq}} \rho + \frac{\sigma_\delta}{\sigma_x} \right) \right]
\]

together with

\[
\frac{n_d}{n_v} = 1 - \tilde{k} \left( \frac{(p-q)}{\sqrt{2pq}} \rho + \frac{\sigma_\delta}{\sigma_x} \right) .
\]

These are similar to the corresponding results obtained in Sections 5.3 and 5.4; \( \tilde{k} \) replaces \( k \) in (5.3.8), \( n_d \) replaces \( n \) and \( k \) replaces \( k \) in (5.4.2). Furthermore, since

\[
\tilde{k} - \frac{k}{k_m} = \frac{1}{2} - \frac{n_f}{n_m + n_f} + k_f \frac{1}{2} - \frac{n_m}{n_m + n_f}
\]

\( \tilde{k} \approx \frac{k}{k_m} \) unless \( n_m \) and \( n_f \) are very different. Thus in this case, the ratio \( n_v/n_d \) will be well approximated by \( n_v/n \) discussed in Section 5.4 and depicted in Figure 5.1.
For trait mean responses the comparison is between $\mu_x$ (or $\mu_y$) evaluated in $\Pi_r = \Pi_0(p + \delta p, U + \delta U)$ and in $\Pi_0(p, U)$. Pre-test expectations can be written in several forms; after some tedious algebra one of these for trait $X$ is

$$E(\delta u_x) = \bar{k}\rho\sigma_a [1 - \frac{k}{k + k_f} \frac{(1+U)}{U_x} \rho \frac{\sigma_x}{\sigma_y} - \frac{\sigma_x}{2n_v}]$$

(5.11.6)

which can be compared with (5.7.1). The first term contains the usual projection of response to directional phenotypic selection together with a correction for identification interference effects of dominance as affected by allelic disequilibrium in $\Pi_0(p, U)$. The corresponding expression for correlated trait response is

$$E(\delta u_y) = \bar{k}\rho\sigma_y [1 - \frac{k}{k + k_f} \frac{(1+U)}{U_y} \rho \frac{\sigma_y}{\sigma_x} - \frac{\sigma_y}{2n_v}]$$

(5.11.7)

In the special case of directional selection in both sexes, $\bar{k}$ and $\bar{k}$ can be approximated using approximation $\hat{k}$ at (2.4.2):

$$\hat{k}_m = k_{0m} - \frac{(N_m - n_m)}{2n_m(N_m + 1)k_{0m}}, \quad k_{0m} = \frac{f(x_{0m})}{n_m}$$

$$G(x_{0m}) = \frac{n_m}{N_m}$$

$$\hat{k}_f = k_0 - \frac{(N_f - n_f)}{2n_f(N_f + 1)k_{0f}}, \quad k_{0f} = \frac{f(x_{0f})}{n_f}$$

$$G(x_{0f}) = \frac{n_f}{N_f}$$

and then

$$\hat{k} = \bar{k} - \frac{1}{4} \left( \frac{(N_m - n_m)}{n_m(N_m + 1)k_{0m}} + \frac{(N_f - n_f)}{n_f(N_f + 1)k_{0f}} \right), \quad \bar{k} = \frac{1}{2} (k_{0m} + k_{0f})$$

$$\hat{k} = \bar{k} - \frac{1}{2(n_m + n_f)} \left( \frac{n_f(N_m - n_m)}{n_m(N_m + 1)k_{0m}} + \frac{n_m(N_f - n_f)}{n_f(N_f + 1)k_{0f}} \right), \quad \bar{k} = \frac{(n_fk_{0m} + n_mk_{0f})}{(n_m + n_f)}.$$
If, in addition, \( \frac{n_m}{N_m} = \frac{n_f}{N_f} \), implying equal selection pressure but not necessarily equal parental numbers selected,

\[
x_{0m} = x_{0f}, \quad k_{0m} = k_{0f} = k_0 \text{ say,} \quad \frac{N_m}{n_m} = \frac{N_f}{n_f} = (1 + \alpha_0) \text{ say}
\]

and

\[
\hat{k} = k_0 - \frac{\alpha_0}{4k_0} \left( \frac{1}{(N_m+1)} + \frac{1}{(N_f+1)} \right)
\]

\[
\hat{\kappa} = k_0 - \frac{\alpha_0}{2(n_m+n_f)k_0} \left( \frac{n_f}{(N_m+1)} + \frac{n_m}{(N_f+1)} \right)
\]

5.12 Consequences of Disruptive Selection

A basic component of experimental disruptive selection is the mating of zygotes deliberately selected from opposite phenotypic extremes of the same test sample (see e.g., Thoday, 1959). Consequences of one such cycle of selection and matings, within a sequence of such generations, can be compiled from the general results in Section 5.10 and those just derived in Section 5.11.

Since matings are across parental groups selected in the style of divergent selection from the same test material (see Section 5.8), the reference population for sampling of test material cannot be expected to be in Hardy-Weinberg equilibrium; \( \Pi_0(p, U) \) is therefore assumed with

\[
p = \frac{1}{2} (p_h + p_l), \quad Upq = -\frac{1}{4} (p_h - p_l)^2
\]

where \( p_h \) and \( p_l \) are gene \( a^+ \) frequencies of previously selected "high" and "low" parental groups. Following sampling of \( N \) test zygotes from \( \Pi_0(p, U) \), high and low selected groups are defined as for divergent selection:
\[ R \equiv \{1,2,\ldots,n_k\}, \quad R \equiv \{N-n_h+1,N-n_h+2,\ldots,N\} \]

with \((n_h+n_k) \leq N\). Standardized selection differentials are \(k_h < 0\) and \(k_h > 0\). Gene frequencies in selected groups are denoted by \((p + \delta p_h)\) and \((p + \delta p_k)\) with pre-test moments as at (5.10.5).

When matings across these selected groups are arranged at random to produce a large offspring population, and thus complete the cycle of disruptive selection, reference population \(\Pi_r\) is \(\Pi_0(p + \delta p, U + \delta U)\) with

\[ \delta p = \frac{1}{2} (\delta p_h + \delta p_k) \]

and changed allelic disequilibrium

\[ (U + \delta U)(p + \delta p)(q - \delta p) = -\frac{1}{4} (\delta p_h - \delta p_k)^2 \]

as at (5.11.2). Thus the situation is analogous to that just described for the general dioecious system with superscripts and subscripts \(m\) and \(f\) replaced by subscripts \(h\) and \(k\) respectively. The important differences are that it is known that \(k_h < 0\) and \(k_h > 0\), and that there is just one test sample of size \(N\) rather than two test samples, one for each sex. This latter difference has no effect on the subsequent algebra since the property that pre-test \(C(\delta p_h, \delta p_k)\) is \(O(g)^2\) and therefore negligible, holds for the case of allelic disequilibrium in just the same way as it does for the case of Hardy-Weinberg genotypic composition and shown at (5.8.1).

Defining \(\overline{k} = \frac{1}{2} (k_h + k_k), \quad \hat{k} = (n_h k_h + n_k k_k)/(n_h + n_k), \) pre-test moments of gene frequency change, as assessed in \(\Pi_r\), are those at (5.11.3) with \(n_d\) replaced by \(4n_h n_k/(n_h + n_k)\). In the special case of symmetric disruptive selection, as is usually the intent in experimental studies, \(n_h = n_k = \frac{1}{2} n\) say, \(n_d = n\) and \(k_h = -k_h = -k\) say. Then using
(5.10.5), pre-test moments are

$$E\{\delta_{p_h} \} = -\rho_k \sqrt{\frac{1}{2} pq(1+U)} = -E\{\delta_{p_h} \}$$

and

$$V\{\delta_{p_h} \} = \frac{2pq(1+U)}{n} - V\{\delta_{p_h} \}$$

so that in \( \Pi_r \)

$$E\{\delta p\} = 0 \quad \text{and} \quad V\{\delta p\} = \frac{pq}{2n} (1+U). \quad (5.12.1)$$

In comparison with the case of no selection \((N = n)\) and no allelic
disequilibrium \((U = 0)\) in Section 5.6, \(E\{\delta p\}\) is the same (namely zero)
but drift variance is smaller for the same census number, \(n\), of parents
of \( \Pi_r \) because \(Upq = -\frac{1}{4} (p_h - p_h^2)^2 < 0. \) The extreme case is when \(p_h = 1\)
and \(p_h = 0\) so that \(p = q = 1/2, \) \(U = -1\) and \(V\{\delta p\} = 0\) which is the obvious
result since then all matings are between \((++)\) and \((--\) types and all
offspring are heterozygotes \((+-)\). More generally, repeated cycles of
experimental disruptive selection will tend to induce parental groups
with \(p_h > p_h^2\), \(U < 0\) and reduced drift, thus prolonging life of poly-
morphism in comparison with no selection, a phenomenon discussed by
Simpson (1944) and Mather (1953) in relation to maintenance of genetic
polymorphism in natural populations.
6. SINGLE POLYGENSES AND RECURRENT SELECTION

Repeating cycles of a particular style of selection, using selected zygotes of one generation as parents of test material in the next generation, is referred to as recurrent selection. In particular, repeating cycles of directional phenotypic selection is referred to as recurrent mass selection. By maintaining critical parameters of the same selection style (such as N, n, k, etc.) constant over all cycles, the resulting sequence of generations may be modeled as a stationary stochastic process some properties of which may be used to characterize consequences of recurrent selection.

For any particular style of selection reference populations \( \Pi_0 \) and \( \Pi_r \) were deliberately specified in Chapter 5 so that a single cycle considered there may be regarded as one of a repeated sequence of such cycles; \( \Pi_r \) in one cycle serves as \( \Pi_0 \) for sampling of test zygotes in the next cycle.

6.1 Review

Three different (but not unrelated) analytical methods, each treating gene frequency \( p \) as a stationary stochastic process, have been used to analyze consequences of recurrent selection, particularly recurrent mass selection. Two of these, namely the Fokker-Planck diffusion equation (see Kimura, 1964, for review) and the Beta-distribution method of Curnow and Baker (1968) are approximations assuming \( p \) is a continuous random variable with appropriate boundary considerations (\( p = 0 \) or \( 1 \), the fixation conditions). Assuming \( n \) parental zygotes per generation the third method models gene frequency as a stationary Markov chain on a discrete state space of \( (2n + 1) \)
states including two absorbing barriers (fixation states), and has been used in a variety of ways by Feller (1951, 1957), Ewens (1963), Allan and Robertson (1964), Hill (1969a, 1969b), Narain and Robertson (1969), Carr and Nassar (1970a, 1970b) and Hedrick (1970). These three methods can also be grouped another way: the diffusion equation approximation assumes continuous generations whereas Markov chain and Curnow-Baker methods are worked over discrete, nonoverlapping generations. When applied to diploid zygotes as in the current context, all three methods involve some form of approximation that reduces a two-dimensional stochastic process (three genotypic frequencies summing to unity) to a one-dimensional stochastic process (two allelic frequencies summing to unity). The only exception to this aspect arises in the case of no selection, just drift.

Three major properties of the stochastic process have been used to characterize recurrent selection and drift. First, the probability of ultimate fixation of a specified allele, usually the favorable allele. Denoted by \( u(p_0) \), indicating dependence on initial frequency \( p_0 \) of the specified allele, this probability may be regarded as the selection limit (Robertson, 1960). Second, the rate of approach to fixation which is analogous to the rate of increase in homozygosity, or of decrease in heterozygosity, in regular inbreeding systems with no selection (Wright, 1921). Asymptotic rate is a function of the largest eigenvalue of the transition probability matrix for the Markov chain as used by Robertson (1962), Karlin (1968) and others. In general the rate can be represented by an effective number analogous to an inbreeding effective number of Kimura and Crow (1963). Third, some measures of longevity of the process such as average number of generations to homozygosity (Feller, 1954; Carr
and Nassar, 1970b) and average number of generations to fixation of the
Half-life of the selection process was defined by Robertson (1960) as
that number of generations by which expected gene frequency change
attains half of the selection limit. His approximation for the half-
life has been shown to be an overestimate particularly in case of larger
parental census numbers (Hill, 1969a; Hedrick, 1970). In addition to
gene frequency response, trait mean changes are of interest and have
been analyzed by Hill (1969b), Baker and Curnow (1969) and Robertson
among independent but identically selected replicate populations.

The three analytical methods have been compared in several ways by
several workers for some or all of the properties just described. In
particular, Pike (1969) compared the probability transition matrix
method described by Robertson (1960) with the computationally simpler
method of Curnow and Baker (1968). He concluded that the only serious
inaccuracy in the Curnow-Baker method was a tendency to fix too high a
proportion of genes particularly when gene frequency was close to a
fixation boundary and population census number was small (less than
n = 8). He also asserts that transition matrix methods cannot be
utilized in dioecious situations with unequal numbers of parental males
and females.

In this chapter attention is restricted to five specific aspects:
(1) A comparison of selection limits obtained by diffusion equation
approximations and transition matrix methods for recurrent mass
selection in a monoecious system (Section 6.2).
(ii) A demonstration that dioecious systems can be analyzed by transition matrix methods (in contrast to the assertion of Pike, 1969) and a comparison of selection limits so obtained with those provided by diffusion equation approximations (Section 6.3).

(iii) An examination of rate effective numbers as a means of characterizing rates of approach to fixation (Section 6.4).

(iv) An examination of variation in trait mean advance among independent replicate populations using the transition matrix method rather than the Curnow-Baker method (Section 6.5).

(v) An examination of median life of polymorphism with particular reference to experimental disruptive selection (Section 6.6).

Discrete generations are indexed by \( t = 0, 1, 2, \ldots \), throughout, and \( p_t \) denotes gene \( a^+ \) frequency in generation \( t \).

6.2 Selection Limits in a Monoecious System

By assuming \( n \) sufficiently large that \( p \) can be treated as a continuous random variable on the interval \([0,1]\), and assuming that the change \( \delta p_t \) over time \([t,t+\delta t]\) is sufficiently small, Kimura (1957) modeled the population as a continuous stochastic process. He obtained the following approximation for \( u(p_0) \) by the Kolmogorov backward solution to a general Fokker-Planck diffusion equation under appropriate boundary conditions (flux into \( p_t = 0 \) and \( p_t = 1 \)):

\[
\hat{u}(p_0) = \frac{\int \frac{p_0}{K(p)} \, dp}{\int \frac{1}{K(p)} \, dp}
\]

(6.2.1)
\[ K(p) = \exp\{-2\int \frac{E(\delta p)}{V(\delta p)} \, dp\} , \]

where \(E(\delta p)\) and \(V(\delta p)\) are the mean and variance of change \(\delta p\) per generation. In the context of Chapter 5 \(E(\delta p)\) and \(V(\delta p)\) are identified as pre-test moments at (5.3.8).

Assuming a relative fitness scale of (1+s):(1+hs):1 for genotypes \((++):(+-):(--)\) respectively, Kimura (1962) chose the moments of \(\delta p\) as

\[ E(\delta p) = spq[h + p(1-2h)], \quad V(\delta p) = \frac{pq}{2n} \]  \hspace{1cm} (6.2.2)

obtaining

\[ K(p) = \exp\{-2ns[(2h-1) \, pq + p]\} \]  \hspace{1cm} (6.2.3)

and noted that \(n\) could be replaced by an effective number if necessary.

In the special case of additivity on the fitness scale, \(h = 1/2\), \(\hat{u}(p_0)\) at (6.2.1) evaluates explicitly as

\[ \hat{u}(p_0) = \frac{1 - e^{-2np_0}}{1 - e^{-2ns}} , \quad \text{(Kimura, 1957)} \]  \hspace{1cm} (6.2.4)

Choice of \(E(\delta p)\) as at (6.2.2) omits mean fitness \([1 + sp(2h + p(1-2h))]\) as the denominator included in the transition matrix studies of Carr and Nassar (1970a, 1970b) and the omission of which was studied numerically by Hedrick (1970).

For the single polygene model developed in Chapter 5, pre-test moments at (5.3.8) apply for any rank set \(R_n\) defining selection style:

\[ E(\delta p) = kpq \left(\frac{A_{\delta p}}{A_{x}}\right) , \quad A = [a - (p-q) \, d] \]

\[ V(\delta p) = \frac{pq}{2n_v} , \]  \hspace{1cm} (6.2.5)

Comparing \(E(\delta p)\) at (6.2.5) and at (6.2.2), the parametric correspondence is
\[ s = k \left( \frac{2a}{\sigma_x^2} \right), \quad (2h-1) = \frac{d}{a}. \] 

(6.2.6)

In the case of mass selection and no dominance the equivalent correspondence between \( s \) and \( k(2a/\sigma_x^2) \) was derived from a consideration of trait mean advance by Falconer (1964) and was used by Robertson (1960), in combination with a curtailed series expansion of \( u(p_0) \) at (6.2.4), in developing his theory of selection limits to recurrent mass selection. Here it has been shown that the correspondence at (6.2.6) applies for any style of selection defined by \( R_n \). Robertson's expression for the selection limit is derived for the additive case \( (h = 1/2) \) and can be written in the form

\[ u(p_0) - p_0 = \sum_{t=1}^{\infty} E[\delta p_t] = \lim_{\tau \to \infty} \sum_{t=0}^{\tau} \frac{1}{2} s E[p_t q_{t-1}] \approx snp_0 q_0. \]  

(6.2.7)

the approximation for the limit being derived by assuming that

\[ E[p_t q_{t-1}] = (1 - 1/2n) E[p_{t-1} q_{t-1}] \] as predicted by drift theory alone (see Section 5.6), which ignores changes in \( p_t q_t \) attributable to selection induced changes in \( p_t \); again \( n \) can be replaced by an effective number as required.

Although this approximation does not perform very well quantitatively when compared with the diffusion equation and transition matrix methods (Robertson, 1960; Hill, 1969a) it does provide some qualitative predictions that have been verified by these alternative methods. In particular, writing \( s = 2ka/\sigma_x^2 \) in (6.2.7) and using the approximation

\[ k = k_0, \quad nk = Nf(x_0) \] for \( x_0 \) satisfying \( G(x_0) = n/N \), \( snp_0 q_0 \) is proportional to \( nk \) and is maximized when \( x_0 = 0 \) or \( n = \frac{1}{2} N \) (Robertson, 1960).

Additionally, \( snp_0 q_0 \) is a function of \( n \) and \( N \) only through the product \( nk \) and since \( nk^{(n)} = (N-n)k^{(N-n)} \) as at (2.8.3), the selection limit is
predicted to be symmetric in choice of \( n \), given \( N \), about this maximum at \( n = \frac{1}{2} N \). Precisely the same argument applies to the diffusion equation approximation (6.2.1, 6.2.3) which is also a function of \( n \) and \( N \) only through product \( nk \).

This approximate selection limit can be extended to accommodate dominance effects by the following argument. Using (6.2.2), with general degree of dominance \( h \), provides

\[
\begin{align*}
  u(p_0) - p_0 &= s h \sum_{t=0}^{\infty} E[p_t^2 q_t] + s(1-2h) \sum_{t=0}^{\infty} E[p_t^2 q_t] \\
  \end{align*}
\]

and the same drift theory (Binomial sampling) yielding \( E[p_{t+1}^2 q_{t+1}] = (1 - 1/2n) E[p_t^2 q_t] \) also yields

\[
\begin{align*}
  E[p_{t+1}^2 q_{t+1}] &= (1 - \frac{1}{2n})(1 - \frac{1}{n}) E[p_t^2 q_t] + (1 - \frac{1}{2n}) \frac{1}{2n} E[p_t^2 q_t] \\
  \end{align*}
\]

as can be verified directly or obtained from gene frequency sampling moments given by Crow and Kimura (1970). As before,

\[
\begin{align*}
  \sum_{t=0}^{\tau} E[p_t^2 q_t] &= p_0 q_0 \sum_{t=0}^{\tau} (1 - \frac{1}{2n})^t = 2np_0 q_0 \{ 1 - (1 - \frac{1}{2n})^{\tau+1} \} \\
  \end{align*}
\]

and after some algebra,

\[
\begin{align*}
  \sum_{t=0}^{\tau} E[p_t^2 q_t] &= \frac{1}{2} p_0 q_0 \sum_{t=0}^{\tau} (1 - \frac{1}{2n})^t \{ 1 + (p_0 - q_0)(1 - \frac{1}{n})^t \} \\
  &= np_0 q_0 \{ 1 - (1 - \frac{1}{2n})^{\tau+1} \} + \frac{n(p_0 - q_0)}{3n-1} \{ 1 - (1 - \frac{1}{2n})^{\tau+1} (1 - \frac{1}{n})^{\tau+1} \}.
  \end{align*}
\]

Then taking the limit as \( \tau \to \infty \) and collecting terms provides

\[
\begin{align*}
  u(p_0) - p_0 &\approx snp_0 q_0 \{ 1 + (1-2h)(p_0 - q_0) \frac{n}{3n-1} \} \\
  &\approx snp_0 q_0 \{ 1 + \frac{1}{3} (1-2h)(p_0 - q_0) \}
\end{align*}
\]

as the extension of Robertson's selection limit to include dominance.
Thus qualitatively, dominance is not predicted to change the maximizing property of choice \( n = \frac{1}{2} N \) but in comparison with \( s n p_0 q_0 \) at \( h = 1/2 \), dominance effects are predicted to increase the selection limit

\[
\begin{align*}
\text{if } p_0 < \frac{1}{2} & \text{ when } h > \frac{1}{2} \\
\text{if } p_0 > \frac{1}{2} & \text{ when } h < \frac{1}{2} .
\end{align*}
\] (6.2.9)

The discrete state space for modeling the process as a stationary Markov chain consists of \((2n+1)\) gene frequencies that are possible in \( n \) selected zygotes in generation \( t \):

\( p_t = \frac{i}{2n} , \ i = 0,1,2,\ldots,2n \),

where \( i \) is the number of \( a^+ \) alleles in parental generation \( t \). States \( i = 0 \) and \( i = 2n \) are the absorbing barriers corresponding to fixation states \( p_t = 0 \) and 1 respectively. The \((2n+1) \times (2n+1)\) matrix of transition probabilities is denoted by \( W \) with typical element \( w(i,j) \) as the conditional probability

\[
w(i,j) = \Pr[p_{t+1} = \frac{i}{2n} | p_t = \frac{i}{2n}] , \quad \sum_{j=0}^{2n} w(i,j) = 1 \ \text{all } i.
\]

The \((2n+1) \times 1\) vector whose \( i \)th element \( v(i;t) \) is the probability that the process is in state \( i \) in generation \( t \), conditional only on the foundation generation \( t = 0 \), is denoted by \( v_t^- \):

\[
v(i;t) = \Pr[p_t = \frac{i}{2n} | p_0] , \quad \sum_{i=0}^{2n} v(i;t) = 1 \ \text{all } t.
\]

Then the probability distribution of gene frequency decays over discrete generations according to

\[
v_t^- = v_{t-1}^- W = v_0^- W^t
\]
with an ultimate two-point distribution defined by
\[ u(p_0) \equiv \lim_{t \to \infty} v(2n; t) = 1 - \lim_{t \to \infty} v(0; t). \]

Hill (1969a) computed the elements \( w(i, j) \) of \( W \) using quadrature to deal with distributions of order statistics from the distribution at (5.1.3) without neglecting terms \( O(g^2) \); his results therefore are not limited to the polygenic case. He compared this method with the computationally simpler procedure of assuming a Binomial distribution for \( j \) in a sample of \( 2n \) Bernoulli trials with parameter \( \theta_i = \theta_t + E(\delta p | p_t) \), where \( p_t = i/2n \) and \( E(\delta p | p_t) \) is the expected change in gene frequency from a cycle of directional selection in test material sampled from parents with gene frequency \( p_t \). He found excellent agreement between the two methods under conditions corresponding to the polygenic assumption of Chapter 5. Using the pre-test moment \( E(\delta p) \) at (6.2.5) for \( E(\delta p | p_t) \), this Binomial assumption is adopted here; \( w(i, j) \) is calculated as
\[ w(i, j) = \binom{2n}{j} \theta_i^j (1 - \theta_i)^{2n-j} \]

\[ \theta_i = \frac{i}{2n} + k \left( \frac{1}{2n} \right) (1 - \frac{i}{2n}) \left( \frac{\sigma_x}{\sigma} - \frac{X}{n} - 1 \right) \left( \frac{d}{\sigma x} \right) \]  \hspace{1cm} (6.2.10)

with \( k \), defined by directional selection of \( n \) from \( N \), calculated from the tabulation of \( u_i|N \) given by Harter (1960).

The reference population adopted for assessment of response in any generation \( t > 0 \) is \( \Pi_0(p_t) \) the large Hardy-Weinberg offspring population obtained by random mating the group of \( n \) zygotes selected in generation \( t \) with gene frequency \( p_t \). Parameters of \( \Pi_0(p_t) \) calculated from \( v_t \) for each generation \( t \) in the studies described here, are defined with reference to their counterparts in Chapter 5 as follows:
\[ \bar{p}_t = \sum_{i=0}^{2n} \left( \frac{i}{2n} \right) v(i; t) , \quad \text{limit} \quad \bar{p}_t \rightarrow u(p_0) \quad t \rightarrow \infty \]

\[ \sigma_t^2 = \sum_{i=0}^{2n} \left( \frac{i}{2n} \right)^2 v(i; t) - \left( \bar{p}_t \right)^2 \]

\[ F_t = \frac{\sigma_t^2}{\bar{p}_t (1 - \bar{p}_t)} \]  

\[ n_{et} = \frac{(1 - F_{t-1})}{2(F_t - F_{t-1})} \]

\[ \Pr[p_t = 1 | p_0] = v(2n; t) , \quad \text{limit} \quad v(2n; t) \rightarrow u(p_0) \quad t \rightarrow \infty \]

\[ \Pr[p_t = 0 \text{ or } 1 | p_0] = v(2n; t) + v(0; t) \]

\[ \bar{\mu}_n (t) = \sum_{i=0}^{2n} \left( \frac{i}{n} - 1 \right) a + \frac{i}{n} (1 - \frac{i}{2n}) d \quad v(i; t) \]

\[ = (2p_t - 1) a + 2p_t (1 - p_t) (1 - F_t) d \]

\[ \sigma_t^2(t) = \sum_{i=0}^{2n} \left( \frac{i}{n} - 1 \right) a + \frac{i}{n} (1 - \frac{i}{2n}) d \quad \frac{2}{n} \quad v(i; t) - \left( \bar{\mu}_n (t) \right)^2 \]

Parameters \( \bar{p}_t \) and \( \sigma_t^2 \) are simply the mean and variance of the distribution of \( p_t \). \( F_t \) is the fixation index of Wright (1951), a parameter of the expected genotypic array of the large offspring population \( \Pi_0 (p_t) \). In regular inbreeding systems, with no selection, the approximate steady state condition \( (1 - F_t) = \lambda (1 - F_{t-1}) \) characterizes asymptotic rate of increase in homozygosis as \( (1 - \lambda) \) although only in the case of random mating (including selfing with probability \( 1/n \)) is the rate \( (1 - \lambda) = 1/2n \) for all \( t \geq 1 \). More generally the rate at generation \( t \)
is \( (1-\lambda_t) = (F_t - F_{t-1})/(1-F_{t-1}) \) defining the rate effective number \( n_{et} \) formulated from \((1-\lambda_t) = 1/2n_{et} \) as at (6.2.12). In the case of no selection \((k = 0)\), \( n_{et} = n \) all \( t \geq 1 \). Parameters \( Pr[p_t = 1|p_0] \) and \( Pr[p_t = 0 \ or \ 1|p_0] \) at (6.2.13) are self-explanatory, the sequence of the latter through \( t = 1,2,3,... \), defining the probability distribution of life of polymorphism. Parameters \( \mu_x(t) \) and \( \sigma^2(t) \) are the expectation and variance, over many independent replicate populations with the same stochastic history, of trait mean \( \mu_x \) defined at (5.1.1); \( \mu_x(t) - \mu_x(0) \) is the expected trait mean advance after \( t \) cycles of selection as assessed in \( \Pi_0(p_t) \).

All calculations were performed on an IBM 360/75 computer in double precision. Except in the case \( d = 0 \) when \( \hat{u}(p_0) \) was calculated directly from (6.2.4), values of \( \hat{u}(p_0) \) at (6.2.1) were obtained by iterative quadrature, of the closed Newton-Cotes type using Bode's five point rule, the quadrature interval being halved in each iteration. For the transition matrix method, \( v_t \) and parameters (6.2.11) through (6.2.14) were obtained iteratively via (6.2.10) through sufficient generations to stabilize all parameters up to the number of decimal digits printed. Later, when it was found that \( F_t \) (to 5 decimal digits) was invariably the last parameter to stabilize, this stopping rule was modified to one that monitored \( F_t \) alone.

All these calculations were performed for \((a/\sigma_x) = 0.1\) at each of \( d = 0 \ (h = 1/2, \text{additivity}), \ d/a = 1 \ (h = 1, \text{dominance}) \) and \( d/a = 3 \ (h = 2, \text{overdominance}) \) for all combinations of \( n \) and \( N \) in the array \( N = 16, n = 2(2)16 \) and \( N = 32, n = 2,4(4)28,30 \) and \( 32 \), and for several different initial frequencies \( p_0 \). Only the results for \( u(p_0 = 0.5) \) are given in Table 6.1 as the objective here is to compare selection limits.
Table 6.1. Comparison of selection limits $\hat{u}(0.5)$ and $v(2n; \infty)$ obtained by diffusion equation approximation and transition matrix methods for recurrent mass selection in a monoecious system

<table>
<thead>
<tr>
<th>n</th>
<th>$h = \frac{1}{2}$</th>
<th>$h = 1$</th>
<th>$h = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{u}(0.5)$</td>
<td>$v(2n; \infty)$</td>
<td>$\hat{u}(0.5)$</td>
</tr>
<tr>
<td>2</td>
<td>0.6480</td>
<td>0.6479</td>
<td>0.6551</td>
</tr>
<tr>
<td>4</td>
<td>0.7233</td>
<td>0.7231</td>
<td>0.7386</td>
</tr>
<tr>
<td>6</td>
<td>0.7603</td>
<td>0.7599</td>
<td>0.7801</td>
</tr>
<tr>
<td>8</td>
<td>0.7714</td>
<td>0.7711</td>
<td>0.7926</td>
</tr>
<tr>
<td>10</td>
<td>0.7603</td>
<td>0.7600</td>
<td>0.7801</td>
</tr>
<tr>
<td>12</td>
<td>0.7233</td>
<td>0.7231</td>
<td>0.7386</td>
</tr>
<tr>
<td>14</td>
<td>0.6480</td>
<td>0.6480</td>
<td>0.6551</td>
</tr>
</tbody>
</table>

---

obtained by diffusion equation approximation and transition matrix methods, and the maximum (with respect to $p_0$) discrepancies between them are found to occur at $p_0 = 0.5$ or a little less than 0.5 depending on the dominance ratio $d/a$.

Considering $v(2n; \infty)$ first, that the selection limit is almost perfectly symmetrical with respect to $n$ about the maximum at $n = \frac{1}{2} N$ in the case of additivity is a feature predicted above in relation to the theory of Robertson (1960) and is also discussed by Hill (1969a); in fact that panel on Table 6.1 defined by $N = 16$, $h = 1$ corresponds to the situation depicted in Figure 2 of Hill (1969a), his parameter $\alpha$ being $(2a/\sigma_x)$. The results in Table 6.1 indicate that this symmetry property progressively deteriorates as the dominance ratio increases through
complete dominance \((h = 1)\) to overdominance \((h = 2)\). The diffusion equation approximation necessarily provides symmetric results irrespective of dominance ratio for the reason described beneath (6.2.7). Since \(\hat{u}(p_0) \geq v(2n;\infty)\) for all variations in \(n, N\) and \(p_0\) examined, this enforced symmetry of \(\hat{u}(p_0)\) inevitably imposes an asymmetric pattern to discrepancies between selection limits determined by the two methods. In absolute magnitude there is generally good agreement between the two methods, and excellent agreement in the case of additivity where \(\hat{u}(p_0) - v(2n;\infty)\) is always less than 0.001 for \(N = 16\) and 32 at \((a/\sigma_x) = 0.1\). As the dominance ratio increases, discrepancies increase to maxima of 0.022 and 0.032 for \(N = 16\) and 32 respectively at \(h = 2\) and \((a/\sigma_x) = 0.1\). As a function of selection intensity, maximum discrepancies generally occur for \(\frac{1}{4} N < n < \frac{1}{3} N\) in the case of additivity and for \(n = \frac{1}{8} N\) in the cases of complete dominance and overdominance.

Examination of results for other values of \(p_0\) and \((a/\sigma_x)\) indicates that discrepancies generally decrease as \(|p_0 - 0.5|\) increases and as \((a/\sigma_x)\) decreases, the pattern of discrepancies described above being maintained throughout. The prediction at (6.2.9) regarding effects of dominance and initial gene frequency is also verified by the transition matrix method. For example, when \(p_0 = 0.25\), \(N = 16\) and \(a/\sigma_x = 0.1\) the following values of \(v(2n;\infty)\) are obtained for \(h \geq \frac{1}{2}\):

<table>
<thead>
<tr>
<th>(n)</th>
<th>(h = \frac{1}{2})</th>
<th>(h = 1)</th>
<th>(h = 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0.4451</td>
<td>0.4926</td>
<td>0.5757</td>
</tr>
<tr>
<td>8</td>
<td>0.4979</td>
<td>0.5628</td>
<td>0.6735</td>
</tr>
<tr>
<td>12</td>
<td>0.4462</td>
<td>0.4966</td>
<td>0.5903</td>
</tr>
</tbody>
</table>
6.3 Selection Limits in a Dioecious System

For a dioecious population under recurrent mass selection maintained by \( n_m \) and \( n_f \) parental males and females selected from test populations of \( N_m \) and \( N_f \) males and females respectively, moments of gene frequency changes per cycle of selection are given in Section 5.11. In using the diffusion equation approximation \( \hat{u}(p_0) \) at (6.2.1), effects of allelic disequilibrium are neglected (\( U \) is assumed zero) and moments \( E(p) \), \( V(p) \) defining \( K(p) \) are chosen with reference to (5.11.3), and analogous to (6.2.2), as

\[
E(\delta p) = spq[1 + q(1-2h)] \quad V(\delta p) = \frac{pq}{2n_d}
\]

(6.3.1)

\[
s = (k_m + k_f) \frac{a}{\sigma_x^2} \quad (2h-1) = \frac{d}{a} \quad n_d = \frac{4n_m n_f}{(n_m + n_f)}
\]

Formulation of the transition matrix method is similar to that for a monoecious population described in the previous section but with the major difference of an explosive increase in dimensions of the transition matrix. The discrete state space consists of \( (2n_m+1)(2n_f+1) \) factorially arranged pairs of gene frequencies that are possible in \( n_m \) selected males and \( n_f \) selected females in generation \( t \):

\[
p^m_0 = \frac{i_m}{2n_m}, \quad i_m = 0,1,2,\ldots,2n_m
\]

\[
p^f_0 = \frac{i_f}{2n_f}, \quad i_f = 0,1,2,\ldots,2n_f
\]

\[
p_t = \frac{1}{2} (p^m_t + p^f_t)
\]

where \( i_m \) (\( i_f \)) is the number of \( a^+ \) alleles in \( n_m \) (\( n_f \)) selected males (females) of generation \( t \). States \( (i_m = 0, i_f = 0) \) and \( (i_m = 2n_m, i_f = 2n_f) \) are absorbing barriers corresponding to fixation states \( p_t = 0 \) and 1 respectively. The \( (2n_m+1)(2n_f+1) \times (2n_m+1)(2n_f+1) \) matrix of transition
probabilities is $W$ with typical element $w(i_m, i_f; j_m, j_f)$ as the conditional probability
\[
w(i_m, i_f; j_m, j_f) = \Pr[p_t^{m} = \frac{j_m}{2n_m} \& p_t^{f} = \frac{j_f}{2n_f} | p_t^{m} = \frac{i_m}{2n_m} \& p_t^{f} = \frac{i_f}{2n_f}]
\]
(6.3.2)
\[
\sum_{j_m} \sum_{j_f} w(i_m, i_f; j_m, j_f) = 1, \text{ all } (i_m, i_f).
\]
In calculations reported here, rows of $W$ are organized hierarchically with $i_f = 0, 1, \ldots, 2n_f$ within each of $i_m = 0, 1, \ldots, 2n_m$; ordering of elements $v(i_m, i_f; t)$ in the $(2n_m + 1)(2n_f + 1) \times 1$ vector $v_t$ is necessarily the same, with
\[
v(i_m, i_f; t) = \Pr[p_t^{m} = \frac{i_m}{2n_m} \& p_t^{f} = \frac{i_f}{2n_f} | p_0^{m}, p_0^{f}]
\]
\[
\sum_{i_m} \sum_{i_f} v(i_m, i_f; t) = 1, \text{ all } t.
\]
Then the joint probability distribution of gene frequencies $p_t^{m}$ and $p_t^{f}$ decays over discrete generations according to
\[
v_t = v_{t-1} W = v_0 W^t
\]
(6.3.3)
with limit $v(2n_m, 2n_f; t) = 1 - \lim_{t \to \infty} v(0, 0; t) = u(p_0^{m}, p_0^{f})$.

In any generation $t > 1$, $p_t^{m}$ and $p_t^{f}$ are correlated--very highly so in advanced generations. But conditional on $p_t^{m}$ and $p_t^{f}$ in generation $t$, independence of sampling $N_m$ male test zygotes and $N_f$ female test zygotes, together with independence of selection practices, ensures that $p_{t+1}^{m}$ and $p_{t+1}^{f}$ are conditionally independent (see Section 5.11).

Thus $w(i_m, i_f; j_m, j_f)$ at (6.3.2) can be written as the product
\[ \Pr[p^m_{t+1} = \frac{j^m}{2n^m}, p^m_t = \frac{i^m}{2n^m} \& p^f_t = \frac{i^f}{2n^f}] \]

\[ \cdot \Pr[p^f_{t+1} = \frac{j^f}{2n^f}, p^m_t = \frac{i^m}{2n^m} \& p^f_t = \frac{i^f}{2n^f}] \]

and using the same Binomial assumption as for the monoecious case, \( w(i^m, i^f; j^m, j^f) \) can be calculated as the product of two Binomial probabilities:

\[ w(i^m, i^f; j^m, j^f) = \binom{2n^m}{j^m} \binom{2n^f}{j^f} \frac{j^m}{m} \left(1 - \frac{j^m}{m}\right) \frac{j^f}{f} \left(1 - \frac{j^f}{f}\right) \]

(6.3.4)

with \( \theta^m, \theta^f \) accommodating the allelic disequilibrium when provided by (5.10.3):

\[ \theta^m = p + k_m pq(1+U)(\frac{A}{\sigma_x}), \theta^f = p + k_f pq(1+U)(\frac{A}{\sigma_x}) \]

where

\[ p = \frac{1}{2} \left( p^m + p^f \right) = 1 - q, \quad p^m = \frac{i^m}{2n^m}, \quad p^f = \frac{i^f}{2n^f} \]

\[ U p q = -\frac{1}{4} \left( p^m - p^f \right)^2, \quad A = a - \frac{(1-U)}{(1+U)} (p-q) d \]

The reference population adopted for assessment of response in any generation \( t > 0 \) is \( \Pi_0(p^t, U^t) \) the large offspring population obtained by random mating the \( n^m \) males and \( n^f \) females selected in generation \( t \) with gene frequencies \( p^m_t \) and \( p^f_t \) respectively. Parameters of \( \Pi_0(p^t, U^t) \) calculated from \( v^t \) for each generation \( t \) in the studies described here, are defined as follows:

\[ \frac{i^m}{2n^m} = \frac{1}{\Pi^t} \sum \frac{v}{m, i^f, t} \]
\[
\sigma_{mt}^2 = \Sigma \left( \frac{i_m}{2n_m} \right) v(i_m, i_f; t) - (p_m t)^2 \\
- p_t^f = \Sigma \left( \frac{i_f}{2n_f} \right) v(i_m, i_f; t) \\
\sigma_{ft}^2 = \Sigma \left( \frac{i_f}{2n_f} \right) v(i_m, i_f; t) - (p_f t)^2 \\
\sigma_{mft}^2 = \Sigma \left( \frac{i_m}{2n_m} \right) \left( \frac{i_f}{2n_f} \right) v(i_m, i_f; t) - (p_m t) (p_f t) \\
- p_t = \frac{1}{2} \left( p_m t + p_f t \right) \\
\sigma_t^2 = \frac{1}{4} (\sigma_{mt}^2 + 2 \sigma_{mft}^2 + \sigma_{ft}^2) \\
F_t = \frac{\sigma_{mft}^2 - \frac{1}{4} (p_m t - p_f t)^2}{p_t (1 - p_t)} \\
n_{et} = \frac{(1 - F_{t-1})}{2(F_t - F_{t-1})} \\
Pr[p_t = 1|p^m_0, p^f_0] = v(2n_m, 2n_f; t) \\
Pr[p_t = 0 \text{ or } 1|p^m_0, p^f_0] = v(2n_m, 2n_f; t) + v(0, 0; t) \\
\overline{\mu_x}(t) = \Sigma \left( \frac{i_m}{2n_m} + \frac{i_f}{2n_f} \right) a \\
+ \left[ \frac{i_m}{2n_m} (1 - \frac{i_f}{2n_f}) + \frac{i_f}{2n_f} (1 - \frac{i_m}{2n_m}) d \right] v(i_m, i_f; t) \\
\sigma_{\mu}^2(t) = \Sigma \left( \frac{i_m}{2n_m} + \frac{i_f}{2n_f} \right) - 1) a \\
+ \left[ \frac{i_m}{2n_m} (1 - \frac{i_f}{2n_f}) + \frac{i_f}{2n_f} (1 - \frac{i_m}{2n_m}) d \right] v(i_m, i_f; t) - \{ \overline{\mu_x}(t) \}^2 ,
\]
where $\Sigma \Sigma$ denotes the double summation

$$\sum_{i_m=0}^{2n_m} \sum_{i_f=0}^{2n_f} \mu_{i_m} \mu_{i_f} \sigma^2_{i_m} \sigma^2_{i_f}$$

Parameters $\bar{p}_t$ and $\sigma^2_t$ are simply the mean and variance of the distribution of $p_t$ in $\Pi_0(p_t, U_t)$. $F_t$ is the fixation index parameterizing the expected genotypic array in $\Pi_0(p_t, U_t)$ and $n_{et}$ the corresponding rate effective number. In the case of no selection ($k_m = k_f = 0$), $n_{et} = n_d$ all $t \geq 1$. $Pr[p_t = 1|p_0^m, p_0^f]$ is the probability that gene $a^+$ is fixed at generation $t$ conditional on initial $a^+$ frequencies $p_0^m$ and $p_0^f$. In the results reported here $p_0^m = p_0^f = p_0$ say, because preliminary examination of cases with $(p_0^m + p_0^f) = 2p_0$ and $p_0^m \neq p_0^f$ showed no significant effect on the ultimate distribution of $p_t^m$ and $p_t^f$. Thus

$$u(p_0^m, p_0^f) = \lim_{t \to \infty} v(2n_m, 2n_f; t) = \lim_{t \to \infty} \bar{p}_t$$

is written as $u(p_0^m, p_0^f)$ as before. Parameters $\mu_x(t)$ and $\sigma^2(t)$ have the same meaning as at (6.2.14).

All calculations were performed as described for the monoecious case in Section 6.2 but for a smaller array of $N = N_m + N_f$ and $n = n_m + n_f$ combinations; namely,

for $h = \frac{1}{2}, 1; n_m = n_f = 1(1)6, 2N_m = 2N_f = N = 16$

for $h = \frac{1}{2}, 1; n_f = N_f = 8, n_m = 1(1)6, N_m = 8$

for $h = \frac{1}{2}, 1; \text{all other possible } (n_m, n_f) \text{ such that } n = 4(2)12, N = 16$

with $(a/\sigma_x) = 0.1$ and several values of $p_0$. Thus cases of equal
(k_{m} = k_{f}) and unequal (k_{m} \neq k_{f}) selection pressures and parental numbers are encompassed, the latter case including situations with selection of one sex only. Selection limits are given in Tables 6.2 and 6.3; again only \( p_0 = 0.5 \) is included for the same reason as in the monoecious system.

Consider first the results obtained by transition matrix methods, \( v(2n_m, 2n_f; \infty) \). When \( 2N_m = 2N_f = N = 16 \) and \( n_m = n_f = \frac{1}{2} n \), as in Table 6.2, the selection limit is maximized when \( n = \frac{1}{2} N \) as was the case for monoecious systems. When \( N_m = N_f = 8 \) and there is no selection in females, \( n_f = 8 \), as in the upper panel of Table 6.3, the selection limit is not maximized at \( n_m = \frac{1}{2} N_m = 4 \), as might be a tempting first thought, but at \( n_m = 3 \) instead.

These features are predictable using a modified version of Robertson's approximate selection limit extended to include dominance as at (6.2.8). The argument for derivation is analogous to that used in deriving (6.2.8) but with \( \bar{k} = \frac{1}{2} (k_m + k_f) \) replacing \( k \) and

\[
\frac{n_d}{n_m n_f} = \frac{4n_m n_f}{(n_m + n_f)}
\]

replacing \( n \):

\[
u(p_0) - p_0 \approx (k_m + k_f) n \quad \text{d} \quad \sigma_x^{-1} \quad p q [1 + \frac{1}{3} (1-2h) (p_0 - q_0)] \quad (6.3.9)
\]

Given \( p_0 \) and \( (\alpha/\sigma_x) \) the selection limit is indicated to be maximized when \( n_d (k_m + k_f) \) is maximized. Using the approximations

\[
k_m \approx \frac{f(x_m) N}{n_m}, \quad k_f \approx \frac{f(x_f) N}{n_f}
\]

\[
G(x_m) = \frac{n_m}{N_m}, \quad G(x_f) = \frac{n_f}{N_f}
\]

quantity \( n_d (k_m + k_f) \) can be written as
Table 6.2. Comparison of selection limits $\hat{u}(0.5)$ and $v(2n_m,2n_f;\infty)$ obtained by diffusion equation approximation and transition matrix methods for recurrent mass selection in a dioecious system; $N_m = N_f = 8$, $n_m = n_f$

<table>
<thead>
<tr>
<th>$n_n_d$</th>
<th>$(n_m, n_f)$</th>
<th>$h = \frac{1}{2}$</th>
<th>$h = 1$</th>
<th>$h = 2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$u(0.5)$ $v(2n_m, 2n_f; \infty)$</td>
<td>$u(0.5)$ $v(2n_m, 2n_f; \infty)$</td>
<td>$u(0.5)$ $v(2n_m, 2n_f; \infty)$</td>
<td>$u(0.5)$ $v(2n_m, 2n_f; \infty)$</td>
</tr>
<tr>
<td>2</td>
<td>$(1,1)$</td>
<td>0.6386</td>
<td>0.6053</td>
<td>0.6449</td>
</tr>
<tr>
<td>4</td>
<td>$(2,2)$</td>
<td>0.7131</td>
<td>0.6592</td>
<td>0.6672</td>
</tr>
<tr>
<td>6</td>
<td>$(3,3)$</td>
<td>0.7502</td>
<td>0.7325</td>
<td>0.6672</td>
</tr>
<tr>
<td>8</td>
<td>$(4,4)$</td>
<td>0.7614</td>
<td>0.7478</td>
<td>0.7814</td>
</tr>
<tr>
<td>10</td>
<td>$(5,5)$</td>
<td>0.7502</td>
<td>0.7395</td>
<td>0.7687</td>
</tr>
<tr>
<td>12</td>
<td>$(6,6)$</td>
<td>0.7131</td>
<td>0.7051</td>
<td>0.7272</td>
</tr>
</tbody>
</table>

Table 6.3. Comparison of selection limits $\hat{u}(0.5)$ and $v(2n_m,2n_f;\infty)$ obtained by diffusion equation approximation and transition matrix methods for recurrent mass selection in a dioecious system; $N_m = N_f = 8$, $n_m \neq n_f$

<table>
<thead>
<tr>
<th>$n$</th>
<th>$(n_m, n_f)$</th>
<th>$n_d$</th>
<th>$h = \frac{1}{2}$</th>
<th>$h = 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$u(0.5)$ $v(2n_m, 2n_f; \infty)$</td>
<td>$u(0.5)$ $v(2n_m, 2n_f; \infty)$</td>
<td>$u(0.5)$ $v(2n_m, 2n_f; \infty)$</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>$(1,8)$</td>
<td>3.56</td>
<td>0.6239</td>
<td>0.6071</td>
</tr>
<tr>
<td>10</td>
<td>$(2,8)$</td>
<td>6.40</td>
<td>0.6744</td>
<td>0.6617</td>
</tr>
<tr>
<td>11</td>
<td>$(3,8)$</td>
<td>8.73</td>
<td>0.6899</td>
<td>0.6799</td>
</tr>
<tr>
<td>12</td>
<td>$(4,8)$</td>
<td>10.67</td>
<td>0.6843</td>
<td>0.6763</td>
</tr>
<tr>
<td>13</td>
<td>$(5,8)$</td>
<td>12.31</td>
<td>0.6630</td>
<td>0.6567</td>
</tr>
<tr>
<td>14</td>
<td>$(6,8)$</td>
<td>13.71</td>
<td>0.6272</td>
<td>0.6227</td>
</tr>
</tbody>
</table>

| 4     | $(1,3)$      | 3.00  | 0.6686             | 0.6423  | 0.6778  | 0.6472  |
| 6     | $(1,5)$      | 3.33  | 0.6588             | 0.6363  | 0.6669  | 0.6410  |
| 6     | $(2,4)$      | 5.33  | 0.7298             | 0.7110  | 0.7459  | 0.7222  |
| 8     | $(1,7)$      | 3.50  | 0.6386             | 0.6197  | 0.6449  | 0.6225  |
| 8     | $(2,6)$      | 6.00  | 0.7131             | 0.6972  | 0.7272  | 0.7072  |
| 8     | $(3,5)$      | 7.50  | 0.7502             | 0.7360  | 0.7687  | 0.7505  |
| 10    | $(3,7)$      | 8.40  | 0.7192             | 0.7076  | 0.7340  | 0.7190  |
| 10    | $(4,6)$      | 9.60  | 0.7428             | 0.7319  | 0.7605  | 0.7464  |
| 12    | $(5,7)$      | 11.67 | 0.7065             | 0.6986  | 0.7199  | 0.7098  |
\[
\frac{1}{4} n_d(k_m + k_f) \approx \left( \frac{n_f}{n_m + n_f} \right) N_m f(x_m) + \left( \frac{n_m}{n_m + n_f} \right) N_f f(x_f).
\] (6.3.10)

If any three of the degrees of freedom among \( n_m, n_f, N_m \) and \( N_f \) are considered fixed, maxima of \( n_d(k_m + k_f) \) with respect to the fourth can be investigated analytically. For example, suppose \( N_m, N_f \) and \( n_f \) are fixed; writing \( n_m = N \frac{g(x_m)}{n_m} \) in (6.3.10) and differentiating the resulting \( n_d(k_m + k_f) \) with respect to \( x_m \) yields

\[
n_f f(x_f) [N \frac{g(x_f)}{n_m} + n_f]^{-2} [N \frac{f(x_m)}{n_m} - x \frac{g(x_m)}{n_m} - \frac{n_f x_m + N_f f(x_f)}{n_m}]
\]

which is zero for a maximum of \( n_d(k_m + k_f) \) when

\[
N \frac{L(x_m)}{n_m} = n_f x_m + N_f f(x_f),
\] (6.3.11)

where \( L(x) = f(x) - xG(x) \) is the unit Normal linear loss function defined by Raiffa and Schlaifer (1961). The following special cases are of interest in regard to Tables 6.2 and 6.3. When there is no selection on females, \( n_f = N_f \), (6.3.11) reduces to

\[
L(x_m) = \frac{N_f}{n_m} x_m.
\]

Values of \( x_m \) and the corresponding \( n_m/N_m \) satisfying this identity are listed next:

<table>
<thead>
<tr>
<th>( n_f/N_m )</th>
<th>( x_m )</th>
<th>( n_m/N_m )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \infty )</td>
<td>zero</td>
<td>0.500</td>
</tr>
<tr>
<td>39.40</td>
<td>0.01</td>
<td>0.496</td>
</tr>
<tr>
<td>7.50</td>
<td>0.05</td>
<td>0.480</td>
</tr>
<tr>
<td>3.50</td>
<td>0.10</td>
<td>0.460</td>
</tr>
<tr>
<td>1.53</td>
<td>0.20</td>
<td>0.421</td>
</tr>
<tr>
<td>1.00</td>
<td>0.275</td>
<td>0.392</td>
</tr>
<tr>
<td>0.58</td>
<td>0.40</td>
<td>0.345</td>
</tr>
<tr>
<td>0.40</td>
<td>0.50</td>
<td>0.309</td>
</tr>
<tr>
<td>0.20</td>
<td>0.70</td>
<td>0.242</td>
</tr>
<tr>
<td>0.11</td>
<td>0.90</td>
<td>0.184</td>
</tr>
</tbody>
</table>

(6.3.12)
Thus, when \( N_m = N_f = n_f \) the selection limit is predicted to be maximized at \( n_m \approx (0.39)N_m \); when \( N_m = N_f = n_f = 8 \) as in the upper panel of Table 6.3, \( v(2n_m, 2n_f; \infty) \) is maximum at \( n_m = 3 = (0.375)N_m \), the nearest integer to \( (0.39)N_m \).

When \( (n_m/N_m) = (n_f/N_f) \), (6.3.11) reduces to

\[
(N_m - N_f) f(x_m) = x_m (n_m + n_f), \quad x_f = x_m.
\]  \( (6.3.13) \)

If \( N_m > N_f \), this solution can be written

\[
\frac{(N_m + N_f)}{(N_m - N_f)} = \frac{f(x_m)}{x_m G(x_m)} = \frac{k_m}{x_m^2}.
\]

Then if \( N_m = 2N_f \) for example, the solution is \( x_m \approx 0.345 \) and \( n_m \approx (0.365)N_m \). However if \( N_m = N_f = \frac{1}{2} N \), (6.3.13) yields solution \( x_m = 0 \) and \( n_m = \frac{1}{2} N_m \) the pattern found in Table 6.2. Of course in this case \( n_m = n_f \) also, so that \( n_d = n_m + n_f \) and \( k_m = k_f \); then the approximate selection limit at (6.3.9) reverts to that for the monoecious case with \( n \) replaced by \( (n_m + n_f) \) and \( k \) by \( k_m = k_f \), with the corresponding symmetry properties discussed previously, but with the difference that \( k \) is defined for directional selection of \( \frac{1}{2} n \) from \( \frac{1}{2} N \).

Turning now to the diffusion equation approximation, \( \hat{u}(p_0) \) is a function of \( n_m, n_f, N_m \) and \( N_f \) only through the product \( n_d k \). So again the corresponding symmetry properties apply if \( n_m = n_f \) and \( k_m = k_f \).

An additional identity arises when \( n_m \neq n_f \) but \( 2(n_m + n_f) = N_m = N_f \). In this case \((n_m, n_f)\) combinations \((n_m, n_f), (n_m, n_f), (n_m, n_f)\) and \((n_f, n_m)\) all have the same product \( n_d k \) by virtue of (2.8.3); for example, combinations (2,2), (2,6) and (6,6) yield identical \( \hat{u}(p_0) \) values in Tables 6.2 and 6.3, as do combinations (3,3), (3,5) and (5,5).
Agreement between the two methods of determining \( u(p_0) \) is not as good as that in the monoecious case. The pattern of discrepancies is similar in that \( \hat{u}(p_0) \) is always greater than \( v(2n_m,2n_f;\infty) \) and magnitudes of discrepancies increase with increasing dominance ratio, but differs in other respects. When \( n_m = n_f \) and \( N_m = N_f = 8 \) as in Table 6.2, maximum discrepancy occurs at the highest selection intensity with magnitudes 0.033, 0.039 and 0.051 for \( h = 1/2, 1 \) and 2 respectively when \( (a/\sigma_x) = 0.1 \). A similar pattern is obtained when there is no selection in one sex, \( n_f = N_f \) say, as in the upper panel of Table 6.3; again maximum discrepancies occur at the highest selection intensity with magnitudes 0.017 and 0.020 for \( h = 1/2 \) and 1 respectively when \( (a/\sigma_x) = 0.1 \). Comparing discrepancies in the lower panel of Table 6.3 with those in Table 6.2 yields the following observation: when \( (n_m + n_f) \) is fixed discrepancies decrease as \( n_d \) increases, that is as \( (n_m - n_f) \) decreases. In some cases the discrepancies are sufficiently large to change the order of \( (n_m, n_f) \) combinations when ranked according to calculated selection limit. For example, when \( N_m = N_f = 8 \) and \( h = 1/2 \) in Tables 6.2 and 6.3, combination \( (n_m, n_f) = (5,7) \) has a greater \( v(2n_m,2n_f;\infty) \) value than combinations \( (2,6) \) and \( (2,2) \), whereas the reverse situation applies to \( \hat{u}(0.5) \) values:

\[
\begin{array}{ccc}
(n_m, n_f) & \hat{u}(0.5) & v(2n_m,2n_f;\infty) \\
(5,7) & 0.7065 & 0.6986 \\
(2,6) & 0.7131 & 0.6972 \\
(2,2) & 0.7131 & 0.6892 \\
\end{array}
\]

The overall ranking of \( (n_m, n_f) \) combinations across Tables 6.2 and 6.3 on the basis of \( v(2n_m,2n_f;\infty) \) is of some interest; combination \( (4,4) \) is superior, and \( (1,1) \) is inferior, to all other variations in \( (n_m, n_f) \)
and the maximum of \( v(2n_m, 2n_f; \infty) \) when \( n_m \neq n_f \) occurs for combination \((3, 5)\).

### 6.4 Rate Effective Number \( n_{et} \) under Recurrent Mass Selection

Values of \( n_{et} \) for each generation \( t \geq 1 \) where computed as described in Section 6.2 for monoecious systems and Section 6.3 for dioecious systems. Progressions of \( n_{et} \) for a monoecious system are depicted in Figure 6.1 for \( p_0 = 0.25 \) and \( p_0 = 0.5 \); these particular examples are chosen as representative of the general patterns observed in all monoecious results obtained. For \( h = 1 \) and \( 2 \) qualitative features of these general patterns are stable over variations in parameters \( n, N, p_0, \frac{a}{\sigma_x} \) and \( h \); quantitative features are affected by such variations. In particular, the occurrence of the peak \( n_{et} \) values is earlier and their amplitude greater when \( N \) increases for given selection pressure \( (n/N) \).

In the case of additivity \( (h = 1/2) \) the occurrence of a minor peak in \( n_{et} \) is dependent upon initial frequency \( p_0 \) as shown in Figure 6.1. The same general patterns are found for the dioecious systems, two representative examples of which are shown in Figure 6.2.

In interpreting these \( n_{et} \) values it has to be remembered that the rate of increase in fixation index \( F_t \) is inversely proportional to \( n_{et} \). Asymptotic rate effective number \( \hat{n_e} \) is defined as

\[
\hat{n_e} = \lim_{t \to \infty} n_{et}
\]

and an important finding in all the monoecious and dioecious systems examined is that \( \hat{n_e} \) is not a function of initial frequency \( p_0 \). With \( \frac{a}{\sigma_x} \) fixed at 0.1, values of \( \hat{n_e} \) are shown in Table 6.4 for monoecious systems with \( N = 16 \) and 32.
Figure 6.1. Progression of rate effective number under recurrent mass selection in a monoecious system.
Figure 6.2. Progression of rate effective number under recurrent mass selection in a dioecious system
Table 6.4. Asymptotic rate effective numbers for recurrent mass selection in a monoecious system; \(N = 16\) and \(32\)

<table>
<thead>
<tr>
<th>n</th>
<th>(h = \frac{1}{2})</th>
<th>(h = 1)</th>
<th>(h = 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\hat{n}_e)</td>
<td>(\hat{n}_e/n)</td>
<td>(\hat{n}_e)</td>
</tr>
<tr>
<td>---</td>
<td>------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>2</td>
<td>1.93</td>
<td>0.965</td>
<td>2.06</td>
</tr>
<tr>
<td>4</td>
<td>3.66</td>
<td>0.915</td>
<td>4.21</td>
</tr>
<tr>
<td>6</td>
<td>5.30</td>
<td>0.883</td>
<td>6.33</td>
</tr>
<tr>
<td>8</td>
<td>6.97</td>
<td>0.871</td>
<td>8.47</td>
</tr>
<tr>
<td>10</td>
<td>8.82</td>
<td>0.883</td>
<td>10.69</td>
</tr>
<tr>
<td>12</td>
<td>10.98</td>
<td>0.915</td>
<td>12.98</td>
</tr>
<tr>
<td>14</td>
<td>13.50</td>
<td>0.964</td>
<td>15.09</td>
</tr>
<tr>
<td>---</td>
<td>------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>2</td>
<td>1.90</td>
<td>0.950</td>
<td>2.05</td>
</tr>
<tr>
<td>4</td>
<td>3.45</td>
<td>0.863</td>
<td>4.12</td>
</tr>
<tr>
<td>8</td>
<td>5.77</td>
<td>0.721</td>
<td>7.62</td>
</tr>
<tr>
<td>12</td>
<td>7.73</td>
<td>0.644</td>
<td>10.67</td>
</tr>
<tr>
<td>16</td>
<td>9.92</td>
<td>0.620</td>
<td>13.92</td>
</tr>
<tr>
<td>20</td>
<td>12.87</td>
<td>0.644</td>
<td>17.92</td>
</tr>
<tr>
<td>24</td>
<td>17.29</td>
<td>0.720</td>
<td>23.30</td>
</tr>
<tr>
<td>28</td>
<td>24.13</td>
<td>0.862</td>
<td>30.06</td>
</tr>
<tr>
<td>30</td>
<td>28.43</td>
<td>0.948</td>
<td>32.68</td>
</tr>
</tbody>
</table>

Values of ratio \(\hat{n}_e/n\) are included in Table 6.4 because this ratio appears to conform with the definition of acceleration/retardation factors given by Robertson (1962); \(\hat{n}_e/n > 1\) implies an asymptotic rate under selection slower than that under no selection, referred to as retardation, whereas \(\hat{n}_e/n < 1\) implies the reverse, referred to as acceleration, these comparisons of selective and neutral monoecious systems being made for the same census number \(n\).

In the case of additivity \(\hat{n}_e/n\) is almost perfectly symmetrical with respect to \(n\), about \(n = \frac{1}{2}N\), as shown in Table 6.4. Recalling the same symmetry in regard to selection limits (see Table 6.1), this identifies a contributing component of \(u(p_0)\); maximum selection limit is attained at maximum acceleration of rate of increase in homozygosis.
Asymptotic rates are also affected by magnitude of gene effect as shown in Table 6.5 for \((a/\sigma_x) = 0.05\) and 0.10. The trend indicated in Table 6.5 is general for all systems examined; decreasing magnitude of gene effect within the polygenic range, while keeping selection pressure constant, decreases the asymptotic rate especially in large test populations. The exception to this generality occurs in very small test populations, particularly at low selection pressures, where drift effects dominate the process.

Table 6.5. Asymptotic rate effective numbers for recurrent mass selection in a monoecious system; \(8n = N\), \((a/\sigma_x) = 0.05\) and 0.10

<table>
<thead>
<tr>
<th>((a/\sigma_x) =)</th>
<th>(h = \frac{1}{2})</th>
<th>(h = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(N)</td>
<td>(n)</td>
<td>(\hat{n}_e)</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>1.98</td>
</tr>
<tr>
<td>24</td>
<td>3</td>
<td>2.94</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>3.87</td>
</tr>
<tr>
<td>40</td>
<td>5</td>
<td>4.70</td>
</tr>
<tr>
<td>48</td>
<td>6</td>
<td>5.49</td>
</tr>
<tr>
<td>80</td>
<td>10</td>
<td>7.93</td>
</tr>
<tr>
<td>120</td>
<td>15</td>
<td>9.44</td>
</tr>
</tbody>
</table>

Similar properties of asymptotic rates are found for dioecious systems. Examples are given in Tables 6.6 and 6.7 where \((a/\sigma_x) = 0.1\). In these tables acceleration/retardation is measured by ratio \(\hat{n}_e/n_d\) rather than \(\hat{n}_e/n\); of course in Table 6.6 where \(n_m = n_f = \frac{1}{2} n\), \(n_d = n\). Again in the case of additivity \((h = 1/2)\), comparison with Tables 6.2 and 6.3 indicates the correspondence between attainment of dioecious selection limit and maximum acceleration of rate of approach to fixation.
Table 6.6. Asymptotic rate effective numbers for recurrent mass selection in a dioecious system; \( N_m = N_f = 8, n_m = n_f \)

<table>
<thead>
<tr>
<th>((n_m, n_f))</th>
<th>(n_d)</th>
<th>(h = \frac{1}{2})</th>
<th>(h = 1)</th>
<th>(h = 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\hat{n}_e)</td>
<td>(\hat{n}<em>e/\hat{n}</em>{d})</td>
<td>(\hat{n}_e)</td>
<td>(\hat{n}<em>e/\hat{n}</em>{d})</td>
</tr>
<tr>
<td>(1,1)</td>
<td>2</td>
<td>1.97</td>
<td>0.985</td>
<td>2.13</td>
</tr>
<tr>
<td>(2,2)</td>
<td>4</td>
<td>3.77</td>
<td>0.943</td>
<td>4.40</td>
</tr>
<tr>
<td>(3,3)</td>
<td>6</td>
<td>5.45</td>
<td>0.908</td>
<td>6.60</td>
</tr>
<tr>
<td>(4,4)</td>
<td>8</td>
<td>7.16</td>
<td>0.895</td>
<td>8.77</td>
</tr>
<tr>
<td>(5,5)</td>
<td>10</td>
<td>9.02</td>
<td>0.902</td>
<td>10.97</td>
</tr>
<tr>
<td>(6,6)</td>
<td>12</td>
<td>11.15</td>
<td>0.930</td>
<td>13.19</td>
</tr>
</tbody>
</table>

Table 6.7. Asymptotic rate effective numbers for recurrent mass selection in a dioecious system; \( N_m = N_f = n_f = 8 \)

<table>
<thead>
<tr>
<th>((n_m, n_f))</th>
<th>(n_d)</th>
<th>(h = \frac{1}{2})</th>
<th>(h = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\hat{n}_e)</td>
<td>(\hat{n}<em>e/\hat{n}</em>{d})</td>
<td>(\hat{n}_e)</td>
</tr>
<tr>
<td>(1,8)</td>
<td>3.56</td>
<td>3.49</td>
<td>0.980</td>
</tr>
<tr>
<td>(2,8)</td>
<td>6.40</td>
<td>6.12</td>
<td>0.956</td>
</tr>
<tr>
<td>(3,8)</td>
<td>8.73</td>
<td>8.25</td>
<td>0.945</td>
</tr>
<tr>
<td>(4,8)</td>
<td>10.67</td>
<td>10.11</td>
<td>0.948</td>
</tr>
<tr>
<td>(5,8)</td>
<td>12.31</td>
<td>11.80</td>
<td>0.959</td>
</tr>
<tr>
<td>(6,8)</td>
<td>13.71</td>
<td>13.38</td>
<td>0.976</td>
</tr>
</tbody>
</table>

6.5 Trait Mean Advance and Variation among Replicates

Hill (1969b) considered the progression of trait X mean, \( \mu_X(t) \), for a monoecious system using the transition matrix method. His primary concern was the effect of dominance and the opposing actions of selection and inbreeding depression. Baker and Curnow (1969) considered the effects of subdividing a fixed resource into replicate populations independently and identically manipulated according to the same selection style.

Their interest was to examine the potential trait mean advance utilizing variation among subdivision means in addition to selection within subdivisions. The particular system that they analyzed in detail by the
β-distribution method as corrected by Curnow and Baker (1969), involves recurrent selection among a finite number of females, with an effectively infinite number of males, in each cycle. Here these aspects are treated differently using transition matrix methods described in Sections 6.2 and 6.3 for monoecious and dioecious systems respectively.

Tables 6.8 and 6.9 contain values of $\mu = 100 \frac{\mu_x(t)}{\sigma_x}$ and $\sigma = 100 \frac{\sigma_x(t)}{\sigma_x}$ for $t = 5, 10, 15$ and 100, all with $(a/\sigma_x) = 0.1$ and $p_0 = 0.5$, in a monoecious system with a fixed resource of $rN = 240$ test zygotes consisting of $r$ replicate populations each with $N$ test zygotes and a fixed mass selection pressure of $n = (0.125)N$; in Table 6.8 dominance ratio $(d/a)$ is zero whereas in Table 6.9 $(d/a) = 1$, complete dominance in polygenic effects.

In interpreting these tables it is useful to refer $\mu$ to its possible range; $\mu \geq -10.0$ attained when $p_t = 0$ and $\mu \leq +10.0$ attained when $p_t = 1$. In the case of additivity $(d = 0)$ the results may be briefly summarized as follows. Expected trait mean, $\mu$, over all replications increases with recurrent cycles of mass selection, ultimately reaching the limit predictable from $u(p_0) = \lim_{t \to \infty} p_t$, namely $100 \frac{\mu_x(t)}{\sigma_x} \to 200(u(p_0) - 1/2)(a/\sigma_x)$. Variance of trait mean, $\sigma^2$, among many replicates increases quite rapidly through early cycles of selection except in populations with larger parental census number, and then stabilizes or substantially decreases depending on parental census number through the influence of the latter on $u(p_0)$. When $u(p_0)$ is close to 1.0, $\sigma^2$ is ultimately insignificant relative to $\mu$, a result predictable from the Binomial variance $u(p_0)(1 - u(p_0))$. With regard to exploiting variation among replicates, with the purpose of advancing trait mean given a fixed resource, Baker and Curnow (1969)
Table 6.8. Expectation and standard deviation of trait mean (x 100) in a subdivided monoecious system under recurrent mass selection; \( rN = 240, 8n = N, (a/a_{x}) = 0.1, d = 0 \)

<table>
<thead>
<tr>
<th>( r )</th>
<th>15</th>
<th>10</th>
<th>6</th>
<th>5</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>16</td>
<td>24</td>
<td>40</td>
<td>48</td>
<td>80</td>
<td>120</td>
<td>240</td>
</tr>
<tr>
<td>( n )</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>( t=0 )</td>
<td>( \mu )</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>2.298</td>
<td>7.412</td>
<td>6.012</td>
<td>5.543</td>
<td>4.367</td>
<td>3.590</td>
<td>2.553</td>
</tr>
<tr>
<td>( t=5 )</td>
<td>( \mu )</td>
<td>2.813</td>
<td>3.778</td>
<td>4.860</td>
<td>5.176</td>
<td>5.856</td>
<td>6.210</td>
</tr>
<tr>
<td>( t=10 )</td>
<td>( \mu )</td>
<td>2.929</td>
<td>4.155</td>
<td>5.716</td>
<td>6.201</td>
<td>7.242</td>
<td>7.754</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>9.502</td>
<td>8.812</td>
<td>7.294</td>
<td>6.635</td>
<td>4.781</td>
<td>3.556</td>
<td>2.174</td>
</tr>
<tr>
<td>( t=15 )</td>
<td>( \mu )</td>
<td>2.962</td>
<td>4.375</td>
<td>6.616</td>
<td>7.436</td>
<td>9.227</td>
<td>9.841</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>9.551</td>
<td>8.992</td>
<td>7.499</td>
<td>6.686</td>
<td>3.855</td>
<td>1.775</td>
<td>0.160</td>
</tr>
<tr>
<td>( t=100 )</td>
<td>( \mu )</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>3.491</td>
<td>4.724</td>
<td>5.930</td>
<td>6.251</td>
<td>6.896</td>
<td>7.213</td>
<td>7.522</td>
</tr>
</tbody>
</table>

Table 6.9. Expectation and standard deviation of trait mean (x 100) in a subdivided monoecious system under recurrent mass selection; \( rN = 240, 8n = N, (a/a_{x}) = 0.1, d = a \)

<table>
<thead>
<tr>
<th>( r )</th>
<th>15</th>
<th>10</th>
<th>6</th>
<th>5</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>16</td>
<td>24</td>
<td>40</td>
<td>48</td>
<td>80</td>
<td>120</td>
<td>240</td>
</tr>
<tr>
<td>( n )</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>10</td>
<td>15</td>
<td>30</td>
</tr>
<tr>
<td>( t=0 )</td>
<td>( \mu )</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
<td>5.000</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>8.347</td>
<td>6.965</td>
<td>5.147</td>
<td>4.566</td>
<td>3.241</td>
<td>2.491</td>
<td>1.639</td>
</tr>
<tr>
<td>( t=5 )</td>
<td>( \mu )</td>
<td>3.124</td>
<td>4.602</td>
<td>6.373</td>
<td>6.872</td>
<td>7.832</td>
<td>8.248</td>
</tr>
<tr>
<td>( t=10 )</td>
<td>( \mu )</td>
<td>3.033</td>
<td>4.551</td>
<td>6.623</td>
<td>7.242</td>
<td>8.394</td>
<td>8.819</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>9.470</td>
<td>8.614</td>
<td>6.558</td>
<td>5.655</td>
<td>3.297</td>
<td>2.067</td>
<td>1.105</td>
</tr>
<tr>
<td>( t=15 )</td>
<td>( \mu )</td>
<td>3.002</td>
<td>4.514</td>
<td>6.976</td>
<td>7.861</td>
<td>9.566</td>
<td>9.950</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>9.539</td>
<td>8.923</td>
<td>7.164</td>
<td>6.181</td>
<td>2.905</td>
<td>0.964</td>
<td>0.122</td>
</tr>
</tbody>
</table>
computed the expected trait mean for the best of \( r \) replicates as
\[
\mu_1 = \bar{\mu}_X(t) + \mu_r |_r \sigma_\mu(t),
\]
where \( \mu_r |_r \) is the expectation of the largest order statistic in a sample of size \( r \) from the unit Normal distribution. When this is done in Table 6.8 values of \( \mu_1 \) obtained are almost always in excess of 10.0 which is impossible. The same problem arises in Table 8 of Baker and Curnow (1969) although they do not comment on the phenomenon. Clearly the assumption of Normally distributed replicate means (implied by use of \( \mu_r |_r \)) is not even approximately appropriate in these cases.

In the case of dominance (.d = a, Table 6.9), a similar pattern of results applies but with two important differences. First, the initial mean is 5.0 rather than zero as in the case of additivity; potential for trait mean advance is therefore halved. Second, inbreeding depression effects have to be balanced against advance from selection. Using (6.2.12), \( \bar{\mu}_X(t) \) is greater than \( \bar{\mu}_X(0) \) when \( d = a \), if
\[
(p_t - \frac{1}{2}) + \bar{p}_t q_t (1 - F_t) > (p_0 - \frac{1}{2}) + p_0 q_0
\]
and ultimately, \( \bar{p}_t \to u(p_0), F_t \to 1, \bar{\mu}_x(\infty) \) is greater than \( \bar{\mu}_X(0) \) if
\[
u(p_0) > p_0 + p_0 q_0 \tag{6.5.1}
\]
or \( u(0.5) > 0.75 \) when \( p_0 = 0.5 \) as in Table 6.9. Using the transition matrix method values of \( u(0.5) \) for the \((n,N)\) combinations \((3,24)\) and \((5,40)\) in Table 6.9 are 0.7257 and 0.8488 respectively corresponding to the decreasing and increasing trends in \( \nu \) respectively that are shown there. When \( (a/\sigma_x) = 0.05 \), half of its value in Table 6.9, \( u(0.5) = 0.5763, 0.6173, 0.6970, 0.7344, 0.8571 \) for \((n,N)\) combinations \((2,16), (3,24), (5,40), (6,48)\) and \((10,80)\) respectively when \( d = a; \)
then not until the (10,80) combination is reached does selection
dominate inbreeding depression to yield a positive trait mean response.

More generally, for dominance ratio \((d/a) > 0\), ultimate trait mean
advance \(\mu_x(\infty) - \mu_x(0)\) is expected to be positive when

\[ u(p_0) > p_0 + p_0 q_0 \frac{(d)}{a} \]

which is impossible if \(p_0 > (a/d)\).

In Tables 6.8 and 6.9 selection pressure is fixed at \(8n = N\) and
initial frequency at \(p_0 = 0.5\). Trait mean advance and its variance
are, of course, affected by variations in \((n/N)\) and \(p_0\) as shown in
Table 6.10 where \(N\) is fixed at 16 and \(\mu = 100 \frac{\mu_x(t)}{\sigma_x}\),
\(\sigma = 100 \frac{\sigma_x(t)}{\sigma_x}\) are tabulated for \(n = 4, 8\) and 12 at each of
\(p_0 = 0.25\) and 0.50.

<table>
<thead>
<tr>
<th>(n)</th>
<th>(p_0)</th>
<th>4</th>
<th>0.25</th>
<th>0.50</th>
<th>8</th>
<th>0.25</th>
<th>0.50</th>
<th>12</th>
<th>0.25</th>
<th>0.50</th>
</tr>
</thead>
<tbody>
<tr>
<td>t=0</td>
<td>(\mu)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t=5</td>
<td>(\mu)</td>
<td>1.886</td>
<td>2.311</td>
<td>1.333</td>
<td>1.666</td>
<td>0.714</td>
<td>0.919</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t=10</td>
<td>(\mu)</td>
<td>6.792</td>
<td>6.767</td>
<td>5.000</td>
<td>5.161</td>
<td>4.014</td>
<td>4.354</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t=15</td>
<td>(\mu)</td>
<td>2.917</td>
<td>3.429</td>
<td>2.383</td>
<td>2.831</td>
<td>1.329</td>
<td>1.653</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t=100</td>
<td>(\mu)</td>
<td>8.571</td>
<td>8.048</td>
<td>6.858</td>
<td>6.567</td>
<td>5.601</td>
<td>5.786</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t+∞</td>
<td>(\mu)</td>
<td>3.427</td>
<td>3.966</td>
<td>3.156</td>
<td>3.636</td>
<td>1.841</td>
<td>2.237</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sigma)</td>
<td>9.314</td>
<td>8.544</td>
<td>7.967</td>
<td>7.286</td>
<td>6.681</td>
<td>6.663</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t=100</td>
<td>(\sigma)</td>
<td>9.940</td>
<td>8.949</td>
<td>9.997</td>
<td>8.401</td>
<td>9.891</td>
<td>8.916</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>t+∞</td>
<td>(\sigma)</td>
<td>9.902</td>
<td>4.462</td>
<td>4.958</td>
<td>5.421</td>
<td>3.925</td>
<td>4.463</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\sigma)</td>
<td>9.940</td>
<td>8.949</td>
<td>10.000</td>
<td>8.403</td>
<td>9.942</td>
<td>8.949</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When \( p_0 = 0.5 \), \( \bar{\mu}_x(0) \) is zero; but when \( p_0 = 0.25 \), \( 100 \frac{\bar{\mu}_x(0)}{\sigma_x} = -5. \)

In Table 6.10, 5,000 has been added to each value of \( \mu \) listed for \( p_0 = 0.25 \) so that for both \( p_0 = 0.5 \) and 0.25 it is trait mean advance

\[
100 \frac{\bar{\mu}_x(t)}{\sigma_x} - 100 \frac{\bar{\mu}_x(0)}{\sigma_x}
\]

that is tabulated for \( \mu \). As \( t \to \infty \), symmetry of \( \mu \) and \( \sigma \) with respect to \( n = 4 \) and 12 is a consequence of the symmetry of selection limits \( u(p_0) \) discussed in Section 6.2. Trait mean advance when \( p_0 = 0.25 \) is never greater than that when \( p_0 = 0.5 \) in populations as small as that in Table 6.10 \( (n < N = 16) \) although the potential for advance is greater, being from -5.0 to +10.0 as opposed to 0.0 to +10.0. For both \( p_0 = 0.25 \) and 0.50, increasing the selection pressure (decreasing \( n \)) from that providing the maximum selection limit \( (n = 8) \) results in superior trait mean advance in early generations but only at the expense of an increased variance and therefore less reliable advance, and a sacrifice of potential advance in much later generations. Alternatively, decreasing the selection pressure \( (n > 8) \) reduces variance of trait mean response in early generations but sacrifices more than half the potential mean advance in early generations when \( n = 12 \).

Turning now to dioecious systems the same general patterns are found when \( N_m = N_f, n_m = n_f \). This is predictable in view of the close similarity between the monoecious system \((n,N)\) and the dioecious system with \( 2N_m = 2N_f = N \) and \( 2n_m = 2n_f = n \) found previously in respect to selection limits. Quantitatively the results differ, largely because \( \bar{k} = k_m = k_f \) refers to selection of \( \frac{1}{2} n \) from \( \frac{1}{2} N \) whereas \( k \) corresponds to mass selection of \( n \) from \( N \). This results is a reduced trait mean
advance and an increased variance of trait mean response when compared with a monoecious system having the same census number of parents 
\( n = n_m + n_f \) and the same total number of test zygotes \( N = N_m + N_f \).

Two aspects of trait mean advance in dioecious systems are shown in Table 6.11. The rightmost panel contains \( \mu \) and \( \sigma \) when no selection is practiced on one sex \( n_f = 8 \), the \((n_m, n_f)\) combination \((3, 8)\) corresponding to maximum selection limit in this case. The other panel contains four \((n_m, n_f)\) combinations such that \( n_m + n_f = 8 \), combination \((4, 4)\) corresponding maximum selection limit over all variations in \((n_m, n_f)\) with \( N_m = N_f = 8 \). With \( n_m + n_f \) fixed, not only does combination \( n_m = n_f \) maximize the selection limit, but also trait mean advance increases and variance of trait mean response decreases in all generations as \( n_m - n_f \) decreases. This is a consequence of both selection pressure and drift effects because \( n_d \) increases but \( \bar{k} \) decreases as \( n_m - n_f \) decreases:

<table>
<thead>
<tr>
<th>((n_m, n_f))</th>
<th>( n_d )</th>
<th>( \bar{k} ) (assuming ( N_m = N_f = 8 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>((1, 7))</td>
<td>3.50</td>
<td>0.8135</td>
</tr>
<tr>
<td>((2, 6))</td>
<td>6.00</td>
<td>0.7586</td>
</tr>
<tr>
<td>((3, 5))</td>
<td>7.50</td>
<td>0.7330</td>
</tr>
<tr>
<td>((4, 4))</td>
<td>8.00</td>
<td>0.7253</td>
</tr>
</tbody>
</table>

Thus for the situation in Table 6.11, increases in \( \bar{k} \) achieved by choosing \( n_m < n_f \), with \( n_m + n_f \) fixed, are not sufficient to offset the increased drift with the consequence that trait mean advance is not enhanced in early generations.

6.6 Life of Polymorphism

The number of parental generations through which the population remains unfixed is random variable \( T \) defined formally as
Table 6.11. Expectation and standard deviation of trait mean (x 100) in a dioecious system under recurrent mass selection; $N_m = N_f = 8$, $(a/\sigma_x) = 0.1$, $p_0 = 0.5$, $d = 0$

<table>
<thead>
<tr>
<th>$(n_m,n_f)$</th>
<th>$n_m + n_f = 8$</th>
<th>$n_f = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1,7)</td>
<td>(2,6)</td>
</tr>
<tr>
<td>$t=0$</td>
<td>$\mu$</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>1.338</td>
</tr>
<tr>
<td>$t=5$</td>
<td>$\mu$</td>
<td>7.254</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>1.915</td>
</tr>
<tr>
<td>$t=10$</td>
<td>$\mu$</td>
<td>8.698</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>2.177</td>
</tr>
<tr>
<td>$t=15$</td>
<td>$\mu$</td>
<td>9.267</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>2.395</td>
</tr>
<tr>
<td></td>
<td>$\sigma$</td>
<td>2.395</td>
</tr>
<tr>
<td>$t \to \infty$</td>
<td>$\mu$</td>
<td>9.709</td>
</tr>
</tbody>
</table>

$T = t$ if $p_{t+1} = 0$ or 1 but $0 < p_t < 1$ \hspace{1cm} (6.6.1)

and referred to as life of polymorphism; \( T \) is distributed over the non-negative integers. Note that \( T = 0 \) if \( p_1 = 0 \) or 1 so that the foundation parents (\( t = 0 \)) are not counted as a generation in the stochastic life of polymorphism.

For the Markov chain model of a monoecious system expected life of polymorphism, $E(T)$, can be found by the following method. Let $U$ denote the matrix of transition probabilities, corresponding to just transient: transient transitions, obtained from $W$ by deleting the appropriate two rows and two columns corresponding to fixation states; $U$ is the $(2n-1) \times (2n-1)$ matrix $[w(i,j)]$ where

$$w(i,j) = Pr[p_t = \frac{1}{2n} | p_{t-1} = \frac{1}{2n}], \quad i,j = 1,2,\ldots,2n-1.$$
Similarly let \( z_t \) denote the vector obtained from \( v_t \) by deleting the two elements corresponding to \( p_t = 0 \) and 1; \( z_t \) is a \( (2n-1) \times 1 \) vector \([v(i;t)]\) where

\[
v(i;t) = \Pr[p_t = \frac{i}{2n} | p_0], \quad i = 1, 2, ..., 2n-1.
\]

The recursion at (6.2.10) also applies to \( z_t \) and \( U \):

\[
z_t^* = z_{t-1}^* \ U = z_0^* U^t
\]

but with the difference that all elements of \( z_t \) decay to zero as \( t \to \infty \) because ultimately all probability is concentrated in the two fixation states. Thus

\[
\lim_{s \to \infty} U^s = 0 \ (\text{null}).
\]

Let \( \beta_t \) denote the probability that \( T = (t-1), \ t = 1, 2, ... \):

\[
\beta_t = \sum_{i=1}^{2n-1} \{ w(i,0) + w(i,2n) \} \ v(i,t-1)
\]

\[
= \sum_{i=1}^{2n-1} \sum_{j=1}^{2n-1} \{ 1 - \sum_{j=1}^{2n-1} w(i,j) \} \ v(i,t-1)
\]

\[
= \sum_{i=1}^{2n-1} v(i,t-1) - \sum_{j=1}^{2n-1} v(j,t)
\]

\[
= (z_{t-1}^* - z_t^*) \ 1,
\]

where \( 1 \) is the \( (2n-1) \times 1 \) vector of unities. Then using (6.6.2),

\[
\beta_t = z_0^* (I-U) U^{t-1} \ 1,
\]

where \( I \) is the \( (2n-1) \times (2n-1) \) identity matrix. By definition, \( \beta_t \) is the (discrete) density function for the distribution of \( T \) and so
\[ E(T) = \sum_{t=0}^{\infty} t \beta_t e^{t+1} \]

\[ = z_0^{-1} (I-U) \left( \sum_{t=0}^{\infty} t U^t \right) \\
= z_0^{-1} \text{limit} \left\{ (I-U) \sum_{t=0}^{s} t U^t \right\} \]

\[ = z_0^{-1} \text{limit} \left\{ \sum_{t=1}^{s} U^t - s U^{s+1} \right\} \]

\[ = z_0^{-1} (I-U)^{-1} \text{limit} \left\{ (I-U)^{-1} \sum_{t=1}^{s} U^t - s U^{s+1} \right\} \]

and using (6.6.3),

\[ E(T) = z_0^{-1} (I-U)^{-1} U_1. \] (6.6.5)

For a specific foundation parental group with \( p_0 = \frac{i}{2n}, \) \( z_0 \) consists of all zero elements except \( v(i,0) = 1. \) Thus if \( T \) is the \((2n-1) \times 1\) vector with \( i^{th} \) element as \( T \) specific to initial gene frequency \( \frac{i}{2n}, \) (6.6.5) yields

\[ E(T) = (I-U)^{-1} U_1. \] (6.6.6)

The matrix \((I-U)^{-1}\) also finds application in other parameters of the stochastic process. If \( u \) denotes the \((2n-1) \times 1\) vector with \( i^{th} \) element as \( u(p_0 = \frac{i}{2n}), \) the probability of ultimate fixation of the favorable allele given initial frequency \( p_0, \) Feller (1957) shows that

\[ u = (I-U)^{-1} w, \]

where \( w \) is the \((2n-1) \times 1\) vector of transition probabilities in \( W \) corresponding to just transitions from transient states to the \( p_t = 1 \)
fixation state. Additionally, if \( \tau \) denotes the \((2n-1) \times 1\) vector with
ith element \( \tau_i \) as the number of generations to fixation or loss of allele
\( a^+ \) given initial frequency \( p_0 = \frac{1}{2n} \), Feller (1954) shows that

\[
E(\tau) = (I-U)^{-1} \ 1 .
\]  
(6.6.7)

Carr and Nassar (1970b) use (6.6.7) to calculate \( E(\tau_i) \) and define a
retardation factor as the ratio of \( E(\tau_i) \) under selection to that for
neutral genes instead of the ratio of asymptotic rates, \( \hat{n}_e/n \), as in
Section 6.4. Since \( (I-U)^{-1} U \ 1 = (I-U)^{-1} \ 1 - \ 1 \), (6.6.6) and (6.6.7)
can be combined to yield

\[
E(\tau) = E(\tau) - 1
\]

and so expected life of polymorphism as defined here is always one
generation less than expected number of generations to fixation as
defined by Feller (1957).

In terms of numerical results obtained by computations described
in Sections 6.2 and 6.3, the probability distribution of \( T \) is defined by

\[
Pr[T \leq t] = v(2n,t+1) + v(0,t+1)
\]
in the monoecious case, and

\[
Pr[T \leq t] = v(2n_m,2n_f,t+1) + v(0,0;t+1)
\]
in the dioecious case. Since these values were computed every gener-
ation for all monoecious and dioecious systems studied here, percentiles
of the distribution of \( T \) are available. In this section median life of
polymorphism is examined in preference to \( E(T) \) partly because of its
availability just noted and also because of its indirect interpretation
as that number of generations by which 50 percent of replicate popula-
tions are still polymorphic. Formally the 100\( \alpha \) percentile of the
discrete distribution of \( T \) is taken to be that integer \( t_{100\alpha} \) satisfying
\[
\text{Pr}[T + 1 \leq t_{100\alpha}] \leq \alpha \text{ and Pr}[T \leq t_{100\alpha}] > \alpha .
\] (6.6.8)

In terms of calculated probabilities, \( t_{100\alpha} \) is the largest integer for which
\[
v(2n, t_{100\alpha}) + v(0, t_{100\alpha}) \leq \alpha
\] (6.6.9)
in a monoecious system, and
\[
v(2n_m, 2n_f; t_{100\alpha}) + v(0,0; t_{100\alpha}) \leq \alpha
\] (6.6.10)
in a dioecious system. Specification of just one particular percentile, such as the "median" \( t_{50} \), is unique but as usual with discrete distributions joint specification of two or more adjacent percentiles does not
sometimes occur. Even when limiting to single percentiles, derivation
always result in unique percentiles. For example, if in a monoecious
system it happens that
\[v(2n, 37) + v(0, 37) = 0.483 \text{ and } v(2n, 38) + v(0, 38) = 0.512\]
then 37 generations is \( t_{49} \) and \( t_{51} \) as well as median life of polymorphism,
the two \( t_{50} \) are essentially indistinguishable but \( t_{4,8} \) and \( t_{6} \) match at \( t_{50} \).

Table 6.12 contains median lives of polymorphism and \( t_{50} \) for
recurrent mass selection in a monoecious system with \( p_0 = 0.5 \) and
\( (a/s_x) = 0.1 \). For purposes of comparison some values of \( t_{50} \) and \( t_{95} \)
in the case of no selection (\( N = n \)) are included in Table 6.12. Life
of polymorphism is affected by initial gene frequency and magnitude of
polygenic effects. Table 6.13 contains \( t_{50} \) and \( t_{95} \) for systems with
\( p_0 = 0.25 \) and \( (a/s_x) = 0.1 \) at \( N = 16 \), and can be compared with the
upper panel of Table 6.12. Table 6.14 contains \( t_{50} \) and \( t_{95} \) for a
system with \( p_0 = 0.5 \) and \( 8n = N \) at \( (a/s_x) = 0.05 \) and \( 0.10 \).

When parental census number is small, \( n = 2, 3 \) and \( 4 \), the distribution of \( T \) is little affected by variations in selection pressure, initial
Table 6.12. Median life of polymorphism and \( t_{95} \) for recurrent mass selection in a monoecious system; \( p_0 = 0.5 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( h = \frac{1}{2} )</th>
<th>( h = 1 )</th>
<th>( h = 2 )</th>
<th>( N = n )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( t_{50} )</td>
<td>( t_{95} )</td>
<td>( t_{50} )</td>
<td>( t_{95} )</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>10</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>22</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>33</td>
<td>11</td>
<td>39</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>45</td>
<td>16</td>
<td>54</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>57</td>
<td>21</td>
<td>69</td>
</tr>
<tr>
<td>12</td>
<td>22</td>
<td>72</td>
<td>25</td>
<td>84</td>
</tr>
<tr>
<td>14</td>
<td>27</td>
<td>89</td>
<td>30</td>
<td>98</td>
</tr>
<tr>
<td>16</td>
<td>33</td>
<td>105</td>
<td>33</td>
<td>105</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>10</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>21</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>38</td>
<td>15</td>
<td>49</td>
</tr>
<tr>
<td>12</td>
<td>18</td>
<td>52</td>
<td>22</td>
<td>71</td>
</tr>
<tr>
<td>16</td>
<td>24</td>
<td>69</td>
<td>30</td>
<td>94</td>
</tr>
<tr>
<td>20</td>
<td>31</td>
<td>89</td>
<td>39</td>
<td>121</td>
</tr>
<tr>
<td>24</td>
<td>40</td>
<td>119</td>
<td>50</td>
<td>156</td>
</tr>
<tr>
<td>28</td>
<td>53</td>
<td>163</td>
<td>62</td>
<td>199</td>
</tr>
<tr>
<td>30</td>
<td>61</td>
<td>191</td>
<td>67</td>
<td>216</td>
</tr>
<tr>
<td>32</td>
<td>67</td>
<td>214</td>
<td>67</td>
<td>214</td>
</tr>
</tbody>
</table>

Table 6.13. Median life of polymorphism and \( t_{95} \) for recurrent mass selection in a monoecious system; \( p_0 = 0.25 \)

<table>
<thead>
<tr>
<th>( n )</th>
<th>( h = \frac{1}{2} )</th>
<th>( h = 1 )</th>
<th>( h = 2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( t_{50} )</td>
<td>( t_{95} )</td>
<td>( t_{50} )</td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>13</td>
<td>45</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>71</td>
<td>24</td>
</tr>
<tr>
<td>16</td>
<td>24</td>
<td>96</td>
<td>24</td>
</tr>
</tbody>
</table>
Table 6.14. Median life of polymorphism and $t_{95}$ for recurrent mass selection in a monoecious system; $(a/\sigma_x) = 0.05$ and $0.10$

<table>
<thead>
<tr>
<th>n = $\frac{1}{8}N$</th>
<th>(a/\sigma_x) = 0.05</th>
<th>(a/\sigma_x) = 0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$h = \frac{1}{2}$</td>
<td>$h = 1$</td>
</tr>
<tr>
<td></td>
<td>$t_{50}$</td>
<td>$t_{95}$</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>5</td>
<td>8</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>10</td>
<td>34</td>
</tr>
<tr>
<td>10</td>
<td>17</td>
<td>52</td>
</tr>
<tr>
<td>15</td>
<td>22</td>
<td>65</td>
</tr>
<tr>
<td>30</td>
<td>31</td>
<td>74</td>
</tr>
</tbody>
</table>

frequency, degree of dominance or magnitude of gene effects within the polygenic range; drift effects dominate the process and only a high degree of overdominance can significantly effect life of polymorphism.

For larger parental census numbers and selection with no dominance, life of polymorphism is stochastically shortened (both $t_{50}$ and $t_{95}$ decrease) in comparison with similar sized systems of neutral genes, and also by increasing selection pressure ($N$ increasing while $n$ constant). With selection and overdominance, $h = 2$, median life of polymorphism is increased in comparison with similar size systems of neutral genes, and $t_{95}$ may be as such as doubled, provided that $p_0$ is intermediate (say $0.2 < p_0 < 0.8$). Unlike retardation factor $\hat{n}_e/n$ which is independent of initial gene frequency, median life of polymorphism is sensitive to extremal values of $p_0$ and may be less under selection with overdominance than under a neutral system with the same parental census. With selection pressure held constant (as in Table 6.14) decreasing the magnitude of gene effect stochastically increases
life of polymorphism for all except the smallest \( n \), and markedly so for large \( n \) in the case of complete dominance.

Similar patterns are found for dioecious systems provided that findings dependent upon \( n \) are understood to be dependent upon \( n_d \). Tables 6.15 and 6.16 contain \( t_{50} \) and \( t_{95} \) for assorted dioecious systems with \( (a/\sigma_x) = 0.1 \) and \( p_0 = 0.5 \).

Table 6.15. Median life of polymorphism and \( t_{95} \) for recurrent mass selection in a dioecious system; \( N_m = N_f = 8, n_m = n_f \)

<table>
<thead>
<tr>
<th align="left">((n_m, n_f))</th>
<th>(n_d)</th>
<th align="left">(h = \frac{1}{2})</th>
<th align="left">(h = 1)</th>
<th align="left">(h = 2)</th>
<th align="left">(N = n)</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">:----------------</td>
<td>---------</td>
<td align="left">:-----------------</td>
<td align="left">:-----------------</td>
<td align="left">:-----------------</td>
<td align="left">:-----------------</td>
</tr>
<tr>
<td align="left">((1,1))</td>
<td>2</td>
<td align="left">2</td>
<td align="left">10</td>
<td align="left">3</td>
<td align="left">11</td>
</tr>
<tr>
<td align="left">((2,2))</td>
<td>4</td>
<td align="left">6</td>
<td align="left">23</td>
<td align="left">7</td>
<td align="left">26</td>
</tr>
<tr>
<td align="left">((3,3))</td>
<td>6</td>
<td align="left">10</td>
<td align="left">34</td>
<td align="left">12</td>
<td align="left">41</td>
</tr>
<tr>
<td align="left">((4,4))</td>
<td>8</td>
<td align="left">14</td>
<td align="left">46</td>
<td align="left">16</td>
<td align="left">56</td>
</tr>
<tr>
<td align="left">((5,5))</td>
<td>10</td>
<td align="left">18</td>
<td align="left">59</td>
<td align="left">21</td>
<td align="left">71</td>
</tr>
<tr>
<td align="left">((6,6))</td>
<td>12</td>
<td align="left">23</td>
<td align="left">73</td>
<td align="left">26</td>
<td align="left">86</td>
</tr>
</tbody>
</table>

Table 6.16. Median life of polymorphism and \( t_{95} \) for recurrent mass selection in a dioecious system; \( N_m = N_f = 8, n_m \neq n_f \)

<table>
<thead>
<tr>
<th align="left">(n)</th>
<th align="left">((n_m, n_f))</th>
<th>(n_d)</th>
<th align="left">(h = \frac{1}{2})</th>
<th align="left">(h = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td align="left">:-----</td>
<td align="left">:----------------</td>
<td>---------</td>
<td align="left">:-----------------</td>
<td align="left">:-----------------</td>
</tr>
<tr>
<td align="left">9</td>
<td align="left">(1,8)</td>
<td>3.56</td>
<td align="left">7</td>
<td align="left">22</td>
</tr>
<tr>
<td align="left">10</td>
<td align="left">(2,8)</td>
<td>6.40</td>
<td align="left">12</td>
<td align="left">39</td>
</tr>
<tr>
<td align="left">11</td>
<td align="left">(3,8)</td>
<td>8.73</td>
<td align="left">17</td>
<td align="left">54</td>
</tr>
<tr>
<td align="left">12</td>
<td align="left">(4,8)</td>
<td>10.67</td>
<td align="left">21</td>
<td align="left">66</td>
</tr>
<tr>
<td align="left">13</td>
<td align="left">(5,8)</td>
<td>12.31</td>
<td align="left">24</td>
<td align="left">77</td>
</tr>
<tr>
<td align="left">14</td>
<td align="left">(6,8)</td>
<td>13.71</td>
<td align="left">27</td>
<td align="left">88</td>
</tr>
</tbody>
</table>

| 4    | (1,3)          | 3.00    | 5                | 17              | 5               | 20              |
| 6    | (1,5)          | 3.33    | 7                | 20              | 7               | 23              |
| 6    | (2,4)          | 5.33    | 9                | 31              | 10              | 36              |
| 8    | (1,7)          | 3.50    | 7                | 22              | 7               | 24              |
| 8    | (2,6)          | 6.00    | 11               | 36              | 12              | 42              |
| 8    | (3,5)          | 7.50    | 13               | 44              | 15              | 52              |
| 10   | (3,7)          | 8.40    | 16               | 50              | 18              | 59              |
| 10   | (4,6)          | 9.60    | 18               | 57              | 20              | 68              |
| 12   | (5,7)          | 11.67   | 22               | 71              | 25              | 83              |
Turning now to recurrent disruptive selection, a single cycle of which was analyzed in Section 5.12, the transition matrix method derived for recurrent selection in a dioecious system applies directly; it is only necessary to change from $n_m$ and $n_f$ directionally selected from $N_m$ and $N_f$, to $n_h$ and $n_g$ directionally selected from $N$ with corresponding standardized selection differentials $k_h (>0)$ and $k_g (<0)$ respectively. Allelic disequilibrium induced by disruptive selection is then properly accounted through repeated cycles of disruptive selection with $k_h$ and $k_g$ held constant.

Only symmetric disruptive selection ($k_g = -k_h$) in an additive model ($d = 0$) is considered here. Results were obtained for $N = 32$, $n_h = n_g = 2, 4$ and $6$, $p^h_0 = p^g_0 = 0.5$ and $(a/\sigma_X) = 0.1$ and $0.2$.

The capacity of disruptive selection to maintain polymorphism should be reflected in the distribution of $T$, the life of polymorphism. Comparison of median life $t_{50}$, under disruptive selection with $(a/\sigma_X) = 0.1$, with $t_{50}$ for neutral genes in a monoecious system having the same parental census number $(n_h + n_g) = n = N$, yields identical integer numbers of generations. Increasing $(a/\sigma_X)$ to $0.2$, close to the limit of the polygenetic range, yields the same result. Thus, as measured by $t_{50}$, symmetric disruptive selection does not prolong life of polymorphism when compared with no selection for census numbers $n = 4, 8$ and $12$ which are not atypical in experimental disruptive selection (see e.g., Thoday, 1959).

The notion that disruptive selection is a mechanism for maintenance of polymorphism in finite populations is a reduced drift variance per generation as at (5.12.1). Values of $\sigma^2_t = \mathcal{V}(p_t)$ are given in Table 6.17 for symmetric disruptive selection and for neutral genes. The small
Table 6.17. Values of $\sigma^2_t$ for symmetric disruptive selection and for neutral genes; $(a/\sigma_X) = 0.2$, $p_0 = 0.5$, $d = 0$

<table>
<thead>
<tr>
<th>Gen. $t$</th>
<th>$N = 4$</th>
<th>$N = 32$</th>
<th>$N = 8$</th>
<th>$N = 32$</th>
<th>$N = 12$</th>
<th>$N = 32$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 4$</td>
<td>$n_h = n_k = 2$</td>
<td>$n = 8$</td>
<td>$n_h = n_k = 4$</td>
<td>$n = 12$</td>
<td>$n_h = n_k = 6$</td>
</tr>
<tr>
<td>1</td>
<td>0.03125</td>
<td>0.03017</td>
<td>0.01562</td>
<td>0.01523</td>
<td>0.01042</td>
<td>0.01022</td>
</tr>
<tr>
<td>2</td>
<td>0.05859</td>
<td>0.05703</td>
<td>0.03027</td>
<td>0.02961</td>
<td>0.02040</td>
<td>0.02004</td>
</tr>
<tr>
<td>3</td>
<td>0.08252</td>
<td>0.08062</td>
<td>0.04401</td>
<td>0.04312</td>
<td>0.02997</td>
<td>0.02947</td>
</tr>
<tr>
<td>4</td>
<td>0.10345</td>
<td>0.10134</td>
<td>0.05688</td>
<td>0.05581</td>
<td>0.03913</td>
<td>0.03852</td>
</tr>
<tr>
<td>5</td>
<td>0.12177</td>
<td>0.11953</td>
<td>0.06895</td>
<td>0.06773</td>
<td>0.04792</td>
<td>0.04719</td>
</tr>
<tr>
<td>10</td>
<td>0.18423</td>
<td>0.18207</td>
<td>0.11888</td>
<td>0.11722</td>
<td>0.08665</td>
<td>0.08554</td>
</tr>
<tr>
<td>20</td>
<td>0.23270</td>
<td>0.23159</td>
<td>0.18124</td>
<td>0.17955</td>
<td>0.14327</td>
<td>0.14189</td>
</tr>
<tr>
<td>50</td>
<td>0.24968</td>
<td>0.24963</td>
<td>0.24008</td>
<td>0.23948</td>
<td>0.22023</td>
<td>0.21929</td>
</tr>
<tr>
<td>100</td>
<td>0.25000</td>
<td>0.25000</td>
<td>0.24961</td>
<td>0.24956</td>
<td>0.24646</td>
<td>0.24623</td>
</tr>
<tr>
<td>$\infty$</td>
<td>0.25000</td>
<td>0.25000</td>
<td>0.25000</td>
<td>0.25000</td>
<td>0.25000</td>
<td>0.25000</td>
</tr>
</tbody>
</table>

Reduction in $\sigma^2_t$ induced by disruptive selection translates into little difference in the actual distribution function of $T$ obtained as

$$Pr[T < t] = v(2n_h, 2n_k; t+1) + v(0, 0; t+1)$$

and

$$Pr[T < t] = v(2n; t+1) + v(0; t+1)$$

for disruptive selection and for neutral genes respectively, and listed in Table 6.18.

Examination of the fractiles in Table 6.18 indicates that life of polymorphism is stochastically increased in symmetric disruptive selection but so small is this increase when $n = 4, 8$ and $12$ that percentiles of the discrete distribution of $T$ do not differ significantly from those for neutral genes in a monoecious system with the same parental census number.
Table 6.18. Distribution function $\Pr[T \leq t]$ for symmetric disruptive selection and for neutral genes; $(a/\sigma_x) = 0.2$, $p_0 = 0.5$, $d = 0$

<table>
<thead>
<tr>
<th>Life t</th>
<th>$N = 4$</th>
<th>$N = 32$</th>
<th>$N = 8$</th>
<th>$N = 32$</th>
<th>$N = 12$</th>
<th>$N = 32$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 4$</td>
<td>$n_h = n_x = 2$</td>
<td>$n = 8$</td>
<td>$n_h = n_x = 4$</td>
<td>$n = 12$</td>
<td>$n_h = n_x = 6$</td>
</tr>
<tr>
<td>5</td>
<td>0.3341</td>
<td>0.3238</td>
<td>0.0694</td>
<td>0.0663</td>
<td>0.0143</td>
<td>0.0136</td>
</tr>
<tr>
<td>10</td>
<td>0.6562</td>
<td>0.6455</td>
<td>0.2842</td>
<td>0.2770</td>
<td>0.1209</td>
<td>0.1174</td>
</tr>
<tr>
<td>15</td>
<td>0.8236</td>
<td>0.8154</td>
<td>0.4760</td>
<td>0.4675</td>
<td>0.2654</td>
<td>0.2600</td>
</tr>
<tr>
<td>20</td>
<td>0.9095</td>
<td>0.9039</td>
<td>0.6198</td>
<td>0.6113</td>
<td>0.3998</td>
<td>0.3933</td>
</tr>
<tr>
<td>25</td>
<td>0.9536</td>
<td>0.9500</td>
<td>0.7246</td>
<td>0.7168</td>
<td>0.5131</td>
<td>0.5063</td>
</tr>
<tr>
<td>30</td>
<td>0.9762</td>
<td>0.9740</td>
<td>0.8005</td>
<td>0.7937</td>
<td>0.6059</td>
<td>0.5932</td>
</tr>
<tr>
<td>35</td>
<td>0.9878</td>
<td>0.9864</td>
<td>0.8555</td>
<td>0.8497</td>
<td>0.6813</td>
<td>0.6749</td>
</tr>
<tr>
<td>40</td>
<td>0.9937</td>
<td>0.9929</td>
<td>0.8954</td>
<td>0.8906</td>
<td>0.7424</td>
<td>0.7364</td>
</tr>
<tr>
<td>45</td>
<td>0.9968</td>
<td>0.9963</td>
<td>0.9242</td>
<td>0.9203</td>
<td>0.7918</td>
<td>0.7863</td>
</tr>
<tr>
<td>50</td>
<td>0.9984</td>
<td>0.9981</td>
<td>0.9451</td>
<td>0.9419</td>
<td>0.8317</td>
<td>0.8267</td>
</tr>
<tr>
<td>75</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9891</td>
<td>0.9881</td>
<td>0.9419</td>
<td>0.9393</td>
</tr>
<tr>
<td>100</td>
<td>a</td>
<td>a</td>
<td>0.9978</td>
<td>0.9976</td>
<td>0.9800</td>
<td>0.9787</td>
</tr>
</tbody>
</table>

*Indicates $\Pr[T \leq t] > 0.99995$. 

---

**Note:**
- Table values are approximate and may vary slightly from the exact calculations.
- The table entries are based on theoretical calculations and may not reflect exact probabilities due to rounding and approximation methods.
7. THE DOUBLE POLYGENE MODEL

Low moments of average response, means and variances, derived in Chapter 3 in the general context of linear responses, are applied here to models of two polygenes with additive, dominance and epistatic effects for a single cycle of phenotypic selection. Algebraic methods are analogous to those used for single polygenes in Chapter 5 so that completeness of algebraic details in the derivations is unnecessary.

7.1 Reference Population $\Pi_0$

Parameters of the two loci are indicated by subscripts 1 and 2. At locus $i$ three diploid genotypes are generated by gene $a_i^+$ and its alleles $a_i^-$, $i = 1, 2$. The model for $\Pi_0$ adopted as a reference population for $N$ randomly sampled test zygotes assumes random mating without selection among parents of test material; $p_i$ denotes the frequency of gene $a_i^+$ in this parental group and $q_i = 1 - p_i$, $i = 1, 2$. Irrespective of the genotypic composition of the parental group, or its size, genotypic composition of $\Pi_0$ is defined to follow Hardy-Weinberg proportions marginally at each locus, but linkage equilibrium within randomly united gametes is not assumed. Recombination fraction is denoted by $r$, $0 \leq r \leq \frac{1}{2}$, and linkage disequilibrium characterized by $\Lambda$ the covariance between genes within gametes. Then denoting gametic frequencies, in gametes uniting to form zygotes of $\Pi_0$, by $\theta_{11}$, $\theta_{10}$, $\theta_{01}$, $\theta_{00}$ for $(++)$, $(+-)$, $(-+)$ and $(--)$ types respectively,

$$\theta_{11} = p_1p_2 + \Lambda, \quad \theta_{10} = p_1q_2 - \Lambda$$

$$\theta_{01} = q_1p_2 - \Lambda, \quad \theta_{00} = q_1q_2 + \Lambda$$

(7.1.1)
For some purposes it is convenient to express covariance $\Lambda$ as

$$\Lambda = \lambda \sqrt{p_{1}q_{1}} \sqrt{p_{2}q_{2}},$$

(7.1.2)

where $\lambda$ is the correlation between genes within uniting gametes.

Effects at the two loci are assumed to influence trait $X$ the basis for phenotypic ranking in test material. Marginal trait $X$ means per genotype are shown in Table 7.1.

**Table 7.1. Trait $X$ means per genotype and indicator variables for the two loci model**

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Frequencies</th>
<th>$X$-means</th>
<th>$Z_1$</th>
<th>$Z_2$</th>
<th>$Z_3$</th>
<th>$Z_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(++)$</td>
<td>$\theta^{2}_{11}$</td>
<td>$a_1 + a_2 - e$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>$(+-)$</td>
<td>$\theta^{2}_{10}$</td>
<td>$a_1 - a_2 + e$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$(+\mp)$</td>
<td>$\theta^{2}_{01}$</td>
<td>$-a_1 + a_2 + e$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$(\mp\mp)$</td>
<td>$\theta^{2}_{00}$</td>
<td>$-a_1 - a_2 - e$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$(++)$</td>
<td>$2^{6} \theta^{10}_{11}$</td>
<td>$a_1 + d_2$</td>
<td>1</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
</tr>
<tr>
<td>$(+-)$</td>
<td>$2^{6} \theta^{10}_{10}$</td>
<td>$d_1 + a_2$</td>
<td>1/2</td>
<td>1</td>
<td>1/2</td>
<td>0</td>
</tr>
<tr>
<td>$(+\mp)$</td>
<td>$2^{6} \theta^{10}_{00}$</td>
<td>$d_1 - a_2$</td>
<td>1/2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$(\mp\mp)$</td>
<td>$2^{6} \theta^{10}_{00}$</td>
<td>$-a_1 + d_2$</td>
<td>0</td>
<td>1/2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$(++)$</td>
<td>$2^{6} \theta^{01}_{11}$</td>
<td>$d_1 + d_2$</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>$(+-)$</td>
<td>$2^{6} \theta^{01}_{10}$</td>
<td>$d_1 + d_2$</td>
<td>1/2</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
</tr>
<tr>
<td>$(+\mp)$</td>
<td>$2^{6} \theta^{01}_{00}$</td>
<td>$d_1 + d_2$</td>
<td>1/2</td>
<td>1/2</td>
<td>0</td>
<td>-1/2</td>
</tr>
</tbody>
</table>
Included in this model, which is a simple mapping of that given by Kuehl et al. (1968), are additive and dominance effects at each locus together with additive x additive epistatic effects. Without loss of generality it is assumed that $a_i$ is non-negative identifying "favorable" allele $a_i^+$ with frequency $p_i$, $i = 1, 2$.

Genotypic mean of $\Pi_0$ is obtained from Table 7.1 as

$$\mu_x = \sum_{i=1}^{2} \{(p_i - q_i) a_i + 2p_i q_i d_i\} - \{(p_1 - q_1)(p_2 - q_2) + 2\Lambda\} e.$$  \hfill (7.1.3)

In the case of linkage equilibrium ($\Lambda = 0$) genotypic variance may be orthogonally partitioned using the orthogonal scales of Cockerham (1954):

- **additive component:** $\sigma_a^2 = \sigma_{a_1}^2 + \sigma_{a_2}^2$, $\sigma_{a_1}^2 = 2p_i q_i A_i^2$

- **dominance component:** $\sigma_d^2 = \sigma_{d_1}^2 + \sigma_{d_2}^2$, $\sigma_{d_1}^2 = (2p_i q_i d_i)^2$  \hfill (7.1.4)

- **epistatic component:** $\sigma_{aa}^2 = 4p_1 p_2 q_1 q_2 e^2$,

where $A_i = [a_i - (p_i - q_i) d_i - (p_j - q_j) e]$, $i \neq j$, is the allelic substitution effect for locus $i = 1, 2$.

With linkage disequilibrium ($\Lambda \neq 0$) orthogonal partition of genotypic variance is **not** possible. However, genotypic variance may be partitioned into the components at (7.1.4) together with a residual composed of covariance components attributable to correlation between genes within gametes. For example, the additional component attributable to covariance between additive effects at the two loci is $2\sigma_{a_1 a_2}^2 = 4\Lambda A_1 A_2$. Since gametes are united at random not all of the effects are correlated with each other. Additive effects at locus $i$ are not correlated with dominance effects at locus $j$, but dominance effects at the two loci are
correlated. Additionally, additive effects are correlated with those epistatic effects expressed within gametes whereas dominance effects are correlated with those epistatic effects expressed between gametes.

Using these concepts total genotypic variance may be written as

\[
\{\sigma_\alpha + 2\sigma_\alpha a_1 a_2 + 2\sigma_\alpha e^2\} \\
+ \{\sigma_\delta^2 + 2\sigma_\delta a_1 a_2 + 2\sigma_\delta e^2\} \\
+ \{\sigma_{\alpha\alpha} + 2\sigma_{\alpha e}\},
\]

where \(\sigma_\alpha^2\), \(\sigma_\delta^2\) and \(\sigma_{\alpha\alpha}^2\) are the orthogonal components of an equilibrium population as defined at (7.1.4) and the remaining terms are covariance components defined as follows:

\[
\sigma_\alpha a_1 a_2 = \lambda \sigma_{a_1} \sigma_{a_2}, \quad \sigma_\delta a_1 a_2 = \lambda^2 \sigma_{\delta_1} \sigma_{\delta_2} \\
\sigma_{ae} = (\sigma_{a_1 e} + \sigma_{a_2 e}) \quad \text{where} \quad \sigma_{a_1 e} = \lambda \sigma_{a_1} \sigma_{e_1} \\
\sigma_{\delta e} = (\sigma_{\delta_1 e} + \sigma_{\delta_2 e}) \quad \text{where} \quad \sigma_{\delta_1 e} = \lambda \sigma_{\delta_1} \sigma_{\alpha} \\
\sigma_{ee} = \frac{1}{2} \lambda \sigma_{e_1} \sigma_{e_2} \quad \text{where} \quad \sigma_{e_1}^2 = 2p_jq_j(p_i-q_i)^2 \cdot e^2.
\]

Although \(a_1\) and \(a_2\) have been defined to be non-negative, parameters \(d_1\) and \(e\) may be positive, zero or negative. In writing the genotypic standard deviations as \(\sigma_{\delta_1}\), \(\sigma_{\alpha}\), \(\sigma_{e_1}\) in (7.1.6) the convention is adopted that they take the sign of their corresponding parameter. Thus, \(\sigma_{\delta_1}\) is positive or negative as \(d_1\) is positive or negative, and \(\sigma_{\alpha}\), \(\sigma_{e_1}\) are all positive or all negative as \(e\) is positive or negative.

The same assumptions regarding Normal phenotypic distributions per genotype, and regarding magnitudes of polygenic effects, are made for
this two loci model as were made for the single polygene model in Section 5.1, and with similar consequences. The overall phenotypic distribution in $\Pi_0$ has standardized density $f(x)$ treated as the unit Normal density on neglecting terms analogous to $O(g)^2$ as before, and expansions similar to those at (5.1.6) may be curtailed to provide density functions for the genotypic distributions analogous to those in Section 5.1; for example

$$f(\frac{x - a_1 - a_2 + e}{\sigma_x}) = f(x)[1 + x(a_1 + a_2 - e - \mu_x)]$$  \hspace{1cm} (7.1.7)

for $^{++}$ genotypes.

Henceforth it is algebraically convenient to set $\sigma_x = 1$ as was done by Kojima (1961) for the single polygene model. Then it is necessary to remember that genotypic standard deviations have been expressed as proportions of the phenotypic standard deviation $\sigma_x$, for example $\sigma_{a_1}$ is then a shortened notation for $\sigma_{a_1}/\sigma_x$.

Thus far nothing has been said about response variable $Y$ in $\Pi_r$. As was the case for single polygenes, consequences of selection are completely characterized by specifying $Y$ as a vector of genotypic indicator random variables. Different aspects of response to selection may then be defined as functions of $Y$, incorporating utilization variations as necessary. Here it is sufficient to work with four such functions, denoted by $Z_1$, $Z_2$, $Z_3$ and $Z_4$, taking the values indicated in Table 7.1 which is read as follows: a random zygote from $\Pi_0$ is assigned that value of $Z_1$ indicated by its genotype.

$Z_1$ is gene $a^+_1$ frequency indicator and $Z_2$ is gene $a^+_2$ frequency indicator, both being analogous to $Z$ in Section 5.1. $Z_3$ is the gametic
frequency indicator for gametes of type (++) \( Z_4 \) contrasts frequencies of the two types of double heterozygotes, "coupling" \( \frac{\cdot \cdot}{\cdot \cdot} \), and "repulsion" \( \frac{\cdot \cdot}{\cdot .} \). Using the frequencies in Table 7.1,

\[
\begin{align*}
E[Z_3] &= \theta_{11} = p_1p_2 + \Lambda, \quad E[Z_4] = \Lambda \\
E[Z_1] &= p_1, \quad E[Z_2] = p_2 \\
V[Z_1] &= \frac{1}{2} p_1q_1, \quad V[Z_2] = \frac{1}{2} p_2q_2 \\
\lambda &= \frac{C[Z_1, Z_2]}{\left[ V[Z_1] \cdot V[Z_2] \right]^{1/2}}.
\end{align*}
\]

(7.1.8)

Conditional densities, \( \pi(Z_i | X) \), for \( Z_i \) given \( X, i = 1, 2, 3, 4 \) are obtained analogous to those at (5.1.12) for \( Z \), and provide the following conditional expectations and conditional variances and covariance:

\[
\begin{align*}
E[Z_3 | X] &= \theta_{11} + \theta_{11} x[q_1A_1 + q_2A_2 - (q_1q_2 - \Lambda) e] \\
E[Z_4 | X] &= \Lambda - \Lambda x \left[ \sum_{i=1}^{2} (p_i - q_i)A_1 - 2p_1q_1d_1 + \frac{1}{2} (p_i - q_i)(p_j - q_j)e \right] \\
&\quad - 2\Lambda e; \\
E[Z_1 | X] &= p_1 + x[p_1q_1A_1 + \Lambda A_2 + (p_1 - q_1)e] \\
V[Z_1 | X] &= \frac{1}{2} p_1q_1 - \frac{1}{2} x(p_1 - q_1)[p_1q_1A_1 + \Lambda A_2 + (p_1 - q_1)e] \\
&\quad - \frac{1}{2} p_1q_1 x\left(2p_1q_1d_1 + 2\Lambda^2 p_2q_2d_2 \right) \\
&\quad + \left((p_1 - q_1)(p_2 - q_2) + 2\Lambda\right)e \\
C[Z_1, Z_2 | X] &= \frac{1}{2} \Lambda - \frac{1}{2} x(2p_1q_1p_2q_2 e) \\
&\quad - \frac{1}{2} x \left( \sum_{i=1}^{2} (p_i - q_i)A_1 + 2p_1q_1d_1 + \frac{1}{2} (p_i - q_i)(p_j - q_j)e \right).
\end{align*}
\]

(7.1.11)
As was the case for single polygenes, the effect of the polygenic assumption is to linearize response functions $Z_i$ of $Y$, when regressed on $X$ neglecting terms $0(g)^2$. Identifying genotypic standard deviations at (7.1.4) and (7.1.6), and standardizing $Z_i$ to $z_i = (Z_i - p_i)/\sqrt{\frac{1}{2} p_i q_i}$ in order to provide a direct comparison with the linearity assumptions at (3.1.1) and (3.1.3), (7.1.10) provides

$$E(z_i | X) = \{\sigma_{a_1} + \lambda(\sigma_{a_2} + \sigma_{e_1})\} x$$

$$= \rho_i x, \text{ say}$$

$$\lambda = \frac{\sigma_{a_1}}{\sqrt{2p_i q_i}}$$

$V(z_i | X) = 1 - x\left[\frac{(p_i - q_i)}{\sqrt{2p_i q_i}}\right]^2 + \lambda^2\sigma_{a_2} + \lambda \sigma_{e_1} + \frac{2\sigma_{e_1}}{\lambda \sigma_{a_2}}$

with symmetrical results for moments $E(z_2 | X)$ and $V(z_2 | X)$ referring to $p_2$, and

$$C(z_i, z_2 | X) = (\lambda - x\sigma_{a_2}) - x \lambda \sum_{i=1}^{2} \left[\frac{(p_i - q_i)}{\sqrt{2p_i q_i}}\right] \left(\sigma_{a_1} + \frac{1}{2} \sigma_{e_1} + \sigma_{\delta_1}\right).$$

### 7.2 Frequencies in Selected Zygotes

Section 7.1 contains all the ingredients necessary for use of results in Chapters 2 and 3 relating to linear responses to selection: $f(x)$ is the unit Normal density providing parameters $k$ and $v$ for any rank set $R_n$ defining selected zygotes and (7.1.9), (7.1.12), (7.1.13) provide the required conditional expectations, variances and covariance in $\Pi_0$. It remains to apply the recipe for pre-test moments at (3.4.3) and (3.5.5) for "average response." Assigning $Z_{jr}$ to test zygote ranked $r$th on the basis $X(r)$, average responses

$$\frac{1}{n} \sum_{r \in R_n} Z_{jr}, \quad j = 1, 2, 3, 4,$$
are the required frequencies (random variables) in the group of \( n \) selected zygotess. The following notation is adopted:

\[
(p_1 + \delta p_1) = \frac{1}{n} \sum_{R} Z_{1r} = \text{gene } a^+ \text{ frequency in selected group,}
\]

\[
(p_2 + \delta p_2) = \frac{1}{n} \sum_{R} Z_{2r} = \text{gene } a^+ \text{ frequency in selected group,}
\]

\[
(\theta_{11} + \delta \theta_{11}) = \frac{1}{n} \sum_{R} Z_{3r} = \text{gamete } (++ \text{ frequency in selected group,}
\]

\[
\theta_{cr} = \frac{1}{n} \sum_{R} Z_{4r} = \text{contrast of frequencies of coupling and repulsion double heterozygotes in selected group.}
\]

Then applying (3.4.3) to (7.1.12) yields

\[
E(\delta p_1) = kp_1 \sqrt{\frac{1}{2} p_1 q_1} , \rho_1 = \{\sigma_{a_1} + \lambda(\sigma_{a_2} + \sigma_{e_1})\}
\]

\[
E(\delta p_2) = kp_2 \sqrt{\frac{1}{2} p_2 q_2} , \rho_2 = \{\sigma_{a_2} + \lambda(\sigma_{a_1} + \sigma_{e_2})\}
\]

(7.2.1)

Applying (3.5.5) to (7.1.12) and introducing variance effective numbers \( n_1 \) and \( n_2 \),

\[
V(\delta p_i) = \frac{p_i q_i}{2n_i} , \ i = 1,2
\]

\[
\frac{n}{n_i} = 1 - k[\frac{(p_i - q_i)}{\sqrt{2p_i q_i}} \rho_i + (\sigma_{\delta_{1j}} + \lambda^2 \sigma_{\delta_{jj}} + \lambda \sigma_{a_a} + \frac{2\sigma_{ee}}{\lambda \sigma_{a_a}}) , \ i \neq j
\]

(7.2.2)

Applying (3.4.3) to (7.1.9) yields

\[
E(\theta_{cr}) = \Lambda(1 - k(u_x - d_1 - d_2))
\]

(7.2.3)

\[
E(\delta \theta_{11}) = k\theta_{11}(q_1 A_1 + q_2 A_2 - (q_1 q_2 - \Lambda)e)
\]

(7.2.4)

Also, using (7.1.13)
\[ C(\delta p_1, \delta p_2) = E[C(\frac{1}{n} \sum_{r} Z_{1r}, \frac{1}{n} \sum_{r} Z_{2r} | X_r; rz_{R_n})] + O(\gamma)^2 \]

\[ = \frac{\sqrt{p_1 q_1 p_2 q_2}}{2n^2} E[\sum_{r} C(z_{1r}, z_{2r} | X_r)] \]

\[ = \frac{\sqrt{p_1 q_1 p_2 q_2}}{2n} E[\lambda - \Delta \sum_{r} \alpha = \Delta \lambda \sum_{i=1}^{2} \frac{(p_i - q_i)}{\sqrt{2p_i q_i}} \]

\[ \cdot (\sigma_{\alpha_i} + \frac{1}{2} \sigma_{\epsilon_i} \sigma_{\delta_i}) ] \]

\[ = \frac{\Lambda}{2n_c} - k \frac{\sqrt{p_1 q_1 p_2 q_2}}{2n} \sigma_{\alpha} , \quad (7.2.5) \]

where \( n_c \) is introduced as a "covariance effective number" defined by

\[ \frac{n}{n_c} = 1 - k \sum_{i=1}^{2} \frac{(p_i - q_i)}{\sqrt{2p_i q_i}} (\sigma_{\alpha_i} + \sigma_{\epsilon_i} + \sigma_{\delta_i}) , \quad i \neq j . \]

This representation is motivated as follows. Under nonselective sampling of a group of \( n \) random zygotes from \( \Pi_0 \), gene frequencies are correlated only by virtue of linkage disequilibrium as at (7.1.8):

\[ C(\delta p_1, \delta p_2) = \frac{\Lambda}{2n} = \frac{1}{2n} E[\theta_{cr}] . \quad (7.2.6) \]

Thus in (7.2.5), \( n_c \) effectively accounts for that portion of \( C(\delta p_1, \delta p_2) \) attributable to linkage disequilibrium as modified by selection. The second term of \( C(\delta p_1, \delta p_2) \) at (7.2.5) is a consequence of selection that occurs even if \( \Pi_0 \) is in linkage equilibrium (\( \Lambda = 0 \)), and is therefore omitted from the formulation of \( n_c \).

Neeley (1971) provides a detailed description of various forms of disequilibria among genes, gametes and zygotes that are induced by truncation selection in infinite populations for several loci with only
additive effects. Here it is sufficient to replace (7.2.6) by

$$E(\theta_{cr}) = 2n \ C(\delta p_1, \delta p_2) + k \ \sqrt{p_1 q_1 p_2 q_2} \ \{(1+\lambda^2) \sigma_{aa} + 2\lambda (\sigma_1 + \sigma_2)\}$$

$$= \Lambda \ \frac{n}{n_c} + k \ \Lambda \{(\lambda \sigma_{aa} + 2(\sigma_1 + \sigma_2))\}$$

(7.2.7)

obtained from (7.2.3) and (7.2.5), and to consider just $\Lambda^S$ the linkage disequilibrium in gametes as constituted in zygotes of the selected group:

$$\Lambda^S = (\theta_{11} + \delta \theta_{11}) - (p_1 + \delta p_1)(p_2 + \delta p_2)$$

$$= \Lambda - (\delta p_1)(\delta p_2) + \{\delta \theta_{11} - p_1 (\delta p_2) - p_2 (\delta p_1)\}.$$

Thus,

$$E(\Lambda^S) = \Lambda - E(\delta p_1) \cdot E(\delta p_2) - C(\delta p_1, \delta p_2)$$

$$+ E(\delta \theta_{11} - p_1 (\delta p_2) - p_2 (\delta p_1))$$

(7.2.8)

and after some algebraic reduction the last term is obtained from (7.2.4) and (7.2.1) in the form

$$\Lambda + E(\delta \theta_{11} - p_1 (\delta p_2) - p_2 (\delta p_1)) = \frac{1}{2} \ E(\theta_{cr}) + n \ C(\delta p_1, \delta p_2)$$

$$= \Lambda \ \frac{n}{n_c} - \frac{1}{2} \ k \ \sqrt{p_1 q_1 p_2 q_2} \ \{(1-\lambda^2) \sigma_{aa} - 2\lambda (\sigma_1 + \sigma_2)\}.$$

(7.2.9)

Then in (7.2.8),

$$E(\Lambda^S) = -E(\delta p_1) \cdot E(\delta p_2) - \frac{1}{2} \ k \ \sqrt{p_1 q_1 p_2 q_2} \ \{(1-\lambda^2) \sigma_{aa} - 2\lambda (\sigma_1 + \sigma_2)\}$$

$$+ \Lambda (1 - \frac{1}{2n}) \ \frac{n}{n_c} + k \ \frac{\sqrt{p_1 q_1 p_2 q_2}}{2n} \ \sigma_{aa}$$

(7.2.10)

the first term of which is negligible, being $O(g)^2$. Thus for moderately large $n$, $E(\Lambda^S - \Lambda)$ is dominated by a single term which is zero in the absence of dominance and epistatic polygenic effects.
7.3 Reference Population \( \Pi_r \)

A convenient reference population for assessment of consequences of phenotypic selection defined by \( R_n \), is the large offspring population obtained by random mating among zygotes in the selected group. Then \( \Pi_r \) is formulated similar to \( \Pi_0 \) in Section 7.1 but with gene frequencies \( (p_1 + \delta p_1), (p_2 + \delta p_2) \) and linkage disequilibrium \( \Lambda^r \). Thus it is convenient to write \( \Pi_0 \equiv \Pi_0(p_1, p_2, \Lambda) \) and \( \Pi_r \equiv \Pi_0(p_1 + \delta p_1, p_2 + \delta p_2, \Lambda^r) \).

Moments of gene frequencies in \( \Pi_r \) are those provided at (7.2.1) but \( E(\Lambda^r) \) has yet to be derived with an appropriate accounting of recombination in the meiotic formation of gametes output by the selected group of parents of \( \Pi_r \). For any general genotypic array used as parents of a large gametic pool, linkage disequilibrium in the pool is equal to linkage disequilibrium in gametes as constituted in the parental array minus the product of the recombination fraction and the contrast of frequencies of coupling and repulsion double heterozygotes in the parental array (Crow and Kimura, 1970). Thus in the gametic pool forming \( \Pi_r \),

\[
\Lambda^r = \Lambda^s - r \theta_{cr}
\]

with expectation provided by (7.2.10) and (7.2.7):

\[
E(\Lambda^r) = -E(\delta p_1) \cdot E(\delta p_2) + (1 - r - \frac{1}{2n}) \Lambda \frac{n}{n_c}
\]

\[
- \frac{1}{2} k \sqrt{p_1 q_1 p_2 q_2} \left[ (1 - \frac{1}{n}) \sigma_{\alpha \alpha} - (1 - 2r) \{ \lambda^2 \sigma_{\alpha \alpha} + 2\lambda (\sigma_{\delta_1} + \sigma_{\delta_2}) \} \right]
\]

(7.3.1)

the first term of which is negligible being \( O(g)^2 \). With free recombination, \( r = 1/2 \), this reduces to

\[
E(\Lambda^r) = -E(\delta p_1) \cdot E(\delta p_2) + (1 - \frac{1}{2} (1 + \frac{1}{n}) \Lambda \frac{n}{n_c} - \frac{1}{2} k \sqrt{p_1 q_1 p_2 q_2} (1 - \frac{1}{n}) \sigma_{\alpha \alpha}
\]
\[ E(\Lambda^R) = -E(\delta p_1) \cdot E(p_2) + (1 - \frac{1}{2n}) \Lambda \frac{n}{n_c} \]
\[ - \frac{1}{2} k \sqrt{p_1 q_1 p_2 q_2} \left\{ (1 - \lambda^2 - \frac{1}{n})\sigma_{aa} - 2\lambda(\sigma_{d_1} + \sigma_{d_2}) \right\} \]
as \( r \to 0 \). Finally if \( \Pi_0 \) is initially in linkage equilibrium, \( \Lambda = 0 \),
\[ E(\Lambda^R) = -E(\delta p_1) \cdot E(\delta p_2) - \frac{1}{2} k \sqrt{p_1 q_1 p_2 q_2} (1 - \frac{1}{n})\sigma_{aa} \]
irrespective of recombination.

### 7.4 Trait Mean Response

Trait X mean for \( \Pi_r = \Pi_0(p_1 + \delta p_1, p_2 + \delta p_2, \Lambda^R) \) is derived from (7.1.3):
\[ (\mu_x + \delta \mu_x) = \frac{2}{\Sigma} \left\{ p_1 - q_1 + 2\delta p_1 a_1 + 2(p_1 + \delta p_1)(q_1 - \delta p_1) d_1 \right\} \]
\[ \quad - \{(p_1 - q_1 + 2\delta p_1)(p_2 - q_2 + 2\delta p_2) + 2\Lambda^R\} e \]
so that
\[ \delta \mu_x = 2\{(\delta p_1)A_1 + (\delta p_2)A_2\} - 2\{(\delta p_1)^2 d_1 + (\delta p_2)^2 d_2\} \]
\[ - 4e(\delta p_1)(\delta p_2) - 2e(\Lambda^R - \Lambda) \tag{7.4.1} \]
and expected trait X mean response
\[ E(\delta \mu_x) = 2[A_1 E(\delta p_1) + A_2 E(\delta p_2)] \]
\[ - 2[d_1 E^2(\delta p_1) + d_2 E^2(\delta p_2) + 2eE(\delta p_1) E(\delta p_2)] \]
\[ - 2[d_1 V(\delta p_1) + d_2 V(\delta p_2) + 2eC(\delta p_1, \delta p_2)] \]
\[ - 2eE(\Lambda^R - \Lambda) \tag{7.4.2} \]
for which all required moments have been derived in Sections 7.2 and 7.3.
7.5 Consequences of No Selection

For the case of no selection defined by \( R_n \equiv \{1,2,\ldots,N\}, k = 0 \) and \( n_1 = n_2 = n_c = n = N \). A compilation of all moments follows:

\[
E(\delta p_1) = E(\delta p_2) = 0, \text{ from (7.2.1)}
\]

\[
V(\delta p_1) = \frac{p_1q_1}{2n}, \quad V(\delta p_2) = \frac{p_2q_2}{2n}, \text{ from (7.2.2)}
\]

\[
C(\delta p_1, \delta p_2) = \frac{A}{2n}, \text{ from (7.2.5)}
\]

\[
E(\Lambda^r) = (1 - r - \frac{1}{2n}) \Lambda, \text{ from (7.3.1)}
\]

\[
E(\delta \mu_x) = -\frac{1}{2n} (\sigma_{\delta_1} + \sigma_{\delta_2}) + \lambda(\gamma - \frac{1}{2n}) \sigma_{aa}, \text{ from (7.4.2) \text{. (7.5.1)}}
\]

The first term of \( E(\delta \mu_x) \) is composed of the inbreeding depressions for the two loci and is unaffected by linkage disequilibrium. The second term is a consequence of linkage disequilibrium, as changed by recombination and drift between \( \Pi_0 \) and \( \Pi_r \), acting on \( \mu_x \) through additive \( x \) additive epistatic effects. Except for tightly linked genes with \( r < \frac{1}{2n} \), if \( e > 0 \) this term is negative when correlation \( \lambda < 0 \) (excess frequency of repulsion gametes) and positive when \( \lambda > 0 \) (excess frequency of coupling gametes). Convenient labels then would be "repulsion depression" and "coupling enhancement." If \( e < 0 \) the situation reverses to one of repulsion enhancement and coupling depression.

Finally, \( E(\delta \mu_x) \) can be written as

\[
E(\delta \mu_x) = -\frac{1}{2n} (\sigma_{\delta_1} + \sigma_{\delta_2} + \lambda \sigma_{aa}) + r \lambda \sigma_{aa} \tag{7.5.2}
\]

so separating effects of drift from those of linkage disequilibrium as changed by recombination.
7.6 Consequences of Directional Selection

When the n top-ranking test zygotes are selected, \( R_n \equiv \{N-n+1, N-n+2, \ldots, N\} \), \( n < N \), with standardized selection differential \( k \) at (2.2.1) and approximation \( \hat{k} \) at (2.4.2). Moments of gene frequency response, and \( \mathcal{E}\{ \lambda^r \} \), are given by direct substitution in (7.2.1), (7.2.2), (7.2.5) and (7.3.1). Pre-test projection of trait mean response, as assessed in \( \Pi_r = \Pi_0(p_1 + \delta p_1, p_2 + \delta p_2, \lambda^r) \) is obtained by substitution of these moments in (7.4.2); in the following expressions phenotypic standard deviation, \( \sigma_x \), is reinstated to facilitate comparisons with (5.7.1).

\( \mathcal{E}\{ \delta u_x \} \) can be written in a variety of forms to display different influences of factors incorporated in this double polygene model. The basic expression is

\[
\mathcal{E}\{ \delta u_x \} = k \left( \rho_1 \left( \frac{1}{\sigma_x} \right)^2 + \rho_2 \left( \frac{1}{\sigma_x} \right)^2 \right) \sigma_x \\
+ k \left( \frac{1}{2} \sigma_x^2 \right)^2 \left\{ (1 + \frac{1}{n}) - \lambda^2 (1-2r) \right\} \sigma_x \\
- k(1-2r) \left\{ \left( \frac{\lambda \sigma_x}{\sigma_x^2} \right) + \left( \frac{\lambda \sigma_x}{\sigma_x^2} \right) \right\} \sigma_x \\
- \frac{1}{2} \left[ k \left( \frac{\sigma_x^2}{\sigma_x^2} \right) + \rho_2 \left( \frac{\sigma_x^2}{\sigma_x^2} \right) + \rho_1 \rho_2 \left( \frac{\sigma_x^2}{\sigma_x^2} \right) \right] \sigma_x \\
- \left( \frac{\sigma_x^2}{\sigma_x^2} \right) \frac{1}{2n} \left\{ \frac{n}{n_1} + \frac{n}{n_2} + \lambda \left( \frac{\sigma_x^2}{\sigma_x^2} \right) \frac{n}{n_c} \{ 1 - 2(n_c - n) \} \right\} \sigma_x \\
+ \left\{ r \lambda \left( \frac{\sigma_x^2}{\sigma_x^2} \right) \frac{n}{n_c} \right\} \sigma_x ,
\]

(7.6.1)

where
\[ \rho_1 = \frac{\sigma_{a1}}{\sigma_x} + \lambda \left( \frac{\sigma_{a2} + \sigma_{e1}}{\sigma_x} \right), \quad \rho_2 = \left( \frac{\sigma_{a2}}{\sigma_x} + \lambda \left( \frac{1}{\sigma_{e2}} \right) \right) \]

in comparison with \( \rho = (\sigma_a/\sigma_x) \) at (5.3.3). All terms of (7.6.1) are readily identified by analogy with results for single polygenes in Section 5.7 and the consequences of no selection in the previous section.

The first term contains contributions from additive effects with selection precision factors \( \rho_1 \) and \( \rho_2 \); it may be rewritten as

\[
k\left( \frac{\sigma_{a1}^2 + \sigma_{a2}^2}{\sigma_x^2} \right) + \lambda \left( \frac{2\sigma_{a1}\sigma_{a2} + \sigma_{a1}\sigma_{e1} + \sigma_{a2}\sigma_{e2}}{\sigma_x^2} \right)\sigma_x
\]

\[
= k\left( \frac{\sigma_{a1}^2 + \sigma_{a2}^2}{\sigma_x^2} \right) + \left( \frac{\sigma_{a2}}{\sigma_x} \right)\sigma_x \quad (7.6.2)
\]

on identifying additive genetic variance and covariance components at (7.1.4) and (7.1.6). This term is a function of \( n \) and \( N \) only through \( k \) and is not affected by variations in recombination fraction \( r \).

The second term of (7.6.1) contains the direct contribution of additive x additive epistatic variance as influenced by recombination, linkage disequilibrium and sampling. This is a positive term with a maximum reached at free recombination and/or linkage equilibrium in \( \Pi_0 \).

The third term of \( E(\delta u_x) \) is a contribution from covariances between dominance and additive x additive effects as influenced by recombination. It can be written as

\[-k(1-2r)\left( \frac{\sigma_{1}e + \sigma_{2}e}{\sigma_x^2} \right)\sigma_x = -k(1-2r)\left( \frac{\sigma_{de}}{\sigma_x} \right)\]
is zero with free recombination and otherwise is positive or negative depending on the signs and magnitudes of \( d_1 \) and the signs of \( e \) and \( \lambda \neq 0 \).

The fourth term of (7.6.1) contains identification interference effects of dominance described previously for the single polygene model and is augmented by similar interference by additive x additive epistasis when \( e > 0 \); in just the same way that \( d > 0 \) results in an interference of phenotypic identification of \( \underset{\text{t}}{N} \) genotypes with mean \( (a) \), as compared to \( \underset{\text{c}}{N} \) genotypes with mean \( (d) \), so does \( e > 0 \) result in an interference of phenotypic identification of \( \underset{\text{t}}{N} \) genotypes with mean \( (a_1 + a_2 - e) \) as compared to less favorable genotypes such as \( \underset{\text{c}}{N} \) with mean \( (a_1 - a_2 + e) \).

The final two terms of \( E(\delta \mu_x) \) may be compared directly with those at (7.5.2) for the case of no selection. The first of these two terms contains the drift effects analogous to inbreeding depressions together with the linkage disequilibrium effect as changed by drift, all as modified by selection, a feature conveniently condensed into the effective number representation: \( n_1 \), \( n_2 \) and \( n_c \). The second of these two terms is the previously labelled coupling/repulsion:depression/enhancement effect again as modified by selection.

Finally, \( E(\delta \mu_x) \) can be rearranged to the form

\[
E(\delta \mu_x) = k\left(\frac{\sigma^2}{2} + \frac{\sigma^2}{\sigma_x^2}\right) \sigma_x
\]

\[
- \frac{1}{2} k\left(\frac{\sigma}{\sigma_x}\right)^2 \left(\frac{\sigma}{\sigma_x}\right) + k\left(\frac{\sigma}{\sigma_x}\right)^2 \left(\frac{\sigma}{\sigma_x}\right) + k\left(\frac{\sigma}{\sigma_x}\right) \left(\frac{\sigma}{\sigma_x}\right)
\]

+ terms involving \( \lambda \) and \( \lambda^2 \)
+ terms involving product (\( \lambda r \))
+ terms of \( O(1/n) \),
the first term of which is the usual projection of mass selection based on twice the parent-offspring covariance: $2\left(\frac{1}{2} \sigma_a^2 + \frac{1}{4} \sigma_{aa}^2 \right)$. Here it should be noted that as far as $E(\delta u_x)$ is concerned, anything less than free recombination is immaterial unless $\lambda \neq 0$ in $\Pi_0$. 
8. DISCUSSION AND SUMMARY

Many points of detail have already been discussed as developed. Here attention is concentrated on aspects of more general significance. Summary statements are inserted at or near the end of some sections as are indications of areas of future research potential.

8.1 Selection Practices

Development of model structures in Section 1.1, although by no means all inclusive, accommodates many applied situations. Selection model III, with which most analysis has been concerned, is particularly appropriate in plant and animal breeding contexts.

Specification of $\Pi_0$, the reference population for sampling of test material, is the essential prerequisite to projection of consequences of selection. It is this specification that determines sampling properties of selection differentials and the stochastic relationship between $X$ and $Y$. Specification of $\Pi_X$, the reference population for response material, is not always so important and in some cases may be so implicit in the definition of response variable $Y$ as not to require formal statement. This is the case in Chapters 3 and 4 where $Y$ is defined to include specification of utilization of selected units.

The assumption of selection criteria based entirely on a single variable $X$ is also a common applied situation. Even when selection criteria involve several measured properties per test unit there is good reason to combine these into a single "index," $X$, as argued by Cochran (1950) in terms of statistical power. Where moments of $X$, or of order statistics in the $X$ dimension, have been used there is no restrictive assumption as to the continuous or discrete nature of
selection basis \( X \). Where distributional properties of \( X \) in \( \Pi_0 \), and of order statistics, have been used attention has been restricted to continuously distributed \( X \). Extension of such results to the case of discretely distributed \( X \) requires further theory of order statistics in samples from discrete density functions such as that described by David and Mishriky (1968). Possible variations in the discrete or continuous, univariate or multivariate nature of response variable \( Y \) have been encountered in previous chapters. In Chapter 4, where a bivariate Normal model is assumed, \( Y \) is univariate continuous; in Chapters 5 and 7, where \( Y \) is defined as a vector of genotypic indicator random variables, \( Y \) is multivariate discrete.

In regard to selection criteria, the analysis of selection model III has been restricted to deterministic censorship. It has been assumed that set \( R_n \), containing ranks defining selected units, is fixed and known which is certainly the case after application of selection criteria but may not be the case prior to sampling of test material. Results obtained in this deterministic setting can be extended to cover stochastic censorship by the following device. There are \( \binom{N}{n} \) distinct possible sets \( R_n \) specifying \( n \) ranks from the integers 1,2,...,\( N \), for each of which pre-test projections can be made. Given the probability distribution \( R \) of \( R_n \) over these \( \binom{N}{n} \) combinations, pre-test projections conditional on \( R_n \) can be averaged over \( R \), a simple matter in the case of linear responses where it is only necessary to average \( k \) over \( R \).

Finally an additional variation arises when there are intended differences in frequency of usage among selected units; for example, top ranking units may be utilized more frequently than lower ranking
units of the same selected group. This feature can be accounted by repeating ranks in $R_n$ in proportion to the intended frequency of usage of units occupying those rank positions, but it is usually simpler to incorporate this feature in the definition of $\Pi_r$.

### 8.2 Parameters of Selection Differentials

Ideally the sampling distribution of selection differential $D_n$, or its standardized equivalent $\Pi_n$, is required for a complete characterization of distributions of average responses in pre-test projections. Unfortunately, it has not proved possible to derive this sampling distribution either generally or for specific distributions of $X$ in $\Pi_0$. However, determination of low moments of $\Delta_n$, as parameterized by $k$, $\nu$ and $w$ defined in Section 2.1, does provide methods for pre-test projection of low moments of average response.

Directional selection is perhaps that selection style forming the most common application and approximations $\hat{k}$ for $k$ and $\hat{\nu}$ for $\nu$, where adequate, serve two purposes. They provide not only numerical determinations of $k$ and $\nu$ for those $X$ distributions for which tabulations or algebraic expressions of moments of order statistics are not available, but also analytical expressions vulnerable to algebraic manipulation of $n$ and $N$ for quite general continuous density functions $f(x)$.

Adequacy of approximation $\hat{k}$ for $k$ is well-established in Chapter 2: $\hat{k}$ takes correct values in the case of no selection, $n = N$, and also asymptotically as $N \to \infty$ with $n/N$ held constant; it satisfies the required symmetry property

$$nk_{(n)} = (N-n)k_{(N-n)}$$

at (2.8.3); it provides exact results for uniformly distributed $X$ and
its calculated values are in excellent agreement with exact values for Normally distributed X and for selection of either top ranking or bottom ranking units from exponentially distributed X, except for extreme selection intensities when \( N \leq 10 \). In regard to Table 2.1 where \( \hat{k} \) is compared with \( k \) in the case of Normally distributed X, only values of \( n \leq \frac{1}{2} N \) are included because approximation errors can be inferred for \( n > \frac{1}{2} N \): since both \( k \) and \( \hat{k} \) satisfy symmetry property (2.8.3),

\[
\frac{100 \hat{k}(N-n)}{k(N-n)} = \frac{100 \hat{k}(n)}{k(n)}, \quad n \leq \frac{1}{2} N.
\]

Additionally, \( \hat{k} \) is simple to obtain analytically and numerically:

\[
\hat{k} = k_0 - C(n,N), \quad C(n,N) = \frac{1}{2(N+1)} \frac{F(x_0)}{f(x_0)},
\]

where \( x_0 \) satisfies \( 1 - F(x_0) = n/N \) and \( k_0 \) is the mean of \( f(x) \) after truncation to \( x > x_0 \).

Adequacy of approximation \( \hat{v} \) for \( v \) has not been so thoroughly investigated. As shown in Chapter 2 it does take the correct value for no selection, \( n = N \), but unlike \( \hat{k} \) for \( k \) there is no asymptotic equivalent. Indeed \( \hat{v} \) itself is the asymptotic form of \( v \) as \( N \to \infty \) with \( n/N \) held constant, this form being chosen because of the empirically found stability of \( v \) in the Normal case and because of its simplicity:

\[
\hat{v} = v_0 + (k_0 - x_0)^2 F(x_0),
\]

where \( v_0 \) is the variance of \( f(x) \) after truncation to \( x > x_0 \). In the Normal case \( \hat{v} \) is readily calculated from standard tabulations of the unit Normal distribution function or of the unit Normal linear loss integral, \( L(u) \), and \( \hat{v} \) satisfies both the required symmetry property at (2.8.4):
\[ n\{1 - v_{(n)}\} = (N-n)\{1 - v_{(N-n)}\} \]
and the known bounds to \( v \) established in Section 10.2:
\[ \frac{n}{N} < v < 1. \]

Since the material in Sections 2.6 and 2.7 and in Section 10.2 was prepared, further empirical evidence for stability of \( v \) over \( N \), given \( n/N \), has been presented by Schaeffer et al. (1970). They also chose to define the standardized variance of \( D_n \) as \( nV[\Delta_n] \), rather than \( V[\Delta_n] \), expressing the hope that \( v \) would be almost constant valued for given \( n/N \), a property convenient for their purpose. They calculated \( v \) for \( (100 \ n/N) = 1(10)(50)(100) \) for all values of \( N \) between 4 and 100 yielding \( n \) an integer. After reporting that \( v \) was found to be nearly constant for all \( N \), they present a tabulation of averaged values, \( \overline{v} \), for each proportion \( n/N \) (their Table 2), a sample from which is compared to \( \hat{v} \) in the following list:

<table>
<thead>
<tr>
<th>( n/N )</th>
<th>( \hat{v} )</th>
<th>( \overline{v} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.3709</td>
<td>0.366</td>
</tr>
<tr>
<td>0.2</td>
<td>0.4679</td>
<td>0.466</td>
</tr>
<tr>
<td>0.3</td>
<td>0.5464</td>
<td>0.542</td>
</tr>
<tr>
<td>0.4</td>
<td>0.6164</td>
<td>0.613</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6817</td>
<td>0.679</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7443</td>
<td>0.742</td>
</tr>
<tr>
<td>0.7</td>
<td>0.8056</td>
<td>0.804</td>
</tr>
<tr>
<td>0.8</td>
<td>0.8670</td>
<td>0.865</td>
</tr>
<tr>
<td>0.9</td>
<td>0.9301</td>
<td>0.929</td>
</tr>
</tbody>
</table>

The maximum difference is 0.005 and \( v \) is always greater than \( \overline{v} \) because \( \hat{v} \) increases to its asymptote \( \hat{v} \) as \( N \to \infty \) with \( n/N \) held constant. The study of Schaeffer et al. (1970) is entirely numerical and limited to the case of Normally distributed \( X \) whereas \( \hat{v} \) as derived here applies to general continuous \( f(x) \). Moreover the symmetry property established here, at
(2.8.4), in the Normal case renders numerical consideration of \( v \) for \( n > \frac{1}{2} N \) unnecessary.

One notion not previously stressed is that of the "typical selected unit." Denoted by \((x^*, y^*)\) this is defined as a randomly sampled member of the selected group:

\[
Pr[(x^*, y^*) \equiv (x_{(1)}, y_{(1)})] = \frac{1}{n}, \text{ all rcR}_n.
\]

Properties of "typical individual response," \( y^* \), introduced in Section 3.7 will be discussed later whereas here it is noted that the procedure used to derive approximation \( \hat{k} \) for \( k \), and \( \hat{\beta}(y_0, n) \) for \( \beta(y_0, n) \) at (3.8.4), can be generalized to a class of expectational operators \( \mathcal{U} \), on \( x^* \), defined by

\[
\mathcal{U}(x^*) = E[u(x^*)],
\]

where \( u \) is some function of \( x^* \) of interest. The generalization is developed in Section 10.3, the main result being the following approximation process:

\[
\mathcal{U}(x^*) = \hat{\mathcal{U}}(x^*) + (N+1)^{-2} \{\text{further terms}\}
\]

\[
\hat{\mathcal{U}}(x^*) = \mathcal{U}(x^*, x_0) - u^-(x_0) C(n, N),
\]

where \( C(n, N) \) and \( x_0 \) are as defined before and \( \mathcal{U}(x^*, x_0) \) is the operator \( \mathcal{U} \) applied to a random unit sampled from \( f(x) \) after truncation to \( x > x_0 \). Some examples follow.

Setting \( u(x^*) = x^* \), \( u^-(x_0) = 1 \) and \( \mathcal{U}(x^*, x_0) = k_0 \), so that

\[
\mathcal{U}(x^*) = E(x^*) \text{ and } \hat{\mathcal{U}}(x^*) = k_0 - C(n, N) = \hat{k}
\]

and in this case the operator \( \mathcal{U} \) translates from \( x^* \) to \( \Delta_n \) directly as is always so when \( u(x^*) \) is a linear function of \( x^* \). This translation
does not, of course, apply when setting \( u(x^*_k) = (x^*_k - k)^2 \) so that
\[
\hat{U}(x^*_k) = V(x^*_k); \text{ instead the procedure used in Section 2.5 is constructed}
\]
to obtain \( V(\Delta_n) \) approximately. Rather than constructing moments of \( x^*_k \)
individually, the moment generating function of \( x^*_k \) is obtained by
setting \( u(x^*_k) = \exp(t x^*_k) \); in the Normal case this yields
\[
U(x^*_k) = E(e^{t x^*_k}), \quad \hat{U}(x^*_k) = e^{x^*_k} \frac{1}{2} t \frac{G(x_0 - t)}{G(x_0)} - t e^{x_0^*} C(n, N),
\]
the first term of \( \hat{U}(x^*_k) \) being the moment generating function for \( f(x) \)
after truncation to \( x > x_0 \).

8.3 Linearity of Response

The assumption of linear response at (3.1.1) finds common application whether appropriate for all \( x \) or just some restricted range of \( x \)
that is of primary interest. In the latter case nonlinear formulations
of response may be linearized over just the restricted range of \( x \) in
the usual manner: if \( E(Y|X) = Q(X) \) for all \( X \) but \( Q(X) \) is approximately
linear in the interval \( I(X) \) containing \( \mu_X \), then
\[
E(Y|X) \approx Q(\mu_X) + (X - \mu_X) Q'(\mu_X), \quad X e I(X)
\]
and
\[
E(y|X) \approx \rho X, \quad X e I(X), \quad \rho = \frac{\sigma}{\sigma_y} Q'(\mu_X).
\]

The assumption regarding conditional variance of response \( Y \), given
\( X \), at (3.1.2) should accommodate most applications, catering for both
homoscedastic \( (b_1 = 0) \) and heteroscedastic \( (b_1 \neq 0) \) situations.

Examples of quadratic response arise in the formulations of trait
mean advance for polygene models of Chapters 5 and 7. The procedure
used there can be generalized for quadratic responses: if, in the
notation of Chapter 3,

\[ E(\delta_n | \Delta_n) = A \Delta_n + B \Delta_n^2 \]

the pre-test expectation is found as

\[ E(\delta_n) = A E(\Delta_n) + B[E^2(\Delta_n) + V(\Delta_n)] \]

\[ = (Ak + Bk^2) + B \frac{v}{n}. \]

### 8.4 Nature of Response

As described in Section 3.2 it is implicit in the definition of
response variable Y that all relevant consequences of selection are
properties of the set

\[ S_n(y) = \{y_r; r \in R_n\} \]

Although the probability distribution of \( S_n(y) \) would provide an ideal
characterization of consequences of selection, its derivation is
inhibited by analytical difficulties discussed in Chapter 3 and in most
cases it is not an essential requisite for projections of interest.

Three summary features of \( S_n(y) \) have been considered in some detail:

(i) "average response," \( \bar{\delta} = \frac{1}{n} \sum_{r \in R_n} y_r \)

(ii) "specified individual response," \( y_r \)

(iii) "typical individual response," \( y_* \)

the last being introduced in Section 3.7 in relation to average
probability of positive response. Methods have been presented for
determining pre- and post-test expectations and variances of these
defined response features.
Post-test distributions of all three of \( y_\ast, y_r \) and \( \delta_n \) are readily obtained, the latter requiring a convolution of conditional density functions that is simple for the bivariate Normal model of \((X,Y)\) in \( \Pi_0 \). However, for none of \( y_\ast, y_r \) and \( \delta_n \) have exact pre-test distributions been derived. The approximation \( \hat{\beta}(y_0,n) \), for pre-test \( \Pr[y_\ast > y_0] \) under directional selection, is derived in Section 3.8 and is a special case of the general class \( \mathcal{U}(x_\ast) \) discussed earlier, namely setting \( u(x_\ast) = \Pr[y_\ast > y_0|x_\ast] = L_y(y_0|x_\ast; \rho) \) in the notation of Chapter 3. This approximation also provides an approximation \( \hat{\alpha}(y_0,i) \), for pre-test \( \Pr[y_1 > y_0] \), via

\[
\hat{\alpha}(y_0,i) = (N-i+1) \hat{\beta}(y_0,N-i+1) - (N-i) \hat{\beta}(y_0,N-i)
\]

as at (3.8.5). Approximations \( \hat{\beta}(y_0,n) \) and \( \hat{\alpha}(y_0,i) \) have been shown to provide exact results for a special bivariate distribution of \( X \) and \( Y \) that has uniform margins and linear regressions. A method for obtaining exact values of \( \alpha(y_0,i) \) is indicated in Section 3.7, namely by numerical integration of

\[
\alpha(y_0,i) = \Pr[y_1 > y_0] = \int_{-\infty}^{\infty} L_y(y_0|t; \rho) z(t;N,i) dt.
\]

Finally, in default of a pre-test distribution of average response, lower bounds for pre-test probability of positive average response are given in Section 3.7 and can be augmented by symmetry considerations as in Section 4.4.

### 8.5 Pre- and Post-test Projections

The relative information contained in pre- and post-test projections has been considered in terms of the corresponding statistical efficiencies.
Detailed results for average response with directional selection under the bivariate Normal model indicate that the pre-test:post-test information ratio, given n/N, decreases quite rapidly as ρ increases beyond 0.5 but is quite stable over increasing N.

An overview of the relative information status in pre- and post-test situations is provided by noting that pre-test projections are based on information contained in rank positions of selected units whereas post-test projections use the observed X values of those rank positions. Thus the relationships between order statistics $X_{(i)}$ and their ranks i specify the relationships between pre- and post-test sampling properties of selection differentials $A_n$. Stuart (1954) presents the following expression for $C_N$, the correlation between $x_{(i)}$ and i in samples of size N from X distributions with standardized continuous densities:

$$C_N = \left( \frac{3(N-1)}{4(N+1)} \right)^{1/2} g,$$

where $g = E(Ginis' standardized mean difference) = 4E(x[F(x) - 1/2])$. When X is Normally distributed $g = 2/\sqrt{\pi}$ and so

$$C_N = \left( \frac{3}{\pi} \frac{(N-1)}{(N+1)} \right)^{1/2}$$

attaining a maximum of $\sqrt{3/\pi} = 0.9772$ as $N \to \infty$. Of course, if the regression of $x_{(i)}$ on i were linear this quantity would provide simple expressions for $\mu_{i|N}$ and k:

$$E(x_{(i)}|i) = \mu_{i|N} = C_N \frac{i - \mu_i}{\sigma_{i}} = \frac{\sqrt{3(2i-N-1)}}{2(N+1)} g.$$

But this regression is only known to be linear in the case of uniformly distributed X except for samples of $N = 2$ and $N = 3$ from symmetric distributions.
8.6 Prediction or Projection?

The important distinction between pre- and post-test projections should not be overemphasized; its importance is provision of the strategy of conditioning and removal of conditioning in translating from post- to pre-test contexts, a strategy followed throughout Chapters 3 and 4. The term projection has been used consistently in the sense of Section 1.1: relating parameters of $\Pi_r$ to those of $\Pi_0$ with a proper accounting of stochastic elements involved. Assigning numerical values to these parameters in specific applications is a separable problem referred to as prediction. Combining the results of projection based on analysis of the selection model assumed, with estimates of parameters based on sample observations, provides predictions some properties of which may be derived. Here this aspect is discussed briefly and by means of example. Nicholson (1960) provides a more general discussion of related prediction problems.

Consider average response in the notation of Chapter 3:

$$\frac{1}{n} \sum_{R_i} Y_r - \mu_y = \delta \sigma_y .$$

Except for selection model I this is a random variable the low moments and distribution of which have been subjected to pre- and post-test projections in Chapter 3 and 4 under a model of linear response. Consider then the parameter $\theta$ defined as

$$\theta = E[\frac{1}{n} \sum_{R_i} Y_r - \mu_y] .$$

Given the sampled test material under selection model III what information is available from which to construct a prediction $\theta$? Assuming linear response the post-test projection is
\[ \theta | \bar{X}_n = \mathbb{E}\left\{ \frac{1}{n} \sum_{r=1}^{n} Y_r - \mu_x | \bar{X}_n \right\} = \beta (\bar{X}_n - \mu_x) \]

as at (3.5.2); here \( \beta \) denotes \( \rho \sigma_y / \sigma_x \). In the trivial case where \( \beta \) and \( \mu_x \) are known exactly, \( \beta (\bar{X}_n - \mu_x) \) is known exactly.

More usually \( \beta \) and \( \mu_x \) are the subject of prior investigation and independent estimation. Suppose then that there exist \( \hat{\beta} \) and \( \hat{\mu}_x \) such that

\[ E(\hat{\beta}) = \beta, \ V(\hat{\beta}) = \sigma^2_\beta, \ E(\hat{\mu}_x) = \mu_x, \ V(\hat{\mu}_x) = \frac{\sigma^2_x}{M} \]

and \( \hat{\beta}, \hat{\mu}_x \) are stochastically independent of each other and of the sampled test material. The latter provides no further information for \( \beta \) but \( \bar{X}_n \) is relevant to \( \mu_x \) and can be combined with \( \hat{\mu}_x \):

\[ \hat{\mu}_x = \frac{(N\bar{X}_N + M\hat{\mu}_x)}{(N+M)} \]

\[ E(\bar{X}_n - \hat{\mu}_x | \bar{X}_n) = (\bar{X}_n - \mu_x), \ V(\bar{X}_n - \hat{\mu}_x | \bar{X}_n) = \frac{\sigma^2_x}{(N+M)} . \]

Then the proposed prediction for \( \theta | \bar{X}_n \) is

\[ \hat{\theta}_n = \hat{\beta}(\bar{X}_n - \hat{\mu}_x) \]

with

\[ E(\hat{\theta}_n | \bar{X}_n) = \beta (\bar{X}_n - \mu_x), \ i.e., \ unbiased \]

and

\[ V(\hat{\theta}_n | \bar{X}_n) = (\bar{X}_n - \mu_x)^2 \sigma^2_\beta + \frac{(\beta^2 + \sigma^2_\beta) \sigma^2_x}{(N+M)} \]

which indicates the importance of possessing a precise estimate \( \hat{\beta} \). It should be noted that the choice \( \hat{\mu}_x \) is that unbiased linear combination of \( \bar{X}_N \) and \( \hat{\mu}_x \) with minimum variance. Also, in the case of Normally
distributed $X$ if $\hat{\mu}_X$ follows a Normal distribution, this can be regarded as a prior distribution on $\mu_X$ and $\hat{\mu}_X$ is then the Bayes estimate of $\mu_X$ being the mean of the posterior distribution on $\mu_X$.

Whereas it may be reasonable to assume that $\beta$ is stable from prior investigation to the particular test sample obtained, it may be known in some applications that $\mu_X$ varies from test sample to test sample. In this situation $(\bar{X}_n - \mu_X)$ may be estimated by the "reach" statistic $(\bar{X}_n - \bar{X}_N) = D_n$ of Section 2.1 corresponding to the case $M = 0$ in the above results:

$$
(\bar{X}_n - \hat{\mu}_X) = (\bar{X}_n - \bar{X}_N) + \frac{M}{(N+M)} (\bar{X}_N - \hat{\mu}_X).
$$

Next, what can be said about the distribution of $\hat{\theta}_n$ from one test sample to another? Low moments are obtained by removing the conditioning on $\bar{X}_n$ from those above:

$$
E(\hat{\theta}_n) = k\rho\sigma_y
$$

which is just the pre-test projection at (3.4.3), and

$$
\nu(\hat{\theta}_n) = \frac{2}{\sigma_X^2} \left[ \frac{2}{\sigma_\beta^2} + (\beta^2 + \sigma_\beta^2) \left( \frac{\nu}{n} + \frac{1}{N+M} \right) \right].
$$

Finally, unbiasedness and variance properties are not the only criteria governing choice of predictions; introduction of appropriate loss functions and translation of the problem to the decision theoretic context is indicated.

8.7 Selection Designs and Strategies

In planning selection practices, in the pre-test context, specifications of design variables $n$ and $N$ and manipulations of selection precision $\rho$ (where possible) are commonly based on projections of expected average responses, particularly in the areas of plant and
animal breeding. The primary application of the developed treatment of pre-test projections for $\delta_n$, $y_1$, and $y_*$ is the provision of additional criteria for use in such planning. Similarly, in applying selection in the post-test context, the developed post-test projections can be used as criteria in the choice of selected units, involving possible modifications to pre-test intentions in the light of post-test information.

Considering the planning stage first, introduction of pre-test projected variances of response as design criteria is conveniently combined with projected expectations of response into $C$, the pre-test coefficient of variation. It is shown in Chapter 3 that decreasing $C$ results in a higher lower bound for pre-test probability of positive response. For the bivariate Normal model under directional selection, $C$ for average response is frequently of the order 100 percent, less than this when $\rho$ is intermediate to high and $N$ is large, but more when $\rho$ is low and $n$ is small.

Pre-test probability statements concerning response can be incorporated as design criteria in a variety of ways. Here attention is restricted to a class of selection designs, appropriate for directional selection of top ranking units, and denoted by

$$\mathcal{D}(N \text{ or } n; Y, Y_0, \gamma | \rho, N \text{ or } n)$$

corresponding to the pre-test probability criterion

$$\Pr[Y > Y_0 | \rho, N, n] \geq \gamma, \quad 0 < \gamma < 1,$$

where $Y$ is some response feature such as $\delta_n$, $y_1$, or $y_*$ and $\gamma$ is a pre-assigned constant. Calculations of $\Pr[Y > Y_0 | \rho, N, n]$ are required in order to apply this design criterion; exact or approximate values can be
obtained as described above. Some simple examples will illustrate usage of this class of selection designs.

Suppose first that $n$ is fixed by demand and that $Y = y_*$ the typical individual response. Then $D(N; y_*, 0, 0.9| \rho, n)$ denotes the design of choosing the smallest integer $N$ of sampled test units from $\Pi_0$ in order that

$$\Pr[y_* > 0| \rho, N, n] \geq 0.9.$$  

A variation of this might be $D(N; y_{N-n+1}, 0, 0.75| \rho, n)$ implying choice of $N$ to ensure that the lowest ranking of the selected units satisfies

$$\Pr[y_{N-n+1} > 0| \rho, N, n] \geq 0.75.$$  

Alternatively suppose that $N$ is fixed by resource limitations. Then $D(n; y_{N-n+1}, 1, 0.5| \rho, N)$ denotes the design of intending to choose the largest integer $n$ to ensure that the lowest ranking of the selected units satisfies

$$\Pr[Y_{N-n+1} > \mu_Y + \sigma_Y| \rho, N, n] \geq 0.5.$$  

Depending on $N$ and $\rho$ there may not be a solution for this design. Other variations can be constructed that manipulate both $n$ and $N$.

Turning now to post-test application of selection criteria, whatever the planned design, the same principles can be applied to selection of test units with the possibility of modifying originally planned intentions in light of the observed order statistics $X_{(1)}, X_{(2)}, \ldots, X_{(N)}$. The difference is that $N$ is now fixed and it is selection strategy rather than selection design that is involved. Additionally, post-test probability statements are more readily available than pre-test statements as discussed in Chapter 3.
For the bivariate Normal model under directional selection the strategy of choosing \( n \) to maximize the probability of positive average response has already been discussed in Section 4.2:

\[
\Pr[\delta_n > 0|X_{(r)}, r \in R_n] = G\left(\frac{2}{\sqrt{1 - \rho^2}}\right)
\]

and is maximized by choosing \( R_n \) such that \( \sqrt{n} \bar{X}_n \) is a maximum. This is a simple strategy, independent of \( \rho \), with an approximate prior expectation of choosing \( n = (0.27)N \) as at (4.2.5).

More generally there is a class of strategies \( S \), analogous to the class of designs \( D \), formulated as \( S\{n; Y, Y_0, \gamma | \rho \} \). For example \( S\{n; y_{N-n+1}, 0, \gamma | \rho \} \) denotes the strategy of choosing the largest integer \( n \) such that the lowest ranking of the selected units satisfies

\[
\Pr[y_{N-n+1} > 0 | \rho, n; X_{(r)}, r \in R_n] \geq \gamma.
\]

Required \( n \), if it exists, is readily found from

\[
L_y(0|X_{(N-n)}; \rho) < \gamma \leq L_y(0|X_{(N-n+1)}; \rho)
\]

in the bivariate notion of Chapter 3.

These design and strategy considerations have been discussed in some detail in order to illustrate the rich possibilities of augmenting traditional projections of expected response with projections of other distributional properties of response features.

8.8 The Single Polygene Model

The assumption regarding the nature polygenic effects renders linear regressions of genotypic and gene frequencies on phenotypic values \( X \). Then pre-test expectations and variances of changes in these frequencies for a single cycle of phenotypic selection under selection model III are provided by the recipe of Chapter 3. Thus the special
case of directional selection, first considered for this model by Kojima (1961), is readily extended to a general set \( R_n \) of ranks defining selected zygotes and to include other traits correlated with the selection basis by pleiotropic effects.

The consequences of different styles of selection, relevant to common selection experiments, are then largely a matter of compounding the general results. The effects of drift in combination with divergent selection, conveniently represented by well-defined variance effective numbers, are discussed in detail in Section 5.8. Briefly, and in the case of symmetric high and low selection \( (k_h = -k_l) \), initial divergence may be enhanced by drift effects:

\[
D(\mu_x) = \left\{ 2k \left( \frac{\sigma_x}{\sigma} \right)^2 + \left( \frac{1}{n} - \frac{1}{n_v} \right) \left( \frac{\sigma}{\sigma} \right)^2 \right\} \sigma_x^2, \quad n_v > n \text{ when } d > 0
\]

as may be initial asymmetry of response:

\[
S(\mu_x) = -\left( \frac{\sigma}{\sigma_x} \right) \left\{ k^2 \left( \frac{\sigma}{\sigma_x} \right)^2 + \frac{1}{n} \right\} \sigma_x.
\]

For subsequent cycles of divergent selection, with reproductive isolation of high and low lines, analysis of increments to divergence and asymmetry is also indicated in Section 5.8.

For the additive model \( (d = 0) \) trait mean response itself is linearly regressed on phenotype \( X \) and projections of changed genetic variance are possible:

\[
E(\delta \sigma^2_\alpha) = \left( \frac{n_v - n + \frac{1}{2}}{n_v} \right) \sigma_\alpha^2
\]

\[
\frac{n_v - n}{n_v} = \left( \frac{k}{\sigma_x} \right) (p-q)
\]
which incorporates both the effect of selective change in gene frequency and the loss due to drift.

When allelic disequilibrium is introduced into model $\Pi_0$ the recipe of Chapter 3 still applies with appropriate modification, and the results enable a proper accounting for dioecious systems and also for a single cycle of disruptive selection, as discussed in Sections 5.11 and 5.12.

Throughout Chapter 5 reference populations $\Pi_0$ and $\Pi_r$ were deliberately specified so that one cycle of a particular style of selection analyzed there, for a particular monoecious or dioecious system, could be considered as one of a sequence of such cycles in the recurrent selection context.

8.9 Selection Limits in a Monoecious System

The concept of selection limits introduced by Robertson (1960) is a useful criterion in the design of recurrent mass selection programs, where a reasonable gain in early cycles must be balanced against retention of long-term potential, as discussed by Rawlings (1970). Here it has been shown that Robertson's original additive formulation can be extended to include degree of dominance $h$:

$$u(p_0) = snp_0q_0 \left(1 + \frac{1}{3}(1-2h)(p_0 - q_0)\right)$$

and that this approximate limit provides qualitative predictions that are verified by transition matrix methods.

The diffusion equation approximation, $\hat{u}(p_0)$ at (6.2.1, 6.2.3), is the computationally simplest of those procedures proposed for determining reasonably precise values of $u(p_0)$. The adequacy of $\hat{u}(p_0)$, as assessed by comparison with transition matrix results, has been studied
by Ewens (1963), Hill (1969a) and Carr and Nassar (1970a); but in none of these comparisons is it possible to infer the adequacy of \( \hat{u}(p_0) \) for specific \((n,N)\) combinations and specific polygenic effects \((a/\sigma_X, d/\sigma_X)\) other than the general finding that \( \hat{u}(p_0) \) is an overestimate of \( u(p_0) \). This situation has been rectified here with the comparisons described in Section 6.2 a representative sample being presented in Table 6.1. In absolute magnitude there is generally good agreement between \( \hat{u}(p_0) \) and \( v(2n;\infty) \), and excellent agreement in the case of additivity where discrepancies are always less than 0.001 for \( N = 16 \) and \( N = 32 \) at \( (a/\sigma_X) = 0.1 \). As the dominance ratio increases, with \( p_0 = 0.5 \), discrepancies increase to a maximum of 0.032 for \( N = 32 \) at \( h = 2 \) and \( (a/\sigma_X) = 0.1 \). Maximum discrepancies generally occur at \( p_0 = 0.5 \), or a little less than 0.5 in the cases of dominance and overdominance, and for \( \frac{1}{4} N < n < \frac{1}{3} N \) in the case of additivity or \( n = \frac{1}{8} N \) in the cases of dominance and overdominance. For a given value of \( ns \), the discrepancies decrease for larger \( N \) and smaller \( s \) in accordance with the theoretical considerations of Kimura (1957) and the numerical results of Carr and Nassar (1970a). Actual values of \( ns \) per combination in Table 6.1 are shown in Section 10.4.

### 8.10 Selection Limits in a Dioecious System

Previous determination of \( u(p_0) \) for a dioecious system has relied on extrapolation from monoecious results or the \( \beta \)-distribution approximations of Curnow and Baker (1968, 1969). Here it has been shown that dioecious systems can be treated by transition matrix methods the only restriction being a practical difficulty of sizes of transition matrices involved; \( W = (2n_m + 1)(2n_f + 1) \) square and nonsymmetric (Section 6.3).
As discussed by Hill (1969a), maximum selection limit corresponds to maximum product ns for monoecious systems when treated by diffusion equation approximations, and no departure from this rule is found by transition matrix methods. For dioecious systems this must be modified so that n is replaced by an effective number such as nd and s is calculated from \( \overline{\kappa} = \frac{1}{2} (k_m + k_f) \) rather than \( \kappa \). Maxima of

\[
\frac{1}{4} n_d (k_m + k_f) = \left( \frac{n_f}{n_m + n_f} \right) N_f f(x_f) + \left( \frac{n_m}{n_m + n_f} \right) N_m f(x_m)
\]

are considered in Section 6.3 for several variations in \( N_m, N_f, n_m \) and \( n_f \). Such analysis predicts that maximum selection limit is attained by \( n_m = (1/2)N_m, n_f = (1/2)N_f \) when \( n_m = n_f \) and \( N_m = N_f \), but by \( n_m = (0.39)N_m \) when \( N_m = N_f \) and there is no selection on females \( (n_f = N_f) \); both of these predictions are verified by transition matrix results which augurs well for other predictions obtained in Section 6.3 from the same source.

Adequacy of diffusion equation approximation \( \hat{u}(p_0) \) as a means of determining \( u(p_0) \) for dioecious systems is assessed in Section 6.3 by comparison with transition matrix results; actual values of \( n_d \) and \( \overline{\kappa} \) used, and the implied product \( s n_d \), are shown in Section 10.4. Although the pattern of discrepancies between \( \hat{u}(p_0) \) and \( v(2n_m, 2n_f; \omega) \) is somewhat similar to that obtained for monoecious systems in analogous cases \( (n_m = n_f, N_m = N_f) \) their magnitudes are greater, reaching maxima of 0.033 at \( h = 1/2 \) and 0.051 at \( h = 2 \) when \( (a/\sigma_x) = 0.1 \) and \( (N_m + N_f) = 16 \). Unfortunately this comparison of monoecious and dioecious discrepancies is made for systems in which the diffusion equation approximation itself provides different results: \( k \) is
calculated for directional selection of \( n \) from \( N \) whereas \( \overline{k} = k_m = k_f \) is calculated for directional selection of \( (1/2)N \) from \( (1/2)N \). This feature is avoided in Table 8.1 where selection limits are presented for monoecious and dioecious systems in cases where the diffusion equation approximation provides the same result for both systems. The dioecious systems included are artificial in that there do not exist integers \( N_m \) and \( N_f \) that would provide the required \( k_m \) and \( k_f \); actual parameters used for systems included in Table 8.1 are listed next \( (a/\sigma_X = 0.1) \):

<table>
<thead>
<tr>
<th>monoecious system</th>
<th>dioecious system</th>
<th>diffusion approximation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N = 16, n = 4 )</td>
<td>( n = n_f = 2 )</td>
<td>( n = n_d = 4 )</td>
</tr>
<tr>
<td>( k = 1.20104 )</td>
<td>( k = k_f = 1.20104 )</td>
<td>( s = 0.24021 )</td>
</tr>
<tr>
<td>( N = 16, n = 8 )</td>
<td>( n = n_f = 4 )</td>
<td>( n = n_d = 8 )</td>
</tr>
<tr>
<td>( k = 0.76018 )</td>
<td>( k = k_f = 0.76018 )</td>
<td>( s = 0.15204 )</td>
</tr>
<tr>
<td>( N = 16, n = 12 )</td>
<td>( n = n_f = 6 )</td>
<td>( n = n_d = 12 )</td>
</tr>
<tr>
<td>( k = 0.40035 )</td>
<td>( k = k_f = 0.40035 )</td>
<td>( s = 0.08007 )</td>
</tr>
</tbody>
</table>

Differences between monoecious and dioecious selection limits, when both are standardized to the same value of \( \hat{u}(0.5) \), decrease as \( n \) increases, are rather stable over different degrees of dominance and are generally large enough to account for the extra discrepancy previously found between \( \hat{u}(0.5) \) and dioecious limits. It remains, however, to account for these differences between monoecious and dioecious limits in these admittedly rather small populations; this is considered next.
Table 8.1. Comparison of selection limits for monoecious and dioecious systems in cases where the diffusion equation approximation provides the same selection limit for both; $(a/\sigma_t^2) = 0.1$, $(k; n, N = 16)$ for monoecious, $(\bar{k} = k, n_m = n_f = \frac{1}{2} n)$ for dioecious, $p_0 = 0.5, d = (2h-1)a$

<table>
<thead>
<tr>
<th>n = 4</th>
<th>n = 8</th>
<th>n = 12</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u(0.5)$</td>
<td>$u(0.5)$</td>
<td>$u(0.5)$</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>$h = 1/2$</td>
<td>0.7233</td>
<td>0.7231</td>
</tr>
<tr>
<td>$h = 1$</td>
<td>0.7386</td>
<td>0.7321</td>
</tr>
<tr>
<td>$h = 2$</td>
<td>0.7687</td>
<td>0.7465</td>
</tr>
</tbody>
</table>

Haploid systems with $2n$ genes, monoecious systems with $n$ diploid zygotes and dioecious systems with $(1/2)n$ of each of male and female diploid zygotes define different modes of packaging genes within zygotes or within zygotes within sexes. In selective systems, where selection is defined to act on these packages of genes, drift behavior is affected differently by different systems of packaging and in a rather obvious manner: the more packaging introduced in relation to applied selection, then the higher the drift component in the drift/selection combination and the lower is the ultimate selection limit for the same basic number, $2n$, of genes. This is demonstrated numerically in Table 8.2 where trends in variance of gene frequency ($\sigma_t^2$) and $\bar{p}_t$ may be compared for monoecious and dioecious systems considered in Table 8.1 in the case of additivity only. Consequences of the first cycle of selection are identical for both monoecious and dioecious systems because of identical initial conditions and selection parameters; but for all $t \geq 2$, $\sigma_t^2$ is larger for the dioecious system and $\bar{p}_t$ is larger for the monoecious system.

Turning now to dioecious systems with $n_m \neq n_f$, which have no direct analogy with monoecious systems, when $(n_m + n_f)$ is fixed...
Table 8.2. Comparison of monoecious and dioecious systems in cases where the diffusion equation approximation provides the same selection limit for both: \((a/d) = 0.1, d = 0,\) 
\((k; n, N = 16)\) for monoecious, \((k = k, n_m = n_f = \frac{1}{2} n)\) for dioecious, \(p_0 = 0.5\)

<table>
<thead>
<tr>
<th></th>
<th>n = 4</th>
<th></th>
<th>n = 8</th>
<th></th>
<th>n = 12</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1(\sigma^2)</td>
<td>0.03114</td>
<td>0.03114</td>
<td>0.01560</td>
<td>0.01560</td>
<td>0.01041</td>
<td>0.01041</td>
</tr>
<tr>
<td>2(\sigma^2)</td>
<td>0.05778</td>
<td>0.05789</td>
<td>0.03009</td>
<td>0.03010</td>
<td>0.02036</td>
<td>0.02037</td>
</tr>
<tr>
<td>5(\sigma^2)</td>
<td>0.11448</td>
<td>0.11592</td>
<td>0.06659</td>
<td>0.06683</td>
<td>0.04740</td>
<td>0.04744</td>
</tr>
<tr>
<td>25(\sigma^2)</td>
<td>0.19624</td>
<td>0.20565</td>
<td>0.15769</td>
<td>0.16216</td>
<td>0.14707</td>
<td>0.14833</td>
</tr>
<tr>
<td>50(\sigma^2)</td>
<td>0.20012</td>
<td>0.21041</td>
<td>0.17383</td>
<td>0.18046</td>
<td>0.18468</td>
<td>0.18729</td>
</tr>
<tr>
<td>100(\sigma^2)</td>
<td>0.20022</td>
<td>0.21055</td>
<td>0.17646</td>
<td>0.18361</td>
<td>0.19874</td>
<td>0.20215</td>
</tr>
<tr>
<td>(n_e)</td>
<td>3.66</td>
<td>3.75</td>
<td>6.97</td>
<td>7.09</td>
<td>10.98</td>
<td>11.06</td>
</tr>
<tr>
<td>(p_1)</td>
<td>0.53003</td>
<td>0.53003</td>
<td>0.51900</td>
<td>0.51900</td>
<td>0.51001</td>
<td>0.51001</td>
</tr>
<tr>
<td>(p_2)</td>
<td>0.55620</td>
<td>0.55246</td>
<td>0.53680</td>
<td>0.53561</td>
<td>0.51960</td>
<td>0.51918</td>
</tr>
<tr>
<td>(p_5)</td>
<td>0.61555</td>
<td>0.60352</td>
<td>0.58332</td>
<td>0.57907</td>
<td>0.54595</td>
<td>0.54439</td>
</tr>
<tr>
<td>(p_{25})</td>
<td>0.71739</td>
<td>0.69321</td>
<td>0.72866</td>
<td>0.72619</td>
<td>0.65327</td>
<td>0.64734</td>
</tr>
<tr>
<td>(p_{50})</td>
<td>0.72296</td>
<td>0.69847</td>
<td>0.76447</td>
<td>0.75088</td>
<td>0.70134</td>
<td>0.69377</td>
</tr>
<tr>
<td>(p_{100})</td>
<td>0.72311</td>
<td>0.69862</td>
<td>0.77091</td>
<td>0.75734</td>
<td>0.72101</td>
<td>0.71298</td>
</tr>
<tr>
<td>(p_{\infty})</td>
<td>0.72311</td>
<td>0.69862</td>
<td>0.77106</td>
<td>0.75751</td>
<td>0.72314</td>
<td>0.71509</td>
</tr>
<tr>
<td>(u(0.5))</td>
<td>0.7233</td>
<td>0.7714</td>
<td>0.7233</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
discrepancies between \( \hat{u}(p_0) \) and \( v(2n_m,2n_f;^\infty) \) decrease as \( n_m - n_f \)
decreases and in some cases discrepancies are sufficiently large to
change the ordering of \((n_m,n_f)\) combinations when ranked according to
selection limits calculated by the different methods.

8.11 Rate Effective Number \( n_{et} \)

Characterization of dynamics of a finite population under recurrent
selection requires parameters developed from two or more cycles of
selection. For the single polygene model, rate effective number \( n_{et} \) as
defined and calculated in Section 6.4 provides such a parameter that can
be followed through progression of the stochastic process whether in a
monoecious or dioecious system. Moreover, asymptotic value \( \hat{n}_e/n \)
corresponds to the retardation/acceleration factor of Robertson (1962).

Here it has been shown numerically that \( \hat{n}_e \) is not a function of
initial gene frequency \( p_0 \), that a component of attainment of maximum
selection limit in additive systems is maximum acceleration of rate of
increase of homozygosis, and that all degrees of overdominance \( h > 1 \)
lead to retardation as measured by \( \hat{n}_e/n \) in monoecious systems and
\( \hat{n}_e/n_d \) in dioecious systems.

In contrast the retardation/acceleration factor defined by Carr and
Nassar (1970b), as the ratio of average number of generations to fixation
under selection to that for neutral genes, is not a parameter of dynamics
as defined above, it is not analytically related to progression of homo-
zygosis, it is sensitive to initial gene frequency and it indicates
"acceleration" for some overdominance systems with extremal \( p_0 \).
8.12 Trait Mean Advance and Variation among Replicates

Trait mean advance after any number of cycles of selection is readily obtained by transition matrix methods for both monoecious and dioecious systems as described in Section 6.5 where influences of polygenic effects, selection pressure, initial frequency and population sizes upon early and ultimate advance and its variance are discussed in some detail. Previous investigation of variance of trait mean advance is restricted to that described by Baker and Curnow (1969) using the $\beta$-distribution approximation. Their assumption of Normally distributed replicate mean in a subdivided resource, is found to be not even approximately appropriate in monoecious and dioecious systems studied here. Moreover, it is not a required assumption when using transition matrix methods; since there is a 1:1 correspondence between gene frequency states in the Markov chain and trait means of large random mating populations with those gene frequencies, vector $v_t$ contains the probability distribution of replicate trait mean advances after cycle $t$. Further study of this feature afforded by transition matrix methods, with particular emphasis on probabilities of positive advances in early cycles of selection, is indicated.

8.13 Life of Polymorphism

Half-life of the selection process, as defined by Robertson (1960) and investigated by Hill (1969a) and Hedrick (1970), is a measure of longevity in relation to attainment of the selection limit. Life of polymorphism, a random variable $T$ defined in Section 6.6, is introduced for different reasons not directly concerned with attainment of selection limits in recurrent mass selection programs; $T$ is defined for both selective and neutral systems.
When recurrent mass selection is of interest, examination of percentiles of the distribution of $T$ provides a stochastic description of utilization of genetic variance, and its eventual exhaustion, under the combined effects of selection and drift. Median life of polymorphism, $t_{50}$, has an appealing practical interpretation in this context: at cycle $t_{50}$ there is a 50:50 chance of residual genetic variance available for further selection progress. Effects of variations in the polygenic model, and in the mass selection parameters for monoecious and dioecious systems, upon life of polymorphism are discussed in detail in Section 6.6 by reference to $t_{50}$ and $t_{95}$.

It is also shown in Section 6.6 that life of polymorphism for polygenes is unlikely to be significantly prolonged by recurrent symmetric disruptive selection in populations with parental census numbers as small as twelve, when compared with neutral genes in systems with the same parental census.

8.14 The Double Polygene Model

Existing studies of recurrent mass selection in finite populations for models of two or more loci are predominantly computer simulations (for review, see Frazer and Burnell, 1970). Some theoretical results are obtained by Hill and Robertson (1968) and Karlin and McGregor (1968), and other results by mixtures of theory and simulation by Latter (1965b) and Hill and Robertson (1966).

Since the original results for a cycle of mass selection for a single polygene, obtained by Kojima (1961), have proved so useful in subsequent studies of recurrent selection using diffusion equation approximations, transition matrix methods and particularly the
\(\beta\)-distribution approximations of Curnow and Baker, it is natural to consider the extension of Kojima's results to one cycle of selection for two polygenes as an initial basis for theoretical studies of recurrent selection for two, or more, polygenes. This extension is provided in Chapter 7 with a full accounting of linkage disequilibrium in gametes and polygenic effects up to and including additive \(\times\) additive epistasis. Derivation of this extension followed directly from the recipe for pre-test moments of average responses developed in Chapter 3.

Detailed consequences of a single cycle of sampling, selection and recombination for this double polygene model are discussed in Sections 7.3 through 7.6. There remains the vexing question of how to use these single cycle developments in deriving theoretical and/or numerical results for recurrent selection without resorting to simulation methods. Theoretically it is possible to construct Markov chain models but dimensions of the required transition matrices appear prohibitive. An attractive alternative would be an extension of the \(\beta\)-distribution approximation to cover two gene frequencies; an initial search for an appropriate bivariate \(\beta\)-like distribution has been unsuccessful.
9. LIST OF REFERENCES


LIST OF REFERENCES (continued)


LIST OF REFERENCES (continued)


LIST OF REFERENCES (continued)


LIST OF REFERENCES (continued)


LIST OF REFERENCES (continued)


LIST OF REFERENCES (continued)


10. APPENDICES

10.1 Variances and Covariances of Normal Order Statistics

The standardized variance of the selection differential in the case of directional selection is defined at (2.2.2) as

$$v = \frac{1}{n} \sum_{i,j} \sigma_{i,j}|N, \ i,j = N-n+1, N-n+2, \ldots, N.$$  

In the case of Normally distributed $X$, for $N = 10$ and $20$, values of $v$ used in Table 2.3 were obtained from the tabulation of $\sigma_{i,j}|N$ given by Sarhan and Greenberg (1956); for $N = 30, 40, 50$ and $100$, values of $v$ in Table 2.3 were obtained by preparing a tabulation of $\sigma_{i,j}|N$ using the method described next.

David and Johnson (1954) present an inverse Taylor series expansion for general $\sigma_{i,j}|N$ which when specialized to the case of the unit Normal distribution provides the following expression:

$$\sigma_{i,j}|N = \sigma_{i,j}|N + \text{terms of order } (N+2)^{-3}, \ i \leq j$$  

$$\sigma_{i,j}|N = \frac{p_{1}q_{1}}{(N+2)} f_{1}^{-1} f_{j}^{-1} [1 + \frac{1}{(N+2)} \{(q_{1}-p_{1})Y_{1} f_{1}^{-1} + (q_{j}-p_{j})Y_{j} f_{j}^{-1} \}$$  

$$+ \frac{1}{2} p_{1}q_{j}Y_{1}Y_{j} f_{1}^{-1} f_{j}^{-1} + \frac{1}{2} p_{1}q_{1}(1 + 2Y_{1}^{2}) f_{1}^{-2} \}$$  

$$+ \frac{1}{2} p_{j}q_{j}(1 + 2Y_{j}^{2}) f_{j}^{-2})$$,

where $Y_{1}$ satisfies $F(Y_{1}) = 1/(N+1) = p_{1} = 1 - q_{1}$ and $f_{1} = f(Y_{1})$.

Calculations were performed on an IBM 360/75 computer in double precision. For given $N$ and $m$, $\sigma_{i,j}|N$ was calculated over the triangular array $1 \leq j = 1, 2, \ldots, m$, with the restriction $m \leq \frac{1}{2} (N+1)$. Values of $Y_{1}$ were obtained by a two-stage procedure. An initial approximation to $Y_{1}$ was calculated by
\[ Y_i = -t + \frac{(c_0 + c_1 t + c_2 t^2)}{(1 + d_1 t + d_2 t^2 + d_3 t^3)}, \]

where \( t^2 = -2\ln(p_i) \), and
\[
\begin{align*}
  c_0 &= 2.515517 & d_1 &= 1.432788 \\
  c_1 &= 0.802853 & d_2 &= 0.189269 \\
  c_2 &= 0.010328 & d_3 &= 0.001308
\end{align*}
\]

For \( 0 < p_i \leq 0.5 \) the absolute error in \( Y_i \) is \( < 0.00045 \) (Hastings, 1955). This initial value was modified iteratively until a value of \( Y_i \) was obtained satisfying
\[
|F(Y_i) - p_i| < 10^{-6}.
\]

\( F(Y_i) \) being calculated by
\[
F(Y_i) = f(Y_i)[a_1 h + a_2 h^2 + a_3 h^3 + a_4 h^4 + a_5 h^5],
\]
where \( h = [1 + (0.2316419)Y_i]^{-1} \), and
\[
\begin{align*}
  a_1 &= +0.319381530 & a_4 &= -1.821255978 \\
  a_2 &= -0.356563782 & a_5 &= +1.330274429 \\
  a_3 &= +1.781477937
\end{align*}
\]

For \( 0 \leq Y_i < \infty \) the absolute error in \( F(Y_i) \) is \( < 7.5 \times 10^{-8} \) (Hastings, 1955).

In this way covariances were obtained for the low order statistics, \( i \leq j = 1, 2, \ldots, m \), whereas for \( v \) the high order statistics, \( i \geq j = N-n+1, \ldots, N \), were required. However, symmetry properties of Normal order statistics, including
\[
\sigma_{i,j|N} = \sigma_{N-i+1,N-j+1|N}
\]
ensure that as far as \( \bar{\nu}(\Delta) \) is concerned it is irrelevant whether the \( n \) top-ranking or \( n \) bottom-ranking order statistics are averaged.

The calculated covariances, which neglect terms of order \((N+2)^{-3}\), were compared with those given by Sarhan and Greenberg (1956) for \( N = 10 \) and 20. Accuracy improved as \( N \) increased from 10 to 20, the proportionate errors halving except for the variance of the extreme order statistic; the actual errors were consistently negative, that is the calculated values were always less than tabulated values. At \( N = 20 \), the proportionate error was always better than \(-1.0\) percent (except for \( \sigma_{1,1|20} \) with \(-2.2\) percent), the results frequently agreed to three decimal places and always differed by less than 0.002 (except for \( \sigma_{1,1|20} \) with 0.006). Additionally, calculated values of \( \sigma_{1,1|N} \) were compared with those given by Ruben (1954) for \( N = 20(10)50 \). The proportionate error decreased from \(-2.2\) percent at \( N = 20 \) to \(-0.9\) percent at \( N = 50 \).

Thus with the exception of \( \sigma_{1,1|N} \), when \( N > 20 \) an accuracy close to three decimal places or better, with a proportionate error better than \(-1.0\) percent, is assured. Values of \( \nu \) given in Table 2.3 are given to four decimal places whereas it is clear from the comparisons above that only the first three decimal places can be cited with confidence when \( N \geq 30 \).

10.2 Some Properties of \( \nu \) under Directional Selection and Normally Distributed \( X \)

Let \( R_n = \{N-n+1, N-n+2, \ldots, N\} \) with standardized variance parameter \( \nu \) defined at (2.2.2) as

\[
\nu = \sum_i \sum_j \sigma_{i,j|N} \quad i, j \in R_n.
\]

(10.2.1)
Some properties of \( v \), that are used in Chapters 2 and 8, are derived here. It is convenient to adjust notation slightly and write \( v(n) \) for \( v \); \( f(x) \) is assumed to be the unit Normal density function throughout.

A basic result for \( \sigma_{i,j}|N \) in Normal samples is

\[
\sum_{j=1}^{N} \sigma_{i,j}|N = \sum_{i=1}^{N} \sigma_{i,j}|N = 1, \text{ all } i \text{ or } j. \tag{10.2.2}
\]

Presumably this property is known elsewhere but I have been unable to find it in standard texts such as Sarhan and Greenberg (1962); it is readily verified by consulting the tabulation of Sarhan and Greenberg (1956). Symmetry properties enable the partitioning of the \( N \times N \) variance-covariance matrix \( [\sigma_{i,j}|N] \) as follows:

\[
\begin{array}{c|c|c}
S_1 & S_2 & S_3 \\
n \times n & n \times (N-2n) & n \times n \\
\hline
S_2 & S_4 & S_2 \\
(N-2n) \times n & (N-2n) \times (N-2n) & (N-2n) \times n \\
\hline
S_3 & S_2 & S_1 \\
n \times n & n \times (N-2n) & n \times n \\
\end{array}
\]

Here each cell of the partition has been assigned \( S_i \) the sum of all elements it contains. Using (10.2.2),

\[
2S_1 + 4S_2 + 2S_3 + S_4 = N \tag{10.2.4}
\]

and by definition (10.2.1),

\[
v(n) = S_1. \tag{10.2.5}
\]

Consider first the assertion that \( 0 < v(n) < 1 \), for \( 0 < n < N \), at (2.7.1). Using (10.2.2) and (10.2.3),
\[ S_1 \leq (S_1 + S_2 + S_3) = n \]

the equality holding only if \( n = N \) (for only then does \( S_2 = S_3 = 0 \)).

Thus in (10.2.5)

\[ v(n) = \frac{1}{n} S_1 < 1 \quad n < N \]

as required at (2.7.2).

Consider next the symmetry property at (2.8.4). Rewriting (10.2.4) provides

\[ S_1 + (S_1 + 2S_2 + S_4) + 2(S_2 + S_3) = N \]

and substituting \( (S_2 + S_3) = n - S_1 \)

\[ (n - S_1) = (N - n) - (S_1 + 2S_2 + S_4) . \]

But \( S_1 = nv(n) \) and \( (S_1 + 2S_2 + S_4) = (N - n) \ v(N-n) \), so that

\[ n\{1 - v(n)\} = (N-n)\{1 - v(N-n)\} \quad (10.2.6) \]

as required for (2.8.4). That approximation \( \hat{v} \) for \( v(n) \) satisfies (10.2.6) is shown next. In the case of Normally distributed \( X \), \( \hat{v} \) at (2.6.1) reduces to

\[ \hat{v}(n) = 1 - \left\{ \frac{Nf(x_0)}{n} - x_0 \right\} \left\{ \frac{Nf(x_0)}{(N-n)} + x_0 \right\} \frac{(N-n)}{N} \quad G(x_0) = \frac{n}{N} \]

and so

\[ n\{1 - \hat{v}(n)\} = n\left\{ \frac{Nf(-x_0)}{n} + (-x_0) \right\} \left\{ \frac{Nf(-x_0)}{(N-n)} - (-x_0) \right\} \frac{(N-n)}{n} \]

\[ = (N-n)\left\{ \frac{Nf(-x_0)}{(N-n)} - (-x_0) \right\} \left\{ \frac{Nf(-x_0)}{n} + (-x_0) \right\} \frac{n}{N} \]

\[ G(-x_0) = \frac{(N-n)}{N} \]

\[ = (N-n)\{1 - \hat{v}(N-n)\} \]

as required at (2.8.5).
In the notation of Section 2.1 a somewhat surprising result is that
\( C(x_n, x_N) \) is not a function of \( n \):
\[
C(x_n, x_N) = \frac{\sigma^2}{N}.
\] (10.2.7)

This follows from
\[
C(x(i), x_N) = \frac{\sigma^2}{N}, \text{ all } i = 1, 2, \ldots, N
\]

obtained directly from (10.2.2).

Finally, consider the variance of the "reach" statistic \( D_n \) defined in Section 2.1:
\[
V(D_n^*) = V(x_n - x_N)
\]
\[
= V(x_n) - 2C(x_n, x_N) + V(x_N)
\]

and so, using (10.2.7),
\[
V(D_n^*) = \left[ v(n) - \frac{n}{N} \right] \frac{\sigma^2}{n}.
\] (10.2.8)

Since \( V(D_n^*) > 0 \) for \( n < N \),
\[
v(n) > \frac{n}{N}, \text{ } n < N
\] (10.2.9)
as required for (2.7.3).

It should be noted that all of these results are based on property (10.2.2) for the Normal distribution. It would be interesting to know if there is a general class of distributions for which (10.2.2) holds. It might be conjectured that the class of symmetric density functions would possess this property, but this is soon found not to be the case because of the following counter-example. The standardized uniform density, \( f(x) = 1/\theta \) for \( |x| < \frac{1}{2} \theta \) with \( \theta^2 = 12 \), is symmetric and Plackett (1960) provides the following expression for \( \sigma_{i,j|N}^2 \):
\[ \sigma_{i,j} | N = \frac{12i}{(N+1)^2} \frac{(N-i+1)}{(N+2)} \]

Thus, for all \( i = 1, 2, \ldots, N \),

\[
\sum_{j=1}^{N} \sigma_{i,j} | N = \frac{12}{(N+1)^2} \left\{ \sum_{j=1}^{N} \frac{i}{N} (N-n+1) + \sum_{j=1}^{N} \frac{i}{N} (N-j+1) - \sum_{j=1}^{N} \frac{i}{N} (N-j+1) \right\}
\]

\[
= \frac{6i}{(N+1)(N+2)} (N-i+1)
\]

and so (10.2.2) is not satisfied. It is still true, of course, that

\[ \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{i,j} | N = N, \text{ because} \]

\[ \sum_{i=1}^{N} i^2 = \frac{1}{6} N(N+1)(2N+1) \]

and so

\[ \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{i,j} | N = \frac{6}{(N+1)(N+2)} \sum_{i=1}^{N} i(N-i+1) \]

\[ = \frac{N}{(N+2)} \{3(N+1) - (2N+1)\} = N \]

as required.

10.3 Generalization of Approximations \( \hat{k} \) and \( \hat{\beta}(y_0, n) \)

in the Case of Directional Selection

The process by which approximations \( \hat{k} \) for \( k \) and \( \hat{\beta}(y_0, n) \) for \( \beta(y_0, n) \) were obtained can be generalized as follows. Let \( R_n \) and \( S_n \) denote the sets

\[ R_n \equiv \{N-n+1, N-n+2, \ldots, N\} \text{, } S_n \equiv \{(x_1, y_i); i \in R_n\} \]

and let \((x_*, y_*)\) denote a single member of \( S_n \) chosen at random, that is

\[ \Pr[(x_*, y_*) = (x_1, y_i)] = \frac{1}{n} \text{, all } i \in R_n. \]
Consider the class of expectational operators $\mathcal{U}$ defined by
\[ \mathcal{U}\{x_\star\} = E\{u(x_\star)\} , \] (10.3.1)

where $u(x_\star)$ is any function with existent expectation and possessing such analytical properties as are required in the following developments.

Under selection model II of Chapter I the analogue of $x_\star$ is $(X - \mu_x)/\sigma_x$ where $X$ is observed on a single randomly sampled unit from $\Pi_0$ after truncation of the standardized marginal distribution of $X$ to $x > x_0$. Define then
\[ \mathcal{U}\{x,x_0\} = E\{u(x)|x_0\} = \frac{1}{G(x_0)} \int_0^\infty u(x) f(x) \, dx . \] (10.3.2)

Now under selection model III let $r = N - n$ and, conditional on $x_{(r)}$, treat $x_{(r+1)}, x_{(r+2)}, \ldots, x_{(N)}$ as a random sample from $f(x)$ truncated at $x > x_{(r)}$. Then using (10.3.2),
\[ E\{u(x_\star)|x_{(r)}\} = \frac{1}{G(x_{(r)})} \int_{x_{(r)}}^\infty u(x) f(x) \, dx = \mathcal{U}\{x_\star,x_{(r)}\} \]
and removing the conditioning on $x_{(r)}$ provides
\[ \mathcal{U}\{x_\star\} = E[E\{u(x_\star)|x_{(r)}\}] \]
\[ = \int_{-\infty}^\infty \mathcal{U}\{x_\star,x_{(r)}\} z(x_{(r)};N,r) dx_{(r)} \]
\[ = \frac{N}{n} \int_{-\infty}^\infty \{\int u(x) f(x) \, dx\} z(t;N-1,r) \, dt , \]

where $z(x_{(r)};N,r)$ is the density function for the $r$th order statistic in a sample of size $N$ and is defined at (2.3.5).
Expanding in a Taylor series about \( t = a \) provides

\[
\int_{t}^{\infty} u(x) f(x) \, dx = \int_{a}^{\infty} u(x) f(x) \, dx - (t-a) u(a) f(a) - \frac{1}{2} (t-a)^2 \{u'(a) f(a) + u(a) f'(a)\} - \ldots .
\]

\[
= G(a) U(x_*, a) - u(a) f(a) \{ (t-a) + \frac{1}{2} (t-a)^2 \frac{f'(a)}{f(a)} \}
\]

\[
- \frac{1}{2} (t-a)^2 u'(a) f(a) - \ldots .
\]

Taking the integration at (10.3.3) term by term and choosing \( a = x_0 \)
satisfying \( G(x_0) = \frac{n}{N} \),

\[
U(x_*) = U(x_*, x_0) - \frac{1}{2} u'(x_0) \frac{f(x_0)}{G(x_0)} \{ (u_r | N-1 - x_0)^2 + \sigma_{r,r} | N-1 \}
\]

\[
- u(x_0) \frac{f(x_0)}{G(x_0)} \{ (u_r | N-1 - x_0) + \frac{1}{2} \sigma_{r,r} | N-1 \frac{f'(x_0)}{f(x_0)} \}
\]

\[
+ \frac{1}{2} (u_r | N-1 - x_0)^2 \frac{f'(x_0)}{f(x_0)} - \ldots .
\]

Then using the moment approximations at (2.3.11) as discussed in Chapter 2,

\[
U(x_*) = \hat{U}(x_*) + (N+1)^{-2} \{ \text{further terms} \}
\]

(10.3.4)

\[
\hat{U}(x_*) = U(x_*, x_0) - u'(x_0) C(n,N)
\]

where \( U(x_*, x_0) \) is defined at (10.3.2) and \( C(n,N) \) is the parameter of \( f(x) \)
defined at (2.3.12). Thus for selection model III, approximation \( \hat{U}(x_*) \)
for \( U(x_*) \) is equal to the analogue of \( U(x_*) \) under selection model II minus
the correction term \( u'(x_0) C(n,N) \).
10.4 Numerical Values of Parameters Used in the Diffusion Equation Approximations for Selection Limits

Table 6.1 contains values of \( \hat{u}(p_0) \) obtained for monoecious systems as

\[
\hat{u}(p_0) = \frac{I(p_0)}{I(1)} \quad I(x) = \int_{0}^{x} \exp\{-2ns[p - (1-2h)p(1-p)]\} \, dp.
\]

Tables 6.2 and 6.3 contain values of \( \hat{u}(p_0) \) for dioecious systems similarly obtained but with \( n \) replaced by \( n_d = 4n_m m_f/(n_m + m_f) \). Actual values of \( k, s \) and implied products \( ns \) or \( n_d s \), used in these tables are given in this section.

For Table 6.1 where \( (a/q_x) = 0.1 \)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>k</th>
<th>s</th>
<th>sn</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>1.52537</td>
<td>0.30507</td>
<td>0.610</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.20104</td>
<td>0.24021</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.96173</td>
<td>0.19235</td>
<td>1.154</td>
</tr>
<tr>
<td>N = 16</td>
<td>8</td>
<td>0.76018</td>
<td>0.15204</td>
<td>1.216</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.57704</td>
<td>0.11541</td>
<td>1.154</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.40035</td>
<td>0.08007</td>
<td>0.961</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>0.21791</td>
<td>0.04358</td>
<td>0.610</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1.85840</td>
<td>0.37168</td>
<td>0.743</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1.58334</td>
<td>0.31667</td>
<td>1.267</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1.23517</td>
<td>0.24703</td>
<td>1.976</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>0.98595</td>
<td>0.19719</td>
<td>2.366</td>
</tr>
<tr>
<td>N = 32</td>
<td>16</td>
<td>0.77867</td>
<td>0.15573</td>
<td>2.492</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.59157</td>
<td>0.11831</td>
<td>2.366</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>0.41172</td>
<td>0.08234</td>
<td>1.976</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>0.22619</td>
<td>0.04524</td>
<td>1.267</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.12389</td>
<td>0.02478</td>
<td>0.743</td>
</tr>
</tbody>
</table>
For Table 6.2 where $N_m = N_f = 8$ and $(a/\sigma_x) = 0.1$

<table>
<thead>
<tr>
<th>$(n_m, n_f)$</th>
<th>$n_d$</th>
<th>$\bar{k}$</th>
<th>$s$</th>
<th>$s_{n_d}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,1)</td>
<td>2</td>
<td>1.42360</td>
<td>0.28472</td>
<td>0.569</td>
</tr>
<tr>
<td>(2,2)</td>
<td>4</td>
<td>1.13791</td>
<td>0.22758</td>
<td>0.910</td>
</tr>
<tr>
<td>(3,3)</td>
<td>6</td>
<td>0.91621</td>
<td>0.18324</td>
<td>1.099</td>
</tr>
<tr>
<td>(4,4)</td>
<td>8</td>
<td>0.72529</td>
<td>0.14506</td>
<td>1.160</td>
</tr>
<tr>
<td>(5,5)</td>
<td>10</td>
<td>0.54973</td>
<td>0.10995</td>
<td>1.099</td>
</tr>
<tr>
<td>(6,6)</td>
<td>12</td>
<td>0.37930</td>
<td>0.07586</td>
<td>0.910</td>
</tr>
</tbody>
</table>

For Table 6.3 where $N_m = N_f = 8$ and $(a/\sigma_x) = 0.1$

<table>
<thead>
<tr>
<th>$(n_m, n_f)$</th>
<th>$n_d$</th>
<th>$k_m$</th>
<th>$k_f$</th>
<th>$\bar{k}$</th>
<th>$s$</th>
<th>$s_{n_d}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1,8)</td>
<td>3.5555</td>
<td>1.42360</td>
<td>0</td>
<td>0.71180</td>
<td>0.14236</td>
<td>0.506</td>
</tr>
<tr>
<td>(2,8)</td>
<td>6.4000</td>
<td>1.13791</td>
<td>0</td>
<td>0.56896</td>
<td>0.11379</td>
<td>0.728</td>
</tr>
<tr>
<td>(3,8)</td>
<td>8.7273</td>
<td>0.91621</td>
<td>0</td>
<td>0.45811</td>
<td>0.09162</td>
<td>0.800</td>
</tr>
<tr>
<td>(4,8)</td>
<td>10.6667</td>
<td>0.72529</td>
<td>0</td>
<td>0.36264</td>
<td>0.07253</td>
<td>0.774</td>
</tr>
<tr>
<td>(5,8)</td>
<td>12.3077</td>
<td>0.54973</td>
<td>0</td>
<td>0.27486</td>
<td>0.05497</td>
<td>0.677</td>
</tr>
<tr>
<td>(6,8)</td>
<td>13.7143</td>
<td>0.37930</td>
<td>0</td>
<td>0.18965</td>
<td>0.03793</td>
<td>0.520</td>
</tr>
<tr>
<td>(1,3)</td>
<td>3.0000</td>
<td>1.42360</td>
<td>0.91621</td>
<td>1.16991</td>
<td>0.23398</td>
<td>0.702</td>
</tr>
<tr>
<td>(1,5)</td>
<td>3.3333</td>
<td>1.42360</td>
<td>0.54973</td>
<td>0.98666</td>
<td>0.19733</td>
<td>0.658</td>
</tr>
<tr>
<td>(2,4)</td>
<td>5.3333</td>
<td>1.13791</td>
<td>0.72529</td>
<td>0.93160</td>
<td>0.18632</td>
<td>0.994</td>
</tr>
<tr>
<td>(1,7)</td>
<td>3.5000</td>
<td>1.42360</td>
<td>0.20337</td>
<td>0.81349</td>
<td>0.16270</td>
<td>0.569</td>
</tr>
<tr>
<td>(2,6)</td>
<td>6.0000</td>
<td>1.13791</td>
<td>0.37930</td>
<td>0.75861</td>
<td>0.15172</td>
<td>0.910</td>
</tr>
<tr>
<td>(3,5)</td>
<td>7.5000</td>
<td>0.91621</td>
<td>0.54973</td>
<td>0.73297</td>
<td>0.14659</td>
<td>1.099</td>
</tr>
<tr>
<td>(3,7)</td>
<td>8.4000</td>
<td>0.91621</td>
<td>0.20337</td>
<td>0.55979</td>
<td>0.11196</td>
<td>0.940</td>
</tr>
<tr>
<td>(4,6)</td>
<td>9.6000</td>
<td>0.72529</td>
<td>0.37930</td>
<td>0.55230</td>
<td>0.11046</td>
<td>1.060</td>
</tr>
<tr>
<td>(5,7)</td>
<td>11.6667</td>
<td>0.54973</td>
<td>0.20337</td>
<td>0.37655</td>
<td>0.07531</td>
<td>0.879</td>
</tr>
</tbody>
</table>