Part III. Empirical Research Evidence Regarding Decision Making

STOCHASTIC LEARNING IN RATS WITH HYPOTHALAMIC IMPLANTS

Merrill M. Flood
Mental Health Research Institute, The University of Michigan, Ann Arbor, Mich.

Introduction

Several investigators have conducted learning experiments with humans and with animals designed to test the validity of various stochastic* models for learning. Estes[1] pioneered in this field. Independently, but at about the same time, very similar theoretical approaches were being developed by Bush and Mosteller,[2] by Miller and McGill,[4] and by Bales *et al.*[5]

The outstanding work of Bush and Mosteller, especially on the important problems of statistical estimation with applications to experimental data of various workers, was reported fully in their book on *Stochastic Models for Learning*.[6] Recent work by Burke, Bush, Estes, Mosteller, and others is reported in a volume entitled *Studies in Mathematical Learning Theory*, edited by Bush and Estes.[7] Especially noteworthy, more recently, is the theoretical approach of Luce,[8] as reported in his book, *Individual Choice Behavior*, and in the work of Suppes[9] and others at Stanford University, Stanford, Calif.

The present paper reports on intermittent work, started in 1951, with one particular stochastic learning model that I have called the "symmetrical model." The characteristics of this mathematical model that are essential for the present paper are summarized in the next section. The symmetrical model has been compared with some other stochastic models in an earlier paper of mine[3] in which a comparison was made particularly with the combining-of-classes model of Bush and Mosteller. Critical aspects of experimental and theoretical procedure and meanings are discussed there also. Suffice it to say that, for present purposes, the symmetrical model seems worth checking experimentally with respect to its range of validity. I have called this range of validity the "scope" of the theory.

Although the work is discussed here in the context of individual learning in rats, my interest in it has been and is primarily with reference to organization theory. A series of earlier papers[10] deal with the use of stochastic models of this kind as models for decision-making groups of interacting individuals. No attempt is made in the present paper to indicate further the use of such stochastic models in organization science, so that it is best read as a straightforward paper on individual learning.

The motivation for using rats, rather than human subjects, is threefold:

(1) Close control, and use of extremes of reward and punishment is possible with rats.

* Estes[1] referred to his model as "statistical" or "probabilistic," and Bush and Mosteller[2] to theirs as "statistical" or "mathematical." I have called[3] the approach "stochastic learning theory," because the mathematical models used are ordinary stochastic processes; "stochastic learning" is now a term quite commonly used.

(2) Implanted electrodes in the brain of the rat can be used to provide stimuli that are either rewarding or punishing, and readings from electrodes can provide useful information regarding the rat's behavior.

(3) Experimental costs and difficulties are less with rats than with human subjects when they are utilized for extensive observation over a considerable period of time.

The experiments reported here were all conducted with rats having chronic implants in the hypothalamic region, so that the rat was rewarded by electrical stimulation at the inner tip of the electrode. The rat made his choices by pressing pedals arranged rather symmetrically within a cage, and stimulation followed immediately upon pressing a pedal if the experimenter's pattern of rewards so required. Details of the apparatus and experimental techniques will be presented in a separate paper; a summary is included in *Appendix A*.

The prime objective of the present experimentation is to determine whether or not a procedure can be devised for measuring the six behavioral parameters of the symmetrical stochastic learning model in a reliable manner for an individual rat. A later objective, granted success with the present objective, is the determination of rough limits on the scope of the theory. For example:

(1) Do the measured behavioral parameters remain reasonably constant for an individual rat with changes in such experimental variables as number of pedals available, strength of electrical stimulation, location of implant, type of reward (such as food versus implant), alteration by surgery or drugs, and age?

(2) Are the measured behavioral parameters adequate to pass predictively from free pace to forced pace conditions, from free pace to latency, and so on?

(3) Do the behavioral parameters vary from rat to rat, and are the parameters related to other psychophysical measures?

Concurrently with the rat experiments, there is parallel experimentation with human subjects utilizing monetary rewards rather than chronic implants; these experiments will be reported elsewhere.

The present paper reports on the analytical techniques used to estimate behavioral parameters, and discusses some preliminary experimental results obtained with two rats. It is concluded by my colleagues and me that the behavioral parameters can be estimated reliably. No attempt has yet been made to determine the range of validity of these estimates, or the scope of the theory, as the experimental conditions are changed.

Symmetrical Model

The symmetrical stochastic learning model is described by the following mathematical relations. We start with a more general stochastic model,

$$p(t+1) = M^{r_t s_t} p(t), \tag{1}$$

where $p(t)$ is the m-dimensional stochastic vector whose r^{th} component is a probability $p_r(t)$, and where M^{rs} is a square stochastic matrix of order m whose elements M_{ij}^{rs} depend only upon r and s. It is supposed that the subject makes a sequence of responses among a fixed finite set of alternatives, and that there is a probability $p_r(t)$, at the end of moment t, that response r will occur before the end of moment $(t+1)$. The notation is such that r_t denotes the response

actually made after moment t, and s_t denotes the outcome (stimulus) following response r_t.

The symmetrical model is the special case of **1**, in which the following conditions are satisfied:*

$$M_{ij}^{rs} = a^s\delta_{ij}\delta_{ri} + c^s\delta_{ij}(1 - \delta_{ri}) + \frac{(1 - a^s)}{m - 1}(1 - \delta_{ri})\delta_{rj} + b^s(1 - \delta_{rj})\delta_{ri} + \mu(1 - b^s - c^s)(1 - \delta_{ij})(1 - \delta_{ri})(1 - \delta_{rj}) \tag{2}$$

for $r = 1, 2, \cdots, m$; $s = 1, 2, \cdots, n$,

where

$$\mu = \begin{cases} 1/(m - 2) & \text{if } m > 2 \\ 0 & \text{if } m = 2, \end{cases}$$

and where

$$\begin{aligned} 0 &\leq a^s \leq 1, \\ 0 &\leq b^s \leq 1, \\ 0 &\leq c^s \leq 1, \\ 0 &\leq b^s + c^s \leq 1. \end{aligned} \tag{3}$$

The quantities a^s, b^s, and c^s are the three behavioral parameters corresponding to outcome s; thus, there are $3n$ behavioral parameters in all, since there are n outcomes possible.

It is easily seen that **1** and **2** are equivalent to the following conditions:

$$\begin{cases} p_{r_t}(t + 1) = A^s p_{r_t}(t) + B^s, \\ p_r(t + 1) = C^s p_{r_t}(t) + D^s p_r(t) + E^s \quad \text{for } r \neq r_t, \end{cases} \tag{4}$$

where

$$\begin{aligned} A^s &\equiv a^s - b^s, \\ B^s &\equiv b^s, \\ C^s &\equiv \left(\frac{1 - a^s}{m - 1} - \frac{1 - b^s - c^s}{m - 2}\right), \\ D^s &\equiv \left(c^s - \frac{1 - b^s - c^s}{m - 2}\right), \\ E^s &\equiv \left(\frac{1 - b^s - c^s}{m - 2}\right). \end{aligned} \tag{5}$$

In other words, the probability of a response r_t being repeated is a constant B^s, plus its previous probability multiplied by a constant A^s; and the probability of a new response r is a constant E^s, plus its previous probability multiplied by a constant D^s, plus the previous probability of the previous response r_t multiplied by a constant C^s.

* δ_{xy} is the familiar Kronecker delta, such that: $\delta_{xy} = 1$ if $x = y$, and $\delta_{xy} = 0$ if $x \neq y$.

It has been shown[3] that the symmetrical learning model satisfies the combining-of-classes condition of Bush and Mosteller if and only if the behavioral parameters satisfy the following additional conditions:

$$C^s = 0, \text{ for } s = 1, 2, \cdots, n. \tag{6}$$

In this special case, the equations **4** reduce to the following:

$$p_{r_t}(t+1) = (a^s - b^s)p_{r_t}(t) + b^s,$$
$$p_r(t+1) = (a^s - b^s)p_r(t) + (1 - a^s)/(m-1), \text{ for } r \neq r_t.$$

However, there are other stochastic learning models of the form **1** that are not symmetrical but do nevertheless satisfy the combining-of-classes condition;[3] these models are not considered in the present paper.

Parameter Estimation

Experimental data always consist of observations on r_t and s_t, for $t = 1, 2, \cdots, N$. The possible values of r_t are $r_t = 1, 2, \cdots, m$; the possible values of s_t are $s_t = 1, 2, \cdots, n$; and N denotes the number of successive responses in one experimental trial. Of course, the trials may be repeated, with N different in each case.

In addition to the behavioral parameters a^s, b^s, and c^s, there are also the m parameters $p(1)$, which we shall refer to as the "state parameters." The state parameters describe the subject at the beginning of any experimental trial. It is equally proper to speak of the "state of the subject at moment t" as the stochastic vector $p(t)$; however, for convenience, we shall use state parameters to indicate the starting state for a particular experimental trial.

Maximum-likelihood estimates would be adequate, for our purposes, but are often difficult to obtain. For data r_t and s_t from a single experimental trial, these maximum-likelihood estimates would be defined by the following relations. First, we define quantities $\hat{p}(t)$ recursively as follows:

$$\hat{p}(t+1) \equiv \hat{M}^{r_t s_t}\hat{p}(t) \quad \text{for} \quad t = 1, 2, \cdots, N, \tag{7}$$

where $\hat{M}_{ij}^{rs}$ and $\hat{p}(1)$ are defined to be the values of the parameters M_{ij}^{rs} and $p_k(1)$ that maximize the "likelihood"

$$L[M_{ij}^{rs}, p_k(1)] \equiv \prod_{t=1}^{N} \hat{p}_{s_t}(t), \tag{8}$$

where the estimated parameters $\hat{M}_{ij}^{rs}$ and $\hat{p}(1)$ are required also to satisfy the restrictions of the symmetrical learning model. A numerical example is given in the next section to illustrate this estimation procedure.

If several trials are available for analysis, perhaps with different starting states and lengths, then the likelihood estimates would be the values of M_{ij}^{rs} and $p_k^h(1)$ that maximize the likelihood

$$L[M_{ij}^{rs}, p_k^h(1)] = \prod_{l=1}^{q} \prod_{t=1}^{N_l} \hat{p}_{r_t l}^{l}(t), \quad h = 1, 2, \cdots, q; \quad k = 1, 2, \cdots, m; \tag{9}$$

where q is the number of separate trials and N_l is the length of trial l. In this case, of course, the quantities $\hat{p}_{r_t l}^{l}(t)$ are defined as follows, analogous to the definition in **7**:

$$\hat{p}^l(t+1) \equiv \hat{M}^{r_t l s_t l}\hat{p}^l(t) \quad \text{for} \quad (t = 1, 2, \cdots, N_l; \quad l = 1, 2, \cdots, q). \tag{10}$$

In many instances, the experimenter will determine the sequences of outcomes of s_t^l in advance of the trials. If these sequences are chosen suitably, the calculations required to obtain usable estimates of the parameters can be reduced greatly. Furthermore, the state parameters may sometimes be determined by the nature of the experiment so that they need not be estimated from the experimental data. We now consider the estimation problem for one design of experiment that takes advantage of such simplification. This method is illustrated by a numerical example in the next section, but is not used extensively in the present paper.

If the experiment is designed to require repeated trials, each with the same sequence of outcomes and each with the starting state parameters all equal, then the parameters may be estimated rather easily. Specifically, for convenience, we start by considering the following design:

$$\begin{aligned} &N_1 = N_2 = \cdots = N_q = 3, \\ &\hat{p}_1{}^l(1) = \hat{p}_2{}^l(1) = \cdots = p_m{}^l = (1/m), \quad \text{for} \quad (l = 1, 2, \cdots, q), \\ &s_t{}^l = 1, \quad \text{for} \quad (l = 1, 2, \cdots, q; \qquad t = 1, 2, 3). \end{aligned} \tag{11}$$

Furthermore, for convenience, we start by considering the case in which

$$m = 3, \qquad n = 2. \tag{12}$$

Thus, the subject chooses one from among three possible choices for a sequence of three responses in each of q trials, and the outcome is always the same from among two possible outcomes (for example, outcome No. 1). As a further matter of notation, since we shall presently be concerned only with the parameters a^1, b^1, c^1, from the matrices M^{r1}, we shall temporarily drop the superscripts corresponding to s and have simply:

$$M^1 \equiv \begin{Vmatrix} a & b & b \\ \frac{1-a}{2} & c & 1-b-c \\ \frac{1-a}{2} & 1-b-c & c \end{Vmatrix}$$

$$M^2 \equiv \begin{Vmatrix} c & \frac{1-a}{2} & 1-b-c \\ b & a & b \\ 1-b-c & \frac{1-a}{2} & c \end{Vmatrix} \tag{13}$$

$$M^3 \equiv \begin{Vmatrix} c & 1-b-c & \frac{1-a}{2} \\ 1-b-c & c & \frac{1-a}{2} \\ b & b & a \end{Vmatrix}.$$

There are actually $3^3 = 27$ possible sequences of responses that the subject can make in one trial with three responses, each from among three possible alternatives represented by all the possible arrangements of the integers 1, 2, and 3. However, because of the symmetry of the stochastic model in use, it is easily seen that there are only five types of sequences that require separate consideration in the analysis; this is so because the first response in a trial may be arbitrarily labeled 1, the first subsequent response that is different (if any) may be arbitrarily labeled 2, and the remaining distinct response (if any) may be labeled 3. Hence, the five distinctive sequences of responses are:

111 (includes 222 and 333),
112 (includes 113, 221, 223, 331, and 332),
121 (includes 131, 212, 232, 313, and 323),
122 (includes 133, 211, 233, 311, and 322),
123 (includes 132, 213, 231, 312, and 321).

The experimental data may therefore be completely described by five frequencies, as follows:

F_{111} = the number of times the first response is twice repeated,
F_{112} = the number of times the first response is once repeated,
F_{121} = the number of times the second response differs from the first response when the third response is the same as the first response, and
F_{121} and F_{123} are defined similarly.

Of course,

$$F_{111} + F_{112} + F_{121} + F_{122} + F_{123} = q.$$

With this notation, by substitution in **9**, we can obtain the likelihood function for this particular design of experiment when it yields observations F_{ijk}. We shall now derive such an expression for $L[a, b, c]$, after some preliminary calculations. Consider first the case F_{111}. The theoretical probability of p_{111} of observing response patterns 111, 222, or 333, is as follows:

$$p_{111} = P_{111} + P_{222} + P_{333} = 3P_{111},$$

where P_{ijk} denotes the theoretical probability of observing the single sequence (i, j, k). Also, since $p_k{}^l(1) = (\frac{1}{3})$, for $(l = 1, 2, \cdots, q; k = 1, 2, 3)$, we shall temporarily omit the superscript on $p^l(1)$.

We adopt the customary notation that a prime on a matrix denotes the transposed matrix. Since $p(l)$ and $p(1)$ are matrices with one column, or equivalently column vectors, then $p'(l)$ and $p'(1)$ denote their transposition; thus, $p'(1) = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ in our present discussion. It will also be convenient to let e_h denote the unit column vector with m rows, having unity in its h^{th} row and zeros elsewhere; and to let J denote the column vector with m rows and all entries unity. Thus, also, $p(1) = J/3$ and $p'(1) = J'(3)$.

The theoretical probability of response i at moment t is therefore $e_i'p(t)$. Consequently,

$$P_{ijk} = [e_i'p(1)][e_j'M^ip(1)][e_k'M^jM^ip(1)] = [p_i(1)][p_j(2)][p_k(3)]. \quad \textbf{(14)}$$

We now proceed to calculate P_{111}, P_{112}, P_{121}, P_{122}, P_{123}, that correspond to the observational data F_{ijk}. It follows easily that:

$$M^1p(1) = \tfrac{1}{3}\begin{Vmatrix} a + 2b \\ (3 - a - 2b)/2 \\ (3 - a - 2b)/2 \end{Vmatrix} = \begin{Vmatrix} p_1(2) \\ p_2(2) \\ p_3(2) \end{Vmatrix} \equiv \begin{Vmatrix} w \\ (1 - w)/2 \\ (1 - w)/2 \end{Vmatrix}, \tag{15}$$

where $p_1(2) \equiv w$ denotes the theoretical probability of response 1 at moment 2. Similarly:

$$\begin{aligned} M^1M^1p(1) &= \tfrac{1}{6}\begin{Vmatrix} 2a^2 - 4b^2 + 2ab + 6b \\ -a^2 + 2b^2 - ab - 3b + 3 \\ -a^2 + 2b^2 - ab - 3b + 3 \end{Vmatrix} \\ &= \begin{Vmatrix} 3w^2 - 3bw + b \\ (-3w^2 + 3bw - b + 1)/2 \\ (-3w^2 + 3bw - b + 1)/2 \end{Vmatrix} = \begin{Vmatrix} p_1(3) \\ p_2(3) \\ p_3(3) \end{Vmatrix}, \end{aligned} \tag{16}$$

$$\begin{aligned} M^2M^1p(1) &= \tfrac{1}{12}\begin{Vmatrix} a^2 + 4b^2 + 4ab + 6ac + 12bc - 6a - 12b - 6c + 9 \\ -2a^2 + 4b^2 - 2ab + 6a + 6b \\ a^2 - 8b^2 - 2ab - 6ac - 12bc + 6b + 6c + 3 \end{Vmatrix} \\ &= \tfrac{1}{4}\begin{Vmatrix} 1 + 3w^2 + 2(1 - 3w)(1 - c) \\ 2 - 6w^2 - 2(1 - 3w)(1 + b) \\ 1 + 3w^2 + 2(1 - 3w)(b + c) \end{Vmatrix} = \begin{Vmatrix} p_1(3) \\ p_2(3) \\ p_3(3) \end{Vmatrix}. \end{aligned} \tag{17}$$

A straightforward calculation shows that:

$$p_{111} = 3P_{111} = 3[\tfrac{1}{3}][w][3w^2 - 3bw + b] = 3w^3 - 3bw^2 + bw. \tag{18}$$

Similarly:

$$\begin{aligned} p_{111} &= w[3w^2 + (1 - 3w)b], \\ p_{112} &= w[(1 - 3w^2) - (1 - 3w)b], \\ p_{121} &= \left(\frac{1 - w}{2}\right)\left[\frac{1 + 3w^2}{2} + (1 - 3w)(1 - c)\right], \\ p_{122} &= \left(\frac{1 - w}{2}\right)[1 - 3w^2 - (1 - 3w)(1 + b)], \\ p_{123} &= \left(\frac{1 - w}{2}\right)\left[\frac{1 + 3w^2}{2} + (1 - 3w)(b + c)\right]. \end{aligned} \tag{19}$$

Here, for example, we have used the fact that $p_{121} = 6P_{121}$ since the six equally likely actual sequences (121, 131, 212, 232, 313, 323) are all included under p_{121}.

The logarithm of the likelihood function **9** is, in this special case:

$$\begin{aligned} ln(L[a, b, c]) &= q\left[f_{111} ln\left(\frac{p_{111}}{3}\right) + f_{112} ln\left(\frac{p_{112}}{6}\right) + f_{121} ln\left(\frac{p_{121}}{6}\right) + f_{122} ln\left(\frac{p_{122}}{6}\right)\right. \\ &\quad \left. + f_{123} ln\left(\frac{p_{123}}{6}\right)\right] \\ &= q[f_{111} ln\, p_{111} + f_{112} ln\, p_{112} + f_{121} ln\, p_{121} + f_{122} p_{122} + f_{123} p_{123} \\ &\quad - ln 6 + f_{111} ln 2], \end{aligned}$$

where $f_{ijk} \equiv (F_{ijk})/q$ are the relative frequencies. The parameter estimates that we seek, therefore, are the quantities $\hat{a}$, $\hat{b}$, $\hat{c}$ that maximize $ln\, L$ but are subject to the restrictions of **3**.

Useful parameter estimates might be found by solving the following system of equations and inequalities, if such a solution exists:

$$\begin{aligned} &p_{111} = f_{111}, \quad p_{112} = f_{112}, \quad p_{121} = f_{121}, \quad p_{122} = f_{122}, \quad p_{123} = f_{123}, \\ &0 \le a \le 1, \qquad 0 \le b \le 1, \qquad 0 \le c \le 1, \qquad 0 \le b + c \le 1. \end{aligned} \tag{20}$$

Of course, the system of relations **20** should have a solution if the type of stochastic model used is the one yielding the observations and if the number of observations is great enough. When the system **20** does not have a solution, because of the inequalities, it can be used to obtain initial values to be used in an iterative procedure for determining actual likelihood estimates.

It is quite easy to find the explicit solution of **20**, if it exists, as follows:

$$\begin{aligned} \hat{w} &= f_{111} + f_{112}, \\ \hat{c} &= 1 + \frac{2}{(1 - \hat{w})(1 - 3\hat{w})}\left[\frac{(1 + 3\hat{w}^2)(1 - \hat{w})}{4} - f_{121}\right], \\ \hat{b} &= -1 + \frac{2}{(1 - \hat{w})(1 - 3\hat{w})}\left[\frac{(1 - 3\hat{w}^2)(1 - \hat{w})}{2} - f_{122}\right], \\ \hat{a} &= 3\hat{w} - 2\hat{b}. \end{aligned} \tag{21}$$

After $\hat{a}$, $\hat{b}$, and $\hat{c}$ are computed according to **21**, it is only necessary to check that the values found satisfy the other requirements of **20** in order to ensure that they constitute the desired solution; this means checking the inequalities and the equations for f_{111} and f_{123}.

Exactly the same formulas apply in any case when a, b, c are abbreviations for a^s, b^s, c^s, whatever the value of s, provided only that s remains constant throughout the set of trials analyzed. Accordingly, independent estimates of a^s, b^s, and c^s may be obtained by selecting from among all the trials just those for which s has the desired value for the first two responses.

Formulas exactly analogous to **21** can also be found, in a straight-forward manner, when there are more than three responses in the replicated trials.

For example, in the case of four responses per trial, the observational data would be the following set of relative frequencies:

$$f_{1111}, \quad f_{1112}, \quad f_{1121}, \quad f_{1122}, \quad f_{1123}, \quad f_{1211}, \quad f_{1212}, \quad f_{1213}, \quad f_{1221}, \quad f_{1222},$$
$$f_{1223}, \quad f_{1231}, \quad f_{1232}, \quad f_{1233}.$$

The 14 theoretical probabilities p_{ijkl}, that correspond to these 14 observed relative frequencies, are each a function of the parameters a, b, c. Some subset of the 14 equations $p_{ijkl} = f_{ijkl}$ could be used to determine estimates of a, b, and c; again, it would be necessary to check that these estimates satisfy the requirements of **20**. Alternatively, the likelihood function could be used to obtain estimates in this case. It is already quite clear that the algebraic and computational complexities mount very rapidly as the number of responses is increased beyond 3 or 4 per trial.

Formulas analogous to **21** can also be found, in a straightforward but tedious manner, if the value of s_t is not constant over the set of responses but the sequential pattern of s_t is the same for each trial. For example, if the values of s_t are (1, 2, 1), then the case of four responses per trial would yield 14 equations between p_{ijkl} and f_{ijkl} from which the six parameters could be estimated. On the other hand, if the values of s_t were (1, 1, 2), then the first three responses would yield estimates of the parameters corresponding to $s = 1$, and the four-response data could be used to estimate the remaining three parameters corresponding to $s = 2$. As always, maximum-likelihood estimates represent a superior alternative procedure when the calculations are feasible.

A fundamentally different method of estimation, whether or not maximum-likelihood techniques are used, would be to make estimates disregarding the inequalities in **20** and then "round off" as necessary to satisfy them. For example, if an estimate of -1.5 were obtained for some parameter, it would be arbitrarily rounded up to 0. Although this method has been used by others,[6] we consider it unsatisfactory and only use it occasionally as the first step in an iterative calculation of likelihood estimates.

We shall discuss other theoretical aspects of the parameter estimation problem as the need arises in treating particular cases. This section is intended only to lay the theoretical foundation; the following section will give some numerical examples of estimation.

Illustrative Estimation Examples

Our first example is intended to illustrate the likelihood estimation method for a simple case. Assume that the experimental data are:

$$r_1 = 2, \quad r_2 = 1, \quad r_3 = 2,$$
$$s_1 = 1, \quad s_2 = 1, \quad s_3 = 2.$$

Assume further, that $p_1(1) = p_2(2) = p_3(3) = \frac{1}{3}$. In this case, from Equation **19**, we have the following likelihood function:

$$p_{121} = \left(\frac{1 - w}{2}\right)\left[\frac{1 + 3w^2}{2} + (1 - 3w)(1 - c)\right].$$

To find the largest value of p_{121} we consider:

$$0 = \frac{\partial p_{121}}{\partial w} = \frac{-9}{4}(w - 1)^2 + c(2 - 3w),$$

$$0 = \frac{\partial p_{121}}{\partial c} = \frac{1}{2}(w - 1)(1 - 3w).$$

Consequently, relative extrema exist when either:

$$(w = 1, c = 0) \quad \text{or} \quad (w = \tfrac{1}{3}, c = 1).$$

The latter values yield the larger likelihood, namely:

$$p_{121}(w = \tfrac{1}{3}, c = 1) = \tfrac{2}{9}.$$

The restrictions of **20** further require:

$$b + c \leq 1, \quad b \geq 0, \quad \text{so} \quad b = 0.$$

Finally, since $a = 3w - 2b$, we have as our tentative maximum-likelihood estimates:

$$\hat{a} = 1, \quad \hat{b} = 0, \quad \hat{c} = 1.$$

However, as we shall now see, these values are not the final maximum-likelihood estimates; the correct values are determined by checking the likelihood function values on all the boundaries.

The boundaries of our parameter region, as required by **20**, are:

$$a = 0, \quad b = 0, \quad c = 0;$$

$$a = 1, \quad b = 0, \quad c = 0;$$

$$a = 0, \quad b + c = 1;$$

$$a = 1, \quad b + c = 1.$$

We consider these four cases in turn. In the first two cases, since $w = (\frac{1}{3})(a + 2b)$, we have

$$p_{121} = \tfrac{3}{4}, \tfrac{2}{9}.$$

In the third case, using $1 - c = b = (3w/2)$, we have:

$$p_{121} = (\tfrac{1}{4})(6w^3 - 9w^2 + 2w + 1).$$

This takes on the value $\frac{1}{4}$ for $w = 0$ and $w = \frac{2}{3}$, the boundaries on w when $a = 0$, and the interior maximum occurs when:

$$0 = \frac{d(p_{121})}{dw} = (1/2)(9w^2 - 9w + 1) \quad \text{or} \quad w = \frac{3 - \sqrt{5}}{6}.$$

It is easily verified that $p_{121} < \frac{3}{4}$ when $w = (3 - \sqrt{5})/6$. A similar analysis of the fourth case also yields a value of p_{121} less than $\frac{3}{4}$. Thus, the maximum-likelihood estimates are $\hat{a} = \hat{b} = \hat{c} = 0$. Exactly similar calculations, for the

outcome patterns ($s_1 = s_2$) and three responses in a single trial, yield the following maximum-likelihood estimates in the five cases.

$$p_{111} = 1 \quad \text{for} \quad \hat{a} = 1, \quad \hat{b} = 1, \quad \hat{c} = 0.$$
$$p_{112} = 4/9 \quad \text{for} \quad \hat{a} = 0, \quad \hat{b} = 1, \quad \hat{c} = 0.$$
$$p_{121} = 3/4 \quad \text{for} \quad \hat{a} = 0, \quad \hat{b} = 0, \quad \hat{c} = 0.$$
$$p_{122} = \sqrt{3}/6 \quad \text{for} \quad \hat{a} = 0, \quad \hat{b} = \sqrt{3}/2, \quad \hat{c} = 0.$$
$$p_{123} = 3/4 \quad \text{for} \quad \hat{a} = 0, \quad \hat{b} = 0, \quad \hat{c} = 1.$$

It is apparent, from these sample calculations, that explicit solutions for maximum-likelihood estimates could be obtained in a straightforward but very tedious manner for trials with more than three responses.

We shall now illustrate the use of our estimation methods by treating two synthetic examples, one of which consists of many repeated trials of three responses each, and the other of which consists of one trial with 75 responses. The examples are "synthetic" because the data for them were obtained by choosing arbitrary values for the behavioral and state parameters, fixing also upon an arbitrary choice of outcome sequences, and then calculating responses that would result by using a table of random numbers. Since this technique of using synthetic data is employed frequently in our work, we shall digress for a moment to show how this is done.

Consider an example of a synthetic trial having four successive responses, the first two of which have the same outcome and the third a different outcome. Our preliminary data are:

$$m = 3; \quad n = 2; \quad s_1 = s_2 = 1, \quad s_3 = 2;$$
$$p_1(1) = p_2(1) = p_3(1) = 1/3; \quad a^1 = 0.11, \quad b^1 = 0.01, \tag{22}$$
$$c^1 = 0.40, \quad a^2 = 0.19, \quad b^2 = 0.01, \quad c^2 = 0.50; \quad N = 4.$$

We next choose four from a table of random probabilities, say:

$$R_1 = 0.862, \quad R_2 = 0.283, \quad R_3 = 0.622, \quad R_4 = 0.461.$$

The first response is, therefore, $r_1 = 3$ since $p_1(1) + p_2(1) < R_1$. The probability vector, after the first response, is now:

$$p(2) = M^{31}p(1) = \begin{vmatrix} 0.40 & 0.59 & 0.445 \\ 0.59 & 0.40 & 0.445 \\ 0.01 & 0.01 & 0.11 \end{vmatrix} \begin{vmatrix} 0.33 \\ 0.33 \\ 0.34 \end{vmatrix} = \begin{vmatrix} 0.478 \\ 0.478 \\ 0.044 \end{vmatrix}.$$

Consequently, $r_2 = 1$ since $R_2 \leq p_1(2)$. The new probability vector, after the second response and outcome is:

$$p(3) = M^{11}p(2) = \begin{vmatrix} 0.11 & 0.01 & 0.01 \\ 0.445 & 0.40 & 0.59 \\ 0.445 & 0.59 & 0.40 \end{vmatrix} \begin{vmatrix} 0.478 \\ 0.478 \\ 0.044 \end{vmatrix} = \begin{vmatrix} 0.058 \\ 0.430 \\ 0.512 \end{vmatrix}.$$

Again, since $p_1(3) + p_2(3) < R_3$ then $r_3 = 3$. Continuing:

$$p(4) = M^{23}p(3) = \begin{vmatrix} 0.50 & 0.49 & 0.405 \\ 0.49 & 0.50 & 0.405 \\ 0.01 & 0.01 & 0.19 \end{vmatrix} \begin{vmatrix} 0.058 \\ 0.430 \\ 0.512 \end{vmatrix} = \begin{vmatrix} 0.447 \\ 0.451 \\ 0.102 \end{vmatrix},$$

and, since $p_1(4) \leq R_4 < p_1(4) + p_2(4)$, we have finally $r_4 = 2$. This illustrates the method of generating synthetic trials.

An IBM 704 Computer was used to generate 600 synthetic trials of three responses each with the following preliminary data:

$$m = 3; \quad n = 2; \quad s_1 = s_2 = 1; \quad p_1(1) = p_2(1) = p_3(1) = \tfrac{1}{3};$$
$$a^1 = 0.05, \quad b^1 = 0.04, \quad c^1 = 0.50, \quad a^2 = 0.13,$$
$$b^2 = 0.01, \quad c^2 = 0.60; \quad N = 3, \quad q = 600.$$

When the 600 sequences were enumerated the results were as follows:

$$F_{111} = 1, \quad F_{112} = 24, \quad F_{121} = 262, \quad F_{122} = 16, \quad F_{123} = 297,$$

or, equivalently:

$$f_{111} = .00166, \quad f_{112} = .04, \quad f_{121} = .43667, \quad f_{122} = .02667, \quad f_{123} = .495.$$

When these values for f_{ijk} are substituted in **21** we obtain:

$$\hat{a} = -0.0216, \quad \hat{b} = 0.0733, \quad \hat{c} = 0.5329.$$

This illustrates the fact that the estimates yielded by **21** cannot always be used without modification because, as in this case, they may violate the restrictions of **3**. Of course, in this example, the true maximum-likelihood estimates would be the values of a, b, c that maximize the quantity.

$$ln\, L(a, b, c) = ln\, p_{111} + 24\, ln\, p_{112} + 262\, ln\, p_{121} + 16\, ln\, p_{122} + 297\, ln\, p_{123}, \quad \textbf{(23)}$$

where the p_{ijk} are as given by **19** and a, b, c are subject to the restrictions of **3**. These estimates are $\hat{a} = 0.0135$, $\hat{b} = 0.0483$, $\hat{c} = 0.5454$, for which $ln\, L(\hat{a}, \hat{b}, \hat{c}) = -0.94655$.

The actual 75 responses (r_t) for a synthetic trial whose parameters were those of **22**, produced by an IBM 704 run, are as follows:

31321	32213	21213	23121	31233	21231	23231	31232
32131	31332	32321	31232	32332	12121	23231	

(24)

The sequence of outcomes (s_t) that was used in this synthetic trial was as follows:

11212	12121	12211	11112	22111	11211	22111	11111
11211	21211	12221	21111	11211	12211	12111	

(25)

In principle, but not in practice, one can easily find the maximum-likelihood estimates that follow from these values of $p(1)$, r_t, and s_t by maximizing L

in **8**. We shall now describe and illustrate a method that yields approximate values for such estimates.

Approximations good to one decimal place, for each of the six parameters, could be obtained by computing the value of the likelihood function **8** for each of the 11^6 possible sets of parameter values between 0 and 1; actually, there are somewhat less than 11^6 possible sets because of the restrictions that $b^1 + c^1 \leq 1$ and $b^2 + c^2 \leq 1$. Then the set of parameter values yielding the largest likelihood value would be chosen as the approximate maximum likelihood estimated parameters. The IBM 704 code we have used for such calculations computes these values at the rate of about 11/sec., for the case represented by the data of **24** and **25**, including the successive choice of sets of parameter values and selection and printout of the largest likelihood value found for all sets. At this rate of calculation, it would require some thirty hours of IBM 704 time to determine the six parameter values accurately to about one decimal place, for a single trial consisting of 75 successive responses among three alternatives; obviously, some more economical scheme of calculation must be found than simply computing and comparing among all points of a fine lattice in the parameter space.

Two schemes have been used for handling data like that of **23** and **24**. One consists in computing a relatively coarse lattice, perhaps starting with 3^6 points, and then computing a somewhat finer lattice centered around the best point found previously, until all points of the lattice give about the same likelihood value. Another method consists in searching directly for the best parameter values by some computational scheme that tries a point, or points, and then moves next to a new trial point according to some systematic criterion. We shall call the first method the "successive lattice method," and the second method the "direct search method." We shall now present some results obtained using a successive lattice method.

An IBM 704 code was written to make possible a lattice computation over a subspace of the 6-dimensional unit cube that represents the parameter space. This "lattice subspace," and the fineness of the "mesh" within it, is chosen by the experimenter; roughly speaking, the lattice subspace can include any closed interval for each of the six parameters, and the fineness of the mesh can be chosen arbitrarily and independently for each parameter. The following example, for data generated synthetically as for **24** and **25**, will illustrate this procedure.

For our present synthetic trial we chose:

$$m = 3; \quad n = 2; \quad s_t \text{ as in } \mathbf{25}; \quad p_1(1) = p_2(1) = p_3(1) = \tfrac{1}{3}; \quad a^1 = .01,$$
$$b^1 = 0, \quad c^1 = .4, \quad a^2 = .19, \quad b^2 = 0, \quad c^2 = .25; \quad N = 75.$$

The sequence of responses (r_t) produced synthetically by the IBM 704 run were:

$$\begin{array}{cccccccc} 21312 & 13231 & 32312 & 12323 & 13131 & 21221 & 21323 & 13232 \\ 32323 & 23232 & 31323 & 22313 & 21213 & 23321 & 23232 & \end{array} \tag{26}$$

The first lattice, FIGURE 1, was used here. Thus, each parameter was allowed

to take on the values 0, ½, and 1 until all allowable combinations were computed. The result was:

$$ln\ L(0, 0, 0; ½, 0, 0) = 75\ (-0.7285) \geq ln\ L(a^1, b^1, c^1, a^2, b^2, c^2).$$

Following this, the lattice in FIGURE 2 was used. The best result was:

$$ln\ L(0, 0, 0; ¼, 0, 0) = 75\ (-0.7207).$$

FIGURES 3, 4, and 5, with three additional lattices and their best results, were then used.

The example was not carried beyond this stage, but it illustrates not only the difficulty in using the lattice method but also its usefulness. Thus, after computing for only about 3600 points (ten minutes of IBM 704 time), we have some idea of the location of the likelihood estimates and also we see that the value of the likelihood function varies little as the parameter values are changed rather substantially. We shall now turn to the direct search type of method.

	a^1	b^1	c^1	a^2	b^2	c^2
Min.	0	0	0	0	0	0
Max.	1	1	1	1	1	1
Mesh	2	2	2	2	2	2

FIGURE 1.

	a^1	b^1	c^1	a^2	b^2	c^2
Min.	0	0	0	0.25	0	0
Max.	0.5	0.5	0.5	0.75	0.5	0.5
Mesh	2	2	2	2	2	2

FIGURE 2.

There are a number of direct search methods that could be used. The steepest ascent methods are a popular variety for this type of problem. One of these methods simply starts at some more or less arbitrary point in the parameter space, then computes the value of the function on some fairly small lattice centered on this point, then determines from these points the direction in the parameter space that ascends most rapidly, then moves step by step in this direction until progress upward ceases, and then repeats this general process until a relative maximum seems to have been reached. Unfortunately, the method of steepest ascent does not seem to work well in our type of problem, where the function is a polynomial of very high degree in six parameters, so that there is no assurance that a relative maximum is also a general maximum. We favor, instead, a method like that of Hooke and Jeeves,[11] in which the ascent is more direct and is done without seeking the most rapid ascent at each move. Their method was coded for the IBM 704, and we have modified this code for our present purposes (*Appendix B*). Its use will be illustrated and discussed in later sections, both for seeking maxima in cases like those just treated by the lattice method and in cases like those represented by **23**.

We turn next to a few questions that arise in considering the statistical significance of some of the results obtained by using maximum-likelihood estimates derived in the manner just described.

Parameter Significance

In the preceding section, two sets of synthetic trials were analyzed to estimate the parameter values that produced them. In both instances, the estimates were apparently not very close to the true values. Our present task is to develop a significance test that will indicate whether or not such discrepan-

	a^1	b^1	c^1	a^2	b^2	c^2
Min.	0	0	0	0.125	0	0
Max.	0.25	0.25	0.25	0.375	0.25	0.25
Mesh	2	2	2	2	2	2

$ln\ L\ (0, 0, 0; 0.375, 0, 0) = 75\ (-0.7191)$

FIGURE 3.

	a^1	b^1	c^1	a^2	b^2	c^2
Min.	0	0	0	0.31	0	0
Max.	0.12	0.12	0.12	0.43	0.12	0.12
Mesh	2	2	2	2	2	2

$ln\ L\ (0, 0, 0; 0.31, 0, 0) = 75\ (-0.7184)$

FIGURE 4.

	a^1	b^1	c^1	a^2	b^2	c^2
Min.	0.001	0.001	0.001	0.28	0.001	0.001
Max.	0.061	0.061	0.061	0.34	0.061	0.061
Mesh	2	2	2	2	2	2

$ln\ L\ (0.001, 0.001, 0.31; 0.31, 0.001, 0.001) = 75\ (-0.7190)$

FIGURE 5.

cies may reasonably be attributed to random fluctuations in our experiments. Such tests will also show how confidence intervals may be chosen for parameter estimates obtained in this manner.

In one case, that of 600 trials each of three responses, the true values and the estimates were as shown in TABLE 1. The first significance test is directed toward a comparison of the observed and theoretical distributions, as shown in TABLE 2.

From the data of TABLE 2 we find $\chi^2 = 5.12$, indicating a degree of discrepancy between theoretical and observed distributions that would arise less than 10 per cent of the time; actually about as large a discrepancy as we would want to tolerate routinely. From a confidence region standpoint, we would

therefore say that a set of parameter values that led to a distribution with no larger discrepancy than this from the observed would be included within a confidence region based on the probability level corresponding to $\chi^2 = 5.12$. In particular, we would want our parameter estimates (0, 0.07, 0.53) to satisfy this criterion, and we now verify that this is a point within our confidence region; the new "theoretical" values in TABLE 3 are based on this set of parameter values, rather than upon the true values used in the synthetic run. From the data of TABLE 3 we find $\chi^2 = 1.7$, indicating that our estimates are well within the critical region determined by our data. Here then is one general method for determining points within the critical region.

TABLE 1
TRUE VALUES AND ESTIMATES

	a^1	b^1	c^1	a^2	b^2	c^2
True	0.05	0.04	0.50	0.13	0.01	0.60
Est.*	0	0.07	0.53		None	

* These are not really maximum-likelihood estimates.

TABLE 2
OBSERVED AND THEORETICAL DISTRIBUTIONS

	p_{111}	p_{112}	p_{121}	p_{122}	p_{123}
Theoretical*	0.00175	0.04160	0.44900	0.04300	0.46500
Observed	0.00167	0.04000	0.43667	0.02667	0.49500

* Based on the true parameter values.

TABLE 3
THEORETICAL AND OBSERVED PARAMETER VALUES

	p_{111}	p_{112}	p_{121}	p_{122}	p_{123}
Theoretical	0.003	0.044	0.433	0.035	0.487
Observed	0.00167	0.04000	0.43667	0.02667	0.49500

It is already clear that a run of 600 trials does not narrow the critical region down to two-decimal accuracy for our parameters in this particular case. We turn now to our second type of estimate, that based upon a run of one trial with 75 responses among the three alternatives, as displayed in **26**.

We have already noted that:

$$ln\ L(0, 0, 0; 0.31, 0, 0) = 75\ (-0.7184) = ln\ L(\text{est. parameters}),$$

is the largest log-likelihood value obtained from the sequence of lattices used in approximating the maximum-likelihood estimates in our 75-response trial. A calculation made with the IBM 704 shows also that:

$$ln\ L(0.01, 0, 0.4; 0.19, 0, 0.25) = 75\ (-0.7595) = ln\ L(\text{true parameters}).$$

We consider the likelihood-ratio:

$$2\ ln\ \lambda \equiv 2\ ln\left[\frac{L(\text{est.})}{L(\text{true})}\right] = 6.165 = \chi^2.$$

This value of χ^2 is well within our confidence region, taking into account the six degrees of freedom represented by the six parameters.

In general, then, we will consider a set of parameter values θ to be within our confidence region, relative to a maximum-likelihood estimate denoted $\hat{\theta}$, provided the following condition is satisfied:

$$2\ ln\ \lambda = 2[ln\ L(\theta) - ln\ L(\hat{\theta})] = \chi^2 \leq \chi_0^2,$$

where χ_0^2 is chosen so that the probability of observing no larger value is 0.95 if $\hat{\theta}$ were correct, if we work at the 5 per cent confidence level. For example, if an experimental trial with one animal yields maximum-likelihood estimates $\hat{\theta}$, and the likelihood value for another animal's data, computed using the parameter value $\hat{\theta}$, is L^*, then we will not consider the two animals to differ (on the basis of this experimental data) unless

$$2[ln\ L^* - ln\ L(\hat{\theta})] > \chi_0^2.$$

In other words, if we assume that $\hat{\theta}$ estimates the true value common to the two animals, then this test will reject the hypothesis only if the second animal's experimental data yield a likelihood value incompatible with this hypothesis.

An example will illustrate this principle. Two rats were run in a three-choice experiment with the following results:

Rat 1: r_t =	13232	13132	32132	12132	12321	32121	31321	31312
	12213	21211	21213	13132	13131	23232	1313	
s_t =	11212	11112	11111	11111	11111	11111	11111	11111
	11212	11111	11111	11111	11111	11111	1111	
Rat 2: r_t =	13213	13223	21321	32321	32313	21311	21311	32131
	32211	31313	21313	21313	31313	23132	13132	
s_t =	11212	12121	12211	11112	22111	11211	22111	11111
	11211	21211	12221	21111	11211	12211	12111	

After several IBM 704 mesh runs, the following approximate values were found for the six parameters for Rats 1 and 2:

$$\textit{Rat 1:}\ (0.06,\ 0,\ 0.4;\ 0.13,\ 0,\ 0.6) = \hat{\theta}_1\,.$$

$$\textit{Rat 2:}\ (0.09,\ 0,\ 0.5;\ 0.13,\ 0,\ 0.4) = \hat{\theta}_2\,.$$

If we take $\chi_0^2 = 12.6$, corresponding to a 5 per cent confidence level based on six degrees of freedom, then the confidence region for Rat 2 will include any set of parameters θ such that $2[ln\ L(\hat{\theta}_2) - ln\ L(\theta)] \leq 12.6$; in particular, this condition will test whether Rat 1 is distinguished from Rat 2 when we use $\theta = \hat{\theta}_1$.

In this case, we found $[2\ ln\ L(\hat{\theta}_2)] = -140.46$, so our condition becomes $[-ln\ L(\theta) \leq 76.53]$. Actually, $-ln\ (\hat{\theta}_1) = 73.09$, so that Rat 1 is indistinguishable from Rat 2 on this basis.

We also found several sets of parameter values within the confidence region, so defined, for Rat 2. Among these are:

$$b^1 = 0, \quad a^2 = 0.13, \quad b^2 = 0, \quad c^2 = 0.4,$$

and

$$a^1\text{: } 0.08,\ 0.09,\ 0.08,\ 0.07,\ 0.09,\ 0.08,\ 0.07,\ 0.06,\ 0.05,\ 0.04;$$

$$c^1\text{: } 0.5,\ \ 0.4,\ \ 0.4,\ \ 0.4,\ \ 0.3,\ \ 0.3,\ \ 0.3,\ \ 0.3,\ \ 0.3,\ \ 0.3.$$

Two sets outside the confidence region are:

$$\left.\begin{aligned} a^1 &= 0.03,\ 0.02 \\ c^1 &= 0.3,\ 0.3 \end{aligned}\right\} \quad \text{and} \quad b^1 = 0,\ a^2 = 0.13,\ b^2 = 0,\ c^2 = 0.4.$$

On another trial, for Rat 2 with 51 responses, the parameter estimates were as follows:

$$\hat{\theta} = (0.06,\ 0,\ 0.3;\ 0.13,\ 0,\ 0.6) \quad \text{with} \quad -2\ ln\ (\hat{\theta}) = 83.844.$$

These parameter values seem very close to the ones listed above for Rat 1 so, on the basis of these experimental data, the two rats would not seem to be distinguishable. When these parameter estimates $(\hat{\theta})$ for Rat 2 are used to determine the likelihood based on the previous 75-response trial for Rat 2, we find $-ln\ (\hat{\theta}) = 74.37 \leq 76.53$, so that again we would not distinguish any difference in Rat 2 between these two trials.

The examples just discussed should suffice to illustrate our methods for estimating parameters, and for assessing the significance of differences among them. Basically, it is familiar maximum-likelihood estimation with confidence regions based on the likelihood-ratio significance test.

Preliminary Experimental Results

Some preliminary experimental results have been obtained with two rats, here called Rat *I* and Rat *J*. Unfortunately, Rat *J* died before the planned schedule of experiments had been completed. The full schedule has also not yet been completed for Rat *I*.

Two types of experiments were conducted. In one type (Schedule A), a probability of reinforcement was assigned to each alternative, and the number of reinforced responses depended upon the actual responses made by the animal. In the other type (Schedule B), the sequence of reinforcements was entirely independent of the actual responses made.

Schedule A used the probabilities (0.2, 0.6, 0.8) or (0, 0.6, 1). Schedule B normally consisted of repeated blocks of 12 stimuli, each block starting with six reinforced responses and ending with six unreinforced responses. Normally, an animal was kept responding for about one hour and made 100 to 350 responses in that period.

The animal could press any one of six bars at each response. If the animal moved to the center of the cage so as to activate the photocell, after pressing any bar, then all six bars were reset and made active. The apparatus recorded the time at which each bar was pressed and the time when the photocell was first activated after each bar-press. For the present analysis, the responses used were the bars first pressed after each resetting by the photocell; corresponding to each such response r_t, there is either an outcome $s_t = 1$, denoting that the response was reinforced, or an outcome $s_t = 2$, denoting that the response was not reinforced. Also, for the present analysis, bars symmetrically opposite within the cage were treated as though they were identical; thus, there were three alternatives available for each response: a left bar, a middle bar, and a right bar.

Every effort was made to keep the opposite sides of the cage as nearly identical as possible. For example, the walls behind the bars were painted differently for left, middle, and right, but the same for each direction on both sides of the cage. Nevertheless, it became apparent that the experimental animals

TABLE 4
MAXIMUM-LIKELIHOOD ESTIMATES
Starting State Assumed

Trial No.	Rat	$\hat{a}^1$	$\hat{b}^1$	$\hat{c}^1$	$\hat{a}^2$	$\hat{b}^2$	$\hat{c}^2$	N	L^*	d
144	J	0.999	0.085	0.915	0.949	0.007	0.987	231	0.772	0.036
146	J	0.949	0.101	0.899	0.949	0.151	0.849	134	1.017	0.063
148	J	0.999	0.063	0.931	0.999	0.033	0.967	269	0.829	0.031
150	J	0.899	0.100	0.900	0.999	0.100	0.700	144	1.064	0.058
196†	I	0.999	0.001	0.998	0.999	0.050	0.949	206	0.944	0.041

† Search calculation not complete.

behaved differently on opposite sides of the cage, so grouping together two bars symmetrically placed in the cage proved to be a more arbitrary combination of response classes than had been intended. Consequently, the present analysis is offered as an example of results obtained with a somewhat arbitrary grouping of response classes, rather than for three response classes that might more reasonably be expected to satisfy the symmetry conditions upon which our mathematical model is based.

The results shown in TABLE 4 were obtained using Schedule A (0.2, 0.6, 0.8) with Rat J, and using Schedule B with Rat I. The quantity d, in TABLE 4, is the relevant value for χ_0^2 working at the 1 per cent level with six parameters; that is, any set of six parameters yielding a likelihood value less than $(L^* + d)$ is within the 1 per cent confidence region. (In the table, $L^* = [-ln\ L]/N$.) For example, on Trial 148, the following six sets of parameter values include three sets that are within this confidence region ($L^* \leq 0.86$) and three sets that are not.

The first three sets of parameter values shown in TABLE 5 differ from the maximum-likelihood estimates, shown in TABLE 4, by the following maximum amounts: 0.050, 0.100, 0.131, 0.100, 0.072, and 0.072. It is likely that parame-

ter values could be found within the confidence region that would show even greater deviations. On the other hand, the second three sets of parameter values shown in TABLE 5 differ from the maximum-likelihood estimates, shown in TABLE 4, by the following minimum amounts: 0, 0.013, 0.031, 0, 0.018, and 0.032.

Thus only rather slight changes in certain parameter values will sometimes place the set outside the confidence region. It can only be concluded, from the data of TABLE 4, that the maximum-likelihood estimates are reasonably similar from trial to trial and from rat to rat; a more extensive experiment and analysis is required in order to test this apparent stability of parameter estimates and to narrow the confidence regions appreciably.

TABLE 5
TRIAL 148 LIKELIHOODS

a^1	b^1	c^1	a^2	b^2	c^2	L^*
0.949	0.100	0.900	0.999	0.052	0.948	0.860
1.000	0.200	0.800	0.899	0.105	0.895	0.858
0.999	0.100	0.900	0.999	0.001	0.898	0.848
0.999	0.100	0.900	0.999	0.001	0.999	0.920
0.999	0.100	0.700	0.898	0.102	0.898	0.912
0.999	0.050	0.900	0.999	0.051	0.898	0.878

TABLE 6
MAXIMUM-LIKELIHOOD ESTIMATES
Starting State Predicted

Trial No.	$p\hat{L}(1)$	$p\hat{C}(1)$	$p\hat{R}(1)$	$\hat{a}^1$	$\hat{b}^1$	$\hat{c}^1$	$\hat{a}^2$	$\hat{b}^2$	$\hat{c}^2$	L^*
146	0.215	0.739	0.046	0.937	0.112	0.875	0.987	0.086	0.890	1.010
148	0.331	0.562	0.107	0.993	0.070	0.924	0.999	0.026	0.974	0.822
150	0.080	0.356	0.564	0.942	0.076	0.918	0.999	0.063	0.880	1.056

The maximum-likelihood estimates, shown in TABLE 4, were all computed on the assumption that the first response of each trial was equally likely to be any one of the three bars, thus: $p_1(1) = p_2(1) = p_3(1) = \frac{1}{3}$. The computations were done on the IBM 704, using the modified search code described in *Appendix B*.

Another calculation was made, also using the search code, in which the initial state for each trial was taken to be the ending state for the preceding trial; this is an alternative to the assumption of equal probability among bars at the start of each trial. The results are shown in TABLE 6. It is clear, from TABLE 6, that the use of predicted starting states improves the fit of the model to the data, since each value of L^* is smaller than in TABLE 4, where the starting state was assumed to be probability $\frac{1}{3}$ for each bar. Unfortunately, the estimates of TABLES 4 and 6 are not entirely reliable because the search code does not absolutely ensure correct likelihood values.

The starting states shown in TABLE 6, for each trial, were calculated from the data for the preceding trial. More specifically, the starting state for Trial 144 was taken arbitrarily to be $p_1(1) = p_2(1) = p_3(1) = \frac{1}{3}$. This yielded the maximum-likelihood estimates for Trial 144, as shown in TABLE 4. The end state for Trial 144, based upon the estimated parameters, was calculated to be: $p_L(232) = 0.215$, $p_C(232) = 0.739$, $p_R(232) = 0.046$. Since the first response in Trial 146 is the next one made by Rat *J* after the 231st response in Trial 144, with a rest of one hour intervening, the end state of Trial 144 was taken as the starting state for Trial 146. Similar reasoning applies to Trials 148 and 150.

The actual schedule of reinforcements used on Trials 144, 146, 148, and 150 led to the gross results shown in TABLE 7. The results in TABLE 7 followed from

TABLE 7
DISTRIBUTION OF BAR PRESSES

Trial No.	Number of times bar pressed						
	Total	Reinforced			Unreinforced		
		Left	Center	Right	Left	Center	Right
144	231	36	128	2	24	31	10
146	134	12	36	9	40	28	9
148	269	3	97	60	29	67	13
150	144	38	18	20	10	48	10

TABLE 8
COMPARISON OF ACTUAL AND SCHEDULED REINFORCEMENT PROBABILITIES

Trial No.	Scheduled			Actual		
	L	C	R	L	C	R
144	0.6	0.8	0.2	0.60	0.81	0.17
146	0.2	0.6	0.8	0.23	0.56	0.50
148	0.2	0.6	0.8	0.09	0.59	0.82
150	0.8	0.2	0.6	0.79	0.27	0.67

use of Schedule A (0.2, 0.6, 0.8), with a permutation of the probabilities among the bars after Trials 144 and 148. The actual relative frequencies are compared with the scheduled probabilities in TABLE 8. The starting states shown in TABLE 6 clearly show the tendency for the rat to select the responses that have led to reinforcements most regularly during the recent past, but they also show that former good alternatives continue to be selected for a considerable period of time.

It is interesting that $(\hat{b}^1 + \hat{c}^1)$ and $(\hat{b}^2 + \hat{c}^2)$ are close to unity in value for each trial shown in TABLE 6. Furthermore, if the combining-of-classes assumption holds, then necessarily

$$C^s \equiv \frac{1 - a^s}{2} - (1 - b^s - c^s) = 0 \text{ for } s = 1, 2.$$

The values for C^s actually observed are given in TABLE 9. Although the observed values for C^s are all near zero, our experiments and analysis are not extensive enough to determine whether or not the combining-of-classes assumption is acceptable for this test situation. This question is being investigated further.

It now seems that one trial, with $N = 1000$, is about right to yield parameter estimates accurate to two decimal places. With our present computational methods, it requires about 30 min. of IBM 704 time to obtain two separate sets of parameter estimates for one such trial, using the search code described in *Appendix B.* For such a trial, the confidence region, at the 1 per cent level, has a value of $d < 0.01$; this seems to be precise enough to yield reasonably sensitive comparisons between sets of parameter estimates.

The effect of small changes in parameter values, in the general neighborhood of values observed for Rats *I* and *J*, is being investigated by the use of synthetic runs on the IBM 704. This synthetic technique is being used to explore empirically the distribution of *L*, and to explore changes induced by changes in the experimental situation. In other words, the synthetic rat also becomes a subject of experimental investigation and various comparisons between gross

TABLE 9
OBSERVED VALUES FOR C^s

Trial No.	144	146	148	150	196
C^1	0	0.018	−0.003	0.023	0
C^2	−0.019	−0.018	0	−0.007	0

behavior of synthetic rat and actual rat should help to highlight similarities and differences between a real rat and its mathematical model.

Conclusions

(1) Stimulation through electrodes implanted in the hypothalamic region of a rat's brain can be used efficiently to provide reinforcing stimuli in a learning experiment extending over a period of months.

(2) Maximum-likelihood estimates of the six parameters in the symmetry model can be approximated, using a search code, in less than 30 min. of IBM 704 time, for a trial consisting of 1000 successive responses from among three alternatives.

(3) Estimated values of the six symmetry model parameters are reasonably constant from trial to trial for a single rat, when measured in an experiment in which the rat has three alternatives for each response and the schedule of reinforcement depends upon the responses actually made.

(4) The very limited experimental data analyzed are compatible with the hypothesis that estimated parameter values are the same for the two rats tested, even though the parameters are measured using two quite different types of reinforcement schedules.

(5) The data analyzed are also compatible with the hypothesis that the combining-of-classes condition is satisfied.

Acknowledgments

I am greatly indebted, for assistance and advice, to several of my colleagues in the Mental Health Research Institute of the University of Michigan. James G. Miller and Ralph W. Gerard encouraged me to resume these investigations, and provided ideal facilities for this type of research in the Institute's laboratories. James Olds introduced me to the use of hypothalamic implants as a reinforcement procedure in learning experiments with rats. I have also benefited greatly from the experience with electrode implantation of Gerard H. Rothschild, of the Institute, and of Robert W. Doty, R. L. Isaacson, and others at the University of Michigan, who have worked with implanted electrodes in rats.

The central concept for the experiment was discussed with R. R. Bush, Eugene Galanter, and R. D. Luce in the summer of 1958; they very generously provided information regarding their current work and the recent work of others that helped in planning the experiment and analysis. Similar more recent conversations with Patrick Suppes and Richard C. Atkinson, and with J. David Birch, have helped in choosing among experimental possibilities.

The extensive computations were made possible by the generous support of the Computing Center of the University of Michigan, through R. C. F. Bartels, where expert assistance with programming and the IBM 704 Computer were always readily available.

C. C. Craig and R. V. Evans have given valuable advice regarding various statistical questions. John Dwyer has helped with the analysis, and has checked several of the calculations.

I am especially grateful to Fred E. Horvath, who supervised the design and construction of the experimental apparatus, implanted the electrodes, trained and ran the experimental animals, and made the experimental trials go smoothly. Above all, I am indebted to him for many pointers regarding the psychology of rats and their behavior in learning experiments of this kind. A summary of his report on the experimental apparatus and techniques is included here as *Appendix A*; a more detailed report will be published elsewhere.

References

1. Estes, W. K. 1950. Toward a statistical theory of learning. Psychol. Rev. **57**(2): 94–107.
2. Bush, R. R. & C. F. Mosteller. 1951. A mathematical model for simple learning. Psychol. Rev. **58**(5): 313–323.
3. Flood, M. M. 1952. On stochastic learning theory. Paper P-353. The Rand Corporation. Santa Monica, Calif.
4. Miller, G. A. & W. J. McGill. 1952. A statistical description of verbal learning. Psychometrika. **17**: 369–396.
5. Bales, R. F., M. M. Flood & A. Householder. 1952. Some group interaction models. Research Memorandum RM-953. The Rand Corporation. Santa Monica, Calif.
6. Bush, R. R. & C. F. Mosteller. 1955. Stochastic Models for Learning. Wiley. New York, N. Y.
7. Bush, R. R. & W. K. Estes, Eds. 1959. Studies in Mathematical Learning Theory. Stanford Univ. Press. Stanford, Calif.
8. Luce, R. D. 1959. Individual Choice Behavior. Wiley. New York, N. Y.
9. Suppes, P. 1959. A linear model for a continuum of responses. *In* Studies in Mathematical Learning Theory. : 400–414. R. Bush and W. Estes, Eds. Stanford Univ. Press. Stanford, Calif.
10. Flood, M. M. 1954. A stochastic model for social interaction. Trans. N. Y. Acad. Sci. **16**(4): 202–205.

11. HOOKE, R. & T. A. JEEVES. 1959. Direct search solution of numerical and statistical problems. Scientific Paper 10-1210-1-P1. Westinghouse Research Laboratories. Pittsburgh, Pa.
12. OLDS, J. & P. MILNER. 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. J. Comp. Physiol. Psychol. **47:** 419–427.
13. OLDS, J. 1956. A preliminary mapping of electrical reinforcing effects in the rat brain. J. Comp. Physiol. Psychol. **49:** 281–285.

APPENDIX A

Experimental Procedure

Subjects and electrode implantation. The majority of the experimental work was carried out with two male, mature albino rats. The electrodes and implantation techniques used were those describe in detail by Olds and Milner.[12] In brief, the electrodes are a pair of twisted silver wires, diamel coated for insulation except at their tips, thus restricting the stimulus to that point. Each pair of electrodes is secured to a lucite pedestal that, in turn, is held in place on the skull by screws through the skull bone. This technique has proven to be very satisfactory for longitudinal studies using implanted rats.

The area selected for implantation was the medial forebrain bundle, lateral to the anterior hypothalamus, an area Olds has rated as positively reinforcing.[13]

Experimental apparatus and procedure. The reinforcing stimulus was an AC pulse of 10 to 25 μamp. intensity, with duration a function of the length of time the bar was depressed, up to an automatic cut-off time of 0.6 sec. Stimulus threshold values for reinforcement in this area are about 12 μamp., while currents in excess of 25 μamp. are best avoided since they make the rats hyperexcitable, disrupting their performance in a choice situation.

Following a one-week postoperative recovery period, the rats first were trained to bar press for a stimulus in an enclosed box with a single bar. The response was shaped up in the usual manner of giving reinforcements for successively closer approximations of the desired response, until the bar finally is depressed. Thereafter reinforcements are given automatically after each bar press, a response both rats were making at rates of 500 to 1000/hour within a few hours of training.

The testing of rats in a choice situation necessitates the rat's return to a neutral choice point between responses. There are advantages to having them proceed to a choice point on their own, so the rats next were trained to go to the opposite side of the box after every bar press. This more complex response also was shaped up by the method of reinforcing successive approximations to the required response sequence. When they had learned to make this total response with a minimum of errors (usually within 10 hours of experimental sessions), they were advanced to the final training phase.

The first apparatus used in the final phase was a large circular cage with three bars equally spaced around its circumference and a disc in the center marking the choice point. Although the rats readily transferred their previous learning to this three-bar condition—going to the choice point after pressing any of the three bars—they also demonstrated a strong bias that practically reduced this to a two-choice situation. Upon reaching the center choice point, they showed

a strong preference for choosing their next response from between the two bars facing them from across the cage, rather than making the inefficient response of turning around to return to the bar just pressed. Reinforcing the bars differentially may have overcome this bias, but it seemed preferable to start with a less biased three-choice situation.

Therefore, a new apparatus was designed to incorporate two desirable features: (1) a choice point at which the rat's visual field would include all of the bars from which he is expected to choose; and (2) ease in converting the apparatus to test the rats on any number of bars. The apparatus is illustrated in FIGURE 6, and consists basically of two fan-shaped center sections, the sides of which can be pivoted at the center point to encompass the desired number of removable U-shaped panels at each end. When the rat passes through the photocell beam at the choice point, the stimulus circuit is reset so he can be

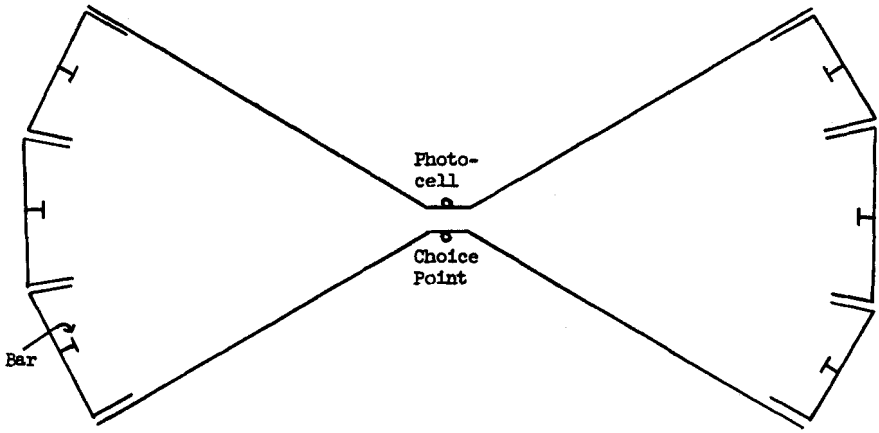

FIGURE 6.

reinforced for his next press of any of the bars. Control circuitry makes it possible to run experiments using patterned or relatively random partial reinforcement schedules, with either the same or differential schedules on each bar.

In the initial sessions in this apparatus, the rats were run with one bar at each end, but they adapted so quickly to this new condition that in subsequent sessions they were run with three bars. Under conditions of 100 per cent reinforcement, there were very few failures to return to the choice point between bar presses. Under partial reinforcement schedules, failures to return increased, indicating emotional outbursts rather than lack of training. The reversal of differential reinforcement schedules has pointed up one problem with this apparatus: a possible nonequivalence of the bars at both ends. Despite our attempts to eliminate all differentiating cues, the rats are discriminating sufficiently between the ends to select different bars at each end. If this nonequivalence proves to be the case, the apparatus will be modified by retaining the bars at one end only. At the other end the bars will be replaced by a blank panel attached to the center section to form a cul-de-sac, with the

photocell relocated close to the end of this blind alley. This will force the rat to enter the alley to reset the device, so that on leaving the alley he can again face all of the bars from which he is expected to choose.

Appendix B

Search Codes

Maximum-likelihood estimation and many other situations require an efficient computational procedure for determining the location of the largest (or smallest) value of a function of several real variables. Some or all of the variables may also be restricted to some region; common restrictions would be nonnegativity or limitation to the unit n-cube.

Our maximum-likelihood estimation problem is essentially that of finding the values for a set of real variables that maximizes the value of a polynomial in the variables, where the variables are restricted to a specified portion of the unit cube.

Mathematically, if we let x denote the set of real variables $(x_1, x_2, \cdots, x_n)$, $f(x)$ a real function of the variables x, and R the domain over which the variables are allowed to range, then our problem is to find a set of values $\hat{x}$ such that $f(\hat{x}) \geq f(x)$ for all x in R. Or, rather, our need is for a computational procedure that will solve this problem economically when $f(x)$ is one of the likelihood functions encountered in analyzing experimental data.

When $f(x)$ is a polynomial, given explicitly as is ours in the likelihood estimation situation, there is no theoretical difficulty in finding the required values $\hat{x}$ by the ordinary methods of calculus. In brief outline, for the case when R is the unit n-cube, all that is necessary is:

(1) Find the roots of the polynomial $f'(x) \equiv df(x)/dx$. For each such root, if it is in R, test to determine whether or not it yields a relative maximum for $f(x)$. Let $M = f(\hat{x})$ be the largest of these relative maxima.

(2) Let $X \equiv (X_1, X_2, \cdots, X_n)$, where X_i is either x_i, 0, or 1. Then $f(X)$ is the value of the function on some bounding hyperplane of the unit n-cube; the dimension of this hyperplane is the number of x_i included in X. By the method outlined in (1), the largest value of $f(X)$ can be determined. This includes the case where all X_i are 0 or 1, representing the values at the corners. Let $M(\bar{X}) = f(\bar{X})$ be the largest of the relative maxima for the bounding hyperplane defined by X.

(3) The required value of $\hat{x}$ is the $\bar{X}$ that satisfies the condition

$$f(\bar{X}) \geq M(\bar{X}) \text{ for all } \bar{X}.$$

In other words, the locations of the largest interior relative maxima are determined on all bounding hyperplanes, and the solution to our problem is the one that yields the largest value of the function. When the degree of the polynomial is high, or the number of variables is large, the techniques of the calculus are generally too inefficient to permit their use, as is normally the case for our likelihood functions. Consequently, we have turned to the approximation method to be described now.

Hooke and Jeeves[11] have developed a procedure for searching through the

domain R, and have used the procedure successfully on various problems. Their procedure involves a combination of two rather different types of searching operations: they have called one type "direct search" and the other type "pattern search." Only the central features of their procedure will be presented here.

Direct search starts with an arbitrary point x^1 in the n-dimensional domain R, where R is defined by the inequalities:

$$d_i \leq x_i \leq D_i \quad \text{for} \quad i = 1, 2, \cdots, n.$$

The first computational step is to compare $f(x^1)$ and $f(x_1^1 + \Delta, x_2^1, \cdots, x_n^1)$, where Δ is some positive real number and where $x_1^1 + \Delta$ is changed to D_1 if $x_1^1 + \Delta > D_1$. If the change in the variable yields a larger value of the function then the new value is retained. If not, x_1^1 is decreased by amount Δ, but not to a value less than d_1, and the changed variable is retained if and only if the function value increases. The procedure is repeated with x_2, then with x_3, and so on through x_n. If no change in the variables yielded an improvement, then the entire process is repeated with $(\Delta/2)$ replacing Δ, unless $(\Delta/2) < \Delta_{\min}$, until the process either terminates or an improved value of the function is located at some stage.

Pattern search follows each instance of success in direct search. If x^o denotes the starting point of a direct search step, and x the improved point at the termination of direct search, then pattern search starts with a direct search from the point $2(\bar{x} - x^o) \equiv \bar{x}^o$; as always, if any coordinate violates the condition

$$d_i \leq 2(\bar{x}_i - x_i^o) \leq D_i$$

then it is replaced by the appropriate value of d_i or D_i. In this direct search, as part of pattern search, the value of $f(x^o)$ is ignored and comparisons are with $f(\bar{x}^o)$ at each step during pattern search; furthermore, the terminal value of $(\Delta/2^g)$ in use at the termination of direct search is used unchanged throughout the direct search portion of pattern search. If direct search from $\bar{x}^o$ yields a terminal point $\bar{\bar{x}}^o$ such that $f(\bar{\bar{x}}^o) > f(\bar{x})$, then a new pattern search starts at the point $2(\bar{\bar{x}}^o - \bar{x})$, and $\bar{\bar{x}}^o$ and $\bar{x}$ become the new values for $\bar{x}$ and x^o in this pattern search step. On the other hand, if $f(\bar{\bar{x}}^o) \leq f(\bar{x})$ then pattern search is terminated and direct search is done about $\bar{x}$ with $(\Delta/2^{g+1})$ replacing $(\Delta/2^g)$.

Eventually the search terminates, when $\Delta(2^{g+1}) < \Delta_{\min}$, and the desired value of $\hat{x}$ is approximated by the final value of x in the search process. This brief description omits several important technical details, and says nothing of the actual computer codes used, for which the reader must be referred to the paper of Hooke and Jeeves.[11]

The search code of Hooke and Jeeves was necessarily modified slightly for our maximum-likelihood problem, because the domain R is not of quite the same form as that treated by Hooke and Jeeves. There are two additional restrictions on the variables, in our case, namely:

$$b^1 + c^1 \leq 1, \qquad b^2 + c^2 \leq 1.$$

Primarily for convenience in modifying the computer code, and because the

change seemed adequate to take proper account of these two added restrictions, the "modified search code" actually used retained the altered value of a variable b (or c) and changed c (or b) as necessary to satisfy the restriction $b + c \leq 1$. For example, if b^1 is to be replaced by $(b^1 + \Delta)$ when $(b^1 + \Delta \leq 1)$ and $(b^1 + \Delta + c^1 > 1)$ then c^1 is replaced by $(1 - b^1 - \Delta)$. The calculations reported upon in this paper, as using a search code, were all done with this simple modification of the Hooke-Jeeves code.

It is evident that the search code does not necessarily and inevitably terminate at a point that is even near the correct one. Nor does it seem likely at present that any computational procedure will soon be found that is both reasonably economical and guaranteed to yield an approximate solution near the correct one. Consequently, all of our empirical results are open to question on this ground.

We have sought protection against this danger of error in several ways. The three principal ways were: (1) comparison of search code approximations with lattice code approximations in sample instances; (2) comparison of search code approximations with others obtained using starting points as distant as possible from the previous solutions; and (3) checks for agreement of solutions by search code techniques with a few simple likelihood estimation problems for which exact solutions are known.

Of course, the most important source of protection against such errors is the comparison of calculated parameters from among independent trials with different animals. The results reported in the present paper were subjected to all of these tests, except for the incomplete calculation of parameters for Rat *J*.