EXPERIMENTAL DESIGN IN FITNESS ESTIMATION

JEROME WILSON¹

Department of Biology, University of Chicago, Chicago, Illinois 60637

Received January 30, 1969

GENOTYPIC fitness is the probable contribution a zygote of a given genotype will make to the gene pool which forms the next generation. This contribution is usually measured relative to that of a standard genotype. Fitness, as determined from changes in genotypic distribution in a population, and the laws of recombination summarize a population's past and future systematic genetic change. Further motives for fitness estimation are these: Certain epistatic relations for fitness and certain recombination rates between loci would lead to long-term maintenance of genetic variability in a population [for example, see LEWONTIN 1964)]. Whether these relations between epistasis and linkage exist awaits fitness estimation in multilocus populations. The relationship between fitness and components of fitness remains tenuous. A physiological component of fitness—such as mating speed or hatchability in vinegar flies—connotes the intuition that the state of the character is systematically related to fitness. (See VAN VALEN (1969) for an attack on this problem.) Demonstration of the relationship requires a measure of fitness independent of the measure of the character state.

A fitness estimation experiment gives results in two areas. The first is the set of fitness estimates and their confidence intervals. The second is the degree of fit of estimates and estimation model to the observations of the experiment. Unless the results in both areas are acceptable, more data must be gathered or a more appropriate estimation model employed (or both). This paper investigates the relationship of confidence intervals to experimental design *assuming the estimation model to be correct*.

A variety of estimation models are used here, but they assume in common a population segregating for two alleles at one autosomal locus, random mating, discrete generations, no mutation, no parental effects, no meiotic drive, and no gametic selection. The generation interval specified by all models starts with zygote formation and ends with gamete pool formation. Each model assumes that the genotypic distribution of the population is observed at the same arbitrary time in each generation interval.

The selection models differ in assuming or not assuming that the selection pattern in the two sexes is identical. That is, selection is sex dependent (D) or sex independent (I). The models differ also in the relationship of the time of observation to the time of selection. One type of model specifies that selection is completed before the time of observation. The second specifies that selection occurs solely between the time of observation and gamete pool formation. The third

¹ Present address: Department of Biology, California State College, Fullerton, California 92631.

allows for selection both before and after observation in a generation interval. Following the usage of PROUT (1965), these models are called Early (E), Late (L), or Early and Late (EL). Thus the six models considered here are IE, IL, IEL, DE, DL, and DEL. The sex-independent models require the estimation of two relative fitnesses, taking the heterozygote as the standard of fitness equal to 1.00. The sex-dependent models require the estimation of four fitnesses, the heterozygote for each sex being the fitness standard of the sex. In the IEL and DEL models each genotype's net fitness is the product of its relative fitness before and after the time of observation in the generation interval. Symbolically, $W(i) = E(i)L(i)$; net genetic survival from zygote formation to observation ($E(i)$) times genetic survival from observation to gene pool formation ($L(i)$).

The selection model is not under the control of the investigator and, at least initially, he is ignorant of the correct selection model for the population. Therefore, he will desire an experimental design which will be efficient (in the broad sense of giving the most information for the least amount of work) for a variety of selection models. For this reason, the relative performance of an experimental design as selection model varies is of interest. The effect of the fitnesses themselves on the efficiency of the experimental design is of interest, also. Two broadly different selection patterns are used in this paper: one pattern is overdominant for fitness, the other is directional.

Two examples of experimental design are offered. One consists of a group of crosses between specific genotypes. It is called the single-cross design. The other full-population design consists of starting a set of populations at high gene frequency for the allele that will eventually be eliminated, given directional selection. If the selection pattern is overdominant, one-half of the set of populations is started at low gene frequency and the other half at high gene frequency. In either case, experiment size is the total of organisms screened over generations and populations. The starting gene frequencies of the full-population design were chosen to minimize the mean coefficient of variation for the IEL and DEL selection models.

The measure of experimental designs used below will be the average of the coefficients of variation of the fitnesses. This measure sacrifices some of the information contained in the vector of fitness estimates and their variances, but has compensating advantages. When a given selection model obtains and a particular experimental design is used, variances of fitness estimates will vary with the fitnesses for a given amount of data. The interest of comparing the effect of the fitness vector on variance of fitness estimates is not the absolute magnitude of the variances; it is the magnitude of the variance relative to the fitness, which the coefficient of variation (C.V.) measures. The C.V. reveals at a glance whether the genotype can be confidently distinguished from a sterile or lethal. Secondly, the relative efficiency of a good experimental design should be high whatever the selection model because the selection model is initially an unknown. I will compare the performance of the designs for each of the selection models. Because one selection model may stipulate more genotypic fitness than the other (e.g., sex-dependent *vs.* sex-independent selection), the average coefficient of variation

TABLE 1

Considerations of fitness estimation discussed in this study

Selection model	Factors intrinsic to the population Genotypic fitnesses	Experimental design	Factors extrinsic to the population Experiment size	Measuring the interaction of factors
IE IL IEL	Overdominant	Full population	The number of individuals tabulated	Sampling variances of fitness estimates
DE DL DEL	Directional	Single cross		↓ Coefficients of variation of fitness estimates
				↓ Mean coefficient of variation of fitness estimates

serves as a comparative measure of an experimental design when selection model varies.

Table 1 diagrams the considerations discussed above as they influence the mean coefficient of variation (M.C.V.).

METHODS

Experiments were simulated deterministically according to the six selection models for directional selection and overdominant selection, giving twelve simulations. The fitness vectors for directional and overdominant selection are listed in Table 2. To maintain comparability for the effect of varying selection models with a given fitness vector, the following rules apply: (1) When selection is sex dependent, the mean of the genotype's fitness in the two sexes equals the genotype's fitness when selection is sex independent. (2) When selection is both early and late, the product of the early and late fitness of a genotype's fitness equals its fitness under either the Early selection model or the Late selection model.

Table 3 describes the two experimental designs. The size of the experiment is increased by fixing population size and increasing replications or *vice versa*. The populations in the full-population regime are the same size in every generation. The fourteen crosses in the single-cross design have the same number of individuals of each sex in every cross in both parents and offspring.

After the data were simulated according to selection pattern, experimental regime, and selection model, they were subjected to the maximum likelihood estimators of the variances of maximum likelihood fitness estimates. The fitness estimates were taken as the true fitnesses (or components of fitness) appropriate to the model under which the data were simulated. This process is analogous to finding the sampling variance of a binomial population when the parameter of the population is known. It is not analogous to estimating the variance of a population empirically from several samples, the composition of which varies stochastically.

The paper of DuMOUCHEL and ANDERSON (1968) lucidly presents the rationale and technique of maximum likelihood estimation of fitnesses and their variances in genetic context. They treat only the case where selection occurs after the time of observation in a generation interval and is sex independent. A description of the estimators of fitness and fitness variances for the cases where selection may be sex dependent and occurring in two stages appears in the APPENDIX. The variance of a genotypic fitness estimate where $W_i = E \times L$ was computed to be:

$$\text{Var}(W) = E^2 \text{Var}(L) + L^2 \text{Var}(E) + 2EL \text{Cov}(E, L) + \text{Var}(E) \text{Var}(L) + \text{Cov}^2(E, L)$$

where E and L are parametric values.

TABLE 2

Genotypic fitness components for six selection models with overdominant and directional selection

Selection pattern	Selection model		Genotypes					
	Sex-dependent, Sex-independent	Early and/or Late	AA♂	Aa♂	aa♂	AA♀	Aa♀	aa♀
Overdominant	Sex-independent	Early	.73	1.0	.73	.73	1.0	.73
Overdominant	Sex-dependent	Early	.81	1.0	.81	.64	1.0	.64
Overdominant	Sex-independent	Late	.73	1.0	.73	.73	1.0	.73
Overdominant	Sex-independent	Early and Late (Early)	.85	1.0	.85	.85	1.0	.85
		(Late)	.85	1.0	.85	.85	1.0	.85
Overdominant	Sex-dependent	Late	.81	1.0	.81	.64	1.0	.64
Overdominant	Sex-dependent	Early and Late (Early)	.90	1.0	.90	.80	1.0	.80
		(Late)	.90	1.0	.90	.80	1.0	.80
Directional	Sex-independent	Early	1.0	.85	.73	1.0	.85	.73
Directional	Sex-dependent	Early	1.0	.90	.81	1.0	.80	.64
Directional	Sex-independent	Late	1.0	.85	.73	1.0	.85	.73
Directional	Sex-independent	Early and Late (Early)	1.0	.92	.85	1.0	.92	.85
		(Late)	1.0	.92	.85	1.0	.92	.85
Directional	Sex-dependent	Late	1.0	.90	.81	1.0	.80	.64
Directional	Sex-dependent	Early and Late (Early)	1.0	.95	.90	1.0	.89	.80
		(Late)	1.0	.95	.90	1.0	.89	.80

This is the sampling variance of the product EL if estimates of E and L are bivariate normal distributed. Given the variances of the net fitnesses, coefficients of variation were computed for each selection model and experimental design over a range of experiment sizes.

RESULTS

The twelve simulations are arranged in several ways. The figures (Figures 1A–1D) displaying mean coefficient of variation of the fitness estimates (as a function of experiment size, experimental design, selection model, and true fitnesses) have logarithmic abscissae and ordinates. Each of the four parts of Figure 1 specifies an experimental design (single cross or full population) and a selection pattern (directional or overdominant). The lines on each of the four graphs are labeled IE, DE, IL, etc. Each line indicates the expected mean coefficient of variation of a selection model as a function of experiment size. The lines IEL and DEL are absent from Figure 1C because the experiment size necessary to give M.C.V.s less than 50% was greater than 10^6 .

General effect of selection model: Adopt the convention that the experiment size required to achieve a given M.C.V. is the "data requirement." As expected, the data requirement of the selection models increases with number of parameters to be estimated. Thus, sex-dependent selection in a population increases an experiment's data requirement over sex-independent selection, all else being equal. More pointedly, the time of selection—early or late—systematically affects the data requirement of the fitness estimation simulations. Regardless of selection model, true fitnesses, or experimental design, late selection always requires more data to achieve a given M.C.V. than early selection. In Figure 1, the data require-

TABLE 3

Experimental designs

Experimental Regime	No. of Populations	No. of Generations	Genotypic distribution of each sex in first generation (%)					
			Males			Females		
			<u>AA</u>	<u>Aa</u>	<u>aa</u>	<u>AA</u>	<u>Aa</u>	<u>aa</u>
Full Population (Overdominant)	1	7	81	18	1	81	18	1
	1	7	1	18	81	1	18	81
Full Population (directional)	1	14	4	32	64	4	32	64
			(Selection against allele <u>a</u> .)					
Single Cross	1	2	100	0	0	0	100	0
	1	2	0	0	100	0	100	0
	1	2	0	50	50	0	50	50
	1	2	50	50	0	50	50	0
	1	2	100	0	0	0	50	50
	1	2	0	0	100	50	50	0
	1	2	50	0	50	100	0	0
	1	2	50	0	50	0	0	100
	1	2	100	0	0	50	0	50
	1	2	0	0	100	50	0	50
	1	2	50	50	0	100	0	0
	1	2	100	0	0	50	50	0
	1	2	0	50	50	0	0	100
	1	2	0	0	100	0	50	50

ment is from 50% to four magnitudes more than early selection, depending on the situation.

Effect of true fitness on data requirement: The effect of change in true fitnesses on data requirements is seen by scanning Figure 1 vertically. Figures 1A and 1C show the data requirement of the full-population design for overdominant and directional selection, respectively. Changing the true fitnesses has almost no effect on the data requirement of the full-population design if the population obeys the IE, DE, IL, or DL models. However, if the population obeys either the IEL or DEL models, the data requirements for directional selection are several magnitudes larger than those for overdominant selection. For a M.C.V. of 10%, the data requirement for the IEL and DEL models exceeds 10^8 individuals under directional selection. Under overdominant selection, 10^6 individuals give an M.C.V. of 10% for the IEL and DEL models.

Turning to the single-cross design (Figures 1B, 1D) the effect of changing true fitnesses is more uniform than for the full-population design. Regardless of the selection model, directional selection requires more data for a given M.C.V.

than overdominant selection. The increase in data requirement from directional to overdominant selection ranges from a few percent (Model DE) to approximately two-fold (Model DEL).

Comparison of experimental design: Figures 1A and 1B or Figures 1C and 1D contrast the effectiveness of the two experimental designs. Whether selection is overdominant or directional is irrelevant to the following findings. (1) The full-population design has an equal or smaller data requirement than the single-cross design if models IE, DE, or IL are appropriate for the population. Depending on the true fitnesses, selection model, and selection pattern, the single-cross design may require up to twice as much data as the full-population design to achieve a given M.C.V. (2) If model DL, IEL, or DEL is appropriate to the population, the single-cross design invariably requires the screening of fewer individuals than the full-population design for a given M.C.V. The difference in data requirement of the two experimental designs ranges from one to three magnitudes. The full-population design, where selection is a directional, is simply inadequate if either selection model DEL or IEL obtains.

Table 4 presents these findings in terms of amount of data required to give M.C.V.s of 50%, 5%, and 0.5%. The magnitude of confidence intervals implied by these M.C.V.s allows the investigator to generally (a) detect lethality or sterility, (b) characterize the selection pattern where selection is moderately intense, and (c) detect selection of the low intensity thought to be important in large natural populations.

DISCUSSION

If it is given that a population is going to run for, say, seven generations, there is an optimal starting gene frequency for minimizing the variances of fitness estimates from that population. The starting frequencies chosen for the full-population design were optimal in this respect. Clearly, this optimization of starting gene frequency is not possible when the true fitnesses are unknown; the optimization of starting gene frequency was done so that full-population design would appear at its best when compared to the single-cross design. The particular crosses of the single-cross design were chosen arbitrarily and there was no attempt made to optimize the design with respect to M.C.V.

It was generally seen that where the full-population design was superior to the single-cross design, it was moderately so. Where the single-cross design was superior, it was greatly so. Under two selection models (IEL and DEL) with di-

FIGURE 1.—Mean coefficient of variation of fitness estimation as a function of experiment size according to experimental design, selection pattern, and selection model. The selection models are: IE = sex independent, early selection; IL = sex independent, late selection; IEL = sex independent, early and late selection; DE = sex dependent, early selection; DL = sex dependent, late selection; DEL = sex dependent, early and late selection. A. Full-population experimental design, overdominant selection. B. Single-cross experimental design, overdominant selection. C. Full-population experimental design, directional selection. D. Single-cross experimental design, directional selection.

Experiment size is the number of individuals screened in the experiment.

TABLE 4
 "Data Requirement" of fitness estimation for mean coefficient of variation of 5% as a function of
 experimental design, selection pattern, and selection model

Experimental Design	Selection Pattern	MCV	Experiment size required for a given MCV as a function of selection model					
			I E	I L	IEL	D E	D L	DEL
Full Population	Over-dominant	(50%) 5% (0.5%)	1.9x10 ³	2.7x10 ³	(6.9x10 ³) 7.5x10 ⁴ (8.7x10 ⁶)	3.8x10 ³	1.4x10 ⁶	6.4x10 ⁶
	Directional	(50%) 5% (0.5%)	1.8x10 ³	5.0x10 ³	(1.0x10 ⁸) > 10 ⁹ (> 10 ¹⁰)	3.6x10 ³	2.4x10 ⁶	(1.7x10 ⁸) > 10 ⁹ (> 10 ¹⁰)
Single Cross	Over-dominant	5%	3.8x10 ³	1.1x10 ⁴	1.2x10 ⁴	7.5x10 ³	2.2x10 ⁴	2.6x10 ⁴
	Directional	5%	2.7x10 ³	3.6x10 ³	5.0x10 ³	5.4x10 ³	7.2x10 ³	2.3x10 ⁴

Unless otherwise noted, MCV's of 50% and 0.5% are calculated by multiplying with the factors 10⁻² and 10², respectively.

rectional selection, screening 10⁶ individuals would not reduce confidence intervals to less than 100% of the fitness estimates using the full-population design. The single-cross regime is expected to give confidence intervals of $\pm 10\%$ by

screening 25,000 individuals in the same situation. The single-cross design is superior when there is early and late selection or when there is sex-dependent late selection. Thus, the relevance of the single-cross design depends (1) on the existence of late and/or sex-dependent selection and (2) the consequences of assuming that late and/or sex-dependent selection does not exist when it does in fact.

PROUT (1965) has partially answered (2). By assuming that late selection does not exist (when it does), the estimation process gives the spurious appearance of frequency-dependent selection. A genotype will appear more fit as its frequency declines. Presumably this conclusion holds when selection is sex dependent as well as in the sex-independent case PROUT considered.

In one sense, sex-dependent selection is a trivial consideration. Where the difference in selection pattern between the sexes is small, fitness estimation assuming sex-independent selection will give an estimate near the mean of corresponding male and female genotypes. When an experiment is not designed to give highly precise fitness estimates anyway, the systematic but small error in ignoring sex-dependent selection may not be fatal to goodness of fit. On the other hand, to ignore the possibility of sex-dependent selection insures that it will not be detected; the maintenance of genetic variability by different selection patterns in the two sexes would then remain largely a curiosity of theory.

In cases where the marker being studied is expressed very early in the generation interval, late selection must necessarily exist. For example, the morphs of *Drosophila* larval protein polymorphism have competition for development survival and for mates between the time the population is sampled and the time the gene pool is formed. But late selection is not confined to such obvious instances. WILSON (1968), estimating early and late fitnesses in *D. melanogaster*, found that late selection against the less fit morphs was usually much more intense than early selection even though the flies were observed very late in the generation. Furthermore, no systematic relationships were found between the intensity of early selection and the intensity of late selection. The case, then, for estimating late selection is: (1) There is no way of eliminating the existence of late selection from experimental design unless the genotypes' contributions to the gamete pool are measured directly. (2) There has not been demonstrated a systematic relationship between early and late selection. (3) The character being screened may be expressed transiently early in the life cycle. (4) Cases exist where late selection is demonstrably important. Thus, an experimental design like the single-cross design must be the conservative choice for discrete generation fitness estimation where precision of estimation is desired. The alternative choice, estimation from a running population, requires an impractically large amount of data where selection is early and late or late and sex-dependent; its modest advantage if the other models obtain does not compensate for its certain inadequacy if they do not.

Fitness estimation in natural populations is militated against by environmental variability and the low likelihood of encountering large genotypic distribution changes in time. These considerations aside, a natural population's resemblance to the full-population experimental design implies the practical impossibility of

precise fitness estimation when late selection is combined with early or sex-dependent selection.

Another implication of this study is the cost of abandoning the random mating assumption from the fitness estimation model. In dioecious organisms nonrandom mating would be expressed as mating interaction parameters applied to pairs of male and female genotypes. Abandonment of random mating, therefore, means that more parameters must be estimated with attendant loss of precision for a given amount of data. The number of mating interaction parameters would increase geometrically with the number of segregating loci considered. It is also important to note that mating interaction estimation has a data requirement of the same magnitude as late selection parameters. The effects of mating interactions in one generation, like the effects of late selection, are not observed until a generation later; when observed, the mating interaction effects are confounded with the effects of recombination and early selection on the next generation's genotypic distribution. When matings are not observed directly, precise estimation of mating interactions would require unusually careful experimental design. If the above considerations hold, mating interaction parameters in natural populations are inaccessible to precise estimation.

SUMMARY

Fitness estimation in a population running for several generations is examined as an experimental design. It is found inadequate where selection occurs after the time of genotype screening in a generation interval if (a) selection is also sex dependent or (b) selection occurs between the time of zygote formation and genotype screening. Because of the resemblance between this experimental design and natural populations, the same conditions would preclude fitness estimation in natural populations without reference to complicating ecological factors. An alternative experimental design utilizing backcrosses, F_1 crosses, and pairwise mating competition is offered that is adequate under conditions (a) and (b); this alternative design was recommended for general use because the investigator cannot know whether the above conditions hold before the fact.

LITERATURE CITED

- DUMOUCHEL, W. H. and W. W. ANDERSON, 1968 The analysis of selection in experimental populations. *Genetics* **58**: 435-449.
- LEWONTIN, R. C., 1964 The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* **49**: 49-67.
- PROUT, T., 1965 The estimation of fitnesses from genotypic frequencies. *Evolution* **19**: 546-551.
- VAN VALEN, L., 1968 Fitness and phenotype in *Drosophila melanogaster*. (Abstr.). *Proc. 12th Intern. Congr. Genet.* **1**: 230.
- WILSON, J., 1968 Experimental determination of fitness interactions in *Drosophila melanogaster* by the method of marginal populations. *Genetics* **59**: 501-511.

APPENDIX

Notation:

Let i = genotype number, j = sex (say 1 for male and 2 for female), k = generation number, l = population or cross number, m = stage of selection (1 = Late, 2 = Early)

and $G(i, j, k, l)$ = number of genotype i of sex j in generation k of population l .

$EG(i, k, l)$ = expected frequency of genotype i of generation k of population l at zygote formation. This is equal in the two sexes.

$w(i, j, m)$ = fitness of genotype i , sex j , selection stage m .

$\bar{w}(j, k, l) = \sum_i w(i, j, 2) EG(i, k, l)$ = mean fitness of sex j in generation k of population l , computed from the time of zygote formation to the time of observation in the generation interval.

Sums of observations over a subscript are indicated by a dot in that subscript's position, e.g.,

$G(\cdot, 1, 3, 7)$ = number of males in generation 3 of population 7. $G(4, \cdot, 2, 6)$ = sum of genotype 4 in both sexes of generation 2, population 6.

When partial derivatives are taken with respect to fitness component $w(i, j, m)$

it will be denoted $\frac{\partial(\quad)}{\partial w(i', j', m')}$ for a first partial derivative and

$\left[\frac{\partial^2(\quad)}{\partial w(i', j', m') \partial w(i'', j'', m'')} \right]$ for a second partial derivative.

Assuming that the observed genotypic distributions are sampled from multinomial populations, the likelihood function (LF) of experimental outcomes is

$$LF = \text{Constant} \times \prod_{i,j,k,l} \left[\frac{w(i,j,2)EG(i,k,l)}{\bar{w}(j,k,l)} \right]^{G(i,j,k,l)} \quad (2)$$

The maximum likelihood estimates of the fitness components are those which maximize the likelihood function as well as the logarithm of the likelihood function ($\ln LF$)

$$\ln LF = \text{Constant} + \sum_{i,j,k,l} G(i,j,k,l) \ln \left(\frac{w(i,j,2)EG(i,k,l)}{\bar{w}(j,k,l)} \right) \quad (3)$$

To maximize the $\ln LF$ with respect to the vector of fitness components, first derivatives are taken and set equal to zero. The derivatives with respect to those fitness components taken as standards (to which the other genotypes' fitnesses are compared) are eliminated. This process leaves a set of simultaneous independent equations—presuming the data contain sufficient degrees of freedom—which may be solved by iterative methods. Summations over k indicated below are from the second generation to the last.

$$\begin{aligned} \frac{\partial(\ln LF)}{\partial w(i', j', m')} &= \sum_l \sum_k \left[\sum_j \sum_i \frac{G(i,j,k,l)}{w(i,j,2)} \cdot \frac{\partial w(i,j,2)}{\partial w(i', j', m')} \right. \\ &\quad \left. + \sum_i \frac{G(i, \cdot, k, l)}{EG(i, k, l)} \frac{\partial EG(i, k, l)}{\partial w(i', j', m')} - \sum_j \frac{G(\cdot, j, k, l) \bar{w}(j, k, l)}{\bar{w}(j, k, l) \partial w(i', j', m')} \right] \end{aligned} \quad (4)$$

Considering the right-hand terms of (4):

- (a) The first term vanishes except when $(i', j', m') = (i, j, 2)$
 (b) The second term vanishes when $m' = 2$.
 $\mathbf{EG}(i, k, l)$ is product of gene frequencies calculated from the observed genotypic array of the previous generation and the Late fitness components (the $w(i, j, 1)$).
 (c) The third term varies according to the value of m' (i.e., according to whether the component is Late or Early).

$$\frac{\partial \bar{w}(j, k, l)}{\partial w(i', j', 1)} = \sum_i w(i, j, 2) \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i', j', 1)} \quad (5a)$$

$$\frac{\partial \bar{w}(j, k, l)}{\partial w(i', j', 2)} = \mathbf{EG}(i', k, l) \quad (5b)$$

Hence,

$$\begin{aligned} \frac{\partial \ln LF}{\partial w(i', j', 1)} &= \sum_i \sum_k \left[\sum_i \frac{G(i, j, k, l)}{\mathbf{EG}(i, k, l)} \cdot \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i', j', 1)} \right. \\ &\quad \left. - \sum_j \frac{G(j, k, l)}{\bar{w}(j, k, l)} \sum_i w(i, j, 2) \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i', j', 1)} \right] \end{aligned} \quad (6a)$$

and

$$\frac{\partial \ln LF}{\partial w(i', j', 2)} = \frac{G(i', j', 2)}{w(i', j', 2)} - \sum_i \sum_k \frac{G(j', k, l)}{\bar{w}(j', k, l)} \mathbf{EG}(i', k, l) \quad (6b)$$

Second partial derivatives are taken to form the information matrix and solve the system of first partial derivatives which have been equated to zero.

We first note that

$$\frac{\partial^2 \mathbf{EG}(i, k, l)}{\partial w(i'', j'', 1) \partial w(i', j', 1)} = 0 \text{ when } j' = j''. \quad (7)$$

i.e., $\mathbf{EG}(i, k, l)$ is the function of a male gene frequency and a female gene frequency, not two male gene frequencies or two female gene frequencies. This term is underlined below where it occurs in the "Late-Late" second derivatives.

$$\begin{aligned} \frac{\partial^2 \ln LF}{\partial w(i'', j'', 1) \partial w(i', j', 1)} &= \sum_i \sum_k \left[\sum_i \frac{G(i, j, k, l)}{\mathbf{EG}^2(i, k, l)} \cdot \right. \\ &\quad \left(\mathbf{EG}(i, k, l) \frac{\partial^2 \mathbf{EG}(i, k, l)}{\partial w(i', j', 1) \partial w(i'', j'', 1)} - \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i', j', 1)} \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i'', j'', 1)} \right) \\ &\quad - \sum_j \frac{G(j, k, l)}{\bar{w}^2(j, k, l)} \left(\bar{w}(j, k, l) \sum_i w(i, j, 2) \frac{\partial^2 \mathbf{EG}(i, k, l)}{\partial w(i', j', 1) \partial w(i'', j'', 1)} \right. \\ &\quad \left. \left. - \sum_i w(i, j, 2) \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i', j', 1)} \sum_i w(i, j, 2) \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i'', j'', 1)} \right) \right] \end{aligned} \quad (8)$$

"Early-Late" second derivatives:

$$\begin{aligned} \frac{\partial^2 \ln LF}{\partial w(i'', j'', 2) \partial w(i', j', 1)} &= \sum_i \sum_k \frac{G(j'', k, l)}{\bar{w}^2(j'', k, l)} \cdot \\ &\quad \left(\mathbf{EG}(i'', k, l) \sum_i w(i, j'', 2) \frac{\partial \mathbf{EG}(i, k, l)}{\partial w(i, j', 1)} - \bar{w}(j'', k, l) \frac{\partial \mathbf{EG}(i'', k, l)}{\partial w(i', j', 1)} \right) \end{aligned} \quad (9)$$

“Early–Early” second derivatives:

$$\frac{\partial^2 \ln LF}{\partial w(i'', j'', 2) \partial w(i', j', 2)} = 0 \text{ when } j' \neq j''. \quad (10a)$$

Inspection of (6b) shows that $\frac{\partial \ln LF}{\partial w(i', j', 2)}$ is a function of the fitness components of one sex only (denoted by j'). Hence the second derivative with respect to the other sex ($j'' \neq j'$) vanishes.

Otherwise, when $j'' = j'$ and $i'' \neq i'$

$$\frac{\partial^2 \ln LF}{\partial w(i'', j'', 2) \partial w(i', j', 2)} = \sum_i \sum_k \frac{G(\cdot, j', k, l)}{\bar{w}^2(j', k, l)} \mathbf{E}G(i', k, l) \mathbf{E}G(i'', k, l) \quad (10b)$$

And when $j'' = j'$; $i'' = i'$ (the second derivative with respect to one Early fitness component),

$$\frac{\partial^2 \ln LF}{\partial^2 w(i', j', 2)} = - \frac{G(i', j', \cdot, \cdot)}{w^2(i', j', 2)} + \sum_i \sum_k \frac{G(\cdot, j', k, l)}{\bar{w}^2(j', k, l)} \mathbf{E}G^2(i', k, l) \quad (10c)$$

This development *in toto* applies to the situation where selection is sex-dependent, Early and Late. The formulae are easily applied to more constrained models by two rules of thumb.

- A. When selection is sex independent, sum over j where indicated and delete subscripts j , j' , and j'' . Equations (7) and (10a) do not hold for sex-independent selection.
- B. If selection is Early only, (6b), (10a), (10b), and (10c) are used. If selection is Late only, (6a) and (8) are used.