Selection Models*

BU-145-M Walter T. Federer May, 1962

I. INTRODUCTION

Since advance in the mean value of a character under selection depends upon the residual genetic variance from the preceding generation and upon the selection pressure, a simple recursive model may be used to describe the population mean at the i^{th} step in a selection program. Four forms of such a model are presented below together with estimators and variances of the estimators for various parameters. In addition, an application is made to one set of data obtained from the experiment described by Papa [1961]. Further detail on this experiment is given by Federer, Robson, and Srb [1959] and by Papa and Federer [1960].

For the models described herein, it is assumed that a large number of factors affect the character under consideration and that the factors affecting a character are similar in expression. (If few factors control the expression of the character under consideration, then some such model as described by Federer, Robson, and Srb [1959] would suffice.) The genotypic effects are assumed to be random variables identically and independently distributed with mean zero and common variance σ_g^2. The environmental effects are also assumed to be identically and independently distributed random variables with zero mean and common variance σ_e^2. In the bivariate distribution of environmental and genotypic effects, a zero covariance is postulated. This implies no genotype x environment interactions for the environments encountered in the collection of a set of data. Whether or not this assumption is justifiable depends upon the genetic material and the environments encountered.

---

* In the Biometrics Unit Mimeograph Series, Cornell University, Ithaca, New York

II. SELECTION MODELS

1. Selection model for known difference between mean of selected population and mean of unselected population

Let the yield of the j^{th} individual in generation i be expressed as:

$$Y_{ij} = \alpha + \sigma_g^2 \sum_{h=1}^{i} \frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}\sigma_g^2+\sigma_e^2} + \epsilon_{ij}$$

$$= \alpha + \sum_{h=1}^{i} \frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\beta} + \epsilon_{ij} \quad , \qquad (1)$$

where α = population mean in generation zero (i.e., in the generation in which selection is first practiced), σ_e^2 = environmental variance, σ_g^2 = genetic variance, $\beta=\sigma_e^2/\sigma_g^2$, b_{i-1} are known coefficients from an inbreeding series (e.g., $b_{i-1} = 2^{-i+1}$ in the selfing series for $i=1,2,\cdots$), $\bar{x}_{is}$ = true mean of population of selected individuals in generation i, $\bar{x}_{i.}$ = true mean of entire unselected population in generation i, $\bar{x}_{is}-\bar{x}_{i.}$ represents the true difference between means of selected individuals and the unselected population in generation i, ϵ_{ij} are identically and independently distributed random variables with zero mean and common variance σ_e^2, $i=1,2,\cdots,v$, and $j=1,2,\cdots,n_i$ = number of observations in generation i. The total phenotypic variance in the generation in which selection is first practiced (the zeroth generation of selection) is $\sigma_g^2+\sigma_e^2$, and the phenotypic variance in generation i is $b_{i-1}\sigma_g^2+\sigma_e^2$. For i=0, $b_{0-1}=0$; thus, $Y_{0j}=\alpha+\epsilon_{0j}$.

In practice, the value $\bar{x}_{is}-\bar{x}_{i.}$ is usually unknown. The experimenter could, however, set $\bar{x}_{is}-\bar{x}_{i.}=(c=\text{a constant})^{1/i}$. Then, the selected progeny in generation i would be all those whose mean $\bar{x}_{is}$ exceeded the unselected population mean $\bar{x}_{i.}$ by $c^{1/i}$. This would require that a large number of individuals be observed in each generation. Because the value for $\bar{x}_{is}-\bar{x}_{i.}$ may be difficult or impossible to obtain, the model in the following section was proposed.

The meaning of the statement that the ϵ_{ij} have common variance σ_e^2 requires amplification. Basically, this statement implies that for a true breeding population the expected value of the variance among v observations taken singly in each of the environments in which the $\bar{y}_{i.}$ are obtained is equal to the expected value of the variance among v observations in generation i. In the analysis of variance table, this would mean:

Source of variance	d.f.	Average value of mean square
Among generations	v-1	σ_e^2
Within generations	v(n-1)	σ_e^2

Also, this implication could be expressed symbolically as:

$$V(Y_{ij}/i) = V(Y_{ij}) \ ,$$

where

$$V(Y_{ij}) = E(Y_{ij}-EY_{ij})^2 = \sigma_e^2 \ .$$

2. Selection model when genotypic and environment effects are independently and normally distributed

If the genotypic and environmental effects are independently and normally distributed, then equation (1) may be rewritten as:

$$Y_{ij} = \alpha + \sigma_g^2 \sum_{h=1}^{i} \frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\sigma_g^2+\sigma_e^2}} + \epsilon_{ij} \quad , \tag{2}$$

where $\bar{z}_{mh}$ are constants obtained from Table XX of Fisher and Yates [1938], Tables 2 and 3 of Federer [1951], or Table 1 of Harter [1961]. (The last reference is more extensive and contains more significant figures than do the first two references.) and where the other symbols are as defined for (1) except that the normality condition on the distribution of the random variables is imposed. The constant $\bar{z}_{mi}$ is the average value of the largest member from a sample of size m from a normal population with zero mean and unit variance. In other words it is the expected value of the largest rank order statistic from a sample of size m. $\bar{z}_{mi}$ is a constant for each generation if m is constant from generation to generation; otherwise, $\bar{z}_{mi}$ will vary with m and reflects the selection pressure in generation i. Also, it may be noted that $(\bar{x}_{hs}-\bar{x}_{h.})/\sqrt{b_{h-1}\sigma_g^2+\sigma_e^2}$ in equation (1) is replaced by $\bar{z}_{mh}$ in equation (2).

3. Selection model when $(\bar{x}_{is}-\bar{x}_{i\cdot})$ is unknown and when the normality assumption for equation (2) is not tenable

In certain situations the assumption of normal distribution of the genotypic and the environmental effects may be untenable. For this case we proceed by rewriting equation (1) as:

$$Y_{ij} = \alpha + \delta \sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\beta}} + \epsilon_{ij} \quad , \tag{3}$$

where $\delta=\sigma_g(\bar{x}_{hs}-\bar{x}_{h\cdot})/\sqrt{b_{h-1}\sigma_g^2+\sigma_e^2}$, where a_h are known constants reflecting changes in selection pressure from generation to generation (a_h=1 if selection pressure is constant throughout all generations of selection), and where the remaining symbols are defined in the same manner as for equation (1). Also, if the normality assumption holds $\delta a_h=\sigma_g \bar{z}_{mh}$ from equation (2). The parameters to be estimated are α, δ, and β.

4. Model for mean progress under selection from generation i to generation i+1

Suppose that at generation i in a selection program it is desired to know (using parameters) or to estimate (using estimates) the mean advance to be made in the i+1st generation for a given selection pressure. From equation (1) the following results:

$$W_{i+1} = \bar{y}_{i+1\cdot} - \bar{y}_{i\cdot} = \frac{b_i(\bar{x}_{i+1,s}-\bar{x}_{i+1,\cdot})}{b_i+\beta}$$

$$+ \frac{1}{n_{i+1}} \sum_{j=1}^{n_{i+1}} \epsilon_{i+1,j} - \frac{1}{n_i} \sum_{j=1}^{n_i} \epsilon_{ij} \quad . \tag{4}$$

For this simple Markovian process the mean advance from generation i to generation i+1 is:

$$E[W_{i+1}=\bar{y}_{i+1,\cdot}-\bar{y}_{i\cdot}] = \frac{b_i(\bar{x}_{i+1,s}-\bar{x}_{i+1,\cdot})}{b_i+\beta} \quad . \tag{5}$$

In the above $\bar{y}_{i}.$ and $\bar{y}_{i+1}.$ are the arithmetic means of the observations in generation i and in i+1 and the remaining symbols are as defined for equation (1).

If the genotypic and environmental effects are normally and independently distributed, equation (5) may be written in the form:

$$E[\bar{y}_{i+1}.-\bar{y}_{i}.] = \frac{b_i \sigma_g \bar{z}_{m,i+1}}{\sqrt{b_i+\beta}} \tag{6}$$

In order to simplify results and minimize variances the n_i should be equal.

III. ESTIMATION OF PARAMETERS

1. Estimation of α, σ_e^2, and σ_g^2 from a separate experiment

If the progeny from the different generations of selections are all compared in one experiment, estimates of α, σ_e^2, and σ_g^2 may be obtained from v individuals (spores, strains, etc.) from the unselected population each replicated n times. Then $\hat{\alpha}=\sum_{i=1}^{v}\sum_{j=1}^{n} Y_{ij}/nv=\bar{y}$, and $\hat{\sigma}_e^2$ and $\hat{\sigma}_g^2$ may be obtained from the following analysis of variance:

Source of variation	d.f.	Mean square	E[m.s.]
Among individuals	v-1	A	$\sigma_e^2+n\sigma_g^2$
Within individuals	v(n-1)	B	σ_e^2

as $\hat{\sigma}_e^2=B$ and $\hat{\sigma}_g^2=(A-B)/n$.*

If the generation means are obtained from a series of experiments, the unselected population will necessarily be included in each experiment if the variance of observations among experiments is different from the variation among

* There are several estimators for σ_e^2 and σ_g^2 in the literature, but a discussion of these is not pertinent to this paper [see Federer, 1962].

individuals within experiments. If the variance of observations among experiments is equal to the variance within experiments, then a single experiment could be conducted at the beginning of the selection program and estimates of α, σ_e^2, and σ_g^2 obtained at this stage.

For any of the models postulated by equations (1) through (3), or any variation of them, an estimator for σ_e^2 is simply the sum of squares of deviations between the generation mean and the fitted point on the curve divided by the degrees of freedom. Whether or not the variance among the n_i observations in generation i is an estimator for σ_e^2 depends upon the conditions stated in the preceding paragraph.

2. Estimation of α and β from equation (1) for $(\bar{x}_{is}-\bar{x}_{i.})$ and b_{i-1} known

To obtain the estimators in this and following sections, the least squares procedure will be used. Four reasons for adopting this procedure are (i) for the ϵ_{ij} normally and independently distributed with mean zero and common variance σ_e^2, the least squares estimators are also maximum likelihood estimators; (ii) the form of the distribution need not be specified; (iii) solutions are possible by least squares procedure which may not be feasible using other procedures; and (iv) theoretical results from standard regression theory are applicable.

When the residual sum of squares from equation (1),*

$$\sum_{i=1}^{v}\sum_{j=1}^{n_i}\left[Y_{ij}-\alpha-\sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\beta}\right]^2 , \tag{7}$$

is differentiated with respect to α and β, when the resulting equations are equated to zero, and when some algebraic manipulations are performed the following equations result:

$$\hat{\alpha} = \bar{y} - \frac{1}{n_.}\sum_{i=1}^{v} n_i \sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\hat{\beta}} \tag{8}$$

$$\sum_{i=1}^{v}(Y_{i.}-n_i\bar{y})\sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{(b_{h-1}+\hat{\beta})^2} - \sum_{i=1}^{v} n_i \sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{(b_{h-1}+\hat{\beta})^2} \cdot$$

$$\left\{\sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\hat{\beta}} - \frac{1}{n_.}\sum_{i=1}^{v} n_i \sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\hat{\beta}}\right\} = 0 \ . \tag{9}$$

---

* For certain values of Y_{ij} and b_{i-1} equation (7) has a minimum. This may not be true for values of Y_{ij} outside a given range and, therefore, it would not be possible to obtain estimators for α and β as described in this section (see section III-7 for further detail). When the range of Y_{ij} for which equation (7) has a minimum is determined, the Y_{ij} may be transformed to fall in this range.

Equation (9) in $\hat{\beta}$ is solved iteratively, and then $\hat{\alpha}$ is obtained by inserting the solution for $\hat{\beta}$ from (9) in (8). If normality holds asymptotic variances for $\hat{\alpha}$ and $\hat{\beta}$ may be obtained from standard maximum likelihood procedures.

As explained previously an estimator for σ_e^2 may be obtained as:

$$\frac{1}{(n_{.}-2)}\sum_{i=1}^{v}\sum_{j=1}^{n_i}\left[Y_{ij}-\hat{\alpha}-\sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\hat{\beta}}\right]^2 \tag{10}$$

with $n_{.}-2$ degrees of freedom. If the variation from generation to generation is larger then within generation variance, then σ_e^2 is estimated as:

$$n_0\hat{\sigma}_e^2 = \frac{1}{v-2}\sum_{i=1}^{v} n_i\left(\bar{y}_{i.}-\hat{\alpha}-\sum_{h=1}^{i}\frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\hat{\beta}}\right)^2 \tag{11}$$

where $n_0=(n_{.}-\Sigma n_i^2/n_{.})/(v-2)$ and where v-2= degrees of freedom for $\hat{\sigma}_e^2$.

Having estimates of β and σ_e^2, an estimate of σ_g^2 is obtained as:

$$\hat{\sigma}_g^2 = \hat{\sigma}_e^2/\hat{\beta} \quad .$$

In analysis of variance terminology the results may be summarized as:

Source of variation	d.f.	Sum of squares
Among generation totals	v	$\sum_{i=1}^{v} Y_{i.}^2/n_i$
Due to fitted regression	2	subtraction
Deviation from fitted regression	v-2	Equation (11)
Within generations	$n_{.}-v$	$\sum_{i=1}^{v}\left\{\sum_{j=1}^{n_i} Y_{ij}^2-\frac{Y_{i.}^2}{n_i}\right\}$
Total	$n_{.}$	$\sum_{i=1}^{v}\sum_{j=1}^{n_i} Y_{ij}^2$

The sum of squares for β eliminating the effect of α is obtained as:

$$SS(\hat{\alpha},\hat{\beta})-SS(\tilde{\alpha})$$

$$=\sum_{i=1}^{v} Y_{i.}^2/n_i - \text{equation (11)}$$

$$-\tilde{\alpha}\ (Y_{..}-\sum_{i=1}^{n} n_i \sum_{h=1}^{i} (\bar{x}_{hs}-\bar{x}_{h.}))\quad , \tag{12}$$

where $\tilde{\alpha}$ is obtained from equation (8) by setting β=0 in equation (7). Likewise, a sum of squares for α (eliminating the effect of β) could be obtained as:

$$SS(\hat{\alpha},\hat{\beta})-SS(\tilde{\beta})\quad , \tag{13}$$

where $\tilde{\beta}$ is obtained from minimizing equation (7) after setting α=0; thus,

$$\sum_{i=1}^{v} Y_{i.} \sum_{h=1}^{i} \frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{(b_{h-1}+\tilde{\beta})^2}$$

$$=\sum_{i=1}^{v} n_i \left(\sum_{h=1}^{i} \frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{b_{h-1}+\tilde{\beta}}\right)\left(\sum_{h=1}^{i} \frac{b_{h-1}(\bar{x}_{hs}-\bar{x}_{h.})}{(b_{h-1}+\tilde{\beta})^2}\right)\ . \tag{14}$$

3. <u>Estimation of α, σ_e^2, and σ_g^2 from equation (2) for $\bar{z}_{mi}$ and b_{i-1} known</u>

Given that the $\bar{z}_{mi}$ and b_{i-1} are known and that the effects are normally and independently distributed, either α, σ_e^2, and σ_g^2 could be estimated from minimization of equation (15) below or α, σ_g and β could be estimated by minimization of equation (16) below (provided a minimum exists):

$$\sum_{i=1}^{v} \sum_{j=1}^{n_i} \left[Y_{ij}-\alpha-\sigma_g^2 \sum_{h=1}^{i} \frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\sigma_g^2+\sigma_e^2}}\right]^2\quad ; \tag{15}$$

$$\sum_{i=1}^{v} \sum_{j=1}^{n_i} \left[Y_{ij}-\alpha-\sigma_g \sum_{h=1}^{i} \frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}+\beta}}\right]^2\quad ; \tag{16}$$

If equation (15) is minimized with respect to α, σ_e^2, and σ_g^2, we obtain:

$$n_{..}\check{\alpha} + \check{\sigma}_g^2 \sum_{i=1}^{n} n_i \sum_{h=1}^{i} \frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2}} = Y_{..} \quad ; \tag{17}$$

$$\sum_{i=1}^{v}\left\{Y_{i.}-n_i\check{\alpha}-\check{\sigma}_g^2 n_i \sum_{h=1}^{i}\frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2}}\right\}\left\{\sum_{h=1}^{i}\frac{b_{h-1}\bar{z}_{mh}}{(b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2)^{3/2}}\right\} = 0 \; ; \tag{18}$$

and

$$\sum_{i=1}^{v}\left\{Y_{i.}-n_i\check{\alpha}-n_i\check{\sigma}_g^2 \sum_{h=1}^{i}\frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2}}\right\}\left\{\sum_{h=1}^{i}\frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2}} - \frac{\check{\sigma}_g^2}{2}\sum_{h=1}^{i}\frac{b_{h-1}^2\bar{z}_{mh}}{(b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2)^{3/2}}\right\} = 0 \; . \tag{19}$$

Something could be done about asymptotic variances for $\check{\alpha}$, $\check{\sigma}_g^2$, and $\check{\sigma}_e^2$ by maximum likelihood procedures.

If σ_e^2 is known or is estimated from another experiment, equations (17) and (19) would be solved to obtain estimates of α and σ_g^2.

The analysis of variance for fitting equations (15) (or equation (16)) is:

Source of variation	d.f.	Sum of squares
Among generations	v	$\Sigma Y_{i.}^2/n_i$
Due to regression	3	subtraction
Deviations from regression	$v-3$	$\sum_{i=1}^{v} n_i \left[\bar{y}_{i.}-\check{\alpha}-\check{\sigma}_g^2 \sum_{h=1} \frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2}}\right]^2$
Within generations	$n_.-v$	$\sum_{i=1}^{v}\left\{\sum_{j=1}^{n_i} Y_{ij}^2-Y_{i.}^2/n_i\right\}$
Total	$n_.$	$\sum_{i=1}^{v}\sum_{j=1}^{n_i} Y_{ij}^2$

As in the previous section, the following sums of squares could be computed to eliminate the effects of the other fixed effects:

$$SS(\check{\alpha},\check{\sigma}_e^2,\check{\sigma}_g^2)-SS(\ddot{\alpha})=\sum_{i=1}^{v} Y_{i.}^2/n_i-\sum_{i=1}^{v} n_i\left[\bar{y}_{i.}-\check{\alpha}-\check{\sigma}_g^2\sum_{h=1}^{i}\frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\check{\sigma}_g^2+\check{\sigma}_e^2}}\right]^2$$

$-Y_{..}^2/n_.$ = sum of squares for σ_e^2 and σ_g^2 eliminating the effect of α. (20)

$SS(\check{\alpha},\check{\sigma}_e^2,\check{\sigma}_g^2)-SS(\dddot{\alpha},\dddot{\sigma}_g^2)$ = sum of squares due to σ_e^2 alone. (21)

$SS(\check{\alpha},\check{\sigma}_e^2,\check{\sigma}_g^2)-SS(\overset{+}{\sigma}_g^2)$ = sum of squares due to α and σ_e^2 eliminating the effect of σ_g^2. (22)

$SS(\check{\alpha},\check{\sigma}_e^2,\check{\sigma}_g^2)-SS(\bar{\sigma}_e^2,\bar{\sigma}_g^2)$ = sum of squares due to α alone. (23)

4. <u>Estimation of α, β, and δ from equation (3) for b_{h-1} and a_h known</u>

The residual sum of squares from equation (3) is:

$$\sum_{i=1}^{v}\sum_{j=1}^{n_i}\left[Y_{ij}-\alpha-\delta\sum_{h=1}^{i}\frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\beta}}\right]^2 \quad , \tag{24}$$

where a_i and b_{i-1} are known constants. Upon differentiating (24) with respect to α, δ, and β, and equating the resulting equations to zero, the following equations are obtained:

$$Y_{..}-n_.\check{\alpha}-\check{\delta}\sum_{i=1}^{i} n_i\sum_{h=1}\frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\check{\beta}}}=0 \quad ; \tag{25}$$

$$\sum_{i=1}^{v}\sum_{j=1}^{n_i}\left[Y_{ij}-\check{\alpha}-\check{\delta}\sum_{h=1}^{i}\frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\check{\beta}}}\right]\left[\sum_{h=1}^{i}\frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\check{\beta}}}\right]=0 \quad ; \tag{26}$$

and

$$\sum_{i=1}^{v} \sum_{j=1}^{n_i} \left[Y_{ij} - \breve{\alpha} - \breve{\delta} \sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right] \left[\sum_{h=1}^{i} \frac{b_{h-1}a_h}{(b_{h-1}+\breve{\beta})^{3/2}}\right] = 0 \quad . \tag{27}$$

From equation (25)

$$\breve{\alpha} = \bar{y} - \frac{\breve{\delta}}{n_.} \sum_{i=1}^{v} n_i \sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}} \quad . \tag{28}$$

Substitution for $\breve{\alpha}$ in equation (26) results in:

$$\breve{\delta} = \frac{\sum_{i=1}^{v} \left[Y_{i.} - n_i \bar{y}\right] \left[\sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right]}{\sum_{i=1}^{v} n_i \left(\sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right)^2 - \frac{1}{n_.} \left(\sum_{i=1}^{v} n_i \sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right)^2} \quad . \tag{29}$$

Substituting for $\breve{\alpha}$ and $\breve{\delta}$ in equation (27) results in the following equation in $\breve{\beta}$:

$$\sum_{i=1}^{v} \sum_{j=1}^{n_i} \left[Y_{ij} - \bar{y} - \frac{\sum_{i=1}^{v} (Y_{i.} - n_i \bar{y}) \sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}}{\sum_{i=1}^{v} n_i \left(\sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right)^2 - \frac{1}{n_.} \left(\sum_{i=1}^{v} n_i \sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right)^2} \left\{\sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}} \right.\right.$$

$$\left.\left. - \frac{1}{n_.} \sum_{i=1}^{v} n_i \sum_{h=1}^{i} \frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right\}\right] \cdot \left[\sum_{h=1}^{i} \frac{b_{h-1}a_h}{(b_{h-1}+\breve{\beta})^{3/2}}\right] = 0 \quad . \tag{30}$$

The value for $\breve{\beta}$ satisfying equation (30) is obtained iteratively. Since asymptotic variances for $\breve{\alpha}$, $\breve{\delta}$, and $\breve{\beta}$ appear rather formidable and tedious, we should note that the $V(\breve{\alpha}|\breve{\beta}=\beta)$ and $V(\breve{\delta}|\breve{\beta}=\beta)$ follow from standard linear regression theory.

Also, the analysis of variance fitting regression equation (3) to the data is:

Source of variation	d.f.	Sum of squares
Among generations	v	$\Sigma Y_{i.}^2/n_i$
Due to regression	3	subtraction
Deviations from regression	$v-3$	$\sum_{i=1}^{n} n_i\left(\bar{y}_{i.}-\breve{\alpha}-\breve{\delta}\sum_{h=1}^{i}\frac{b_{h-1}a_h}{\sqrt{b_{h-1}+\breve{\beta}}}\right)^2$
Within generations	$n_.-v$	$\sum_{i=1}^{v}\left\{\sum_{j=1}^{n_i} Y_{ij}^2-Y_{i.}^2/n_i\right\}$
Total	$n_.$	$\sum_{i=1}^{v}\sum_{j=1}^{n_i} Y_{ij}^2$

Here again the various sums of squares given by equations (20) to (23) are possible for partitioning the 3 degrees of freedom due to regression in the above analysis of variance.

5. Moment estimators for α and β from equation (1) with b_{i-1} and $\bar{x}_{is}-\bar{x}_{i.}$ known

From equations (4) to (5) we note that equation (1) may be put in simpler form by using differences of successive generation means. Thus,

$$E(W_1=\bar{y}_{1.}) = \alpha + b_0(\bar{x}_{1s}-\bar{x}_{1.})/(b_0+\beta) \quad , \tag{31}$$

$$E(W_2=\bar{y}_{2.}-\bar{y}_{1.}) = b_1(\bar{x}_{2s}-\bar{x}_{2.})/(b_1+\beta) \quad , \tag{32}$$

$$E(W_3=\bar{y}_{3.}-\bar{y}_{2.}) = b_2(\bar{x}_{3s}-\bar{x}_{3.})/(b_2+\beta) \quad , \tag{33}$$

$$\vdots$$

Equating the above differences of observed means to their expected values results in the following moment estimators:

$$\mathring{\beta} = \frac{1}{v}\sum_{i=1}^{v-1}\frac{b_i(\bar{x}_{i+1,s}-\bar{x}_{i+1.}-\bar{y}_{i+1.}+\bar{y}_{i.})}{\bar{y}_{i+1.}-\bar{y}_{i.}} \tag{34}$$

and

$$\overset{\circ}{\alpha} = \bar{y}_{1.} - \frac{b_0(\bar{x}_{1s} - \bar{x}_{1.})}{b_0 + \overset{\circ}{\beta}} \tag{35}$$

There is a positive probability that W_{i+1} will be zero or negative resulting in the conclusion that σ_g^2 is zero in the ratio $\sigma_e^2/\sigma_g^2 = \beta$.

It should be noted that both σ_e^2 and σ_g^2 cannot be estimated from equation (1), since this equation depends only upon the ratio σ_e^2/σ_g^2.

6. Least squares estimators using mean differences between successive generations

If equation (1) holds and if the n_i are equal, least squares estimates could be obtained by minimizing the following sum of squares:

$$\left(W_1 - \alpha - \frac{b_0(\bar{x}_{1s.} - \bar{x}_{1.})}{b_0 + \beta}\right)^2 + \sum_{i=2}^{v}\left(W_i - \frac{b_{i-1}(\bar{x}_{is} - \bar{x}_{i.})}{b_{i-1} + \beta}\right)^2 \quad , \tag{36}$$

resulting in

$$\acute{\alpha} = \bar{y}_{1.} - \frac{b_0(\bar{x}_{1s} - \bar{x}_{1.})}{b_0 + \acute{\beta}} \tag{37}$$

and

$$\sum_{i=2}^{v} \frac{b_{i-1}^2(\bar{x}_{is} - \bar{x}_{i.})^2}{(b_{i-1} + \acute{\beta})^3} \sum_{i=2}^{v} \frac{W_i b_{i-1}(\bar{x}_{is} - \bar{x}_{i.})}{(b_{i-1} + \acute{\beta})^2} = 0 \quad . \tag{38}$$

Equation (38) is solved iteratively for $\acute{\beta}$. Even in this form, variances for the estimators are not straightforward.

From equation (3), successive differences of generation means could be obtained and the following sum of squares could be minimized:

$$\left(\bar{y}_{1.} - \alpha - \frac{b_0\delta}{\sqrt{b_0 + \beta}}\right)^2 + \sum_{i=2}^{v}\left(W_i - \frac{b_{i-1}\delta}{\sqrt{b_{i-1} + \beta}}\right)^2 \quad .$$

The resulting equations are:

$$\grave{\alpha} + \frac{b_0 \grave{\delta}}{\sqrt{b_0 + \grave{\beta}}} = \bar{y}_{1\cdot} = W_1 \quad , \tag{39}$$

$$\grave{\delta} \sum_{i=2}^{v} \frac{b_{i-1}^2}{b_{i-1} + \grave{\beta}} = \sum_{i=2}^{v} \frac{W_i b_{i-1}}{\sqrt{b_{i-1} + \grave{\beta}}} \quad , \tag{40}$$

and

$$\sum_{i=1}^{v} \frac{W_i b_{i-1}}{(b_{i-1} + \grave{\beta})^{3/2}} - \frac{\grave{\alpha} b_0}{(b_0 + \grave{\beta})^{3/2}} - \grave{\delta} \sum_{i=1}^{v} \frac{b_{i-1}^2}{(b_{i-1} + \grave{\beta})^2}$$

$$= \sum_{i=2}^{v} \frac{W_i b_{i-1}}{(b_{i-1} + \grave{\beta})^{3/2}} - \grave{\delta} \sum_{i=2}^{v} \frac{b_{i-1}^2}{(b_{i-1} + \grave{\beta})^2}$$

$$= \sum_{i=2}^{v} \frac{W_i b_{i-1}}{(b_{i-1} + \grave{\beta})^{3/2}} - \sum_{i=2}^{v} \frac{b_{i-1}^2}{(b_{i-1} + \grave{\beta})^2} \left\{ \frac{\sum_{i=2}^{v} W_i b_{i-1} / \sqrt{b_{i-1} + \grave{\beta}}}{\sum_{i=2}^{v} b_{i-1}^2 / (b_{i-1} + \grave{\beta})} \right\} = 0 \quad . \tag{41}$$

Equation (41) is solved iteratively for $\grave{\beta}$ and then solutions for $\grave{\alpha}$ and $\grave{\delta}$ are obtained from equations (39) and (40), respectively.

7. Discussion of estimators

The preceding algebra was developed without taking a careful look at the residual sums of squares and the resulting estimators. This was done to illustrate some difficulties encountered in non-linear estimation which are not immediately apparent using the usual procedure for obtaining least squares estimators. A limited empirical, intuitive, and theoretical investigation of the results in section III-4 was pursued, and the findings apply to a number of the remaining sections in much the same manner as for section III-4.

The first fact observed was that equation (30) in $\breve{\beta}$ was equal to zero when $\breve{\beta}=\infty$; it is near zero for $\breve{\beta}=512$ and approaches zero asymptotically as β approaches infinity. The second fact noted for a numerical example was that δ increased as $\breve{\beta}$ increased; by the nature of these two parameters $\breve{\delta}$ should stay constant or should decrease as $\breve{\beta}$ increases. The third item noted was that equation (24) with $\breve{\alpha}=\alpha$, $\breve{\delta}=\delta$, and $\breve{\beta}=\beta$ attained the lowest value for $\breve{\beta}=0$ for a particular set of experimental data. This means that the sum of squares of the residuals does not have a unique minimum for some values of Y_{ij} and b_{i-1}. This, however, could be overcome by an appropriate transformation of the Y_{ij} values and would vary with the range and values of the data obtained.

For β known, the ordinary least squares estimators for the intercept and the slope are estimators for α and δ. Since the $\sum_{h=1}^{\infty} \frac{b_{h-1}}{\sqrt{b_{h-1}+\beta}}$ = a constant = k_1, since the b_{h-1} are of the order of 2^{-h+1}, and since the h^{th} term of this series approaches zero, it appears that the estimators for α and δ are not even consistent. That this is so can be observed from the variance of $\breve{\delta}$ given β where the denominator is of the form $\sum_{i=1}^{v} V_i^2 - \left(\sum_{i=1}^{v} V_i\right)^2/v$ for $V_i = \sum_{h=1}^{i} \frac{b_{h-1}}{\sqrt{b_{h-1}+\beta}}$. The V_i are ordered and rapidly approach a constant, say C_1. This means that the $\sum_{i=1}^{v} V_i^2 - \left(\sum_{i=1}^{v} V_i\right)^2/v$ does not become larger as v increases, but is always less than $\sum_{i=1}^{N}(V_i - C_1)^2$, which does not increase in value for a specified number of significant figures after i = some number N.

Therefore, in order to have consistent estimators for α and δ given β, it would be necessary to replicate experiments for a fixed number of generations, i.e., increase the n_i at the expense of the number of generations v. In fact, the first few observations, say generation 0, 1, and 2, are much more important generations for estimating the parameters α and δ than the later generations; after i=N a specified number of additional generations are essentially of no value in estimating α and δ. The more efficient statistical procedure must, of course, be viewed in light of biological considerations. One of the more important biological considerations is to determine if the postulated model fits for an "adequate" number of generations ("adequate" is defined here to mean until the biologist becomes tired of conducting experiments). Thus, from a statistical point of view the most efficient sampling procedure would be to use replicated observations from two generations, 0 and 1, to fit the model postulated by equation (3) for a specified β. From certain biological points of view it would appear that 10 to 15 or more generations would suffice to observe the appropriateness of the models postulated herein; certain types of experiments may require additional generations, say 30 to 100 generations.

For models of the nature postulated by equations (1) to (3), careful thought must be given to the nature of the parameters being estimated in relation to the sampling plan and the observations. There appears to be a redundancy for some of the estimators obtained. For example, consider the following sum of squares:

$$\sum_{i=1}^{v} \sum_{j=1}^{n_i} \left[Y_{ij} - \alpha - \sigma_g^2 \sum_{h=1}^{i} \frac{b_{h-1}\bar{z}_{mh}}{\sqrt{b_{h-1}\sigma_g^2 + \sigma_e^2}} \right]^2$$

In the above there is a temptation to estimate α, σ_g^2, and σ_e^2 as suggested in section III-3. But, the above sum of squares divided by v-2 and with the parameters replaced by estimates of parameters is defined to be an estimator for σ_e^2. Since σ_e^2 is contained inside the summation, it appears that an estimate of σ_e^2 must be obtained in another manner; then, σ_e^2 is replaced by its estimate and estimators for σ_g^2 and α are obtained.

The genetic basis for models such as those given by equations (1) and (2) is given in various places (e.g., see Falconer [1960]; Searle [1961]; references at end of chapter 23 in Kempthorne [1957]; etc.). However, the estimation problem and the model testing problem appear to have received little discussion in published literature. Results from several long term selection experiments are available, but models for response due to selection follow equations (1) and (2) given the values of the parameters.

IV. A NUMERICAL EXAMPLE

As explained previously by Papa and Federer [1960] and Papa [1961], the selection program for each of several inter- and intra-strain crosses and their reciprocals was carried out at each of three different temperature levels (18°C., 25°C, and 35°C.); each cross at each temperature was replicated. Ten spores or individuals of each mating type, A and a, were grown in duplicate growth tubes for each cross at each temperature in each generation. Occasionally, fewer individuals were obtained due to accidents. The fastest growing A individual was crossed with the fastest growing a individual to obtain the population for the next generation. In addition, growth measurements from eight tubes were obtained for the two selected individuals in each generation.

Since a minimum of ten generations of selection from each replicate at each temperature level was available for the intra-strain cross of the laboratory stocks of Neurospora crassa (77a/74A), these data were selected to illustrate the procedure for comparing experimental data with a theoretical model. The selection summary data are presented in Table 1. The analyses of variance for the two replicates and 10 generations are presented in Table 2. The individual analyses of variance for the variation among 20 individuals and between duplicate growth tubes for each individual are presented in Table 3. In some cases not all ten spores were recovered for one or both of the mating types.

For the experimental conditions encountered the variation between duplicate growth tubes obtained at one time appears to be considerably different from duplicate growth tubes grown at different times. Therefore, the within mean square is defined to have the expectation σ_d^2, a component of variance due to duplicate determinations obtained at one time. Since the degrees of freedom are essentially equal, for all mean squares a simple average of the 20 within mean squares for 18°C. from Table 3, equal to .00114, is an estimate of σ_d^2. The estimates of σ_d^2 for 25°C. and 35°C. are .00262 and .00374, respectively. If the degrees of freedom vary, one could pool the within sums of squares and divide by the pooled within degrees of freedom, but this was not done here.

The expectation of the among individuals mean squares is $\sigma_d^2+2\sigma_g^2(2^{-i+1})$ $=\sigma_d^2+2^{-i+2}\sigma_g^2$, where 2 is the number of growth tubes for each individual and 2^{-i+1} is the coefficient from an inbreeding series for generation i. From the 20 analyses of variance for one temperature an estimator for σ_g^2 is obtained by minimizing the following sum of squares with respect to σ_g^2:

$$\sum_{f=1}^{2} \sum_{i=1}^{10} [(A_{fi}-W_{fi})(2^{i-2})-\sigma_g]^2 \quad ,$$

with the result

Table 1. Mean growths in mm./hr. for cross 77a/74A.

Replicate and generation	Temperature level 18°C. Mean of 20	18°C. Mean A	18°C. Mean a	25°C. Mean of 20	25°C. Mean A	25°C. Mean a	35°C. Mean of 20	35°C. Mean A	35°C. Mean a
I- 1	1.97	2.27	2.27	3.46	3.96	3.94	3.65	4.50	4.68
2	2.30	2.40	2.40	3.98	4.06	4.09	5.23	5.28	5.37
3	2.43	2.46	2.48	3.89	3.94	3.98	5.05	5.12	5.13
4	2.43	2.44	2.51	3.74	3.84	3.87	4.88	4.94	4.94
5	2.48	2.53	2.56	4.08	4.17	4.21	5.03	5.10	5.17
6	2.39	2.42	2.41	4.00	4.03	4.10	4.91	5.01	5.05
7	2.39	2.42	2.42	4.16	4.17	4.28	5.24	5.26	5.29
8	2.46	2.51	2.50	4.05	4.17	4.12	5.39	5.44	5.46
9	2.33	2.50	2.30	4.13	4.18	4.24	5.15	5.20	5.33
10	2.48	2.52	2.50	4.00	4.08	4.04	5.15	5.20	5.26
II- 1	2.35	2.39	2.42	4.01	4.07	4.11	2.92	4.83	4.97
2	2.40	2.48	2.42	3.77	3.95	3.92	5.03	5.12	5.16
3	2.30	2.39	2.39	3.98	4.05	4.04	4.77	4.80	4.88
4	2.39	2.43	2.41	4.07	4.11	4.14	5.25	5.31	5.32
5	2.46	2.50	2.51	4.13	4.17	4.26	5.02	5.14	5.09
6	2.42	2.46	2.47	4.24	4.34	4.31	5.27	5.33	5.36
7	2.55	2.60	2.59	4.19	4.24	4.24	5.29	5.43	5.34
8	2.52	2.57	2.56	4.09	4.30	4.20	5.17	5.22	5.24
9	2.54	2.57	2.62	4.12	4.15	4.17	4.93	5.08	5.01
10	2.45	2.50	2.53	4.35	4.42	4.42	4.75	4.78	4.84
I+II- 1	2.16	-	-	3.74	-	-	3.28	-	-
2	2.35	-	-	3.88	-	-	5.13	-	-
3	2.36	-	-	3.94	-	-	4.91	-	-
4	2.41	-	-	3.90	-	-	5.06	-	-
5	2.47	-	-	4.10	-	-	5.02	-	-
6	2.40	-	-	4.12	-	-	5.09	-	-
7	2.47	-	-	4.18	-	-	5.26	-	-
8	2.49	-	-	4.07	-	-	5.28	-	-
9	2.44	-	-	4.12	-	-	5.04	-	-
10	2.46	-	-	4.18	-	-	4.95	-	-

Table 2. Anlaysis of variance for means of 20 from Table 1.

Source of variation	d.f.	Mean squares 18°C.	25°C.	35°C.
Generations	9	.01883	.04477	.67550
Within generations	10	.01242	.03244	.05886

Table 3. Analyses of variance for data of cross 77a/74A.

		Replicate I*				Replicate II*			
		Among individuals		Within individuals		Among individuals		Within individuals	
Temperature	Generation	d.f.	m.s.	d.f.	m.s.	d.f.	m.s.	d.f.	m.s.
18°C.	1	18	.26528	19	.00347	19	.00163	20	.00079
	2	19	.01705	20	.00145	19	.00146	20	.00131
	3	19	.00221	20	.00075	19	.00291	20	.00077
	4	19	.00348	20	.00079	19	.00272	20	.00084
	5	19	.00320	20	.00070	19	.00231	20	.00103
	6	18	.00072	19	.00042	19	.00322	20	.00219
	7	19	.00070	20	.00031	19	.00365	20	.00202
	8	19	.00325	20	.00126	19	.00121	20	.00046
	9	19	.02488	20	.00051	19	.00379	20	.00141
	10	19	.00052	20	.00050	19	.00165	20	.00179
25°C.	1	18	.65422	18	.00489	19	.00844	20	.00173
	2	19	.01928	20	.00312	19	.00939	20	.00692
	3	19	.00405	19	.00334	19	.00331	20	.00166
	4	19	.01168	20	.00268	19	.00219	20	.00194
	5	19	.03965	20	.00245	19	.00703	20	.00284
	6	19	.00574	20	.00236	19	.00616	19	.00163
	7	19	.00634	20	.00214	19	.00222	20	.00291
	8	19	.01083	20	.00181	19	.01991	20	.00169
	9	19	.00385	19	.00246	18	.00098	19	.00189
	10	19	.00506	20	.00093	19	.00753	20	.00308
35°C.	1	12	3.34077	12	.00602	19	6.12452	20	.00833
	2	19	.01074	20	.00224	19	.00890	20	.00492
	3	19	.00998	20	.00380	19	.00675	20	.00165
	4	19	.00637	20	.00221	19	.00217	20	.00158
	5	19	.00445	20	.00349	19	.02333	20	.00891
	6	19	.01189	20	.00426	18	.00626	18	.00196
	7	19	.00220	20	.00086	19	.00401	20	.00522
	8	19	.00403	20	.00183	19	.00574	20	.00178
	9	19	.00983	20	.00458	19	.00640	20	.00312
	10	18	.00936	19	.00483	19	.00454	20	.00314

* d.f. = degrees of freedom; m.s. = mean square

$$\overset{*}{\sigma}_g{}^2 = \frac{1}{2(10)} \sum_{f=1}^{2} \sum_{i=1}^{10} (A_{fi}-W_{fi})(2^{i-2})$$

where A_{fi} and W_{fi} represent the among and within mean squares for the i^{th} generation in the f^{th} replicate and 2 = number of replicates and 10 = number of generations. The estimate of $\sigma_g{}^2$ for the 18°C. data is $\overset{*}{\sigma}_g{}^2 = .193$. For 25°C. and 35°C., respectively, the estimates of $\overset{*}{\sigma}_g{}^2$ are .248 and .405.

Another estimator giving more weight to the earlier generations would be

$$\overset{\times}{\sigma}_g{}^2 = \sum_{f=1}^{2} \sum_{i=1}^{10} (A_{fi}-W_{fi})2^{-i+2} \Big/ \sum_{f=1}^{2} \sum_{i=1}^{10} 2^{-2i+4} \quad .$$

The estimates of $\overset{*}{\sigma}_g{}^2$ and $\overset{\times}{\sigma}_g{}^2$ for each replication at 18°C., 25°C., and 35°C. are presented in the following table:

	18°C.			25°C.			35°C.		
	Rep I	Rep II	Rep I & II	Rep I	Rep II	Rep I & II	Rep I	Rep II	Rep I & II
$\overset{*}{\sigma}_g{}^2$	.345	.041	.193	.268	.228	.248	.385	.425	.405
$\overset{\times}{\sigma}_g{}^2$	.101	.001	.051	.248	.003	.126	1.253	2.295	1.774

From the results it is apparent that quite different estimates of $\sigma_g{}^2$ may be obtained from the two estimators. In addition, estimates of $\sigma_g{}^2$ using the same estimator varied considerably between the replications. On observation of the data, it becomes apparent that deviant results from only one generation are sufficient to considerably alter estimates of $\sigma_g{}^2$. For example, at 18°C. and 25°C., generation 1 in replicate I produced by far the greatest contribution to $\overset{\times}{\sigma}_g{}^2$, and generation 9 of replicate I at 18°C. produced the largest contribution to $\overset{*}{\sigma}_g{}^2$. Similarly, at 35°C. generation 1 of both replicates produced major contribution to $\overset{\times}{\sigma}_g{}^2$.

From regression theory we could compute the following estimated variances:

$$\hat{V}(\overset{*}{\sigma}_g{}^2) = \sum_{f=1}^{r} \sum_{i=1}^{v} [2^{i-2}(A_{fi}-W_{fi}) - \overset{*}{\sigma}_g{}^2]^2/rv(rv-1)$$

$$\hat{V}(\overset{\times}{\sigma}_g{}^2) = \sum_{f=1}^{r} \sum_{i=1}^{v} [A_{fi}-W_{fi}-2^{i+2}\,\overset{\times}{\sigma}_g{}^2]^2/(rv-1) \sum_{f=1}^{r} \sum_{i=1}^{v} 2^{-2i+4} \quad .$$

Applying these equations to the selection data, the following variances are obtained:

	18°C.	25°C.	35°C.
$V(\overset{*}{\hat{\sigma}}_g{}^2)$	.0240	.0079	.0290
$V(\hat{\sigma}_g{}^2)$	.0030	.00187	2.1550

The large difference between the variances of two estimates at the same temperature is to a large extent due to an extremely large estimate of $\sigma_g{}^2$ in one of 10 generations.

An analysis of variance on the means of the 20 spores for each generation is given in Table 2. The within generation mean square is an estimate of $\sigma_e{}^2+\sigma_d{}^2/40$ since each mean is obtained from 40 observations. Thus, an estimate of $\sigma_e{}^2$ is:

$$\overset{*}{\sigma}_e{}^2 = \text{within generation m.s.} - \overset{*}{\sigma}_d{}^2/40$$

$$= .01242 - .00114/40 = .01214 \quad ,$$

where .01242 is the within generation mean square for 18°C. Similarly, $\overset{*}{\sigma}_e{}^2$ equals .03237 and .05877 for 25°C. and 35°C., respectively. The estimated environmental variance increases with temperature as might be expected. Therefore, the variances over temperatures should not be pooled.

The variance of the mean of 20 individuals at a given temperature level in duplicate tubes is estimated by the within generation mean squares in Table 3. For the experimental data at 18°C. then $\breve{\beta}$ for equations (28) and (29) is computed as:

$$\widehat{\sigma_e{}^2 + \sigma_d{}^2/40} \Big/ \overset{*}{\sigma}_g{}^2 = .01242/.193 = .0644 \quad .$$

For 25°C. and 35°C., $\breve{\beta}$ = .1305 and .1451, respectively. These relatively large genetic variances were unexpected.

Since normality of environmental and genetic effects may be unrealistic and since the selection data were thought to follow equation (3), the parameters α, β, and δ were estimated and theoretical curves were fitted to the data for the three temperatures.

$$W_i = \sum_{h=1}^{i} \frac{2^{-h+1}}{\sqrt{2^{-h+1}+\breve{\beta}}} = \sum_{h=1}^{i} \frac{1}{\sqrt{2^{h-1}+\breve{\beta}2^{2h-2}}} \quad , \tag{42}$$

where the W_i are sums of values from 0, 1, 2, up to i for $\breve{\beta}$ equal to a specified value.

$$\breve{\delta} = \sum_{i=1}^{v} (Y_i.-2\bar{y})W_i/2\left\{\sum_{i=1}^{v} W_i^2-(\Sigma W_i)^2/v\right\} \quad , \tag{43}$$

and

$$\breve{\alpha} = \bar{y} - \breve{\delta} \sum_{i=1}^{v} W_i/v \quad . \tag{44}$$

The computed values for $\breve{\alpha}$ and $\breve{\delta}$ for each of the three temperatures are given in Tables 4, 5, and 6. The computed curves using equation (3) are given in Figures 1, 2, and 3 for 18°C., 25°C., and 35°C., respectively.

Table 4. Computations for $\breve{\beta}$=.0644 and for 18°C. data of Table 1.

Generation of selection = i	W_i	Total for Rep I & II = $Y_i.$	$Y_i.-2\bar{y}$	$\breve{\alpha}+\breve{\delta}W_i$	$\frac{Y_i.}{2}-(\breve{\alpha}+\breve{\delta}W_i)$
0	000000	4.32	-.484	2.199426	-.039426
1	.969274	4.70	-.104	2.295618	.054382
2	1.634818	4.73	-.074	2.361667	.003333
3	2.080678	4.82	.016	2.405914	.004086
4	2.367902	4.94	.136	2.434419	.035581
5	2.543350	4.81	.006	2.451830	-.046830
6	2.644393	4.94	.136	2.461858	.008142
7	2.699627	4.98	.176	2.467340	.022660
8	2.728699	4.87	.066	2.470225	-.035225
9	2.743645	4.93	.126	2.471708	-.006708
	20.412386	48.04	000		-.000005

$\bar{y}$=2.402

$2\bar{y}$=4.804

$$2\left[\sum W_i - \frac{(\Sigma W_i)^2}{10}\right] = 15.209136396$$

$$\sum_{i=1}^{10} (Y_i.-2\bar{y})W_i = 1.509369796$$

$$\breve{\delta} = \frac{1.509369796}{15.209136396} = .099241$$

$$\breve{\alpha} = 2.402 - .099241 \frac{(20.412386)}{10} = 2.402 - .202574 = 2.199426$$

Table 5. Computations for $\breve{\beta}$ = .1305 and for 25°C. data of Table 1.

Generation of selection = i	W_i	Total for Rep I & II = $Y_i.$	$Y_i.-2\bar{y}$	$\breve{\alpha}+\breve{\delta}W_i$	$\frac{Y_i.}{2}-(\breve{\alpha}+\breve{\delta}W_i)$
0	000000	7.47	-.574	3.708406	.026594
1	.940152	7.75	-.294	3.863412	.011588
2	1.570203	7.87	-.174	3.967290	-.032290
3	1.975490	7.81	-.234	4.034111	-.129111
4	2.222784	8.21	.166	4.074883	.030117
5	2.365050	8.24	.196	4.098339	.021661
6	2.442751	8.35	.306	4.111149	.063851
7	2.483626	8.14	.096	4.117889	-.047889
8	2.504633	8.25	.206	4.121352	.003648
9	2.515288	8.35	.306	4.123109	.051891
	19.020337	80.44	000		.000060
		$2\bar{y}$=8.044			
		$\bar{y}$=4.022			

$$2\left[\sum W_i^2 - \frac{(\Sigma W_i)^2}{10}\right] = 12.689720042$$

$$\sum_{i=1}^{10} W_i(Y_i.-2\bar{y}) = 2.092189702$$

$$\breve{\delta} = \frac{2.092189702}{12.689720042} = .164873$$

$$\breve{\alpha} = 4.022 - .164873\,\frac{(19.020337)}{10} = 4.022 - .313594 = 3.708406$$

Table 6. Computations for $\check{\beta}$ = .1451 and for 35°C. data of Table .

Generation of selection = i	W_i	Total for Rep I & II = $Y_{i\cdot}$	$Y_{i\cdot}-2\bar{y}$	$\check{\alpha}+\check{\delta}W_i$	$\frac{Y_{i\cdot}}{2}-(\check{\alpha}+\check{\delta}W_i)$
0	000000	6.57	3.238	3.842264	-.557264
1	.934498	10.26	.452	4.370913	.759087
2	1.557023	9.82	.012	4.723078	.186922
3	1.954751	10.13	.322	4.948074	.116926
4	2.195269	10.05	.242	5.084136	-.059136
5	2.332441	10.18	.372	5.161735	-.071735
6	2.406856	10.53	.722	5.203832	.061168
7	2.445830	10.56	.752	5.225880	.054120
8	2.465805	10.08	.272	5.237180	-.197180
9	2.475924	9.90	.092	5.242904	-.292904
	18.768397	98.08	.000		.000004

$2\bar{y}=9.808$

$\bar{y}=4.904$

$$2\left[\sum W_i^2 - \frac{(\Sigma W_i)^2}{10}\right] = 12.276614316$$

$$\sum W_i(Y_{i\cdot}-2\bar{y}) = 6.944928504$$

$$\check{\delta} = \frac{6.944928504}{12.276614316} = .565704$$

$$\check{\alpha} = 4.904 - .565704\,\frac{(18.768397)}{10} = 4.904 - 1.061736 = 3.842264$$

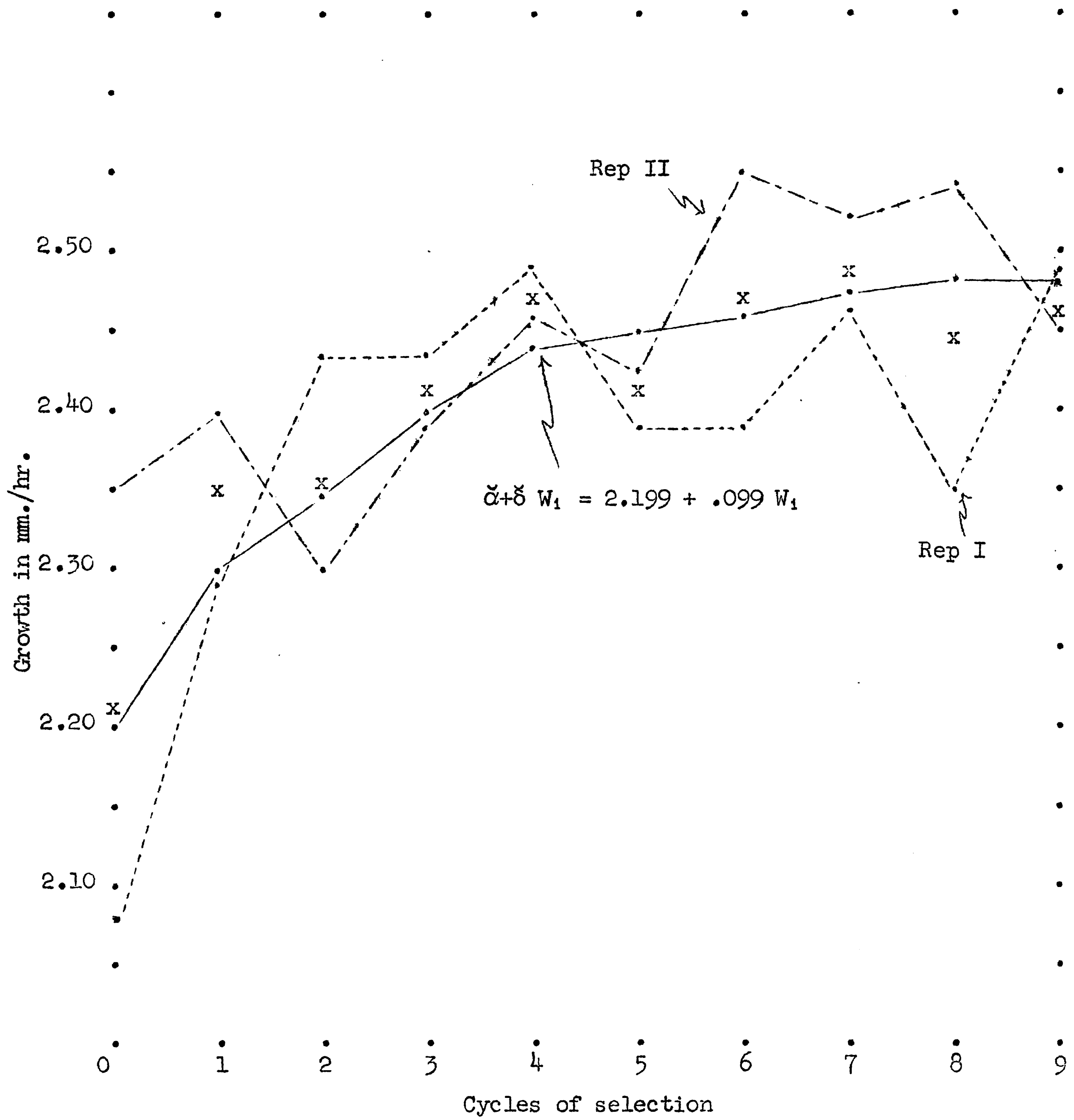

Figure 1. Observed results for each of two replicates and for the mean of two replicates (X's) for data of Table 1 for 18°C., for β=.0644, and for equation (3) as computed in Table 4.

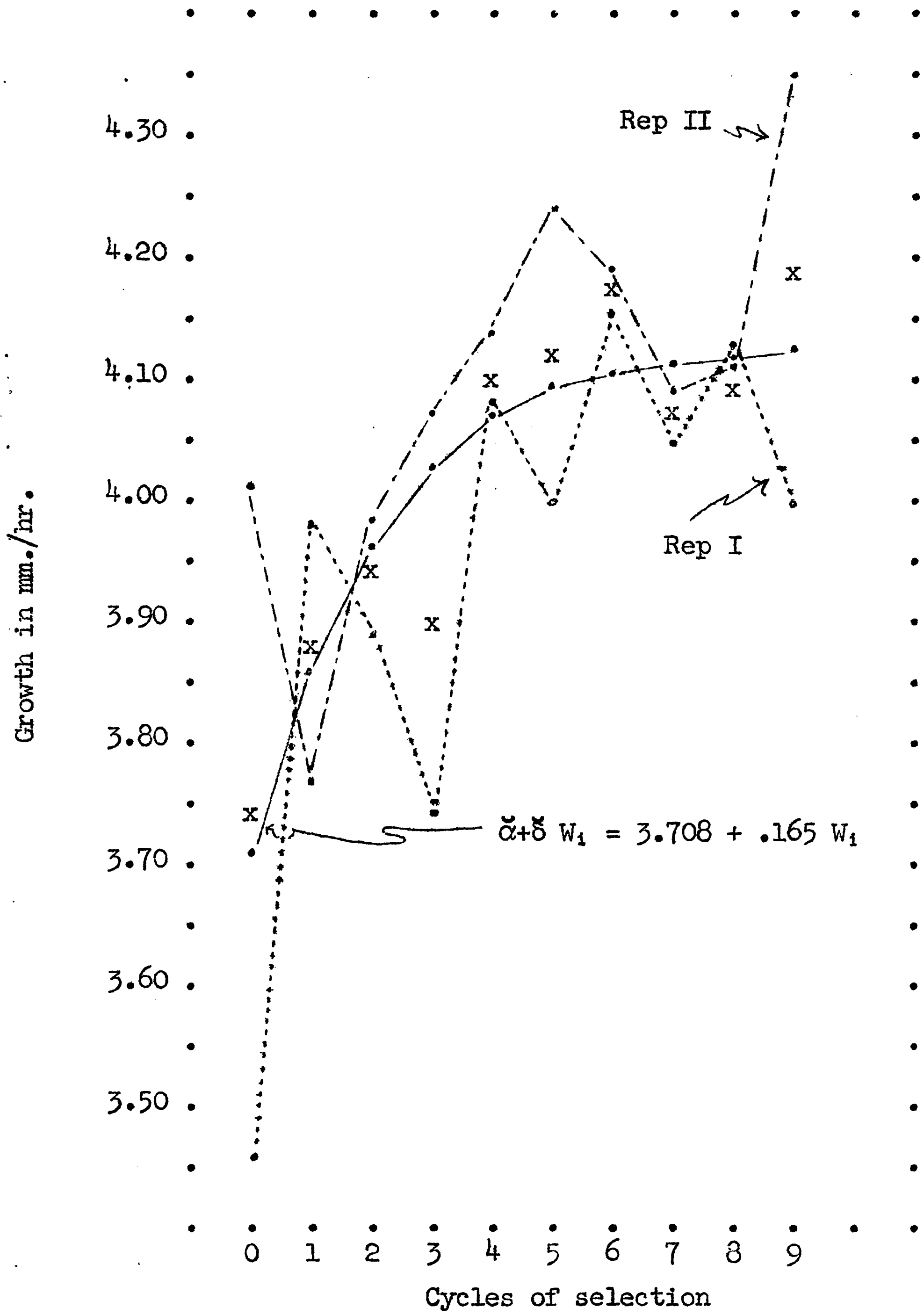

Figure 2. Observed results for each of two replicates and mean of two replicates (X's) for data of Table 1 for 25°C., for β=.1305, and for equation (3) as computed in Table 5.

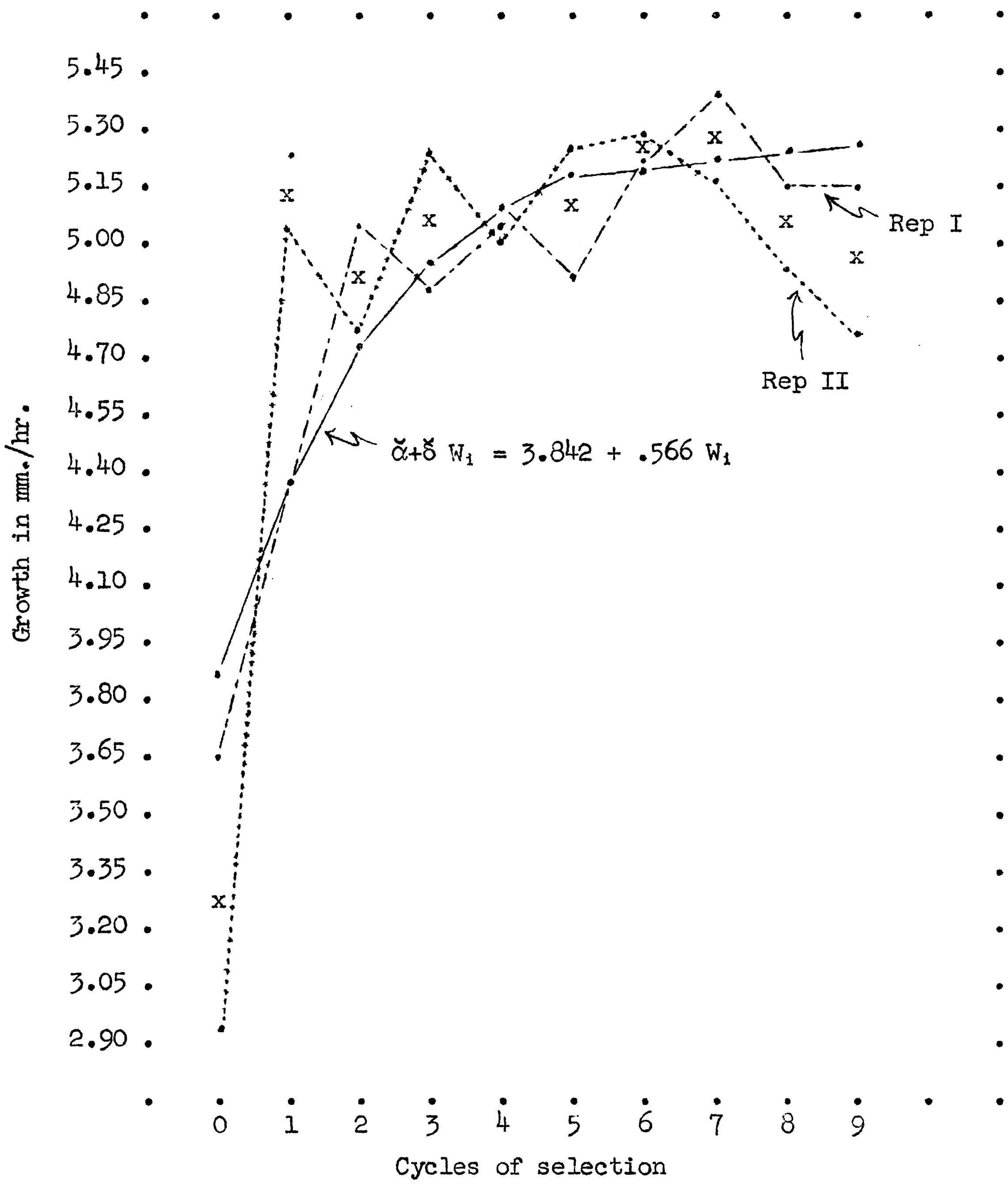

Figure 3. Observed results for each of two replicates and mean of two replicates (X's) for data of Table for 35°C., for β=.1451, and for equation (3) as computed in Table 6.

Although the results are variable fairly good fits were obtained for the data for 18°C. and for 25°C. The 35°C. data do not appear to fit well. It may be that temperature sensitivity is being encountered here and this may be controlled by a few genes. There appears to be little or no progress from selection after the first cycle of selection pressure.

An estimate of $\sigma_e^2+\sigma_d^2/2$ may be obtained from these data from equation (26):

$$\sum_{i=1}^{10} \sum_{j=1}^{2} [Y_{ij}-\breve{\alpha}-\breve{\delta}W_i]^2/(20-2)$$

with 20-2=18 degrees of freedom.

Also, if normality holds and if selection pressure is constant an estimate of σ_g could be obtained by dividing $\breve{\delta}$ by the expected value of the largest rank order statistic from a sample of size 10, i.e., the largest one out of 10 was selected. This results in

$$\frac{.099}{1.54} = .06 \quad \text{for } 18^\circ\text{C., or } \mathring{\sigma}_g^2 = .004$$

$$\frac{.165}{1.54} = .17 \quad \text{for } 25^\circ\text{C., or } \mathring{\sigma}_g^2 = .029$$

$$\frac{.566}{1.54} = .37 \quad \text{for } 35^\circ\text{C., or } \mathring{\sigma}_g^2 = .137 \quad .$$

The agreement with previous estimates of σ_g^2 is poor. However, definite conclusions must await the outcome of all the data yet to be analyzed.

V. LITERATURE CITED

Falconer, D. S. [1960] Introduction to quantitative genetics, Chapters 11, 12, and 13. Ronald Press Co., N. Y.

Federer, W. T. [1951] Evaluation of variance components from a group of experiments with multiple classifications. Iowa Agricultural Experiment Station Bulletin 380.

Federer, W. T. [1962] Estimators for components of variance. Biometrics Unit, Cornell University, Mimeograph Series BU-144-M.

Federer, W. T., Robson, D. S., and Srb, A. M. [1959] A model testing problem in genetics. Biometrics Unit, Cornell University, Mimeograph Series BU-103-M.

Fisher, R. A., and Yates, F. [1938] Statistical tables for biological, agricultural and medical research, 1st ed. Oliver and Boyd, Edinburgh.

Harter, H. L. [1961] Expected values of normal order statistics. Biometrika 48:151-165.

Kempthorne, O. [1957] An introduction to genetic statistics, Chapter 23. Wiley, N. Y.

Papa, K. [1961] Effectiveness of selection based on variability achieved through recombination uncomplicated by heterotic effects in Neurospora. Biometrics Unit, Cornell University, Mimeograph Series BU-138-M.

Papa, K. E., and Federer, W. T. [1960] Crossing plan for Neurospora experiment to determine effect of selection through recombination free of heterotic effects. Biometrics Unit, Cornell University, Mimeograph Series BU-128-M.

Searle, S. R. [1961] Estimating herd improvement from selection programs. Journal of Dairy Science 44:1103-1112.