A STOCHASTIC MODEL WITH APPLICATIONS TO LEARNING[1]

ROBERT R. BUSH AND FREDERICK MOSTELLER

Harvard University

Summary. A stochastic model designed for analyzing data with changing probabilities is presented. On each of a series of trials one of two *alternatives* occurs and the probabilities of occurrence are changed from time to time by *events*. Corresponding to each class of events is an operator which represents a linear transformation on the probabilities of the two alternatives. Cases of fixed event probabilities and of changing event probabilities are considered. Recurrence formulas for moments of the resulting distributions of probabilities are provided. These formulas are often tedious to apply, but for the first and second moments several bounds are provided; these bounds are relatively easy to compute.

The problem of estimating the parameters of the model is discussed. No general solution is obtained but simplifying assumptions lead to interesting special cases for which detailed procedures of parameter estimation are presented. One such special case arises when there are two event operators which commute, implying that the operators have equal limit points or that one operator is the identity operator. The method of maximum likelihood is applied to this case. Another special case, which arises when the slope parameters of the two operators are equal, is discussed in Section 8.

Applications of the model and estimation procedures to certain kinds of data on animal and human learning are described. The examples given are experiments on verbal learning, the avoidance training of dogs, the reward training of rats in a simple T-maze, and the behavior of human subjects in a two-choice situation.

1. Basic concepts and definitions. During the past three years, the authors have been developing a mathematical model for describing a variety of experiments on animal and human learning [2], [3]. This model is closely related to the one developed by Estes [5] and to the more recent work of Miller and McGill [9]. These models have led quite naturally into a study of a class of stochastic processes, which may be viewed as Markov chains with an infinite number of states. In applying the model to the analysis of experimental data, a number of problems in statistical estimation have arisen. In this paper, therefore, we present a summary of the main mathematical results obtained and a discussion of some estimation procedures we have found useful.

Received 6/13/52, revised 6/29/53.

[1] Part of this paper was presented at the meeting of the Institute of Mathematical Statistics, December 27, 1951. The research was facilitated by the authors' membership in the Inter-University Summer Seminar of the Social Science Research Council titled Mathematical Models for Behavior Theory, held at Tufts College, June 28 to August 24, 1951. Except for the period of that Seminar, this research was facilitated by support from the Laboratory of Social Relations, Harvard University.

A learning process, as the term is used here, involves systematic changes in behavior; one type of behavior may become more frequent and another type of behavior may become less so. We shall describe this learning process in a situation where a choice of a number of given alternatives occurs periodically. Each occasion on which there is an opportunity for making a choice will be called a *trial*. Typically, one observes that a particular alternative occurs more and more frequently—this we call learning—until the system stabilizes so that no more average changes in behavior occur—this we call the completion of learning. In later sections we discuss applications of our model to problems in learning, but we will describe the basic structure of the model in somewhat more general terms.

We consider a set of mutuallyl excusive and exhaustive alternatives, A_1, $A_2, \cdots, A_r$. On each trial one and only one of these alternatives will occur. On each *trial* we define a set of r probabilities, $p_1, p_2, \cdots, p_r$, corresponding to the r alternatives. The probability p_i is then the probability that the ith alternative will occur on the trial in question. We assume that all the available information about what alternative occurs on that trial is given by the set of r probabilities. The alternatives which occur on trials previous to trial n, for example, do not influence the outcome of trial n except insofar as they may have determined the probabilities for that trial. On each trial the probabilities must sum to unity since we have taken the r alternatives to be mutually exclusive and exhaustive.

The set of probabilities p_i are altered from time to time by certain *events*, $E_1, E_2, \cdots, E_t$. Corresponding to each event there is a mathematical operator T_j $(j = 1, 2, \cdots, t)$ which operates on the set of r probabilities whenever event E_j occurs. We next give particular representations of these operators.

2. The event operators. It is explicitly assumed that the operators which correspond to the t events are linear. Thus we may represent the set of r probabilities by a column vector and each operator T_j by an $r \times r$ matrix. In the remainder of this paper we will discuss only the case of two alternatives A_1 and A_2 and so we need only a single probability variable, p, the probability associated with A_1, because the probabilities of the two alternatives always sum to unity. Thus we may dispense with the matrix machinery and write for the upper element of the transformed probability vector,

$$Q_j p = a_j + \alpha_j p, \qquad j = 1, 2, \cdots, t \tag{1}$$

where the a_j and α_j are parameters which are restricted only by the requirement that the probabilities must always be in the closed interval from zero to unity. This means that $0 \leqq a_j \leqq 1$ and $-a_j \leqq \alpha_j \leqq 1 - a_j$. The operators Q_j are not (homogeneous) linear operators, though derived from linear matrix operators.

These operators may be applied to an operand p more than once. When Q_j is applied twice we obtain

$$Q_j^2 p = a_j + \alpha_j(Q_j p) = a_j(1 + \alpha_j) + \alpha_j^2 p. \tag{2}$$

When Q_j is applied to p a total of n times, it may easily be shown that

$$Q_j^n p = \lambda_j - (\lambda_j - p)\alpha_j^n, \tag{3}$$

where

$$\lambda_j = a_j/(1 - \alpha_j). \tag{4}$$

When the magnitude of α_j is less than unity, the term in α_j^n of equation (3) approaches zero as n gets large, and so in this case, λ_j is the asymptotic value of the operation $Q_j^n p$.

The t operators Q_j may be applied to p in various orders, corresponding to the orders of occurrence of the t events. If we know in advance the particular sequence of events, it is a simple matter to compute the successive values of probability. Our main interest, however, is in problems in which the precise sequence of events is unknown. For any sequence of the t events and for any initial value of probability, it may be shown that the probability will ultimately lie between two limits; this we have called the trapping theorem. Corresponding to the t operators will be a set of t limits, λ_j, given by equations (4). The trapping theorem states that the asymptotic value of probability from any sequence will lie in the interval including min (λ_j) to max (λ_j), as the least and largest cluster points, respectively, provided only that $0 \leqq \alpha_j < 1$ for all j. Our proof of this theorem is elementary but rather lengthy and will not be given here. The point is that if the starting value of p is between the limits, the sequence of p's will forever remain there. If the starting value of p is outside the limits, the sequence will ultimately be trapped inside the interval or else tend to one of the limits monotonically.

3. Fixed event probabilities. In this section we will describe the process when we do not know the precise sequence of events but know only the probabilities π_j that when an event occurs it will be event E_j. The set of t event probabilities π_j are constant and sum to unity. After n event occurrences there will be at most t^n possible sequences and hence at most t^n possible values of p. The probabilities of occurrences of these sequences will depend upon the π_j, of course. We are interested in the properties of the distribution of values of p for all n. We may order the t^n possible values of p and label them $h = 1, 2, \cdots, t^n$. We denote the hth value after n events by $p_{h,n}$ and the probability that this value will obtain by $P_{h,n}$. The mean of the distribution after n events will be denoted by $V_{1,n}$ and it is defined by

$$V_{1,n} = \sum_{h=1}^{t^n} p_{h,n} P_{h,n}, \qquad (n = 0, 1, \cdots). \tag{5}$$

Now after the $(n + 1)$st event each of the t^n values of p will split into t new values of p. In particular, the hth value will split into t new values $Q_j p_{h,n}$, with corresponding weights $\pi_j P_{h,n}$. Thus, the mean will be given by

$$V_{1,n+1} = \sum_{h=1}^{t^n} \sum_{j=1}^{t} \pi_j P_{h,n}\, Q_j\, p_{h,n}. \tag{6}$$

Using equations (1) and (5) and the fact that the sum of the weights $P_{h,n}$ over all h is unity, it follows that

$$V_{1,n+1} = \bar{a} + \bar{\alpha} V_{1,n}, \tag{7}$$

where the average or actuarial values of the parameters are given by

$$\bar{a} = \sum_{j=1}^{t} \pi_j a_j \tag{8}$$

and

$$\bar{\alpha} = \sum_{j=1}^{t} \pi_j \alpha_j. \tag{9}$$

Equation (7) is a well known linear difference equation and has the solution

$$V_{1,n} = \frac{\bar{a}}{1 - \bar{\alpha}} - \left[\frac{\bar{a}}{1 - \bar{\alpha}} - V_{1,0}\right] \bar{\alpha}^n. \tag{10}$$

The correspondence between equation (10) and equation (3) is clear if $\bar{a}/(1 - \bar{\alpha})$ is regarded as $\bar{\lambda}$. (The expected operator, discussed in the next section, will yield the correct value of $V_{1,n+1}$ from $V_{1,n}$.) Higher raw moments of the distribution may be obtained in a similar way. The kth raw moment after n events is

$$V_{k,n} = \sum_{h=1}^{t^n} P_{h,n} (p_{h,n})^k. \tag{11}$$

After the $(n + 1)$st event the kth raw moment is

$$V_{k,n+1} = \sum_{h=1}^{t^n} \sum_{j=1}^{t} \pi_j P_{h,n} (Q_j p_{h,n})^k. \tag{12}$$

After inserting the expressions for $Q_j p_{h,n}$ from equations (1) into expression (12) and expanding the resulting expressions by the binomial theorem, we obtain

$$V_{k,n+1} = \sum_{j=1}^{t} \pi_j \sum_{i=0}^{k} \binom{k}{i} a_j^{k-i} \alpha_j^i V_{i,n}, \tag{13}$$

where the $\binom{k}{i}$ are binomial coefficients, and where $V_{0,n} = 1$ for all n. Note that the kth raw moment after $n + 1$ events is given in terms of all the raw moments up through the kth after n events. With the expressions (13) we may compute as many moments of the distributions as we choose.

One may also compute the moments of the distribution of $p_{h,n}$ from a moment generating function $\phi_n(\theta)$ defined by

$$\phi_n(\theta) = \sum_{h=1}^{t^n} e^{\theta p_{h,n}} P_{h,n}. \tag{14}$$

It is easy to show that

$$\phi_{n+1}(\theta) = \sum_{j=1}^{t} \pi_j e^{\theta a_j} \phi_n(\theta \alpha_j). \tag{15}$$

4. Identification of alternatives and events. In the preceding sections we have intentionally avoided imposing any direct connection between the events and the occurrences of alternatives from trial to trial. In many problems of practical interest, the events which alter the probabilities immediately follow occurrences of an alternative in a predetermined manner. Hence, in this section we will identify the events with the alternatives; the event probabilities are then equal to the values of p and so are no longer constant. Since we are considering only the case of two alternatives, we will have but two events.

After n event occurrences, that is, after n trials, there will be 2^n possible values of p. The mean of these is still defined by equation (5), but the mean on the next trial is

$$V_{1,n+1} = \sum_{h=1}^{2^n} \{p_{h,n} P_{h,n}(Q_1 p_{h,n}) + (1 - p_{h,n}) P_{h,n}(Q_2 p_{h,n})\}. \tag{16}$$

Using the operations defined by equations (1) and the definitions of the raw moments, equations (5) and (11), we have after simplifications

$$V_{1,n+1} = a_2 + (a_1 - a_2 + \alpha_2) V_{1,n} + (\alpha_1 - \alpha_2) V_{2,n} . \tag{17}$$

We observe at once that the mean on the $(n + 1)$st trial depends upon the second raw moment, $V_{2,n}$, for the preceding trial. Analogous to equation (13) we have for the recurrence formula for the higher moments

$$V_{k,n+1} = \sum_{i=0}^{k} \binom{k}{i} \{(a_1^{k-i} \alpha_1^i - a_2^{k-i} \alpha_2^i) V_{i+1,n} + a_2^{k-i} \alpha_2^i V_{i,n}\}. \tag{18}$$

It is straightforward to write the recursion relation for the moment generating function $\phi_n(\theta)$ of p_n . The derivation is equivalent to that for $V_{k,n+1}$. We get

$$\phi_{n+1}(\theta) = e^{\theta a_2} \phi_n(\theta\alpha_2) + \frac{e^{\theta a_1}}{\alpha_1} \phi_n'(\theta\alpha_1) - \frac{e^{\theta a_2}}{\alpha_2} \phi_n'(\theta\alpha_2), \tag{19}$$

where the primes on the ϕ's refer to derivatives with respect to θ. Thus far we have found this relation more tantalizing than useful.

We see that the kth moment on the $(n + 1)$st trial depends on the $(k + 1)$st moment on the nth trial. This fact makes computations exceedingly difficult. As a result we have developed some approximations and bounds which are much easier to compute.

The first approximation is an obvious one. We have called it the *expected operator approximation*. An expected operator, $\bar{Q}$, is defined by

$$V_{1,n+1} \cong \bar{Q}\, V_{1,n} = V_{1,n} Q_1 V_{1,n} + (1 - V_{1,n}) Q_2 V_{1,n} . \tag{20}$$

Using the definitions of equations (1) we would obtain the approximation

$$V_{1,n+1} \cong a_2 + (a_1 - a_2 + \alpha_2) V_{1,n} + (\alpha_1 - \alpha_2) V_{1,n}^2 . \tag{21}$$

If we compare this approximate result with equation (17) we see that $V_{1,n}^2$ has replaced $V_{2,n}$ in the exact equation. This means that the expected operator approximation behaves as if the variance of the distribution were zero. This behavior is clearly wrong since we know that the density is not concentrated at a point except possibly on the zero-th trial. However, as will be shown below, the expected operator will lead to a bound on the mean, $V_{1,n}$.

The first set of bounds on the mean which we present are obtained from bounds on the second raw moment, $V_{2,n}$. The lower bound on the second raw moment follows from the fact that the variance is never negative:

$$V_{2,n} = V_{1,n}^2 + \sigma_n^2 \geqq V_{1,n}^2. \tag{22}$$

The upper limit on $V_{2,n}$ is a bit more trouble to obtain. Consider first a distribution $g(z)$ on the interval $0 \leqq z \leqq 1$, having mean $U_{1,n}$ and second raw moment $U_{2,n}$. We know at once that

$$U_{2,n} \leqq U_{1,n}. \tag{23}$$

We now transform this distribution to the interval $\mu_2 \leqq x \leqq \mu_1$, by letting

$$z = \frac{x - \mu_2}{\mu_1 - \mu_2}, \tag{24}$$

and find that

$$V_{2,n} \leqq (\mu_1 + \mu_2)V_{1,n} - \mu_1\mu_2. \tag{25}$$

We may take $\mu_1 = \lambda_1$ and $\mu_2 = \lambda_2$, provided that $\lambda_2 < \lambda_1$, and obtain the desired upper bound on $V_{2,n}$. These bounds on the second moment, inequalities (22) and (25), may now be used in our recurrence relation (17) to obtain upper and lower bounds on the means. We shall carry this out only for the asymptotic distribution, for which we let $V_{1,n} = V_{1,n+1} = V_1$ and $V_{2,n} = V_{2,n+1} = V_2$. (Harris has demonstrated [8] that the distribution of p_n approaches a limiting distribution independent of p_0 as $n \to \infty$ when $0 < \lambda_1, \lambda_2 < 1$ and $0 < \alpha_1, \alpha_2 < 1$.) If we then introduce the abbreviations

$$A = a_1 - a_2 + \alpha_2 - 1, \qquad B = \alpha_1 - \alpha_2, \tag{26}$$

equation (17) can be written as:

$$AV_1 + BV_2 + a_2 = 0. \tag{27}$$

We next insert the lower bound on V_2 from (22) into equation (27) and obtain for $B > 0$,

$$BV_1^2 + AV_1 + a_2 \leqq 0. \tag{28}$$

When $B < 0$, the direction of the inequality in (28) is reversed. We denote the quadratic expression in (28) by $q(V_1)$ and note that $q(0) = a_2$ which is positive. Further, from definitions (26) we see that $q(1) = \alpha_1 + a_1 - 1$, which from restrictions on the parameters discussed in Section 2 is a negative quantity. Thus, $q(V_1)$ is positive at zero and negative at unity and so has but one root between zero and unity. If this root is called Y, we have $V_1 \geqq Y$ for $B > 0$ and $V_1 \leqq Y$

for $B < 0$. The bound Y is obtained from the expected operator approximation discussed above.

The upper bound on V_2, given by inequality (25), with $\mu_1 = \lambda_1$ and $\mu_2 = \lambda_2$, may also be inserted in equation (27). For $B > 0$ we have

$$[A + (\lambda_1 + \lambda_2)B]V_1 \geqq B\lambda_1\lambda_2 - a_2 . \tag{29}$$

From the definitions (26) it may be shown that the coefficient of V_1 on the left side of (29) is always negative or zero and so we have

$$V_1 \leqq \frac{B\lambda_1\lambda_2 - a_2}{A + (\lambda_1 + \lambda_2)B}. \tag{30}$$

When $B < 0$, the directions of the inequalities in (29) and (30) are reversed.

An improved pair of bounds on the means $V_{1,n}$ may be obtained from upper and lower bounds on the third raw moments $V_{3,n}$. We consider again a distribution $g(z)$ on the interval $0 \leqq z \leqq 1$. Using the Schwarz inequality

$$\left(\int \eta^2\, dz\right) \cdot \left(\int \xi^2\, dz\right) \geqq \left(\int \eta\xi\, dz\right)^2, \tag{31}$$

and letting $\eta^2 = zg(z)$ and $\xi^2 = z^3g(z)$, one can readily show that

$$U_{3,n} \geqq U_{2,n}^2/U_{1,n} , \tag{32}$$

where $U_{k,n}$ is the kth raw moment of the distribution $g(z)$ on the nth trial. After transforming this distribution to the interval $\lambda_2 \leqq x \leqq \lambda_1$ by equation (24) we obtain

$$V_{3,n} \geqq \lambda_2 V_{2,n} + \frac{(V_{2,n} - \lambda_2 V_{1,n})^2}{V_{1,n} - \lambda_2}. \tag{33}$$

The upper bound on $V_{3,n}$ may be found in the same manner. We let $\eta^2 = (1 - z)\, g\, (z)$ and $\xi^2 = (1 - z)^3\, g\, (z)$ in the Schwarz inequality and then transform to the interval $\lambda_2 \leqq x \leqq \lambda_1$ to obtain

$$V_{3,n} \leqq \lambda_1 V_{2,n} - \frac{(V_{2,n} - \lambda_1 V_{1,n})^2}{\lambda_1 - V_{1,n}}. \tag{34}$$

Inequalities (33) and (34) may now be used in the recurrence relations (17) and (18) to obtain bounds on $V_{1,n}$. Equation (18) for $k = 2$ is

$$\begin{aligned} V_{2,n+1} = a_2^2 &+ (2\alpha_2 a_2 + a_1^2 - a_2^2)V_{1,n} \\ &+ (\alpha_2^2 + 2a_1\alpha_1 - 2a_2\alpha_2)V_{2,n} + (\alpha_1^2 - \alpha_2^2)V_{3,n} . \end{aligned} \tag{35}$$

The second moment $V_{2,n}$ may be eliminated from this expression and equation (17). Further, equation (17) may be used to eliminate the second moment from inequalities (33) and (34). In this way one can write down a recurrence formula in the means alone for the two desired bounds. The reader will be spared the sight of the final result.

Numerical computations of the two sets of bounds on $V_{1,n}$ discussed above

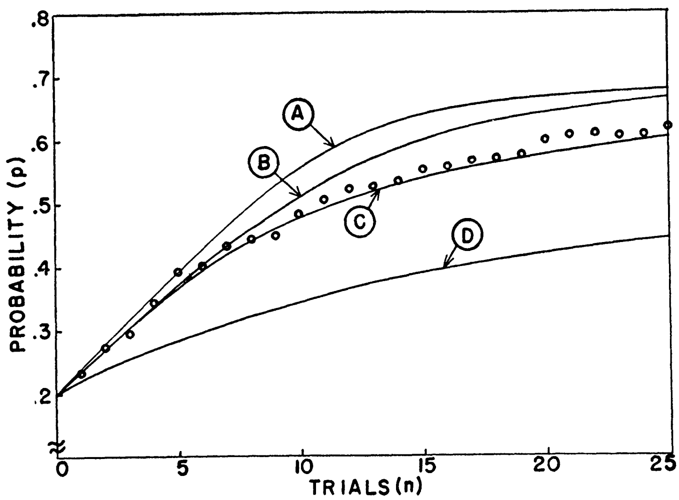

FIG. 1

Bounds on the probability distribution means, $V_{1,n}$, versus trials with the parameter values, $a_1 = 0.3, \alpha_1 = 0.6, a_2 = 0.01, \alpha_2 = 0.9$. Curve A is the expected operator bound which is in this case an upper bound. Curves B and C are the upper and lower bounds, respectively, obtained from maximizing and minimizing the third moment about the mean. Curve D is the lower bound obtained by maximizing the variance. The small circles represent the mean probabilities of 84 Monte Carlo runs made with the above parameter values.

TABLE I

Bounds on the asymptotic mean, $V_{1,\infty}$, for seven numerical examples. The limits λ_1 and λ_2 are the asymptotes obtained by applying one operator only. The bounds on $V_{1,\infty}$ were obtained by maximizing and minimizing the second raw moment, V_2, and the third raw moment, V_3

Parameter Values				Limits		Bounds on $V_{1,\infty}$	
a_1	a_2	α_1	α_2	λ_2	λ_1	From V_2	From V_3
.300	.010	.6	.9	.10	.75	.500–.682	.655–.676
.300	.001	.6	.9	.01	.75	.112–.668	.418–.658
.300	.0001	.6	.9	.001	.75	.013–.667	.093–.654
.396	.001	.6	.9	.01	.99	.394–.987	.967–.986
.396	.003	.6	.7	.01	.99	.184–.961	.723–.878
.360	.03	.6	.7	.10	.90	.557–.718	.637–.657
.300	0	.6	.9	0	.75	0–.667	0–.656

have been carried out for 25 trials with assumed values of the parameters. In Fig. 1 we show the results. Also in Table I we show the results of several such computations of bounds on the asymptotic mean.

In addition to the computations of bounds, we have used the Monte Carlo method of making approximate calculations. This involves using a random number table for making decisions about which operator, Q_1 or Q_2 to apply to the probability. The means of 84 such runs for 25 trials are also shown in Fig. 1.

5. The estimation problem. Most mathematical models involve one or more unknown parameters; these parameters may be related to experimental variables but data from experiments must provide information about the values of the parameters. The model described in Section 4 has a total of five parameters when two alternatives and two events are considered: an initial probability, p_0, and the parameters a_1, α_1, a_2, and α_2 contained in the definitions of the operators Q_1 and Q_2. When the model is applied to a particular experimental problem, one must estimate these five parameters from the data. In this and following sections we will discuss some of these estimation problems. We shall restrict our attention to two alternatives and two events, and to the case when the events are identified with the alternatives as in Section 4.

There is a crucial question as to how many parameters the model can tolerate in the face of particular kinds of data. It appears to us that five parameters are too many for the kinds of data we have been studying. The obvious approach is to avoid using the model in full generality but to make special assumptions for specific applications, that is, to let some of the parameters be zero or unity or to set up relations between certain parameters from considerations such as symmetry. In making such special assumptions in our applications to learning experiments, we have been guided by current psychological theories of learning such as reinforcement theory and association theory. However, we are interested in workable methods for estimating parameters in the general case. Our experience to date has led us to believe that the estimation problem is a very untidy one if there are more than two parameters involved. Therefore, in the sections which tollow we will discuss only special cases where three of the five parameters are feliminated or are assumed to be known.

6. Estimation procedures when the operators commute. The estimation problem is much simplified when the two operators, Q_1 and Q_2, commute. From equation (1), it is easily shown that Q_1 and Q_2 commute if and only if

$$a_1(1 - \alpha_2) = a_2(1 - \alpha_1). \tag{36}$$

This condition is fulfilled when one or both of the two operators is the identity operator, that is, if $a_1 = 0$ and $\alpha_1 = 1$, or if $a_2 = 0$ and $\alpha_2 = 1$, or both. Otherwise the operators commute only if the trapping theorem limits, λ_1 and λ_2, defined by equation (4) are equal to one another. By setting $\lambda_1 = a_1/(1 - \alpha_1) = \lambda_2 = a_2/(1 - \alpha_2) = \lambda$, the two operators become

$$\begin{aligned} Q_1p &= \lambda(1 - \alpha_1) + \alpha_1 p, \\ Q_2p &= \lambda(1 - \alpha_2) + \alpha_2 p. \end{aligned} \tag{37}$$

The cases for which Q_1 or Q_2 is the identity operator, or both, may be considered as special cases of equations (37), since one may set $\alpha_1 = 1$ or $\alpha_2 = 1$ or both. Hence, the most general operations for which $Q_1Q_2 = Q_2Q_1$ are described by equations (37).

We now make one further restriction and then develop a scheme for estimating the remaining parameters. We take $\lambda = 1$ to obtain

$$\begin{aligned} Q_1p &= 1 - \alpha_1 + \alpha_1 p, \\ Q_2p &= 1 - \alpha_2 + \alpha_2 p. \end{aligned} \tag{38}$$

We wish to estimate α_1, α_2, and p_0, the initial value of p, from actual data. Now on trial n there will have been some number k of previous occurrences of A_1 and so the probability $q_{n,k}$ of an A_2 occurrence on trial n is

$$q_{n,k} = 1 - Q_1^k Q_2^{n-k} p_0 . \tag{39}$$

From equations (38) we may easily show that

$$q_{n,k} = \alpha_1^k \alpha_2^{n-k} q_0 , \tag{40}$$

where $q_0 = 1 - p_0$ is the initial probability of an A_2 occurrence. In applications to learning data, two further sets of restrictions are of interest. First, if $\alpha_2 = 1$, we have

$$q_{n,k} = \alpha_1^k q_0 , \qquad (\alpha_2 = 1). \tag{41}$$

Second, if we consider only the data for which $k = 0$, we have

$$q_{n,0} = \alpha_2^n q_0 . \tag{42}$$

Both sets of restrictions yield equations of the form

$$q_\nu = \alpha^\nu q_0 , \tag{43}$$

and so we will be interested in the estimation problems which arise from equations of this type. In general terms, q_ν is the probability of an A_2 occurrence for the specified value of ν.

The data will be represented by random variables $x_{i\nu}$ which are considered to be binomial observations (0 or 1). The index i is used to indicate the ith observation for the specified value of ν. If A_1 occurs on the ith observation for the given value of ν, then $x_{i\nu} = 1$, while if A_2 occurs then $x_{i\nu} = 0$. The total number of binomial observations available in the data for a particular value of ν may not be the same for all values of ν and so we denote that number by N_ν. We further define x_ν to be the number of times A_1 occurs during the N_ν observations, that is,

$$x_\nu = \sum_{i=1}^{N_\nu} x_{i\nu} . \tag{44}$$

The expected value of $x_{i\nu}$ is of course $1 - q_\nu$, that is, q_ν is the probability that $x_{i\nu} = 0$.

The discussion above might be compared with a special bioassay model in which doses are given at levels $D_0, D_1, \cdots, D_k$ (perhaps equally spaced in the logarithm), with a model for proportion living at dose D_ν to be $q_0\alpha^\nu$. We want to estimate q_0 and α. In the bioassay case the number of animals at a dose is N_ν, usually fixed in advance unless some sequential approach is used. In the stochastic model presented here, the N_ν are random variables.

6a. Maximum likelihood estimates of α and p_0. In this section we obtain some maximum likelihood estimates of the parameters α and p_0. We have not investigated the general question of the efficiency of maximum likelihood procedures when applied to stochastic processes. Though such investigations are beyond the scope of this paper, they clearly need to be made. We will proceed in the standard manner, setting aside these more general considerations.

We wish to write down an expression for the likelihood of obtaining an observed set of data. First, the likelihood P_ν of obtaining x_ν occurrences of A_1 and $N_\nu - x_\nu$ occurrences of A_2 in a given order for a particular value of ν is

$$P_\nu = (1 - q_\nu)^{x_\nu}(q_\nu)^{N_\nu - x_\nu}. \tag{45}$$

The likelihood, P, of obtaining the entire set of data, for $\nu = 0, 1, \cdots, \Omega$, is then

$$P = \prod_{\nu=0}^{\Omega} P_\nu = \prod_{\nu=0}^{\Omega} (1 - q_\nu)^{x_\nu}(q_\nu)^{N_\nu - x_\nu}. \tag{46}$$

We insert in this expression the value of q_ν from equation (43) and take the logarithm to obtain

$$\log P = \sum_{\nu=0}^{\Omega} \{x_\nu \log (1 - \alpha^\nu q_0) + (N_\nu - x_\nu) \log (\alpha^\nu q_0)\}. \tag{47}$$

We wish to obtain the simultaneous maximum likelihood estimates of α and q_0 and so we maximize $\log P$ with respect to those two parameters. Setting equal to zero the partial derivatives of $\log P$ with respect to α and q_0 leads to the equations

$$\sum_{\nu=0}^{\Omega} (N_\nu - x_\nu) = \sum_{\nu=0}^{\Omega} \frac{\hat{\alpha}^\nu \hat{q}_0}{1 - \hat{\alpha}^\nu \hat{q}_0} x_\nu, \tag{48}$$

$$\sum_{\nu=0}^{\Omega} \nu(N_\nu - x_\nu) = \sum_{\nu=0}^{\Omega} \nu \frac{\hat{\alpha}^\nu \hat{q}_0}{1 - \hat{\alpha}^\nu \hat{q}_0} x_\nu, \tag{49}$$

where $\hat{\alpha}$ and $\hat{q}_0$ are the maximum likelihood estimates of α and q_0, respectively. These equations, then, must be solved for $\hat{\alpha}$ and $\hat{q}_0$, but only numerical methods are available in general. However, in certain applications to learning data we have found convenient short cuts. We will discuss two of these.

In some applications it is possible to choose Ω such that x_ν for $\nu = 0, 1, \cdots, \Omega$ is some constant R independent of ν. The factor x_ν may then be taken out of the sums on the right sides of equations (48) and (49), leaving the functions

$$F(\hat{\alpha}, \hat{q}_0, \Omega) = \sum_{\nu=0}^{\Omega} \frac{\hat{\alpha}^\nu \hat{q}_0}{1 - \hat{\alpha}^\nu \hat{q}_0}, \tag{50}$$

and

$$G(\hat{\alpha}, \hat{q}_0, \Omega) = \sum_{\nu=0}^{\Omega} \nu \frac{\hat{\alpha}^{\nu}\hat{q}_0}{1 - \hat{\alpha}^{\nu}\hat{q}_0}. \tag{51}$$

We have tabulated these functions for ranges of $\bar{\alpha}$, $\hat{q}_0$, and Ω, but these tables are too lengthy to be included here. The values of the functions F and G are computed from the data with the formulas

$$F(\hat{\alpha}, \hat{q}_0, \Omega) = \frac{1}{R} \sum_{\nu=0}^{\Omega} (N_\nu - R), \tag{52}$$

$$G(\hat{\alpha}, \hat{q}_0, \Omega) = \frac{1}{R} \sum_{\nu=0}^{\Omega} \nu(N_\nu - R), \tag{53}$$

and then the tables are used to obtain the corresponding values of $\hat{\alpha}$ and $\hat{q}_0$.

The second short cut we have found useful was developed for data in which q_0 is known to be unity, or for which one is willing to assume $q_0 = 1$. Equation (49) with $\hat{q}_0 = 1$ may then be solved for $\hat{\alpha}$, that is, we have

$$\sum_{\nu=0}^{\Omega} \nu(N_\iota - x_\nu) = \sum_{\nu=0}^{\Omega} \nu \frac{\hat{\alpha}^{\nu}}{1 - \hat{\alpha}^{\nu}} x_\nu. \tag{54}$$

The left side of this equation can be computed directly from the data but the right side depends on both the x_ν from the data and the estimate $\hat{\alpha}$. However, we have computed tables of the quantities $\nu\hat{\alpha}^{\nu}/(1 - \hat{\alpha}^{\nu})$ versus ν for fifty values of $\hat{\alpha} = .50[.01].99$, thereby greatly facilitating computations of the sum on the right side of equation (54) for a given set of x_ν and a range of values of $\hat{\alpha}$. This procedure is especially workable when we have a good preliminary value of the correct $\hat{\alpha}$ and hence know the appropriate range of values to use. For most data we have studied, $\hat{\alpha}$ is near unity and so we can expand $\hat{\alpha}^{\nu}$ in a power series about unity and retain only the linear term, namely, $\hat{\alpha}^{\nu} \cong 1 - \nu(1 - \hat{\alpha})$. Using this approximation in equation (54) and simplifying yields the simple formula

$$\hat{\alpha} \cong 1 - \frac{\sum_{\nu=0}^{\Omega} x_\nu}{\sum_{\nu=0}^{\Omega} \nu N_\nu}, \tag{55}$$

which may be used to estimate α directly from the data without the use of tables, or at least to obtain a preliminary value of $\hat{\alpha}$.

The asymptotic variances and covariances of the estimates $\hat{\alpha}$ and $\hat{q}_0$ may be obtained by analogy with the procedure used in simpler problems. We illustrate only for the case when q_0 is known and only α is being estimated. We take the second partial derivative with respect to α of $\log P$ of equation (47) and obtain

$$\frac{\partial^2 \log P}{\partial \alpha^2} = -\sum_{\nu=0}^{\Omega} \frac{\nu}{\alpha^2} \left\{ (N_\nu - x_\nu) + \frac{(\nu - 1)q_0\alpha^{\nu} + q_0^2\alpha^{2\nu}}{(1 - \alpha^{\nu}q_0)^2} x_\nu \right\}. \tag{56}$$

We need to take the expected value of this second derivative. From the definition of x_ν, equation (44), we see that $E(x_\nu) = (1 - q_\nu)E(N_\nu) = (1 - \alpha^\nu q_0)E(N_\nu)$. Thus, from (56) we have

$$-E\left(\frac{\partial^2 \log P}{\partial \alpha^2}\right) = \sum_{\nu=0}^{\Omega} \frac{\nu^2 \alpha^{\nu-2} q_0}{1 - \alpha^\nu q_0} E(N_\nu). \tag{57}$$

The asymptotic variance is the reciprocal of this quantity, but in order to compute it we need to evaluate the expected value of N_ν and this cannot be done until the problem is more completely specified. In the general formulation of the estimation problem which followed equation (43), we merely defined N_ν as the number of binomial observations available for a particular value of ν, but we left the distribution of N_ν unspecified. However, for the case of one identity operator which led to equation (41), the index ν corresponds to k, the number of previous occurrences of A_1, and so $N_\nu - 1$ is the number of A_2 occurrences between the kth and $(k + 1)$st A_1 occurrences. Hence N_ν has a negative binomial distribution and expected value

$$E(N_\nu) = \frac{R}{1 - q_\nu} = \frac{R}{1 - \alpha^\nu q_0}, \tag{58}$$

where R is the number of independent sequences in the data. When this expression is used in equation (57) the asymptotic variance may be estimated. For the case of commuting operators and data for which $k = 0$, equation (42) is appropriate and ν corresponds to trial number n. If there are R independent sequences, then N_n is the number of those sequences on trial n for which $k = 0$. It is readily shown that for this case

$$E(N_n) = R\alpha_2^{n(n-1)/2} q_0^n . \tag{59}$$

Again, the asymptotic variance may be estimated when this expression is used in equation (57). For some cases it may be difficult to evaluate $E(N_\nu)$. We presume that little violence will be done to the estimate of the variance by replacing $E(N_\nu)$ with the observed N_ν, providing the N_ν are not too small.

6b. The value of ν when A_1 first occurs. It is instructive to consider a quantity h, defined to be the value of ν when alternative A_1 first occurs. The probability of an A_1 occurrence is $1 - q_\nu$ and q_ν is given by equation (43)

$$q_\nu = \alpha^\nu q_0 . \tag{43}$$

The density function for h is

$$\begin{aligned} f(h) &= q_0(\alpha q_0)(\alpha^2 q_0)(\alpha^3 q_0) \cdots (\alpha^{h-1} q_0)(1 - \alpha^h q_0) \\ &= \alpha^{h(h-1)/2} q_0^h (1 - \alpha^h q_0), \qquad h = 1, 2, \cdots . \end{aligned} \tag{60}$$

The latter form of writing equation (60) is also correct for $h = 0$. In words, $f(h)$ is the probability that A_1 will first occur when $\nu = h$. This distribution might be regarded as a more complicated version of the negative binomial, the complication being that the probabilities depend on a variable ν as expressed by equation (43). The expected value of h is

$$E(h) = \sum_{h=0}^{\infty} hf(h) = \sum_{h=0}^{\infty} h\alpha^{h(h-1)/2} q_0^h (1 - \alpha^h q_0) \tag{61}$$

$$= \sum_{h=0}^{\infty} h\alpha^{h(h-1)/2} q_0^h - \sum_{h=0}^{\infty} h\alpha^{h(h+1)/2} q_0^{h+1}$$

This result may be simplified if we let $y = h + 1$ in the last sum, that is,

$$E(h) = \sum_{h=0}^{\infty} h\alpha^{h(h-1)/2} q_0^h - \sum_{y=1}^{\infty} y\alpha^{y(y-1)/2} q_0^y + \sum_{y=1}^{\infty} \alpha^{y(y-1)/2} q_0^y. \tag{62}$$

The first two sums on the right side of equation (62) cancel and so we have

$$E(h) = \sum_{y=1}^{\infty} \alpha^{y(y-1)/2} q_0^y. \tag{63}$$

For known values of α and q_0 the expected value, $E(h)$, may be computed. If the maximum likelihood estimates $\hat{\alpha}$ and $\hat{q}_0$ are used in equation (63) for α and q_0, respectively, an estimate of $E(h)$ is obtained. From certain kinds of data, a set of values of h will be observed and the mean of these sample values can be compared with the estimated $E(h)$.

The variance of h may be obtained in a similar way. The result is

$$\sigma^2(h) = 2 \sum_{y=1}^{\infty} y\alpha^{y(y-1)/2} q_0^y - E(h) - [E(h)]^2. \tag{64}$$

By replacing α and q_0 with their maximum likelihood estimates, $\sigma^2(h)$ may be estimated from equation (64) and the result compared to the variance of an observed set of h's. Conversely, a table of values of $E(h)$ and $\sigma^2(h)$ for $0 < \alpha < 1$ and $0 < q_0 \leqq 1$ may be constructed and α and q_0 estimated from the observed set of values of h (method of moments).

When $q_0 = 1$, equation (63) reduces to

$$E(h) = \sum_{y=1}^{\infty} \alpha^{\frac{1}{2}y(y-1)}. \tag{65}$$

This series has to do with theta functions. According to Whittaker and Watson [12] in their chapter on theta functions, this series was discussed by Jakob Bernoulli, *Ars Conjectandi* (1713), p. 55. Bromwich [1] lists in his table of contents "theta-series" (p. xii) and discusses our series and some related infinite products in some examples (pp. 101, 116, 117).

6c. An unbiased estimate of α_1. We next consider the problem of estimating α_1 of equation (40) when α_2 and q_0 are known. We will utilize only that portion of the data for which $k = 1$ and so we have probabilities

$$q_n = \alpha_1 \alpha_2^{n-1} q_0, \tag{66}$$

and observations x_{in} for $n = 1, 2, \cdots$, and $i = 1, 2, \cdots, N_{n,1}$. We let

$$x_{n,1} = \sum_{i=1}^{N_{n,1}} x_{in}. \tag{67}$$

Thus, an unbiased estimate of q_n is $1 - x_{n,1}/N_{n,1}$, and for each value of n we obtain an unbiased estimate $\bar{\alpha}_{1,n}$ of α_1:

$$\bar{\alpha}_{1,n} = \frac{1 - \dfrac{x_{n,1}}{N_{n,1}}}{\alpha_2^{n-1} q_0}. \tag{68}$$

We next wish to combine these estimates by taking a weighted mean, $\bar{\alpha}_1$, over those values of n for which $N_{n,1}$ is not zero, that is, we let

$$\bar{\alpha}_1 = \frac{\sum_n W_n \bar{\alpha}_{1,n}}{\sum_n W_n}. \tag{69}$$

We choose the weight W_n to be inversely proportional to the variance of the estimate $\bar{\alpha}_{1,n}$. (If the estimates $\bar{\alpha}_{1,n}$ are independent, this procedure minimizes the variance of $\bar{\alpha}_1$.) The unbiased estimate, $1 - x_{n,1}/N_{n,1}$ of q_n has variance

$$\frac{q_n(1 - q_n)}{N_{n,1}} = \frac{\alpha_1 \alpha_2^{n-1} q_0 (1 - \alpha_1 \alpha_2^{n-1} q_0)}{N_{n,1}}. \tag{70}$$

It follows then from equation (68) that the variance of $\bar{\alpha}_{1,n}$ is

$$\sigma^2(\bar{\alpha}_{1,n}) = \frac{\alpha_1(1 - \alpha_1 \alpha_2^{n-1} q_0)}{N_{n,1} \alpha_2^{n-1} q_0}, \tag{71}$$

and so the weight W_n may be taken to be

$$W_n = \frac{N_{n,1} \alpha_2^{n-1}}{1 - \alpha_1 \alpha_2^{n-1} q_0}. \tag{72}$$

The only difficulty in computing the weights from this result is that the expression contains α_1, the parameter being estimated. However, an iteration scheme may be used to compute $\bar{\alpha}_1$; we begin with an unweighted mean of the $\bar{\alpha}_{1,n}$ and compute the weights, use these to obtain $\bar{\alpha}_1$ from equation (69), use this value of $\bar{\alpha}_1$ to recompute the weights, etc. However, we may replace α_1 by $\bar{\alpha}_1$ in equation (72) and substitute the expression for the W_n into equation (69) and obtain, after simplifications,

$$\sum_n \frac{x_{n,1}}{1 - \bar{\alpha}_1 \alpha_2^{n-1} q_0} = \sum_n N_{n,1}. \tag{73}$$

The right side of this equation is obtained at once from the data; the sum on the left may be computed as a function of $\bar{\alpha}_1$, for the given values of $x_{n,1}$, α_2 and q_0, and the correct value of $\bar{\alpha}_1$ obtained by successive approximations.

It will be noted that the sum is a monotonic increasing function of $\bar{\alpha}_1$ for $0 \leqq \bar{\alpha}_1 \leqq 1, 0 \leqq \alpha_2 \leqq 1, 0 \leqq q_0 \leqq 1$.

The variance of the estimate $\bar{\alpha}_1$ is obtained immediately from equation (66) and the relation

$$\sigma^2(\bar{\alpha}_1) = \frac{\sum_n W_n^2 \sigma^2(\bar{\alpha}_{1,n})}{(\sum_n W_n)^2}. \tag{74}$$

The result is

$$\sigma^2(\bar{\alpha}_1) = \alpha_1 / \sum_n \frac{N_{n,1} \alpha_2^{n-1} q_0}{1 - \alpha_1 \alpha_2^{n-1} q_0} = \frac{\alpha_1}{q_0 \sum_n W_n}. \tag{75}$$

This result may be used to estimate $\sigma^2(\bar{\alpha}_1)$ by replacing α_1 on the right side by its estimate $\bar{\alpha}_1$.

The procedure just described for obtaining an unbiased estimate of α_1 from the subset of data for which $k = 1$ may be generalized to obtaining unbiased estimates of α_1^k for any value of k. Analogous to equation (73) we obtain the relation

$$\sum_n \frac{x_{n,k}}{1 - \overline{\alpha_1^k} \alpha_2^{n-k} q_0} = \sum_n N_{n,k}, \qquad (k = 1, 2, \cdots) \tag{76}$$

where $N_{n,k}$ is the number of observations for the specified values of n and k, and $x_{n,k}$ is the number of those observations which yield alternative A_1. This equation may be solved numerically for $\overline{\alpha_1^k}$ and by taking the kth root of this estimate of α_1^k we obtain an estimate of α_1; for $k > 1$ these estimates are biased, of course, but they may be useful in obtaining an improved estimate of α_1 from certain kinds of data.

6d. Monte Carlo checks on the estimates. As a check on the estimation procedures just described, 30 Monte Carlo runs of 22 trials each were made as described in Section 4. The operators of equations (38) were used with $\alpha_1 = 0.70$, $\alpha_2 = 0.95$, and $p_0 = 0$. The parameter α_2 was estimated by obtaining a numerical solution to equation (49) with $\hat{q}_0 = 1$; ν was taken to be trial number n and only that portion of the data on each run up through the first A_1 occurrence was used. The result obtained was $\hat{\alpha}_2 = 0.9509$ and the estimate of the standard deviation of $\hat{\alpha}_2$ obtained from equation (58) was 0.008. The approximate value obtained from equation (55) was 0.956. Next, the procedure described in Section 6c to estimate α_1 for the subset of data for which $k = 1$ gave $\hat{\alpha}_1 = 0.758$ and $\sigma(\bar{\alpha}_1) = 0.08$. The estimates obtained compare favorably with the true values, $\alpha_1 = 0.70$ and $\alpha_2 = 0.95$, used in making the Monte Carlo computations.

6e. A related problem of estimating a binomial parameter. A problem related to the estimation problem discussed in Section 6a, but not a part of the general stochastic model, may be of some interest. The problem arises when we have a choice about the kind of information we can receive from binomial sampling. In a single binomial trial, event A or event $\bar{A}$ occurs. The probability of event A on a single trial is α and we wish to estimate α on the basis of information received. Suppose we have a choice: we can know the outcomes of N single binomial trials, or else we can know the outcome for each of N blocks of ν trials in the forms "all trials in the block were A's," or "not all trials in the block were A's." The probability that all trials in the block are A's is $q_\nu = \alpha^\nu$. The question for the statistician is what value of ν will give an optimum estimate of α. Dorfman mentions a similar problem concerning Wassermann tests [4]. Blood samples from several people could be pooled and the Wassermann test on the pool would be positive or negative. A negative report on the pool implies a negative report on all blood samples in the pool, while a positive report merely implies that one or more blood samples in the pool are positive. If we are to make a fixed number (N) of Wassermann tests, and if costs of increasing the size ν of the pool are very little, what value of ν should be chosen to get the best estimate of α? Dorfman was interested in identifying the positive individuals; we are interested in the estimation problem. The problem is not necessarily restricted to integral values of ν. For example, let α be the probability that a unit surface area of an industrial material such as sheet metal has no defects. We plan to inspect a sample area from each of N sheets of material, but we have a choice about the size of the area to be inspected. The report of the inspection is either that no defect was found or that some defect was found. We may be quite wise to use such an inspection procedure because actual counts of number of defects in an area can be quite untrustworthy while the reports "no defects" or "some defects" are comparatively reliable. Here the question is what size area should be inspected to give the best estimate of α, a measure of the quality of the product.

It turns out that R. A. Fisher [7] has already solved this question. He calls it the dilution problem. In our notation the maximum likelihood estimate of α is

$$\hat{\alpha} = (y/N)^{1/\nu}, \tag{77}$$

where y is the number of blocks which have all A's (e.g., negative blood tests or no defects). We see from equation (57) with $E(N_\nu) = N$ that the asymptotic variance is

$$\sigma^2(\hat{\alpha}) = \frac{1 - \alpha^\nu}{\nu^2 N \alpha^{\nu-2}}. \tag{78}$$

We wish to choose ν such that $\sigma^2(\hat{\alpha})$ has the smallest possible value for given α and N. It turns out that the minimizing value of ν is

$$\nu = \frac{1.594}{-\log_e \alpha}, \tag{79}$$

which agrees, of course, with Fisher's result. This means that if we have any good preliminary notions of the value of α we can improve the method of estimation of α, compared to the ordinary binomial method, by choosing a value of ν based on the preliminary value of α. If α is near unity, the use of blocks of size ν yields an effective binomial sample of size νN.

7. Applications to learning data. We have found the estimation procedures described above to be useful in analyzing various kinds of learning data. We present two examples.

The first illustration we will describe only briefly. It is closely related to the work of Miller and McGill [9]. The experiments analyzed by those authors were of the following sort. A person was presented with a series of R monosyllabic words and was then instructed to repeat all he could recall. This procedure was repeated for many trials, the order of the words being randomized on each trial. For lists of words not too long, Miller and McGill postulated that recall of a word increased the probability of recall on the next trial according to an operator defined by the first of equations (38); it was further postulated that nonrecall of a word left the recall probability unchanged, that is, $\alpha_2 = 1$ in equations (38). The probabilities of nonrecall, q_ν, after ν previous recalls, are then given by equation (43). The maximum likelihood procedures given in Section 6a are thus appropriate for estimating $\alpha = \alpha_1$ and q_0.

The second example, to be described in more detail, is data obtained by Solomon and Wynne [10] from experiments on the avoidance training of dogs. A dog is placed in a jumping stand and may jump over a barrier to avoid an intense electric shock. The shock is turned on 10 seconds after a signal which defines the start of a trial, and so on each trial the dog either avoids shock or escapes shock. We identify avoidance with alternative A_1 and nonavoidance (escape) with alternative A_2. The record of 30 dogs for 20 trials each is given in Table II; avoidance is denoted by a "1" and nonavoidance by a "0." From these raw data we obtained the numbers $N_{n,k}$ and $x_{n,k}$ where n refers to trial number ($n = 0, 1, \cdots$), and k is the number of previous avoidances ($k = 0, 1, \cdots, n$). Thus, $N_{n,k}$ is the number of dogs on trial n which avoided precisely k times previous to trial n, and $x_{n,k}$ is the number of those dogs which avoid on trial n. In Table III we give these quantities, derived from the raw data of Table II, for $k = 0, 1, 2, 3$.

We assume that the operators of equations (38) are appropriate for this experiment, that is, that both avoidance (A_1) and nonavoidance (A_2) increase the probability p of avoidance and tend to make it unity. The data strongly suggest that p_0, the initial probability of avoidance, is very near zero and so we assume $p_0 = 0(q_0 = 1)$. Thus, from equation (40), the probabilities $q_{n,k}$ of nonavoidance on trial n after k previous avoidances are

$$q_{n,k} = \alpha_1^k \alpha_2^{n-k}. \tag{80}$$

We wish to estimate the parameters α_1 and α_2 from the data.

First we consider the data up to and including the first avoidance of each dog ($k = 0$) and apply the maximum likelihood procedures given in Section 6a.

The index ν becomes trial number n, x_ν becomes $x_{n,0}$, and N_ν becomes $N_{n,0}$. Equation (55) gives $\hat{\alpha}_2 \cong 0.93$, and a numerical solution of equation (54) gives $\hat{\alpha}_2 = 0.923$. With the aid of equations (57) and (59) we estimate the standard deviation of $\hat{\alpha}_2$ to be $\sigma(\hat{\alpha}_2) \cong 0.014$. From the analysis given in Section 6b we next compute the mean and variance of h which is here interpreted to be the

TABLE II

Data on 30 dogs obtained by Solomon and Wynne [10]. The entry "1" indicates a dog avoided and the entry "0" indicates it did not avoid

Dog	Trial																			
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
13	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1
17	0	0	0	0	0	1	1	0	1	1	0	0	1	1	0	1	0	1	1	1
18	0	1	1	0	0	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1
21	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
27	0	0	0	0	0	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1
29	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1
30	0	0	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1
32	0	0	0	0	0	1	0	1	0	1	1	0	1	0	0	0	1	1	1	1
33	0	0	0	0	1	0	0	1	1	0	1	0	1	1	1	1	1	1	1	1
34	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	1	1
36	0	0	0	0	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1
37	0	0	0	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1
41	0	0	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1
42	0	0	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
43	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
45	0	1	0	1	0	0	0	1	0	1	1	1	1	0	1	1	1	1	1	1
47	0	0	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1
48	0	1	0	0	0	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1
46	0	0	0	0	1	1	0	1	0	1	1	0	1	0	1	1	1	1	1	1
49	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
50	0	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1
52	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1
54	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	1	1	0	1
57	0	0	0	0	0	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1
59	0	0	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1
67	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
66	0	0	0	1	0	1	0	1	1	1	0	1	0	1	1	1	1	1	1	1
69	0	0	0	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	1	1
71	0	0	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1

trial on which the first avoidance occurs. Using $\alpha_2 = 0.92$ and $q_0 = 1$ in equations (63) and (64) gives $E(h) = 4.39$ and $\sigma(h) = 2.28$. From the raw data of Table II one gets a mean value of $\bar{h} = 4.50$ and a standard deviation of 2.25. The close agreement between the computed expectation and the observed mean of h must be mainly accidental for the standard deviation of the estimate $\bar{h}$ is $\sigma(\bar{h}) \cong 2.3/\sqrt{30} = 0.42$.

Next we consider that subset of the data for which the number of previous avoidances is precisely one ($k = 1$). The estimation procedures described in Section 6c are directly applicable. A numerical solution of equation (73) yields $\bar{\alpha}_1 = 0.732$ and from equation (75) we obtain $\sigma(\bar{\alpha}_1) = 0.095$. Another estimate of α_1 was obtained from the data for $k = 2$ and equation (76); the result was $\bar{\alpha}_1 = 0.801$. Still another estimate, from the data for $k = 3$, was $\bar{\alpha}_1 = 0.705$. An unweighted mean of the three estimates is about 0.75. The three estimates

TABLE III

Values of N_{nk}, the number of dogs on trial n with precisely k previous avoidances, and x_{nk}, the number of those dogs which avoid on trial n, taken from the data given in Table II

n	$N_{n,0}$	$x_{n,0}$	$N_{n,1}$	$x_{n,1}$	$N_{n,2}$	$x_{n,2}$	$N_{n,3}$	$x_{n,3}$
0	30	0	0	—	0	—	0	—
1	30	3	0	—	0	—	0	—
2	27	3	3	1	0	—	0	—
3	24	4	5	1	1	0	0	—
4	20	7	8	5	2	0	0	—
5	13	4	10	5	7	3	0	—
6	9	2	9	6	9	5	3	3
7	7	4	5	3	10	6	5	3
8	3	2	6	4	7	5	8	4
9	1	0	4	2	6	4	9	8
10	1	1	2	0	4	3	5	5
11	0	—	3	1	1	1	3	2
12	0	—	2	1	1	1	2	2
13	0	—	1	0	1	1	1	1
14	0	—	1	1	0	—	1	1
15	0	—	0	—	1	1	0	—
16	0	—	0	—	0	—	1	1
17	0	—	0	—	0	—	0	—

of α_1 are well within one sigma of the mean 0.75, and so this may be taken as some small evidence of the appropriateness of the model.

An inference which may be made from the above analysis of the Solomon-Wynne data is the following. A trial on which nonavoidance occurs reduces the probability of nonavoidance by a factor 0.92, while a trial on which avoidance occurs reduces the probability of nonavoidance by a factor 0.75. Thus an avoidance is worth about 3.5 nonavoidances in teaching the dog to avoid. Such a conclusion may be of theoretical interest to psychologists.

8. The equal α case. In this section we will consider another special case, one in which the mathematical analysis of the model is especially simple. We let $\alpha_1 = \alpha_2 = \alpha$. The two operators become

$$Q_1p = a_1 + \alpha p = \lambda_1(1 - \alpha) + \alpha p \tag{81}$$
$$Q_2p = a_2 + \alpha p = \lambda_2(1 - \alpha) + \alpha p.$$

In equation (17) the term in the second moment drops out and we have left a simple linear difference equation in the means alone. The solution is

$$V_{1,n} = V_{1,\infty} - (V_{1,\infty} - V_{1,0})(a_1 - a_2 + \alpha)^n, \tag{82}$$

where

$$V_{1,\infty} = \frac{a_2}{1 - (a_1 - a_2 + \alpha)} = \frac{\lambda_2}{1 - \lambda_1 + \lambda_2}, \tag{83}$$

where λ_1 and λ_2 are the limits defined by equation (4). The expected operator $\bar{Q}$, discussed in Section 4, now gives the correct means as may be seen by comparing equations (17) and (21). Since the means from trial to trial are obtained by repeated application of this single operator, the changes in the mean are described by a simple two-state Markov chain with constant transition probabilities, a_2 and $(1 - \alpha - a_1)$.

In equation (18) for the higher raw moments, the terms in $V_{k+1,n}$ also vanish, and so the higher moments may be computed readily from that formula. The equation for the second raw moment becomes

$$V_{2,n+1} = a_2^2 + (2\alpha a_2 + a_1^2 - a_2^2)V_{1,n} + (\alpha^2 + 2a_1\alpha - 2a_2\alpha)V_{2,n}\,. \tag{84}$$

After equation (82) is inserted, a difference equation in the second raw moments is obtained. The solution turns out to be

$$V_{2,n} = \frac{a_2^2}{1 - C}(1 - C^n) + \frac{a_2 B}{1 - g}\left[\frac{1 - C^{n-1}}{1 - C} - g\left(\frac{g^{n-1} - C^{n-1}}{g - C}\right)\right] + p_0 B\,\frac{g^n - C^n}{g - C} + p_0^2 C^n, \tag{85}$$

where

$$B = 2\alpha a_2 + a_1^2 - a_2^2 = (1 - \alpha)[(\lambda_1^2 - \lambda_2^2)(1 - \alpha) + 2\alpha\lambda_2],$$
$$C = \alpha^2 + 2a_1\alpha - 2a_2\alpha = \alpha[\alpha + 2(\lambda_1 - \lambda_2)(1 - \alpha)], \tag{86}$$
$$g = a_1 - a_2 + \alpha = 1 - (1 - \alpha)(1 - \lambda_1 + \lambda_2).$$

As unsightly as this solution may appear, it is an exact expression for $V_{2,n}$ as a function of n and the parameters. Such an exact closed expression is not at present available for the general two-operator model except when $\alpha_1 = \alpha_2 = \alpha$.

Four parameters remain to be estimated: p_0, a_1, a_2, and α. It may be noted from equations (82) and (83) that at most three of these could be estimated from the means; these means depend only upon $V_{1,0} = p_0$, a_2 and $(a_1 - a_2 + \alpha)$. One might expect, however, that the variance of the data from trial to trial along with equation (85) could be used to estimate the fourth parameter. We shall indicate how this procedure might be feasible by making further restrictions.

8a. Equal α case, upper limit unity, lower limit zero. We now require that $a_2 = 0$ and $a_1 = 1 - \alpha$ in addition to requiring that $\alpha_1 = \alpha_2 = \alpha$. These further restrictions imply that the limits defined by equation (4) are $\lambda_1 = 1$ and $\lambda_2 = 0$. The two operators become

$$
\begin{aligned}
Q_1 p &= 1 - \alpha + \alpha p \\
Q_2 p &= \alpha p.
\end{aligned}
\tag{87}
$$

We have found this case useful for analyzing T-maze data on rats with identical and equally frequent rewards on the two sides of the maze (Stanley, [11]). Equation (82) then becomes

$$V_{1,n} = p_0, \qquad n = 0, 1, 2, \cdots. \tag{88}$$

The recurrence relation (84) for the second raw moment becomes

$$V_{2,n+1} = (1 - \alpha)^2 p_0 + \alpha(2 - \alpha)V_{2,n}. \tag{89}$$

The solution of this linear difference equation is

$$V_{2,n} = p_0 - p_0(1 - p_0)\beta^n \tag{90}$$

where

$$\beta = \alpha(2 - \alpha) = 1 - (1 - \alpha)^2. \tag{91}$$

This result may be obtained from equation (85) with $a_2 = 0$ and $a_1 = 1 - \alpha$. The variance of the probability distribution on trial n is

$$\sigma_n^2 = V_{2,n} - p_0^2 = p_0(1 - p_0)(1 - \beta^n). \tag{92}$$

From this result we see that the variance is zero for $n = 0$ and approaches the binomial variance, $p_0(1 - p_0)$, as n gets large providing $|\beta| < 1$. It can be shown that the distribution approaches a distribution with density p_0 at unity and density $(1 - p_0)$ at zero.

From equation (88) we see that the observed means from trial to trial may be used to estimate p_0, but that the means provide no information about the parameter α. Equation (92), on the other hand, shows that the variances, σ_n^2, of the distributions of probabilities depend upon β and hence upon α. As a result, one might expect to obtain an estimate of α from these variances. Such a procedure would lead to a simple double estimation problem as will now be indicated.

On trial n, a distribution of probabilities p_{in} exists. If one has data on K subjects, these data correspond to a sample of K probabilities p_{in} from the population of all possible values of p_{in} on trial n. If one knew the values of these K probabilities, then one could readily estimate the population mean and variance from the sample mean and variance. But the K probabilities, p_{in}, are not known, of course. Each p_{in} becomes the mean of a binomial distribution of random variables x_{in}. If the ith subject chooses A_1 on trial n, then $x_{in} = 1$, while if he chooses A_2 on trial n, then $x_{in} = 0$. The set of x_{in}'s provides the only information we have about the K probabilities p_{in}. The problem is clear—we must use the x_{in} to estimate the p_{in} and then use these estimates to estimate the properties of the distribution of all possible p_{in}'s on the nth trial.

8b. Estimation of p_0. From equation (88) we see that the mean of the d tribution of all possible probabilities p_{in} is p_0 on every trial for the case bei considered, namely, $\alpha_1 = \alpha_2$, $\lambda_1 = 1$, and $\lambda_2 = 0$. For a sample of K subjec we have a sample of K probabilities p_{in}, having a mean $\bar{p}_n$. This sample mea $\bar{p}_n$, provides an estimate of $V_{1,n} = p_0$. The sample mean, $\bar{p}_n$, is estimated i turn by the proportion $\bar{x}_n$ of subjects choosing A_1 on trial n:

$$\bar{x}_n = \frac{1}{K} \sum_{i=1}^{K} x_{in}. \tag{93}$$

Thus an estimate of p_0 is

$$(\bar{p}_0)_n = \bar{x}_n. \tag{94}$$

Such an estimate of p_0 is obtained from the data on each trial and so one can combine these estimates to obtain an improved estimate of p_0. We have not worked out an optimum way of combining the trial estimates, but one estimate of p_0 is obtained from a simple average of the individual trial estimates:

$$\bar{p}_0 = \frac{1}{N+1} \sum_{n=0}^{N} \bar{x}_n = \frac{1}{K(N+1)} \sum_{n=0}^{N} \sum_{i=1}^{K} x_{in}. \tag{95}$$

In other words, we estimate p_0 by the proportion of choices of A_1 in the entire set of data.

8c. Estimation of α.[2] Although we do not have an entirely satisfactory method of estimating α, we provide one method and invite (as in all these problems) suggestions for improving the estimation process.

We will break the data up into two subsets S_1 and S_2 such that all sequences in S_1 begin with an occurrence of A_1 and all those in S_2 begin with an occurrence of A_2. Thus, $x_{i0} = 1$ for $i \in S_1$ and $x_{i0} = 0$ for $i \in S_2$. On trial $n = 1$, sequences in S_1 will have probability $Q_1 p_0 = 1 - \alpha + \alpha p_0$ and sequences in S_2 will have probability $Q_2 p_0 = \alpha p_0$. Therefore we may consider all sequences in S_1 to have an initial probability $Q_1 p_0$ and all those in S_2 to have an initial probability $Q_2 p_0$. According to equation (88), then, the means $V_{1,n}$, will equal these initial probabilities on all future trials. Hence, we can estimate $Q_1 p_0$ by the proportion P_1 of A_1 occurrences in S_1 for trials $n = 1, 2, \cdots$, and similarly can estimate $Q_2 p_0$ by the proportion P_2 of A_1 occurrences in S_2 for trials $n = 1, 2, \cdots$. We then observe that

[2] This subsection was revised and considerably simplified in accordance with a suggestion by a referee.

$$Q_1p_0 - Q_2p_0 = 1 - \alpha. \tag{96}$$

Accordingly we can estimate α by

$$\hat{\alpha} = 1 - (P_1 - P_2). \tag{97}$$

This estimate is very easy to obtain from a set of data, but it clearly does not utilize all the information in the data. More efficient procedures are undoubtedly available.

8d. Estimation of α when limits are not zero and unity. Finally we propose one procedure for estimating α from data when the limits are $0 < \lambda_1 < 1$ and $0 < \lambda_2 < 1$. The means, $V_{1,n}$, from trial to trial are given by equation (82), which we write in the form

$$V_{1,n} = V_{1,\infty} - (V_{1,\infty} - V_{1,0})g^n \tag{98}$$

where $V_{1,\infty}$ and g are defined by equations (83) and (86), namely

$$V_{1,\infty} = \frac{\lambda_2}{1 - \lambda_1 + \lambda_2} \tag{99}$$

$$g = 1 - (1 - \alpha)(1 - \lambda_1 + \lambda_2). \tag{100}$$

Now the mean $V_{1,n}$ on trial n may be estimated by the proportion P_n of A_1 occurrences on trial n, that is

$$P_n = \frac{1}{K}\sum_{i=1}^{K} x_{in} \tag{101}$$

where K is the number of available sequences in the data. We then may sum these proportions P_n over all trials $n = 0, 1, \cdots, N - 1$, to obtain

$$P = \sum_{n=0}^{N-1} P_n = \frac{1}{K}\sum_{n=0}^{N-1}\sum_{i=1}^{K} x_{in}. \tag{102}$$

This quantity P is simply the total number of A_1 occurrences in the data divided by K, the number of sequences. Since P_n estimates $V_{1,n}$, we must sum $V_{1,n}$ of equation (98) over trials. We call this sum S_N:

$$\begin{aligned} S_N &= \sum_{n=0}^{N-1} V_{1,n} = \sum_{n=0}^{N-1} \{V_{1,\infty} - (V_{1,\infty} - V_{1,0})g^n\} \\ &= NV_{1,\infty} - (V_{1,\infty} - V_{1,0})\frac{1 - g^N}{1 - g}. \end{aligned} \tag{103}$$

The quantity P of equation (102) estimates S_N, and so we thus can solve for g in terms of $V_{1,\infty}$, $V_{1,0}$, and N. When we know in advance the values of λ_1, λ_2, and $V_{1,0}$, we may estimate α. In particular when $g^N \ll 1$ we have as the estimate of α,

$$\hat{\alpha} \cong 1 - \frac{V_{1,\infty} - V_{1,0}}{(1 - \lambda_1 + \lambda_2)(NV_{1,\infty} - P)}. \tag{104}$$

We have found this estimation scheme useful for analyzing data on human subjects in a two-choice situation.

8e. Applications to learning experiments. In this section we will apply the model of Section 8 and the estimation procedure given in Section 8d to three sets of data on behavior in a choice situation. The first set of data was obtained by Stanley from seven rats in a T-maze experiment [11]. On each trial the rat could turn either left or right in the maze, and for the portion of Stanley's data being considered here, the rat always found food on one side (alternative A_1) and never found food on the other side (alternative A_2). The second set of data was obtained by the authors with the assistance of Miss J. M. Jarrett from five Harvard undergraduates operating a machine called the "two-armed bandit" (work unpublished). On each trial the subject pushed one of two buttons; one choice was always followed by a penny reward and the other side never led to reward. The third set of data was obtained by R. R. Bush, R. L. Davis, and G. L. Thompson on six high school students in Santa Monica, California (work unpublished). In this experiment, the subjects were presented with two ordinary playing cards, face down, on each trial, and they were told to turn over one of the two cards; if the card turned over was a heart or diamond they received a reward of a nickel. All cards in one position were reward cards, and all cards in the other position were nonreward cards.

In all three experiments we identify the choice which leads to reward with alternative A_1 and the other choice with alternative A_2. We assume that the operators of equation (81) are appropriate, and we take $\lambda_1 = \lambda_2 = 1$, that is, we assume that either choice tends to make the probability p of A_1 equal to unity. From equation (83) we see that $V_{1,\infty} = 1$ and we assume that $V_{1,0} = 0.5$. Thus equation (104) becomes

$$\hat{\alpha} \cong 1 - \frac{0.5}{N - P} \tag{105}$$

where P is defined by equation (102) and is the mean number of choices A_1 up through trial N. Thus, $N - P$ is the mean number of errors (choices of A_2) made by the K subjects in each experiment. The results of the three experiments are summarized in Table IV.

9. Discussion. The stochastic processes described in this paper are closely related to Markov chains [6]. In fact, the process we defined can be considered to be a Markov chain if correctly viewed. A Markov chain is characterized by the property of "path independence." The system can exist in a number of *states* and if it is in the ith state, the transition probabilities to all other states are independent of how the system arrived in the ith state. Now if we identify the alternatives in our model with the states of the system, the process is clearly non-Markovian; the transition probabilities change as required by the operators of equations (1). The process defined by our model is a Markov chain, however, if we identify the states of the system with the values of p. Of course, an infinite number of states are then possible. If the system is in state p then the transition probabilities to t other states are given by equations (1) and the probabilities of transition to all other states are zero. In spite of this observation, we have made little use of the theory of Markov chains with an infinite number of states.

A special case which we have not handled satisfactorily is that for which $\lambda_2 = 0$ and $\lambda_1 = 1$, that is, for which $a_2 = 0$ and $a_1 = 1 - \alpha_1$. In this case the bounds on the means described in Section 4 are of little use since they demand only that the asymptotic mean lie between zero and unity. It is easily shown that all the asymptotic raw moments are equal, provided that a stable asymptotic distribution exists at all. This would mean that the density tends to be concentrated at zero and unity. The proportion which would be concentrated at unity is $V_{1,\infty}$. We have shown that $V_{1,\infty}$ would depend upon p_0 as well as upon α_1 and α_2, but we have not obtained a closed expression for $V_{1,\infty}$ as a function

TABLE IV

Data and computations for three experiments on two-choice situations. The rat data are from Stanley's T-maze experiment [11]. *The two groups of Harvard students were studied by Bush, Mosteller, and Jarrett using the "Two-armed bandit"; the group marked "pay" could either lose or break even on each trial, while the group marked "free" could either break even or win. The data on high school students were obtained by Bush, Davis, and Thompson in Santa Monica, California.*

	Rats	Harvard Students		High School Students
		Pay	Free	
Number subjects	7	5	5	6
Number trials	40	75	75	24
Mean errors	11.1	6.2	14.1	3.5
Estimate $\hat{\alpha}$	0.955	0.919	0.965	0.857

of these quantities. Likewise, we have not satisfactorily handled the estimation problem for this case even though this case seems to be of practical interest in some learning problems.

More generally, the outstanding problems seem to be the need for better expressions for the moments, or at least improved bounds for these moments, and more efficient estimation procedures in the cases we have discussed, and estimation procedures for the less special cases we have not discussed. The estimation procedures would no doubt depend on the particular type of data available; the values of some parameters may be known from symmetry considerations or from other experiments. Furthermore, some kinds of data provide more than one binomial observation per subject per trial, while others provide only one such observation. These considerations complicate the issues, so that a model whose parameters cannot be easily estimated for one type of experiment may be satisfactory in another type.

REFERENCES

[1] T. J. I'A. Bromwich, *An Introduction to the Theory of Infinite Series*, 2nd ed., MacMillan, London, 1942.

[2] R. R. Bush, and F. Mosteller, "A mathematical model for simple learning," *Psychological Review*, Vol. 58 (1951), pp. 313-323.

[3] R. R. Bush, and F. Mosteller, "A model for stimulus generalization and discrimination," *Psychological Review*, Vol. 58 (1951), pp. 413-423.

[4] R. Dorfman, "Detection of defective members of large populations," *Ann. Math. Stat.*, Vol. 14 (1943), pp. 436-440.

[5] W. K. Estes, "Toward a statistical theory of learning," *Psychological Review*, Vol. 57 (1950), pp. 94-107.

[6] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, John Wiley and Sons, 1950.

[7] R. A. Fisher, *The Design of Experiments*, 2nd ed., Oliver & Boyd, London, 1937, pp. 229-235.

[8] T. E. Harris, "A method for limit theorems in Markov chains," (Abstract) *Ann. Math. Stat.*, Vol. 23 (1952), p. 141.

[9] G. A. Miller, and W. J. McGill, "A statistical description of verbal learning," *Psychometrika*, Vol. 17 (1952), pp. 369-396.

[10] R. L. Solomon, and L. C. Wynne, "Traumatic avoidance learning: Acquisition in normal dogs," *Psychological Monographs*, No. 354, Vol. 67 (1953).

[11] J. C. Stanley, Jr., "The differential effects of partial and continuous reward upon the acquisition and elimination of a runway response in a two-choice situation," Ed. D. thesis, Harvard University, 1950.

[12] E. T. Whittaker, and G. N. Watson, *A Course of Modern Analysis*, 4th ed., Cambridge University Press, 1935, p. 462.