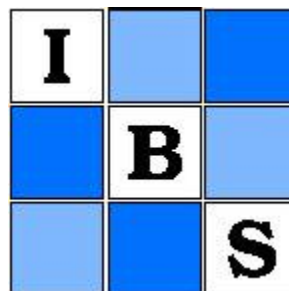


# WILEY



---

A Model for Certain Types of Selection Experiments

Author(s): B. F. J. Manly

Source: *Biometrics*, Vol. 30, No. 2 (Jun., 1974), pp. 281-294

Published by: International Biometric Society

Stable URL: <http://www.jstor.org/stable/2529649>

Accessed: 11-01-2016 18:19 UTC

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Wiley and International Biometric Society are collaborating with JSTOR to digitize, preserve and extend access to *Biometrics*.

<http://www.jstor.org>

## A MODEL FOR CERTAIN TYPES OF SELECTION EXPERIMENTS

B. F. J. MANLY

*Department of Mathematics, University of Otago, P. O. Box 56, Dunedin, New Zealand*

### SUMMARY

A model is suggested for certain types of experiments that involve individuals being selected one-by-one from a population consisting of two or more different classes of individuals. Selection with and without replacement is considered and equations are provided for the estimation of the selectivity of the selection process. The equations are illustrated on data from experiments involving selective predation and food competition.

### 1. INTRODUCTION

The equations given in this paper were first derived for the analysis of certain types of selective predation experiments. However, it seems that the equations are appropriate for analyzing data obtained in other ways as well, and they will, therefore, be developed in a general manner in terms of what will be called Type One and Type Two Selection Experiments. These experiments will now be defined.

Suppose that a population contains  $N$  individuals in  $K$  distinct classes with  $A_i$  individuals in the  $i$ th class ( $\sum A_i = N$ ). Assume that a selection process is observed in which  $n$  individuals are chosen one-by-one from the population in such a way that individuals in the same class have the same probability of being chosen, with the possibility that this probability might vary from class to class. If when an individual is selected it is immediately replaced in the population by a new individual of the same class, so that the  $n$  selected individuals are all chosen from a population with the same composition, then this will be referred to as a Type One Selection Experiment (T.O.S.E.). If selected individuals are not replaced, then it will be referred to as a Type Two Selection Experiment (T.T.S.E.).

A series of predation experiments reported by Popham ([1947] and earlier) were T.O.S.E.'s with the predators operating the selection processes. In these experiments the predators were presented with several specimens of each of two or more forms of prey and were allowed to choose which they would eat. As soon as a prey specimen was eaten, it was replaced by another specimen of the same form. Each experiment was continued until the predators had taken a total of about 100 or more prey specimens. Similar experiments have been conducted more recently (e.g. Allen and Clarke [1968], Mueller [1968; 1971], O'Donald and Pilecki [1970], Alcock [1971], Pilecki and O'Donald [1971]). Descriptions of predation experiments with eaten prey not being replaced, so that they were T.T.S.E.'s, have also been published (e.g. Landenberger [1968], Murdoch [1969], Morrel and Turner [1970]). Later in this paper T.T.S.E.'s involving competitive survival and selective mating will be discussed.

In an earlier paper (Manly *et al.* [1972]) a stochastic model was suggested for the T.T.S.E. with  $K = 2$  in the context of selective predation. Here this model has been generalized for both types of selection experiments with  $k \geq 2$ .

## 2. MEASURES OF SELECTIVITY

It will be assumed that if a selection process is operating on a population that at some stage contains  $x_j$  individuals in the  $j$ th class ( $j = 1, 2, \dots, K$ ) then

$$P_i = \beta_i x_i / \sum_{j=1}^K \beta_j x_j, \quad i = 1, 2, \dots, K, \quad (1)$$

is the probability that the next individual is chosen from the  $i$ th class.

It is convenient to standardize the  $\beta$ 's, which measure the selectivity of the selection process, so that their sum is unity. This is justifiable since it will be seen that if all the  $\beta$ 's are multiplied by the same factor then the  $P_i$ 's remain unchanged. With this form of standardization putting all the  $x_i$ 's equal in equation (1) gives  $P_i = \beta_i$ , and hence  $\beta_i$  has the meaningful interpretation of being equal to the probability that an individual is selected from the  $i$ th class when the selection process has a choice of an equal number of individuals in each of the  $K$  classes. Clearly if all the  $\beta$ 's are equal to  $1/K$ , then selection is completely random with respect to the classes.

A number of measures of selectivity have previously been suggested in connection with selective predation but most of them are unsatisfactory for one reason or another (e.g. see Ivlev [1961] p. 43). Equation (1) is a generalization of Murdoch's [1969] equation (2) and is rather similar to an equation proposed by Tinbergen ([1960] p. 326). With  $K = 2$ ,  $\beta_1$  corresponds to the  $\beta$  of Manly *et al.* [1972].

## 3. ESTIMATION WITH TYPE ONE SELECTION

In this case equation (1) is always

$$P_i = \beta_i A_i / \sum_{j=1}^K \beta_j A_j, \quad (2)$$

and hence if  $e_i$  denotes the total number of individuals selected from the  $i$ th class ( $\sum e_i = n$ ), then the joint distribution of all the  $e_i$ 's is multinomial. It is a simple matter to show that the maximum likelihood (ML) estimators of the  $\beta$ 's are given by

$$\hat{\beta}_i = (e_i/A_i) / \left( \sum_{j=1}^K e_j/A_j \right), \quad i = 1, 2, \dots, K, \quad (3)$$

where these estimators can also be obtained by equating the right hand side of equation (2) to  $\hat{P}_i = e_i/n$ .

Noting that  $\text{var}(e_i) = nP_i(1 - P_i)$  and  $\text{cov}(e_i, e_j) = -nP_iP_j$ ,  $i \neq j$ , it is possible to apply the Taylor series method for approximating the variances etc. of nonlinear estimators (e.g. see Kendall and Stuart [1963] p. 231), and it is found that

$$\text{bias}(\hat{\beta}_i) \simeq \beta_i \{E(x)/E(y) - 1/A_i\}/E(y), \quad (4)$$

$$\text{var}(\hat{\beta}_i) \simeq \beta_i^2 \left\{ \frac{1}{E(e_i)} - \frac{2}{A_i E(y)} + \frac{E(x)}{E(y)^2} \right\}, \quad (5)$$

and

$$\text{cov}(\beta_i, \hat{\beta}_i) \simeq \beta_i \beta_i \{E(x)/E(y) - 1/A_i - 1/A_i\}/E(y) \quad (6)$$

where

$$x = \sum_{i=1}^K e_i/A_i^2 \quad \text{and} \quad y = \sum_{i=1}^K e_i/A_i .$$

In these equations  $E(\ )$  indicates the expected value.

Formulae (4) to (6) can be expected to be accurate providing that all the  $E(e_i)$ 's are large. To test their use in practice, a computer simulation study was carried out, and a number of T.O.S.E.'s were repeated 100 times each. The simulation results indicate that:

- (a) Providing all the  $E(e_i)$ 's are larger than about five, the formulae are accurate and the biases of the  $\hat{\beta}$ 's are insignificant.
- (b) The joint distribution of the  $\hat{\beta}$ 's is well approximated by a multivariate normal distribution providing that none of the intervals  $\beta_i \pm 3 \{ \text{var} (\hat{\beta}_i) \}^{1/2}$  include zero or one. (A near normal distribution is of course expected for ML estimators from large samples.)
- (c) If the  $\beta$ 's in equations (4) to (6) are replaced by  $\hat{\beta}$ 's and the  $E(e_i)$  are replaced by observed values of  $e_i$ 's, then the equations will still give reasonably accurate results, providing that all of the observed  $e_i$ 's are larger than about seven.

For the remainder of this section it will be assumed that the conditions specified in (a) to (c) are satisfied: the vector of estimates  $\hat{\mathfrak{g}}' = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_{k-1})$  will be assumed to have a nonsingular multivariate normal distribution with mean  $\mathfrak{g}' = (\beta_1, \beta_2, \dots, \beta_{k-1})$  and with a variance-covariance matrix  $\Psi$  that is known accurately. (The estimate  $\hat{\beta}_k$  is omitted from the vector  $\hat{\mathfrak{g}}$  so that the distribution is nonsingular.) On this basis:

- (i) The scalar  $(\hat{\mathfrak{g}} - \mathfrak{g})'\Psi^{-1}(\hat{\mathfrak{g}} - \mathfrak{g})$  follows a  $\chi^2$  distribution with  $K - 1$  degrees of freedom (D.F.), and this fact can be used to test whether  $\mathfrak{g}$  is equal to a given vector or to construct a confidence region for  $\mathfrak{g}$  (e.g., see Anderson [1958] p. 54).
- (ii) The statistic

$$g = (\hat{\beta}_i - \beta_i)/\{ \text{var} (\hat{\beta}_i) \}^{1/2} \tag{7}$$

follows a standard normal distribution, and hence the hypothesis that  $\beta_i$  has a particular value can be tested or a confidence interval for  $\beta_i$  alone can be constructed.

- (iii) If  $\beta_i = \beta_j$  then

$$g = \{ \hat{\beta}_i - \hat{\beta}_j \} / \{ \text{var} (\hat{\beta}_i) - 2 \text{cov} (\hat{\beta}_i, \hat{\beta}_j) + \text{var} (\hat{\beta}_j) \}^{1/2} \tag{8}$$

has a standard normal distribution, and the hypotheses that two particular  $\beta$  values are equal can be tested.

- (iv) If the results for  $S$  independent experiments are available where  $\beta_i$  is the same for each experiment then

$$X^2 = \sum_{j=1}^S (\hat{\beta}_{ij} - \beta_i)^2 / \text{var} (\hat{\beta}_{ij}) \tag{9}$$

has a  $\chi^2$  distribution with  $S$  D.F. where  $\hat{\beta}_{ij}$  is the estimate of  $\beta_i$  from the  $j$ th experiment. The hypothesis that  $\beta_i$  has a particular value can therefore be tested. If the weighted mean

$$\bar{\beta}_i = \left\{ \sum_{j=1}^S \hat{\beta}_{ij} / \text{var} (\hat{\beta}_{ij}) \right\} / \left\{ \sum_{j=1}^S 1 / \text{var} (\hat{\beta}_{ij}) \right\} \tag{10}$$

is substituted for  $\beta_i$  in equation (9), then the resulting statistic still has a  $\chi^2$  distribution but with  $S - 1$  D.F. since one D.F. is lost in estimating  $\beta_i$ . The hypothesis that  $\beta_i$  was constant for the  $S$  experiments can therefore be tested. Note that the weighted means  $\bar{\beta}_i$  need not add to unity.

(v) A simple test for completely random selection ( $\beta_i = 1/K$ , for all  $i$ ) is the usual  $\chi^2$  goodness-of-fit test taking  $E(e_i)$  proportional to  $A_i$ .

Of course (i) to (v) do not cover all of the tests that can be applied to the  $\beta$  estimates from one or more experiments. In particular, various tests can be constructed by making use of the fact that a linear combination of the  $\hat{\beta}$  vectors from several experiments will follow a multivariate normal distribution. For example, the estimates from one experiment can be contrasted with the average estimates from several other experiments.

If it is suspected that the vector  $\beta$  was the same for several experiments, then it would be best to calculate the ML estimate of this vector on the basis of the combined results from all the experiments. The hypothesis that  $\beta$  was constant could then be tested using a standard goodness-of-fit test (e.g. see Kendall and Stuart [1961] p. 425). Unfortunately the ML equations have no explicit solution and would have to be solved numerically.

4. AN EXAMPLE OF TYPE ONE SELECTION

Table 1 gives the results of six predation experiments conducted by Popham [1943]. The predators were three rudd (*Scardinius erythrophthalmus*, L.) and the prey were three colors of *Sigara distincta* (Fieb.). In the experiments the rudd were presented with 10 or 12 specimens of the prey, and any specimen eaten was immediately replaced by another specimen of the same color. The experiments were stopped when about 200 prey were eaten. The proportions of the three colors of prey were different for the different experiments, and Popham was interested in seeing how the rudd reacted to changes in these proportions. The  $\chi^2$  values in the last column of Table 1 provide clear evidence that the rudd were not selecting colors completely at random.

Table 2 shows the  $\beta$  estimates calculated from Popham's data using equation (3) together with standard errors estimated using equation (5). Biasses are not given because they are all of the order  $\pm 0.001$  and are therefore insignificant. The high  $\chi^2$  values at the foot of the table provide clear evidence that the selectivity of the rudd varied from experiment to experiment.

TABLE 1  
RESULTS OF SIX PREDATION EXPERIMENTS WITH PREY REPLACEMENT CONDUCTED BY POPHAM [1943]

Experiment	Type 1 prey (light grey)		Type 2 prey (medium grey)		Type 3 prey (dark grey)		$\chi^2$ (2 d.f.)*
	Presented $A_1$	Eaten $e_1$	Presented $A_2$	Eaten $e_2$	Presented $A_3$	Eaten $e_3$	
1	4	105	4	67	4	28	44.47 (sig. 0.1% level)
2	3	56	6	132	3	12	39.84 (sig. 0.1% level)
3	6	144	3	40	3	16	44.48 (sig. 0.1% level)
4	3	55	3	25	6	108	13.74 (sig. 1% level)
5	2	46	2	23	6	131	9.13 (sig. 5% level)
6	2	29	4	62	6	105	1.08 (not sig.)

\* Calculated for test (V) of § 3. The six  $\chi^2$  values are independent.

TABLE 2  
ESTIMATES OF  $\beta$ 's OBTAINED FROM THE DATA OF TABLE 1 WITH ESTIMATED STANDARD ERRORS

Experiment	Type 1 prey		Type 2 prey		Type 3 prey	
	$\hat{\beta}_1$	std.err*	$\hat{\beta}_2$	std.err.	$\hat{\beta}_3$	std.err.
1	0.52	0.035	0.34	0.033	0.14	0.025
2	0.42	0.039	0.49	0.037	0.09	0.025
3	0.56	0.039	0.31	0.038	0.13	0.029
4	0.41	0.040	0.19	0.033	0.40	0.036
5	0.41	0.042	0.20	0.037	0.39	0.035
6	0.31	0.043	0.33	0.035	0.36	0.034
$\chi^2$ (5 d.f.)**	25.01		46.83		112.17	

\* Estimated by substituting  $\hat{\beta}_i$ 's for  $\beta_i$ 's and  $e_i$ 's for  $E(e_i)$ 's in equation (5).

\*\* Calculated from equation (9) with  $\bar{\beta}_i$  substituted for  $\beta_i$  as described in (iv) of § 3. All of the  $\chi^2$  values are significant at the 0.1% level. They are not independent.

There seems little point in carrying out any further tests on Popham's data, and the obvious next step in an analysis would be to attempt to fit a model that relates changes in the  $\beta$  estimates to changes in the  $A_i$ 's. It can be shown that a linear regression model provides an adequate fit, but I have considered this model in a separate paper (Manly [1973]) because the discussion involves the analysis of the results from a number of other predation experiments.

5. ESTIMATION WITH TYPE TWO SELECTION

Consider Table 3 which shows a T.T.S.E. when a total of  $j$  individuals have been selected. Note that  $g_i$  denotes the number of class  $i$  individuals remaining and that

$$\text{Max } (0, A_i - j) < g_i < \text{Min } (A_i, N - j).$$
 (11)

TABLE 3  
A TYPE TWO SELECTION EXPERIMENT WHEN A TOTAL OF  $j$  INDIVIDUALS HAVE BEEN SELECTED

Class of individuals	Number initially available	Number selected	Number remaining
1	$A_1$	$A_1 - g_1$	$g_1$
2	$A_2$	$A_2 - g_2$	$g_2$
.	.	.	.
K	$A_K$	$A_K - g_K$	$g_K$
Totals	N	$j$	$N - j$

Let  $P_i(g_1, g_2, \dots, g_K) = P_i(\mathbf{g})$  denote the probability that the situation shown in Table 3 will arise during an experiment and also let

$${}_iP_{i-1}(\mathbf{g}) = P_{i-1}(g_1, g_2, \dots, g_{i-1}, g_i + 1, g_{i+1}, \dots, g_K). \quad (12)$$

Thus  ${}_iP_{i-1}(\mathbf{g})$  is the probability that when  $j - 1$  individuals are selected there are  $g_i + 1$  of the class  $i$  individuals left and  $g_s$  ( $s \neq i$ ) of the class  $s$  individuals left. Given the situation implied by equation (12), the probability that the  $j$ th individual taken is from class  $i$  is

$$\beta_i(g_i + 1) / \left( \sum_{s=1}^K \beta_s g_s + \beta_i \right).$$

In order for the situation shown in Table 3 to arise it is necessary that when  $j - 1$  individuals were selected there were  $g_i + 1$  of the class  $i$  individuals left and  $g_s$  of the class  $s$  individuals left ( $s \neq i$ ) and the  $j$ th individual taken was from class  $i$ , for just one of the  $K$  possible values of  $i$ . On this basis the recurrence relations

$$P_j(\mathbf{g}) = \sum_{i=1}^K {}_iP_{j-1}(\mathbf{g}) \beta_i(g_i + 1) / \left( \sum_{s=1}^K \beta_s g_s + \beta_i \right), \quad j = 1, 2, \dots, N. \quad (13)$$

are obtained where these are a generalization of equations given earlier for the case when  $K = 2$  (Manly *et al.* [1972] equation 2a).

Equations (13) take a long time to solve on an electronic computer even with  $K = 2$ , and they provide no direct method for estimating the  $\beta$ 's. However in the Appendix to this paper they are shown to lead to the following approximate results that are useful for estimation purposes:

$$\mu_{kj} = \mu_{k,j-1} - \theta_{k,j-1} \quad (14)$$

where  $\mu_{kj}$  is the mean value of  $g_k$  when  $j$  individuals have been selected and

$$\begin{aligned} \theta_{kj} &= \beta_k \mu_{kj} / \mu_j^* \quad \text{where} \quad \mu_j^* = \sum_{i=1}^K \beta_i \mu_{ij}; \\ c_{kkj} &= (1 - 2\beta_k / \mu_{j-1}^*) c_{kk,j-1} + \theta_{k,j-1} \left( 2 \sum_{i=1}^K \beta_i c_{kij-1} / \mu_{j-1}^* + 1 - \theta_{k,j-1} \right); \end{aligned} \quad (15)$$

and

$$\begin{aligned} c_{klj} &= \{1 - (\beta_k + \beta_l) / \mu_{j-1}^*\} c_{kl,j-1} - \theta_{k,j-1} \theta_{l,j-1} \\ &\quad + \left( \theta_{l,j-1} \sum_{i=1}^K \beta_i c_{kij-1} + \theta_{k,j-1} \sum_{i=1}^K \beta_i c_{lij-1} \right) / \mu_{j-1}^* \end{aligned} \quad (16)$$

where  $c_{klj}$  is the covariance between  $g_k$  and  $g_l$  and  $c_{kkj}$  is the variance of  $g_k$  when a total of  $j$  prey have been selected.

It is further shown in the Appendix that by considering a differential equation that is approximately equivalent to equation (14) the relationship

$$\beta_i \simeq \log(\mu_{ij} / A_i) / \sum_{s=1}^K \log(\mu_{sj} / A_s) \quad (17)$$

can be established. This relationship, which depends upon  $N$  being large, has been found to be remarkably accurate even for values of  $N$  as small as 10. It suggests the estimator

$$\hat{\beta}_i = \log(r_i / A_i) / \sum_{s=1}^K \log(r_s / A_s) \quad (18)$$

to be applied when the selection experiment ends.



Making use of the means, variances, and covariances obtained from equations (14) to (16) the Taylor series method (Kendall and Stuart [1963] p. 231) can be applied to obtain approximations for the biases etc. of the estimators given by equation (18). Details of the algebra will not be given here. The results obtained are:

$$\text{bias } (\hat{\beta}_i) \simeq \beta_i \left\{ \sum_{s=1}^K \sum_{l=1}^K \frac{c_{ls}}{\mu_l \mu_s} + \frac{1}{2} \sum_{s=1}^K \ln \left( \frac{\mu_s}{A_s} \right) \sum_{l=1}^K \frac{c_{ll}}{\mu_l^2} - \frac{1}{\beta_i} \sum_{s=1}^K \frac{c_{is}}{\mu_i \mu_s} \right\} / \left\{ \sum_{s=1}^K \ln \left( \frac{\mu_s}{A_s} \right) \right\}^2 \\ - \frac{1}{2} c_{ii} / \left\{ \mu_i^2 \sum_{s=1}^K \ln \left( \frac{\mu_s}{A_s} \right) \right\}, \quad (19)$$

$$\text{var } (\hat{\beta}_i) \simeq \left\{ \beta_i^2 \sum_{s=1}^K \sum_{l=1}^K \frac{c_{ls}}{\mu_l \mu_s} - 2\beta_i \sum_{s=1}^K \frac{c_{is}}{\mu_i \mu_s} + \frac{c_{ii}}{\mu_i^2} \right\} / \left\{ \sum_{s=1}^K \ln \left( \frac{\mu_s}{A_s} \right) \right\}^2, \quad (20)$$

and

$$\text{cov } (\hat{\beta}_i, \hat{\beta}_j) \simeq \left\{ \beta_i \beta_j \sum_{s=1}^K \sum_{l=1}^K \frac{c_{ls}}{\mu_l \mu_s} - \beta_i \sum_{s=1}^K \frac{c_{js}}{\mu_j \mu_s} - \beta_j \sum_{s=1}^K \frac{c_{is}}{\mu_i \mu_s} + \frac{c_{ij}}{\mu_i \mu_j} \right\} / \left\{ \sum_{s=1}^K \ln \left( \frac{\mu_s}{A_s} \right) \right\}^2. \quad (21)$$

In these equations the subscript  $j$  has not been included on the  $c$ 's and  $\mu$ 's. It is to be understood that the value of  $j$  is the value applying at the end of the experiment, i.e.,  $j = n$ . Also  $\ln(x)$  stands for the natural logarithm of  $x$ .

Equations (14) to (21) provide a method for analyzing T.T.S.E.'s. The estimates of  $\beta$  values obtained from equation (18) can be substituted into equations (14) to (16) to evaluate the  $c_{ii}$  values needed for equations (19) to (21), and hence these latter equations can be used to obtain approximations to the biases etc. of the estimators. A listing of a FORTRAN computer program to carry out the calculations is available on request.

To test the use of equations (14) to (21) in practice a simulation study was carried out where a number of T.T.S.E.'s were repeated 100 times by an electronic computer. The simulation results indicate that:

- If all the  $E(e_i)$ 's and  $E(r_i)$ 's are larger than about five, then the formulae are accurate and the biases of  $\hat{\beta}$ 's are insignificant.
- The joint distribution of the  $\hat{\beta}$ 's seems to be well approximated by a multivariate normal distribution, providing that none of the intervals  $\beta_i \pm 3 \{ \text{var } (\hat{\beta}_i) \}^{1/2}$  includes zero or one.
- If biases, variances and covariances are estimated by substituting  $\hat{\beta}$ 's for the  $\beta$ 's in equations (14) to (16) and (19) to (21), then the estimates will usually be reasonably accurate providing that all the observed  $e_i$ 's and  $r_i$ 's are larger than about ten.

Providing that the conditions specified in (a) to (c) are met, the tests (i) to (v) of section 3 can be applied with confidence to the estimates from a T.T.S.E.

## 6. EXAMPLES OF TYPE TWO SELECTION

O'Donald [1971] has described an experiment where the object was to examine whether the ability of *Drosophila melanogaster* to survive food competition is the same for flies with different sternopleural bristle numbers. In the experiment approximately equal numbers of male and female flies were put into a population cage with only eight food tubes instead of the usual 25. After six days nearly 70 percent of the population had died. Most of the deaths were presumably due to lack of food because O'Donald found that in a population of *Drosophila melanogaster* kept in a similar cage but with ample food there were few deaths before three weeks.



The experimental results for males only are given in Table 4. A goodness-of-fit test taking the expected  $r_i$  values to be proportional to the  $A_i$  values gives  $\chi^2 = 17.56$  with seven D.F. and therefore provides evidence against the hypothesis that males with different bristle numbers survive equally well. O'Donald does not give the experimental results for females in his paper, but he says that a similar goodness-of-fit test for females gives a nonsignificant value.

Suppose that at some stage during O'Donald's experiment there were  $x_1$  flies with 14 bristles still alive,  $x_2$  flies with 15 bristles still alive, etc. Then it seems reasonable to suppose that the probability that the next male fly to die was one of the  $x_i$  was given by equation (1). The  $\beta$ 's can be thought of as measures of the "unfitnesses" of flies with different bristle numbers. If it is assumed that the  $\beta$ 's were constant throughout the experiment, then they can be estimated by treating the experiment as a T.T.S.E. The results of doing so are shown in Tables 5 and 6.

Any statistical tests based upon the  $\beta$  estimates are necessarily tentative because the low  $r_i$  values for flies with 14 and 15 bristles do not satisfy condition (c) of section 5. However if tests are carried out it is found that only  $\hat{\beta}_2$  and  $\hat{\beta}_3$  are significantly different from the "random selection" value of 0.125 (see Table 5) while paired comparison tests based upon equation (8) show that only the differences  $\hat{\beta}_2 - \hat{\beta}_3$ ,  $\hat{\beta}_2 - \hat{\beta}_4$ ,  $\hat{\beta}_2 - \hat{\beta}_5$ ,  $\hat{\beta}_2 - \hat{\beta}_6$ ,  $\hat{\beta}_2 - \hat{\beta}_7$ ,  $\hat{\beta}_3 - \hat{\beta}_5$ , and  $\hat{\beta}_3 - \hat{\beta}_8$  are significant at the 5 percent level. The appropriate conclusion would seem to be that flies with 15 bristles had a low ability to survive, flies with 16 bristles had a high ability to survive and flies with other bristle numbers survived about equally well.

In his analysis of the experimental results, O'Donald assumed that there is a quadratic relationship between the fitness of flies and their bristle numbers and estimated the relationship on the basis of the change in the mean and the variance of the bristle number distri-

TABLE 4

THE SURVIVAL OF MALE DROSOPHILA MELANOGASTER IN A FOOD COMPETITION EXPERIMENT REPORTED BY O'DONALD [1971]

Bristle number*	Initial population numbers $A_i$	Survivors $x_i$	Number dying $e_i$
14	25	7	18
15	51	8	43
16	98	45	53
17	104	36	68
18	106	32	74
19	67	22	45
20	54	17	37
21	24)	6)	18)
22	21) 45**	4) 10	17) 35

\* The results for 14 flies with various other bristle numbers have been ignored.

\*\* To maintain reasonably high  $x_i$  values the results for flies with 21 and 22 bristles have been grouped together for all calculations. The results for flies with 14 and 15 bristles have not been combined because of the rather different survival rates.

TABLE 5  
ESTIMATES OF  $\beta$ 's DERIVED FROM THE DATA OF TABLE 4 WITH ESTIMATED STANDARD ERRORS

$i$	Bristle number	$\hat{\beta}_i^*$	Std. err.**	$g^{***}$
1	14	0.128	0.029	0.10
2	15	0.186	0.029	2.10 (sig.5% level)
3	16	0.078	0.011	-4.27 (sig.0.1% level)
4	17	0.107	0.014	-1.29
5	18	0.121	0.015	-0.27
6	19	0.112	0.017	-0.76
7	20	0.116	0.019	-0.47
8	21 or 22	0.152	0.026	1.03

\* The biases in the  $\hat{\beta}$ 's are insignificant being of the order  $\pm 0.001$ .  
\*\* Obtained by substituting the  $\hat{\beta}$ 's for the  $\beta$ 's in equations (14), (15), (16) and (20).  
\*\*\* Calculated according to equation (7). A value significantly different from zero  
Indicates that the particular  $\hat{\beta}$  is significantly different from  $1/K = 0.125$ .

bution caused by the selection process. However, his analysis is not strictly valid because the assumption is made that flies survived independently, and this is clearly not true. The idea that there is a smooth relationship between  $\beta$  values and bristle numbers is intuitively appealing, but Table 5 provides little evidence in its favor, apart from the fact all the estimates higher than 0.125 are at the ends of the range of bristle numbers.

Other examples of T.T.S.E.'s are provided by "male choice" mating experiments where a group of males of one species are put in a cage with a number of females of each of two species. The experiment continues until a certain number of the females have been mated. Levene [1949] discusses the analysis of this type of experiment, and it is interesting to note

TABLE 6  
VARIANCE-COVARIANCE MATRIX FOR THE ESTIMATES OF TABLE 5. THE VALUES BELOW MUST BE MULTIPLIED BY  $10^{-5}$  TO GIVE THE VARIANCES AND COVARIANCES

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$	$\hat{\beta}_7$	$\hat{\beta}_8$
$\hat{\beta}_1$	83							
$\hat{\beta}_2$	-23	82						
$\hat{\beta}_3$	-5	-5	12					
$\hat{\beta}_4$	-8	-7	1	19				
$\hat{\beta}_5$	-9	-8	1	1	23			
$\hat{\beta}_6$	-10	-9	0	0	-1	29		
$\hat{\beta}_7$	-11	-11	-1	-1	-2	-3	36	
$\hat{\beta}_8$	-18	-20	-3	-5	-5	-6	-8	65

The values shown in this table were calculated by substituting  $\hat{\beta}$ 's for  $\beta$ 's in equations (14), (15), (16), (20) and (21).

that although his statistical model is different from that of the present paper his measure of sexual isolation ( $K_{12}$ ) is equal to  $\beta_1 - \beta_2$  in the notation used here.

## 7. DISCUSSION

If an experiment involving a selection process is being planned and the experimenter has a choice between a T.O.S.E. and a T.T.S.E., then statistical considerations suggest that he should choose the former. This is because the analysis of a T.O.S.E. is relatively simple and type two selection tends to minimize the selective effects on a population. It is, however, sometimes possible to avoid heavy calculations in the analysis of a T.T.S.E. by replicating the experiment a number of times sufficient to allow standard errors to be estimated accurately directly from the experimental results (e.g., see Manly *et al.* [1972]). Of course if a T.T.S.E. is stopped before many individuals have been selected, there will be little error in analyzing the experiment as if it were a T.O.S.E. In fact it is a simple matter to show that if all the ratios  $r_i/A_i$  are close to unity, then the estimator of equation (18) is approximately equal to the estimator of equation (3).

One problem with T.T.S.E.'s is knowing when to stop selection, and unfortunately there seems to be no general rule other than that implied by (c) of section 5. With O'Donald's experiment discussed in section 6 it turns out that the standard errors of  $\beta$  estimates were about as small as possible at the point when it was actually stopped.

An interesting problem that has not been discussed in this paper is posed by considering how to decide which of two selection processes is the "most" selective, taking into account the possibility that the two processes might involve different numbers of  $\beta$  values. Work is preceding in this direction.

## ACKNOWLEDGMENTS

I am grateful to Dr. L. M. Cook for bringing to my attention the statistical problems involved in T.T.S.E.'s and to Dr. P. O'Donald for commenting on my analysis of his data in section 6. Also two referees made a number of suggestions that led to many improvements in an earlier version of this paper. Most of the work reported here was carried out in the Department of Mathematics, University of Papua, New Guinea.

## UN MODÈLE POUR DES EXPÉRIENCES COMPORTANT CERTAINS TYPES DE SÉLECTION

### RESUME

Un modèle est suggéré pour certains types d'expériences faisant intervenir des individus qui ont été choisis un à un dans une population comportant deux ou plus de deux classes différentes d'individus. La sélection avec ou sans remplacement est considérée et des équations sont fournies pour l'estimation de la sélectivité du processus de sélection. Les équations sont illustrées sur des données provenant d'expériences faisant intervenir une prédation sélective et une compétition pour la nourriture.

## REFERENCES

- Alcock, J. [1971]. Interspecific differences in the avian feeding behavior and the evolution of Batesian mimicry. *Behavior* 40, 1-9.
- Allen, J. A. and Clarke, B. [1968]. Evidence for apostatic selection on the part of wild Passerines. *Nature* 220, 501-2.

- Anderson, T. W. [1958]. *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.
- Ivlev, V. S. [1961]. *Experimental Ecology of the Feeding of Fishes*. Yale University Press, New Haven.
- Kendall, M. G., and Stuart, A. [1961]. *The Advanced Theory of Statistics*, Vol. 2. Griffin, London.
- Kendall, M. G. and Stuart, A. [1963]. *The Advanced Theory of Statistics*, Vol. 1. Griffin, London.
- Landenberger, D. E. [1968]. Studies on selective feeding in the Pacific starfish *Pisaster* in Southern California. *Ecology* 49, 1062–75.
- Levene, H. [1949]. A new measure of sexual isolation. *Evolution* 3, 315–21.
- Manly, B. F. J., Miller, P., and Cook, L. M. [1972]. Analysis of a selective predation experiment. *Amer. Nat.* 106, 719–36.
- Manly, B. F. J. [1973]. A linear model for frequency-dependent selection by predators. *Res. Popul. Ecol.* 14, 137–50.
- Morrell, G. M. and Turner, J. R. G. [1970]. Experiments on mimicry: I. The response of wild birds to artificial prey. *Behavior* 36, 116–30.
- Mueller, H. C. [1968]. Prey selection: oddity or conspicuousness? *Nature* 217, 92.
- Mueller, H. C. [1971]. Oddity and specific searching image more important than conspicuousness in prey selection. *Nature* 233, 345–6.
- Murdoch, W. W. [1969]. Switching in general predators: experiments on predator specificity and stability of prey populations. *Ecol. Monog.* 39, 335–54.
- O'Donald, P. and Pilecki, C. [1970]. Polymorphic mimicry and natural selection. *Evolution* 24, 395–401.
- O'Donald, P. [1971]. Natural selection for quantitative characters. *Heredity* 27, 137–53.
- Pilecki, C. and O'Donald, P. [1971]. The effects of predation on artificial mimetic polymorphisms with perfect and imperfect mimics at varying frequencies. *Evolution* 25, 365–70.
- Popham, E. J. [1943]. Further experimental studies on the selective action of predators. *Proc. Zool. Soc. Lond.* 112, 105–17.
- Popham, E. J. [1947]. Experimental studies of the biological significance of non-cryptic pigmentation with special reference to insects. *Proc. Zool. Soc. Lond.* 117, 768–83.
- Tinbergen, L. [1960]. The natural control of insects in pinewoods. I. Factors influencing the intensity of predation by songbirds. *Arch. Neer. Zool.* 13, 265–343.

## APPENDIX

The derivation is given here for some of the results on T.T.S.E.'s that were stated in section 5. In addition to the notation already defined, it is useful to have

$$g^* = \sum_{s=1}^K \beta_s g_s$$

and to let  $\mathbf{w}_j$  denote the set of vectors  $\mathbf{g}$  that are possible when a total of  $j$  individuals have been selected.

Also it is useful to note at this point that if  $f(\mathbf{g}) = f(g_1, g_2, \dots, g_K)$  is any function then

$$\sum_{\mathbf{w}_i \cup \mathbf{w}_{j-1}} {}_iP_{i-1}(\mathbf{g})f(g_1, g_2, \dots, g_{i-1}, g_i + 1, g_{i+1}, \dots, g_K) = \sum_{\mathbf{w}_{i-1}} P_{i-1}(\mathbf{g})f(\mathbf{g}), \quad (1a)$$

since  $P_{i-1}(\mathbf{g})$  is zero for any  $\mathbf{g}$  not in  $\mathbf{w}_{i-1}$ .

*Derivation of Equation (14)*

Multiplying both sides of equation (13) by  $g_k$  and adding over  $\mathbf{g}$  vectors in the set  $\mathbf{w}_i \cup \mathbf{w}_{i-1}$  gives

$$\begin{aligned} \sum_{\mathbf{w}_i \cup \mathbf{w}_{i-1}} g_k P_i(\mathbf{g}) &= \sum_{\mathbf{w}_i \cup \mathbf{w}_{i-1}} \left[ \sum_{\substack{i=1 \\ i \neq k}}^K g_k {}_iP_{i-1}(\mathbf{g}) \frac{\beta_i(g_i + 1)}{g^* + \beta_i} \right. \\ &\quad \left. + \{(g_k + 1) - 1\} {}_kP_{i-1}(\mathbf{g}) \frac{\beta_k(g_k + 1)}{g^* + \beta_k} \right]. \end{aligned}$$

Applying result (1a) to the terms on the right hand side and noting that  $P_i(\mathbf{g})$  is zero for any  $\mathbf{g}$  not in  $\mathbf{w}_i$  then leads to

$$\begin{aligned}\sum_{\mathbf{w}_i} g_k P_i(\mathbf{g}) &= \sum_{\mathbf{w}_{i-1}} \left\{ g_k P_{i-1}(\mathbf{g}) \sum_{i=1}^K \frac{\beta_k g_k}{g^*} - P_{i-1}(\mathbf{g}) \frac{\beta_k g_k}{g^*} \right\} \\ &= \sum_{\mathbf{w}_{i-1}} \left\{ g_k P_{i-1}(\mathbf{g}) - P_{i-1}(\mathbf{g}) \frac{\beta_k g_k}{g^*} \right\}\end{aligned}$$

since  $\sum \beta_k g_k$  is equal to  $g^*$ .

But the term on the left hand side of the last equation is  $\mu_{ki}$  while the first term on the right hand side is  $\mu_{ki-1}$ . Hence

$$\mu_{ki} = \mu_{ki-1} - \sum_{\mathbf{w}_{i-1}} P_{i-1}(\mathbf{g}) \frac{\beta_k g_k}{g^*}. \quad (2a)$$

If  $\beta_k g_k / g^*$  is now expanded in a Taylor series about the mean values of the  $g$ 's when  $j-1$  individuals have been selected, then this gives

$$\begin{aligned}\frac{\beta_k g_k}{g^*} &= \frac{\beta_k \mu_{ki-1}}{\mu_{i-1}^*} \left\{ \frac{g_k}{\mu_{ki-1}} - \sum_{s=1}^K \frac{\beta_s (g_s - \mu_{sj-1})}{\mu_{i-1}^*} \right\} \\ &+ \text{higher order terms,}\end{aligned} \quad (3a)$$

where the higher order terms are of order  $(\mu_{i-1}^*)^{-2}$  or smaller. Ignoring the higher order terms and substituting the right hand side of equation (3a) for  $\beta_k g_k / g^*$  in equation (2a) leads to equation (14).

#### Derivation of Equation (15)

Multiplying both sides of equation (13) by  $g_k^2$  and adding over the  $\mathbf{g}$  vectors in the set  $\mathbf{w}_i \cup \mathbf{w}_{i-1}$  gives

$$\begin{aligned}\sum_{\mathbf{w}_i \cup \mathbf{w}_{i-1}} g_k^2 P_i(\mathbf{g}) &= \sum_{\mathbf{w}_i \cup \mathbf{w}_{i-1}} \left[ \sum_{\substack{i=1 \\ i \neq k}}^K g_k^2 P_{i-1}(\mathbf{g}) \frac{\beta_i (g_i + 1)}{g^* + \beta_i} \right. \\ &\quad \left. + \{(g_k + 1)^2 - 2(g_k + 1) + 1\} P_{i-1}(\mathbf{g}) \frac{\beta_k (g_k + 1)}{g^* + \beta_k} \right]\end{aligned}$$

and following the principles used in developing equation (2a) leads to

$$\sum_{\mathbf{w}_i} g_k^2 P_i(\mathbf{g}) = \sum_{\mathbf{w}_{i-1}} \left\{ g_k^2 P_{i-1}(\mathbf{g}) + (1 - 2g_k) P_{i-1}(\mathbf{g}) \frac{\beta_k g_k}{g^*} \right\}.$$

Thus

$$s_{ki} = s_{ki-1} + \sum_{\mathbf{w}_{i-1}} (1 - 2g_k) P_{i-1}(\mathbf{g}) \frac{\beta_k g_k}{g^*}$$

where  $s_{ki}$  is the expected value of  $g_k^2$  when a total of  $j$  individuals have been selected.

If the Taylor series expansion for  $\beta_k g_k / g^*$  (equation 3a) is now substituted into the right hand side of the last equation with the higher order terms being ignored, then this leads to

$$\begin{aligned}s_{ki} &= s_{ki-1} + 2(\theta_{ki-1} / \mu_{i-1}^*) \sum_{i=1}^K \beta_i \sum_{\mathbf{w}_{i-1}} g_k (g_i - \mu_{i-1}) P_{i-1}(\mathbf{g}) - 2(\beta_k / \mu_{i-1}^*) \sum_{\mathbf{w}_{i-1}} g_k^2 + \theta_{ki-1} \\ &= (1 - 2\beta_k / \mu_{i-1}^*) s_{ki-1} + 2(\theta_{ki-1} / \mu_{i-1}^*) \sum_{i=1}^K \beta_i c_{ikj-1} + \theta_{ki-1}.\end{aligned}$$

Equation (15) follows from the last equation by writing

$$c_{kkj} = s_{kj} - \mu_{kj}^2 \quad \text{and} \quad c_{kkj-1} = s_{kj-1} - \mu_{kj-1}^2.$$

*Derivation of Equation (16)*

The derivation of this equation is similar to that for equation (15). Thus from equation (13)

$$\begin{aligned} \sum_{\mathbf{w}_i \cup \mathbf{w}_{i-1}} g_k g_l P_i(\mathbf{g}) &= \sum_{\mathbf{w}_i \cup \mathbf{w}_{i-1}} \left[ \sum_{\substack{i=1 \\ i \neq k \\ i \neq l}}^K g_k g_l P_{i-1}(\mathbf{g}) \frac{\beta_i(g_i + 1)}{g^* + \beta_i} \right. \\ &\quad + \{g_1(g_k + 1) - g_1\}_k P_{i-1}(\mathbf{g}) \frac{\beta_k(g_k + 1)}{g^* + \beta_k} \\ &\quad \left. + \{g_k(g_l + 1) - g_k\}_l P_{i-1}(\mathbf{g}) \frac{\beta_l(g_l + 1)}{g^* + \beta_l} \right] \end{aligned}$$

so that

$$\sum_{\mathbf{w}_i} g_k g_l P_i(\mathbf{g}) = \sum_{\mathbf{w}_{i-1}} \left\{ g_k g_l P_{i-1}(\mathbf{g}) - g_l P_{i-1}(\mathbf{g}) \frac{\beta_k g_k}{g^*} - g_k P_{i-1}(\mathbf{g}) \frac{\beta_l g_l}{g^*} \right\}$$

If the Taylor series expansions for  $\beta_k g_k / g^*$  and  $\beta_l g_l / g^*$  (equation 3a) are substituted into the right hand side of this last equation, with higher order terms being ignored, and the resulting equation is simplified, then equation (16) is obtained.

*Derivation of Equation (17)*

Let  $j/N = t$ ,  $\mu_{kj}/N = U_k(t)$ ,  $\delta t = 1/N$  and  $\delta U_k(t) = U_k(t + \delta t) - U_k(t)$ . Then equation (14) can be written as

$$\frac{\delta U_k(t - \delta t)}{\delta t} = -\beta_k U_k(t - \delta t) / \sum_{s=1}^K \beta_s U_s(t - \delta t)$$

and taking the limit as  $\delta t \rightarrow 0$  ( $N \rightarrow \infty$ ) gives the differential equation

$$\frac{dU_k}{dt} = -\beta_k U_k / \sum_{s=1}^K \beta_s U_s.$$

Hence if  $i$  and  $l$  are two different values of  $k$  then

$$\frac{dU_i}{dU_l} = \frac{\beta_i U_i}{\beta_l U_l}$$

which has the solution

$$\beta_l \log(U_i) = \beta_i \log(U_l) + C$$

where  $C$  is a constant. Since  $U_i = A_i/N$  and  $U_l = A_l/N$  when  $t = 0$ ,  $C$  can be evaluated and equation (4a) becomes

$$\beta_l \log(\mu_{ii}/A_i) = \beta_i \log(\mu_{li}/A_l).$$

Hence

$$\sum_{i=1}^K \beta_i \log(\mu_{ii}/A_i) = \sum_{l=1}^K \beta_l \log(\mu_{li}/A_l)$$

so that

$$\log (\mu_{i i} / A_i) = \beta_i \sum_{l=1}^K \log (\mu_{l i} / A_l)$$

and equation (17) is established.

*Received July 1972, Revised March 1973*

*Key Words:* Selection experiments; Natural selection.