

Foundations of Stimulus-Sampling Theory for Continuous-Time Processes¹

P. SUPPES AND J. DONIO²

Stanford University, Stanford, California

An extension of stimulus-sampling theory is provided for the two-response case where the response is given continuously in time and the reinforcements are given according to some time-dependent process. The particular instance of a Poisson-type noncontingent reinforcement schedule is thoroughly examined. The generalized axioms on which the theory is based are stated and fundamental equations derived. Comparisons between this theory and the discrete-trial theory are provided, with emphasis on the similarities and differences between the results. Some sequential statistics are derived, and the results of a pilot experiment are briefly reported.

1. INTRODUCTION

With few exceptions the variety of stochastic learning theories or models that have been proposed in the literature have been formulated for discrete-time processes. As would be expected, a corresponding emphasis is to be found in the experiments performed to test these theories. Almost always a discrete-trial setup is used. Several reasons for this restriction are apparent. For one thing, it is much easier to perform and record the results of discrete-trial experiments, particularly if response latencies are not measured. Also, from the standpoint of mathematical derivations of predictions, the mathematics of discrete-time stochastic processes is both more elementary and much better developed, particularly in terms of computational methods, than that of continuous-time processes.

This remark about experiments must be rather carefully phrased, for a great deal of the experimental work on operant conditioning should most appropriately be analyzed in terms of continuous-time processes. In fact, however, few such analyses have been attempted, and even a cursory inspection of the experimental literature indicates that scarcely any of the psychologists working in this area recognize the possibility of a mathematical analysis. We hasten to add that this lack of recognition

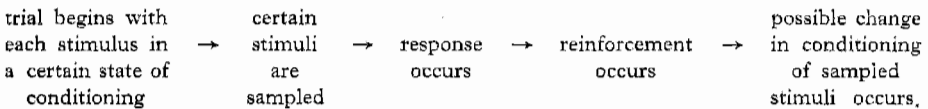
This research was supported by the Office of Scientific Research, United States Air Force Contract AF 49-(638)-1253.

² Present address: Société d'Analyse et de Conception de Systèmes, Paris, France.

has some justification, for the theoretical and mathematical problems besetting the construction of quantitative models adequate to handle the main kinds of operant conditioning schedules of reinforcement seem formidable indeed. We do not solve these problems in this paper.

The approach adopted here is to generalize stimulus-sampling theory, and thus to formulate the basic theory for immediate application to a class of experiments simpler in structure than most operant conditioning ones. The main task of this paper is to formulate the foundations in an explicit fashion. To show how the machinery works, the main results are derived for one of the simplest schedules of reinforcement—the noncontingent case suitably generalized to continuous time. More detailed mathematical analysis as well as some consideration of the statistical problems of estimating parameters is to be found in Donio's dissertation (1964), which will be published in abbreviated form at a later date.

The sequence of events that is postulated to occur on a trial in a discrete-time experiment may be schematized as follows:



In continuous-time experiments the order of events is no longer fixed, for responses may occur at any time in relation to reinforcements. To illustrate the theory as we have formulated it, it will be useful to concentrate on the alternative sequences of events that lead to a change in response. To begin with, it is important to emphasize that we consider in this paper only situations in which the subject is in some response state at all times—thus responses are not conceived as instantaneous events, but as chunks of time. At each time t the subject is making some response, or, in what seems intuitively a better choice of language, is in some one of the response states. For example, the subject might be set the task of monitoring two channels in which intermittent signals appear. At any instant he is monitoring one of the channels but not the other. Roughly speaking, the aim of the theory is to give an account of his behavior in shifting from one channel to the other as a function of the signals he hears. This signal-detection example has one misleading feature with respect to the theory formulated here. For simplicity of results we have restricted ourselves to determinate reinforcement with a correction procedure. This means in the signal-detection case that the subject would need to be told when he missed a signal that came in on the channel to which he was not "tuned." To make this correction procedure a simple matter to handle, in our own initial experiments we have used visual rather than auditory stimuli. The problem for the subject is to have selected at any time either the left or right side of a 16-in. scope. Intermittently, signals in the form of single letters appear in fixed position on the left or right side of the scope. The subject's

job is to predict, by his own selection of position, where the signals are going to appear. In this situation the correction procedure is easy to handle, for the subject is able to observe the signals coming in on either channel, i.e., on either the left or right side. This example in which reinforcing events are intermittent signals should make clear that we conceive of reinforcements as instantaneous events, but response states as chunks of time.

Suppose now the subject is in the left-response state, i.e., has selected the left channel. According to our theoretical formulation, he is in this response state because he has sampled and is holding a stimulus conditioned to this response. [For simplicity again, we also restrict sampling to one element at a time—the sampling assumption of what has come to be called the pattern model after Estes (1959).] At any time t he may change to the right side, i.e., shift response states, for one of two reasons.

One possibility is that a reinforcement may occur on the right side, i.e., a signal may appear on that side of the scope, and he thereby changes the conditioning of the stimulus he is holding. We postulate that this change occurs with constant probability c , conditional on the appearance of a reinforcement.

A second possibility is that he may drop the stimulus he is holding and sample another one. This newly sampled stimulus may be conditioned to the right side. We postulate that the temporal sampling of stimuli is a stationary process independent of the past history of the system, and the selection of the stimulus to be sampled follows the pattern of uniform sampling with replacement.

2. AXIOMS FOR A FINITE NUMBER OF RESPONSES

The axioms are formulated verbally, but with sufficient precision to make an exact mathematical statement fairly routine. For comparative purposes we follow the formulation given in Suppes and Atkinson (1960, Ch. 1) for the discrete-trial case.

CONDITIONING AXIOMS:

- C1. *At every time t , each stimulus is conditioned to one of the response states.*
- C2. *If a stimulus is being held at time t and a reinforcement of a response state occurs at t , then with probability c the stimulus becomes conditioned to this response state if it is not already.*
- C3. *The conditioning of the stimuli which are not held at the time of a reinforcement is not affected by the reinforcement.*
- C4. *The probability c of a change in conditioning is independent of the time t and the preceding pattern of events.*

SAMPLING AXIOMS:

- S1. *Exactly one stimulus is held as a sample at any time t .*
- S2. *The probability of holding a sample once drawn is independent of the time t ; how long it has been held, and the past history of the system.*
- S3. *If there are N stimuli available at time t for sampling, the probability of sampling any one of them is $1/N$, independent of which element was previously sampled, the time t , or the preceding pattern of other events.*

RESPONSE AXIOM:

- R1. *The response state at time t is the response state to which the stimulus sampled at time t is conditioned.*

Because of the extensive comments on these and related axioms in Suppes and Atkinson (1960, Chap. 1) and Suppes (1960), we need only to remark on those aspects that are specifically new to the generalization to a continuous-time process. The only serious modification is Axiom S2, which characterizes the postulated temporal process followed in drawing a new sample. It follows at once from this axiom that if T is the inter-arrival time random variable of the sampling process, then

$$P(T > x + y | T > x) = P(T > y | T > 0).$$

Here $P(T > y | T > 0)$ is the probability that the sample is held longer than y , given that it is drawn at time 0. The equality holds because this probability is independent of t , and the preceding history of the system. It follows from Axiom S2 that the times between the drawing of a new sample are independent and identically distributed, and in particular, follow an exponential distribution, whose parameter we shall take to be λ .

Axioms C2 and C3 express the familiar assumption that the conditioning of a stimulus takes place only when it is being held as a sample. The minor modification required for the present context is the emphasis on the stimulus being *held* rather than on the stimulus drawn on a given trial, for the concept of a given trial has, of course, disappeared. To guard against a possible source of confusion, we remark that the axioms are meant to be interpreted in such a way that the response effects of a change in conditioning are instantaneous. Literally speaking, this is an unrealistic assumption, but it is not an unrealistic first approximation for mathematical convenience.

It should also be remarked that it is completely straightforward to generalize the present axioms to the assumption that more than one stimulus is held as a sampled stimulus at any given time, although we shall not pursue this matter here. In fact, it will be useful to begin the consideration of details by restricting ourselves to a single stimulus element as the entire population of stimuli.

3. ONE-ELEMENT MODEL

To derive some definite predictions for this model, we impose a noncontingent schedule of reinforcement of the following sort. The temporal appearance of reinforcements follows a Poisson law, and thus the inter-arrival times of reinforcements are independently and identically distributed, following an exponential distribution with parameter μ , selected by the experimenter. Moreover, we restrict ourselves for the moment to two response states. With probability π each reinforcement is assigned to response state 1 and with probability $1 - \pi$ to response state 2, just as in the familiar discrete-trial setup.

Let $x_1(t)$ be the mean probability of being in response state 1 at time t . Then, following the usual notation for a "small" interval h , the probability of a reinforcement is approximately μh , or in the usual order-notation, $\mu h + o(h)$, and such a reinforcement has the following effect, according to Axiom C2:

$$\text{With prob. } \mu h + o(h) : x_1(t + h) = (1 - c) x_1(t) + c\pi. \quad (1)$$

And with probability $1 - \mu h - o(h)$ there is no change in $x_1(t)$ because no reinforcement has occurred, for in the one-element case, no sampling process enters. Thus,

$$\text{with prob. } 1 - \mu h - o(h) : x_1(t + h) = x_1(t). \quad (2)$$

Whence, combining probabilities,

$$x_1(t + h) = [(1 - c) x_1(t) + c\pi] \mu h + x_1(t) (1 - \mu h) + o(h), \quad (3)$$

and thus

$$\frac{x_1(t + h) - x_1(t)}{h} = -c\mu x_1(t) + c\pi\mu + \frac{o(h)}{h},$$

hence as $h \rightarrow 0$, we obtain the differential equation

$$\frac{dx_1(t)}{dt} = -c\mu x_1(t) + c\pi\mu, \quad (4)$$

whose solution, for the boundary condition $x_1(t) = p_0$ for $t = 0$, is:

$$x_1(t) = \pi - (\pi - p_0) e^{-c\mu t}. \quad (5)$$

The resemblance between Eq. 1 and 3 and the corresponding expressions for the discrete-trial case is apparent. In the familiar notation $p_{1,n}$ for the mean probability of an A_1 response on trial n , we have as the direct analogy of (1)

$$p_{1,n+1} = (1 - c) p_{1,n} + c\pi.$$

This is the analogue of (1) rather than (3) because in the discrete-trial case a reinforce-

ment does occur at the end of each trial. The solution expressed by (5) corresponds to the familiar explicit solution for the mean learning curve:

$$p_{1,n} = \pi - (\pi - p_{1,0})(1 - c)^{n-1},$$

where the geometric distribution plays the role of the exponential distribution in the continuous case.

It is a familiar fact that in the discrete-trial case the same mean learning curve is obtained for any postulated number N of stimulus elements. Because of the postulated stimulus fluctuation in the sample when $N \neq 1$ this is not so for the continuous case, as we shall see in the next section.

The one-element model has the usual difficulties that beset this model in the discrete case, namely, deterministic predictions that are seldom verified by the data. For example, it is predicted that, with probability one, the response state will not change following a reinforcement of this response state.

4. N -ELEMENT MODEL FOR POISSON REINFORCEMENT DISTRIBUTION

In view of these standard deficiencies of the one-element model, we turn now to the N -element model. For convenience of notation, we shall, however, continue to restrict ourselves to two responses. We must now take account of the sampling process as a source of change in $x_1(t)$. As indicated earlier, we designate the parameter of this process by λ , and, as in the preceding section, we assume a Poisson distribution of reinforcements with parameter μ .

The probability of a joint occurrence of a new sampling and a reinforcement in a small interval h is $\lambda\mu h^2$, and thus can be neglected when we write the transition matrix for t to $t + h$. We shall also ignore the order terms $o(h)$, and write as if the probabilities λh , μh and $1 - \lambda h - \mu h$ were exact instead of approximate, because the order term $o(h)/h$ approaches 0 as $h \rightarrow 0$, and its elimination is routine in all cases. As additional notation, we need $N_i(t)$ for the number of elements conditioned to response state i at time t and $n_i(t)$ for the expectation, $E(N_i(t))$. Also, to suggest the immediate generalization to r responses and to facilitate comparison with the continuum of responses considered in a later section, we write the transition probabilities in terms of transition matrices.

With probability λh a new sample is drawn and we have the two matrix equations:

$$\begin{pmatrix} N_1(t+h) \\ N_2(t+h) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} N_1(t) \\ N_2(t) \end{pmatrix},$$

$$\begin{pmatrix} x_1(t+h) \\ x_2(t+h) \end{pmatrix} = \begin{pmatrix} \frac{N_1(t)}{N} & \frac{N_1(t)}{N} \\ \frac{N_2(t)}{N} & \frac{N_2(t)}{N} \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

With probability μh , a reinforcement occurs and we have the two equations

$$\begin{aligned} \begin{pmatrix} n_1(t+h) \\ n_2(t+h) \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1(t) - x_1(t) \\ n_2(t) - x_2(t) \end{pmatrix} \\ &\quad - \begin{pmatrix} \pi + (1-\pi)(1-c) \\ (1-\pi)c \end{pmatrix} (1-\pi) + \pi(1-c) \begin{pmatrix} \pi c \\ \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} n_1(t) \\ n_2(t) \end{pmatrix} + \begin{pmatrix} -c + c\pi & c\pi \\ c - c\pi & -c\pi \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \end{aligned}$$

and

$$\begin{pmatrix} x_1(t+h) \\ x_2(t+h) \end{pmatrix} = \begin{pmatrix} 1-c-c\pi & \pi c \\ c-c\pi & 1-c\pi \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix},$$

or, in more compact notation

$$\begin{aligned} n(t+h) &= n(t) + Ax(t), \\ x(t+h) &= (I + A)x(t), \end{aligned}$$

where $n(t)$ and $x(t)$ are vectors, I is the identity matrix, and

$$A = \begin{pmatrix} -c + c\pi & c\pi \\ c - c\pi & -c\pi \end{pmatrix}.$$

With probability $1 - \lambda h - \mu h$, nothing happens, i.e.:

$$N(t+h) = N(t), \quad x(t+h) = x(t).$$

Combining these results and letting $h \rightarrow 0$, we obtain the system of differential equations in the expected values

$$\begin{aligned} \frac{dn(t)}{dt} &= \mu Ax(t) \\ \frac{dx(t)}{dt} &= \begin{pmatrix} \frac{\lambda n_1(t)}{N} - \mu c - \mu c\pi - \lambda & \frac{\lambda n_1(t)}{N} + \mu c\pi \\ \frac{\lambda n_2(t)}{N} - \mu c - \mu c\pi & \frac{\lambda n_2(t)}{N} - \mu c\pi - \lambda \end{pmatrix} x(t), \end{aligned} \tag{6}$$

whose solution

$$\begin{aligned} x_1(t) &= \pi + k_1 e^{m_1 t} + k_2 e^{m_2 t}, \\ x_2(t) &= 1 - x_1(t), \\ n_1(t) &= N \left[\pi + \frac{k_1 [m_1 + \lambda + \mu c]}{\lambda} e^{m_1 t} + \frac{k_2 [m_2 + \lambda + \mu c]}{\lambda} e^{m_2 t} \right], \\ n_2(t) &= N - n_1(t), \end{aligned} \tag{7}$$

with the following values for m_1 , m_2 , k_1 and k_2 :

$$\begin{aligned}
 m_1 &= -\frac{\lambda + \mu c}{2} + \frac{1}{2} \left[(\lambda + \mu c)^2 - \frac{4\lambda\mu c}{N} \right]^{1/2}, \\
 m_2 &= -\frac{\lambda + \mu c}{2} - \frac{1}{2} \left[(\lambda + \mu c)^2 - \frac{4\lambda\mu c}{N} \right]^{1/2}, \\
 k_1 &= \frac{\lambda \frac{n_1(0)}{N} - (m_2 + \lambda + \mu c) x_1(0) + (m_2 + \mu c) \pi}{m_1 - m_2}, \\
 k_2 &= \frac{\lambda \frac{n_1(0)}{N} - (m_1 + \lambda + \mu c) x_1(0) - (m_1 + \mu c) \pi}{m_2 - m_1}.
 \end{aligned} \tag{8}$$

We notice that, as $t \rightarrow +\infty$, we have

$$\begin{aligned}
 x_1(\infty) &= \pi & n_1(\infty) &= N\pi \\
 x_2(\infty) &= 1 - \pi & n_2(\infty) &= N(1 - \pi)
 \end{aligned} \tag{9}$$

which is in agreement with the asymptotic mean results for the discrete-trial case.

The prediction of conditional probabilities of the form $P(A_{i,n-1} | E_{j,n})$ is a familiar task in applying a learning model to a discrete-trial experiment. It will be a useful comparison to see how such conditional probabilities are computed for the continuous-time case. As an example, we compute $P(A_{1,t} | E_{i,t_0})$, for $t_0 < t$. For simplicity, we shall consider only the asymptotic results, although the method of analysis is easily extended. For $t \rightarrow \infty$, at time t_0^- , i.e., just before the reinforcement E_i , $P(A_{1,t_0^-}) = \pi$, as may be seen from Eq. 9. Immediately after the reinforcement, at time t_0^+ , the effect of the reinforcement is reflected in the fact that

$$\begin{aligned}
 P(A_{1,t_0^+}) &= \begin{cases} \pi - c(1 - \pi) & \text{if } i = 1, \\ \pi - c\pi & \text{if } i = 2; \end{cases} \\
 n_1(t_0^+) &= \begin{cases} N\pi + c(1 - \pi) & \text{if } i = 1, \\ N\pi - c\pi & \text{if } i = 2, \end{cases}
 \end{aligned} \tag{10}$$

which is an immediate consequence of the axioms.

The problem of finding $P(A_{1,t_1} | E_{i,t_0})$ is just the problem of using the new initial conditions expressed by (10) to find a new solution of the differential Eq. 6 after time $t_1 - t_0$ has elapsed.

We note first that

$$x_1(t, t_0) = \pi + k_1 e^{m_1(t-t_0)} + k_2 e^{m_2(t-t_0)} \tag{11}$$

for

$$0 \leq t_0 < t,$$

and in this case in the expressions for k_1 and k_2 given in (8), we replace $n_1(0)$ by $n_1(t_0)$ and $x_1(0)$ by $x_1(t_0)$. Using (10) we find immediately that

$$P(A_{1,t} | E_{1,t_0}) = \pi + k'_1 e^{m_1(t-t_0)} + k'_2 e^{m_2(t-t_0)} \quad (12)$$

with m_1 and m_2 as defined in (8), and

$$k'_1 = \frac{-c(1-\pi) \left[\left(1 - \frac{1}{N}\right) \lambda - m_2 + \mu c \right]}{m_1 - m_2}$$

$$k'_2 = \frac{-c(1-\pi) \left[\left(1 - \frac{1}{N}\right) \lambda - m_1 + \mu c \right]}{m_2 - m_1}$$

The corresponding expression for $P(A_{1,t}, E_{2,t_0})$ is easily found and is thus omitted here. It differs from (12) only in the coefficients k'_1 and k'_2 . Moreover, by an extension of these methods we may find $P(A_{1,t}, E_{2,t_0}, A_{2,t_1})$ and similar quantities, although the computations are rather tedious.

It is also straightforward but tedious to generalize the results given in this section to the case of r responses. Rather than consider this generalization, in the next and final section we briefly consider the extension to a continuum of responses.

5. EXTENSION TO A CONTINUUM OF RESPONSES

Perhaps a good intuitive example of a continuum of responses is offered by the tuning mechanism of a radio. What we want to emphasize is that the extension of the theory developed in this section is not meant to cover any problems of continuous tracking, but does apply to problems of monitoring signals that may be coming in on what is subjectively conceptualized as a continuum of frequencies. The much more complicated problems of continuous tracking we shall attempt to deal with, or at least begin to deal with, in a subsequent paper.

The axioms for a continuum of responses generalize the axioms of Sec. 2 in much the same manner as those of Suppes (1960) generalize those of Suppes and Atkinson (1960). The axioms stated here follow closely those of Suppes (1960), particularly in using the concept of a smearing distribution to represent the conditioning of any one of the finite set of stimuli. Roughly speaking, the introduction of a smearing distribution amounts to postulating that the conditioning of each stimulus is *smear*ed over a certain interval of responses, possibly the whole continuum available. The conditioning of any stimulus is represented by a probability distribution $K_s(x; z)$ for the response x , where z is the parameter of the distribution affected by the schedule of reinforcement, in a manner made precise by the following axioms,

CONDITIONING AXIOMS:

CC1. For each stimulus s there is at every time t , a unique smearing distribution $K_s(x; \mathbf{z})$ on the interval $[a, b]$ of possible responses such that

- (a) the distribution $K_s(x; \mathbf{z})$ is determined by its mode \mathbf{z} and its variance;
- (b) the variance is a constant σ^2 , which is not time-dependent;
- (c) the distribution $K_s(x; \mathbf{z})$ is continuous and piecewise differentiable in both variables.

CC2. If a stimulus is being held at time t and a reinforcement of a response state occurs at t , then with probability c the mode of the smearing distribution of the stimulus becomes the point of the reinforced response if it is not already.

CC3. The conditioning, i.e., the smearing distribution of any stimulus that is not held at the time of a reinforcement is not affected by the reinforcement.

CC4. The probability c of a change in conditioning is independent of the time t and the preceding pattern of events.

SAMPLING AXIOMS:

Axioms S1-S3 are exactly the same as those of Section 2.

RESPONSE AXIOMS:

CR1. If the held stimulus s and the mode \mathbf{z} of its smearing distribution are given, then the probability of a response in the interval $[a_1, a_2]$ is $K_s(a_2; \mathbf{z}) - K_s(a_1; \mathbf{z})$.

CR2. This probability of response is independent of the time t and the preceding pattern of events.

It should be remarked that certain of the axioms may be weakened without affecting any of the results established in this paper. This is particularly true of Axiom CC1, but we shall pursue this matter here. For a discussion of these matters, the reader is referred to Levine (1965).

As in the case of a finite number of responses, we shall begin by deriving some results for the one-element model. Following the notation of Suppes (1960), let $g_s(\mathbf{z}, t)$ be the marginal density for the parameter \mathbf{z} of the smearing distribution of stimulus s at time t . Because only one element is being considered, we drop the subscript s . Also, as in the finite case, we assume a Poisson law for reinforcements with density $f(y)$ being used to select the point of reinforcement, and with parameter μ for the Poisson process itself. From the axioms on conditioning it follows that when a reinforcement occurs, which has probability μh for a small interval h , we have

$$\text{with prob. } \mu h : g(\mathbf{z}, t + h) = (1 - c)g(\mathbf{z}, t) + cf(\mathbf{z}),$$

and

$$\text{with prob. } 1 - \mu h : g(z, t + h) = g(z, t).$$

Thus, combining probabilities

$$g(z, t + h) = [(1 - c)g(z, t) + cf(z)]\mu h + g(z, t)(1 - \mu h),$$

whence

$$\frac{g(z, t + h) - g(z, t)}{h} = -c\mu g(z, t) + c\mu f(z).$$

Hence as $h \rightarrow 0$, we obtain the differential equation

$$\frac{\partial g(z, t)}{\partial t} = -c\mu g(z, t) + c\mu f(z), \quad (13)$$

whose solution for the initial condition $g(z, t_0)$ is

$$g(z, t) = f(z) - (f(z) - g(z, t_0))e^{-c\mu(t-t_0)},$$

and clearly as $t \rightarrow \infty$

$$g(z) = f(z).$$

It is an immediate consequence of the axioms (see Suppes, 1960) that the marginal response density $r(x, t)$ satisfies the equation

$$r(x, t) = \int_a^b k(x; z)g(z, t) dz,$$

whence

$$r(x, t) = r(x) - (r(x) - r(x, t_0))e^{-c\mu(t-t_0)} \quad (14)$$

and

$$r(x) = \lim_{t \rightarrow \infty} r(x, t) = \int_a^b k(x; z)f(z) dz, \quad (15)$$

which is precisely the mean asymptotic result derived for the discrete-trial case in Suppes (1960). Equation 15 is also the direct analog of the probability matching law

$$\lim_{t \rightarrow \infty} x_1(t) = \pi$$

already derived for the two-response case.

We may also easily find certain conditional distributions for this one-element model. Immediately after a reinforcement y at time t_0 , we have, by the methods of argument already given, that

$$\bar{g}(z, t_0^+; y_{t_0}) = (1 - c)[f(z) - ((f(z) - g(z, 0))e^{-c\mu t_0})] + c\delta(z - y_{t_0}), \quad (16)$$

where $\delta(\mathbf{z} - y_{t_0})$ is the familiar Dirac delta function. The boundary condition expressed by (16) may be used then to find $g(\mathbf{z}, t)$ for $t > t_0$ from the differential Eq. 13. To simplify expressions, we restrict ourselves to the asymptotic case, i.e., when $t_0 \rightarrow \infty$. Then

$$g(\mathbf{z}, t_0^+; y_{t_0}) = (1 - c)f(\mathbf{z}) + c\delta(\mathbf{z} - y_{t_0}), \quad (17)$$

and thus

$$\begin{aligned} r(\mathbf{x}, t_0^+; y_{t_0}) &= (1 - c)r(\mathbf{x}, t_0) + c \int_a^b k(\mathbf{x}; \mathbf{z}) \delta(\mathbf{z} - y_{t_0}) d\mathbf{z} \\ &= (1 - c)r(\mathbf{x}, t_0) + ck(\mathbf{x}; y_{t_0}), \end{aligned} \quad (18)$$

whence for $t > t_0$, since $r(\mathbf{x}, t_0) = r(\mathbf{x})$ at asymptote, the conditional response distribution is given by

$$\begin{aligned} r(\mathbf{x}, t; y_{t_0}) &= r(\mathbf{x}) - [r(\mathbf{x}) - (1 - c)r(\mathbf{x}) - ck(\mathbf{x}; y_{t_0})] e^{-c\mu(t-t_0)} \\ &= r(\mathbf{x}) - c[r(\mathbf{x}) - k(\mathbf{x}; y_{t_0})] e^{-c\mu(t-t_0)}. \end{aligned} \quad (19)$$

For analysis of experimental data, we often want to conditionalize on any reinforcement at t_0 in some fixed interval I of the response continuum.

$$\begin{aligned} g(\mathbf{z}, t_0^+; y_{t_0} \in I_t) &= (1 - c)f(\mathbf{z}) + c \frac{\int_{I_{t_0}} \delta(\mathbf{z} - y)f(y)dy}{\int_{I_{t_0}} f(y)dy} \\ &= (1 - c)f(\mathbf{z}) + \frac{cf^*(\mathbf{z})}{\int_{I_{t_0}} f(y)dy}, \end{aligned}$$

where $f^*(\mathbf{z}) = f(\mathbf{z})$ for $\mathbf{z} \in I_{t_0}$ and $f^*(\mathbf{z}) = 0$ for $\mathbf{z} \notin I_{t_0}$. By a standard argument we then obtain

$$r(\mathbf{x}, t; y \in I_{t_0}) = r(\mathbf{x}) - c \left[r(\mathbf{x}) - \frac{\int_{I_{t_0}} k(\mathbf{x}; y)f(y)dy}{\int_{I_{t_0}} f(y)dy} \right] e^{-c\mu(t-t_0)}. \quad (20)$$

The extension of these results from the one-element model to the N -element model with a continuum of responses is rather complicated and will not be pursued, particularly because the changes considered in the next section permit a considerably simpler development.

6. RECONSIDERATION OF THE SAMPLING AXIOMS

The three sampling axioms formulated earlier and used for both a finite set and a continuum of responses have certainly an inherent conceptual simplicity, but their very simplicity entails a serious divergence from the discrete-trial theory and ensuing complications of analysis. The crux of the matter is this. Because the drawing of a new sample is postulated for each trial in the discrete-trial theory, we may always express the probability of a response, $x_1(n)$, say, in terms of the proportion of stimulus elements conditioned to response A_1 , i.e., on every trial n

$$x_1(n) = \frac{N_1(n)}{N}. \quad (21)$$

Perusal of the derivation of the differential equations of Sec. 4 quickly shows why a corresponding equation does not hold in the continuous-time case. The analog of (21) does not hold immediately after a reinforcement, for at that time the probability of a response state corresponding to the reinforcement is different from $N_1(t)/N$ because of the direct effect of reinforcement on the conditioning of the stimulus held at time t . The natural and obvious change in the continuous-time sampling axioms to bring them into line with the discrete-trial theory is to require that a new sample be drawn immediately after such reinforcement. This assumption also seems psychologically natural in view of the disruptive or discontinuous character of the reinforcements, which approximate instantaneous events. Because of this disruptive character of a reinforcement it seems reasonable to postulate that the organism draws a new stimulus sample of the environment whenever such an event occurs. However, this apparently reasonable postulate does make a very strong prediction about observable changes in response immediately after reinforcement. Because of the theoretical simplifications it introduces, we shall pursue its logical consequences, but we do emphasize that the simplified formal machinery may not be empirically viable. A comparative test of the two models centering around this point is derived below.

Of the three sampling axioms only the second one needs to be revised to accommodate the desired change to a new model.

S2'. The probability of holding a sample once drawn is independent of the time t , how long it has been held, and the past history of the system, except that a new sample is drawn immediately after each reinforcing event.

We first use this new axiom to derive the mean learning curve for the N -element model in the finite-response case.

With probability λh a new sample is drawn and we have

$$n_1(t + h) = n_1(t). \quad (22)$$

With probability μh a reinforcement occurs and we have

$$\begin{aligned} n_1(t+h) &= n_1(t) + \pi c \left(1 - \frac{n_1(t)}{N}\right) - (1-\pi)c \frac{n_1(t)}{N} \\ &= \left(1 - \frac{c}{N}\right) n_1(t) + \pi c. \end{aligned} \quad (23)$$

Combining results, i.e., taking the expectation and letting $h \rightarrow 0$, we obtain the following differential equation from (22) and (23):

$$\frac{dn_1(t)}{dt} = \frac{-c\mu}{N} n_1(t) + \pi c\mu \quad (24)$$

whose solution is

$$n_1(t) = \pi N - (\pi N - n_1(0)) e^{(-c\mu/N)t}, \quad (25)$$

and since for all t in the modified model, which we shall refer to as Model II,

$$x_1(t) = \frac{n_1(t)}{N}, \quad (26)$$

we immediately have from (25)

$$x_1(t) = \pi - (\pi - x_1(0)) e^{(-c\mu/N)t}, \quad (27)$$

which is just the mean learning curve (5) for the one-element model when $N = 1$, and the resemblance between (5) and the discrete-trial curve has already been pointed out. Clearly (27) is very considerably simpler than the corresponding result, Eq. 7, for Model I (which uses Sampling Axiom S2). It is worth noting that the sampling parameter λ does not enter in the mean learning curve (27) of Model II.

The simplicity of (27) is sufficiently encouraging to warrant a deeper look at the properties of Model II. In the discrete-trial theory a familiar technique for estimating parameters and checking goodness of fit is to use the theoretical transition matrix for responses from one trial to the next. Unfortunately, the extension of this approach to the continuous-time theory is not as simple as one might wish, but it will be instructive to carry developments a certain way.

We want to compute the probability of a transition from $A_{1,t}$ to $A_{2,t+h}$, as a typical case. The general method of attack is very similar to that used in the discrete-theory (see Suppes and Atkinson, 1960, pp. 51-55). We first need to compute $E(N_1(t) | A_{1,t})$. To do this, we find $P(N_1(t) = i | A_{1,t})$, for $i = 0, 1, \dots, N$. Restricting ourselves to asymptote, we have at once that

$$\begin{aligned} P(N_1(t) = i | A_{1,t}) &= \pi^{-1} P(A_{1,t} | N_1(t) = i) P(N_1(t) = i) \\ &= \frac{i}{N\pi} \binom{N}{i} \pi^i (1-\pi)^{N-i}, \end{aligned} \quad (28)$$

because $i/N = P(A_{1,t} | N_1(t) = i)$ for all t , and, as in the discrete-theory it is easy to show at asymptote, $P(N_1(t) = i)$ has the indicated binominal distribution. We then have from (28), again at asymptote,

$$\begin{aligned} E(N_1(t) | A_{1,t}) &= \frac{1}{\pi N} \sum_i i^2 \binom{N}{i} \pi^i (1 - \pi)^{N-i} \\ &= (N - 1) \pi + 1, \end{aligned} \quad (29)$$

a result which also holds in the discrete-theory. Similarly,

$$\begin{aligned} E(N_2(t) | A_{2,t}) &= (N - 1) (1 - \pi) + 1, \\ E(N_1(t) | A_{2,t}) &= (N - 1) \pi, \\ E(N_2(t) | A_{1,t}) &= (N - 1) (1 - \pi). \end{aligned} \quad (30)$$

We are now in a position to derive the expression for the transition probability $P(A_{2,t+h} | A_{1,t})$. First of all, with probability λh a new sampling of stimuli may occur, and the probability of drawing a stimulus conditioned to A_2 is just

$$\frac{1}{N} E(N_2(t) | A_{1,t}), \quad (31)$$

which is given by (30). A second possibility that may cause the shift in response state is the occurrence of a reinforcement during the interval h . If an E_1 reinforcement occurs, or if conditioning is not effective, then in accordance with Sampling Axiom S2', the probability of a switch is once again just (31). On the other hand, if an E_2 reinforcement occurs and conditioning is effective, then the number $N_2(t)$ of stimuli conditioned to A_2 at time t will be increased by one. Combining these results and computing the appropriate elementary probabilities, we obtain

$$\begin{aligned} P(A_{2,t+h} | A_{1,t}) &= \lambda h \left(1 - \frac{1}{N}\right) (1 - \pi) \\ &\quad + \mu h \left\{ \pi + (1 - \pi) (1 - c) \right\} \left(1 - \frac{1}{N}\right) (1 - \pi) \\ &\quad \quad \quad + c(1 - \pi) \left\{ \frac{1}{N} - \left(1 - \frac{1}{N}\right) (1 - \pi) \right\} \\ &= \lambda h \left(1 - \frac{1}{N}\right) (1 - \pi) + \mu h \left[\frac{c}{N} (1 - \pi) + \left(1 - \frac{1}{N}\right) (1 - \pi) \right] \\ &= (\lambda + \mu) h \left(1 - \frac{1}{N}\right) (1 - \pi) + \frac{\pi c}{N} (1 - \pi) h, \end{aligned} \quad (32)$$

We have not yet said so, but it should be made explicit even if it is evident in the above derivation that (32) holds only for h small. As h increases other events may intervene between $A_{2,t-h}$ and $A_{1,t}$, and both a reinforcement *and* a new sampling may take place, separate from the one immediately following the reinforcement. In principle the expression for arbitrary h , and not just small h , may be derived, but the expressions are complicated, primarily because the number of intervening reinforcements or response shifts is a random variable whose expectation must be calculated.

For small h we obtain by the above methods the following transition matrix

$$\begin{array}{cc}
 & \begin{array}{c} A_{1,t+h} \\ A_{2,t+h} \end{array} \\
 \begin{array}{c} A_{1,t} \\ A_{2,t} \end{array} & \left[\begin{array}{cc}
 1 - (1 - \pi)h \left[(\lambda + \mu) \left(1 - \frac{1}{N} \right) + \frac{\mu c}{N} \right] & (1 - \pi)h \left[(\lambda + \mu) \left(1 - \frac{1}{N} \right) - \frac{\mu c}{N} \right] \\
 \pi h \left[(\lambda + \mu) \left(1 - \frac{1}{N} \right) + \frac{\mu c}{N} \right] & 1 - \pi h \left[(\lambda + \mu) \left(1 - \frac{1}{N} \right) + \frac{\mu c}{N} \right].
 \end{array} \right]
 \end{array}$$

Given data on this transition matrix we may use the likelihood function derived for the one-element model with discrete-trials in Suppes and Atkinson (1960, p. 49) to provide an estimate of $[(\lambda + \mu) (1 - 1/N) + \mu c/N]$, but obviously other estimates will be needed to identify separately c , λ and N . We may, of course, estimate c/N in a fairly standard way from the mean learning curve.

7. COMPARISON OF MODELS I AND II

It is natural to seek a test that will compare the two basic models proposed. Such a test, by necessity, has to be built around the only feature which differs in the two models, i.e., the occurrence or nonoccurrence of a new stimulus sampling after a reinforcement. Thus, we must consider some statistic referring to what happens after a reinforcement. We consider a time interval of length U from t_0 to $t_0 + U$, in which t_0 is the time at which a reinforcement occurred and we assume that no reinforcements occur in this interval. In $(t_0, t_0 + U)$, a total of H stimulus samplings will occur. H is distributed according to a Poisson law with mean λU . How shall we choose U ? If U is too large, then λU , the expected number of samplings will also be large and will hide the effect that the reinforcement has on the number of changes in the response. If U is too small, we get a time interval for which latency effects are relatively important.

This situation leads us to the following test. In the interval $[t_0, t_0 + U]$ we consider

the random variable $R(U)$ = number of changes in the response position from t_0 to $t_0 + U$.

$$R(U) = R_1 + R_2(U),$$

$$R_1 = \begin{cases} 1 & \text{if the reinforcement at time } t_0 \text{ is} \\ & \text{immediately followed by a change in response,} \\ 0 & \text{otherwise,} \end{cases}$$

and

$$R_2(U) = \text{number of changes in the response position in } (t_0^+, t_0 + U).$$

Since R_1 is a 0-1 random variable, its expectation is also the probability that it is equal to 1. For both models, $R_2(U)$ has the same distribution which will be analyzed later. However, R_1 is different for the two models. The following table tells how much.

TABLE 1
COMPARATIVE DISTRIBUTION OF R_1 FOR THE TWO MODELS

At Asymptote	Model I	Model II
$E(R_1 A_2, t_0, E_{1, t_0}) =$	c	$(1 - \frac{1}{N})\pi + \frac{c}{N}$
$E(R_1 A_2, t_0, E_{2, t_0}) =$	0	$(1 - \frac{1}{N})\pi$
$E(R_1 A_1, t_0, E_{1, t_0}) =$	0	$(1 - \frac{1}{N})(1 - \pi)$
$E(R_1 A_1, t_0, E_{2, t_0}) =$	c	$(1 - \frac{1}{N})(1 - \pi) + \frac{c}{N}$

As for the distribution of $R_2(U)$, it involves the number H of stimulus samplings which occurred in $(t_0, t_0 + U)$, the number of stimuli conditioned to A_1 at time t_0^+ (since this number remains constant throughout the time interval considered and, when divided by N , represents the probability of getting an A_1 response at each new stimulus sampling). The exact distribution has been derived, Donio (1964), and requires detailed computations. Approximations, however, may be used here.

As already indicated, we do not want to choose U in such a way that $R_2(U)$ would be too large or too small. We shall choose it to be of the same order of magnitude as R_1 , i.e., we shall choose U such that $\lambda U \simeq 1$. Then $H(U)$ will be a Poisson random variable with mean 1. In the interval $(t_0, t_0 + U)$, $H(U)$ stimulus samplings occur, each one bringing back an A_1 response with probability $N_1(t_0^+)/N$ and an A_2 response

with probability $1 - N_1(t_0^+)/N$. For small $H(U)$, it is possible to consider all sample paths and the corresponding probabilities. If $H(U) = h$, the number of paths to be considered is 2^h . When $H(U)$ becomes large, this is no longer possible, but because $H(U)$ has a Poisson (with mean 1) distribution, this occurs with a rapidly decreasing probability. It is then possible to give approximations whose order of magnitude will be discussed soon. Table 2 provides a basis for discussion. In this table and subsequently, it is convenient to introduce the response random variable $X(t)$. If $X(t) = 1$ then response A_1 is being made at time C_1 and if $X(t) = 0$, response A_2 is being made at time t . We explicitly note that the table is constructed on the assumption that the interval U is selected so that $\lambda U = 1$. As an example of how the table is computed, under the condition that no reinforcement occurs between t_0 and $t_0 - U$:

$$f(R_2(U) | X_2(t_0^+) = 1, N_1(t_0^+) = k, H(U) = 2) = \begin{cases} 0 & \text{with Prob. } (k/N)^2 \\ 1 & \text{with Prob. } 1 - k/N \\ 2 & \text{with Prob. } (k/N)(1 - k/N). \end{cases}$$

We may average out over the distribution of $H(U)$. Then

$$P\{R_2(U) = 0 | X(t_0^+) = 1, N_1(t_0^+) = k\} = \sum_{i=0}^{\infty} \left(\frac{k}{N}\right)^i e^{-\lambda U} \left(\frac{\lambda U}{i!}\right)^i.$$

Since we have chosen $\lambda U = 1$, this gives

$$P\{R_2(U) = 0 | X(t_0^+) = 1, N_1(t_0^+) = k\} = e^{k/N-1} = e^{-(1-k/N)}.$$

More generally:

$$\begin{aligned} & P\{R_2(U) = j | X(t_0^+) = 1, N_1(t_0^+) = k\} \\ &= \sum_{i=0}^{\infty} P\{R_2(U) = j | X(t_0^+) = 1, N_1(t_0^+) = k, H(U) = i\} P\{H(U) = i\}, \end{aligned}$$

and this quantity lies between $\sum_{i=0}^4$ (same quantity) and

$$\sum_{i=0}^4 (\text{same quantity}) + \sum_{i=5}^{\infty} P\{H(U) = i\},$$

where we see from Table 2 that $\sum_{i=5}^{\infty} P\{H(U) = i\} = 0.009$. This explains why we get a good approximation on the distribution of $R_2(U)$. Once the computations are done, one gets the conditional probabilities shown in Table 3. At asymptote, we know the joint distribution of $X(t_0)$ and $N_1(t_0)$ before the occurrence of the reinforcement.

TABLE 2

PROBABILITIES OF VARIOUS NUMBERS OF SWITCHINGS FOR DIFFERENT NUMBER OF STIMULUS SAMPLINGS IN THE INTERVAL $(t_0, t_0 + U)$

$H(U)$ = No. of Stimulus Samplings in $(t_0^+, t_0 + U)$		Underlying Conditions $N_1(t_0^+) = k$ and			
		$X(t_0^+) = 1$		$X(t_0^+) = 0$	
$P(H(U) = j)$		No. of switchings	probability	No. of switchings	probability
0	0.366	0	1	0	1
1	0.366	0	k/N	0	$1 - k/N$
		1	$1 - k/N$	1	k/N
2	0.183	0	$(k/N)^2$	0	$(1 - k/N)^2$
		1	$1 - k/N$	1	k/N
		2	$k/N(1 - k/N)$	2	$k/N(1 - k/N)$
3	0.061	0	$\left(\frac{k}{N}\right)^3$	0	$\left(1 - \frac{k}{N}\right)^3$
		1	$\left(1 - \frac{k}{N}\right)\left(1 - \frac{k}{N} + \frac{k^2}{N^2}\right)$	1	$\frac{k}{N}\left(1 - \frac{k}{N} + \frac{k^2}{N^2}\right)$
		2	$\frac{k}{N}\left(1 - \frac{k^2}{N^2}\right)$	2	$\frac{k}{N}\left(2 - \frac{3k}{N} + \frac{k^2}{N^2}\right)$
		3	$\frac{k}{N}\left(1 - \frac{k}{N}\right)^2$	3	$\left(1 - \frac{k}{N}\right)\left(\frac{k}{N}\right)^2$

Table continued

$H(U) = \text{No. of Stimulus Samplings in } (t_0^+, t_0 + U)$		Underlying Conditions $N_1(t_0^+) = k$ and			
		$X(t_0^+) = 1$		$X(t_0^+) = 0$	
$P(H(U) = j)$		No. of switchings	probability	No. of switchings	probability
4	0.015	0	$\left(\frac{k}{N}\right)^4$	0	$\left(1 - \frac{k}{N}\right)^4$
		1	$\left(1 - \frac{k}{N}\right)\left(1 - \frac{2k}{N} + \frac{2k^2}{N^2}\right)$	1	$\frac{k}{N}\left(1 - \frac{2k}{N} + \frac{2k^2}{N^2}\right)$
		2	$\frac{k}{N}\left(1 - \frac{k}{N}\right)\left(1 + \frac{2k^2}{N^2}\right)$	2	$\frac{k}{N}\left(1 - \frac{k}{N}\right)\left(3 - \frac{4k}{N} + \frac{2k^2}{N^2}\right)$
		3	$2\frac{k}{N}\left(1 - \frac{k}{N}\right)^2$	3	$2\frac{k^2}{N^2}\left(1 - \frac{k}{N}\right)$
		4	$\frac{k^2}{N^2}\left(1 - \frac{k}{N}\right)^2$	4	$\frac{k^2}{N^2}\left(1 - \frac{k}{N}\right)^2$
> 4	0.009	0	$\sum_{i=0}^{\infty} \left(\frac{k}{N}\right)^i = \frac{(k/N)^5}{1 - k/N}$	0	$\frac{N}{k}\left(1 - \frac{k}{N}\right)^5$
		1	"	0	"
		2	"	2	"
		3	"	3	"
		4	"	4	"
		5	"	5	"
		.	"	.	"
		.	"	.	"
.	"	.	"		

TABLE 3
 CONDITIONAL PROBABILITIES FOR $R_2(U) = j^a$

$R_2(U)$	$X(t_0^+) = 1, N_1(t_0^+) = k$	$X(t_0^+) = 0, N_1(t_0^+) = k$
0	$e^{-(1-k/N)}$	$e^{-k/N}$
1	$\left(1 - \frac{k}{N}\right) \left[0.625 - 0.091 \frac{k}{N} + 0.091 \frac{k^2}{N^2}\right]$	$\frac{k}{N} \left[0.625 - 0.091 \frac{k}{N} + 0.091 \frac{k^2}{N^2}\right]$
2	$\frac{k}{N} \left(1 - \frac{k}{N}\right) \left[0.259 + 0.061 \frac{k}{N} + 0.03 \frac{k^2}{N^2}\right]$	$\frac{k}{N} \left(1 - \frac{k}{N}\right) \left[0.340 - 0.121 \frac{k}{N} + 0.030 \frac{k^2}{N^2}\right]$
3	$\frac{k}{N} \left(1 - \frac{k}{N}\right)^2 \left[0.061 + 0.015 \frac{k}{N}\right]$	$\left(1 - \frac{k}{N}\right) \frac{k^2}{N^2} \left[0.076 - 0.015 \frac{k}{N}\right]$
4	$0.015 \frac{k^2}{N^2} \left(1 - \frac{k}{N}\right)^2$	$0.015 \frac{k^2}{N^2} \left(1 - \frac{k}{N}\right)^2$
> 4	—	—

^a All the quantities given here are the lower bounds. The upper bounds are equal to the lower bounds plus 0.009.

Table 1 allows us to compute what happens at t_0 and from Table 2 we may then deduce the distribution of $R(U)$, the number of switchings in $(t_0, t_0 + U)$. Computations will not be developed any further here. As an example, we shall write down the exact value of

$$\begin{aligned}
 & P\{R(U) = 0 \mid X(t_0) = 1, E_{1, t_0}\} \\
 &= \frac{1}{e} \sum_{j=0}^{N-1} \binom{N-1}{j} e^{(j+1)/N} \pi^j (1-\pi)^{N-1-j} \quad \text{for Model I} \\
 &= \frac{1 + (N-1)\pi}{N} \frac{1}{e} \sum_{j=0}^{N-1} \binom{N-1}{j} e^{(j+1)N} \pi^j (1-\pi)^{N-1-j} \quad \text{for Model II,}
 \end{aligned}$$

The probability for Model II is equal to that of Model I, multiplied by the factor $(\pi + (1 - \pi)/N)$. Also, for Model I, this probability is greater than π and smaller than $\pi + \pi^2/2 + \pi(1 - \pi)/2N$ as may be shown by elementary computations.

At any rate, provided we are in a more than one-element model, the following is true:

$$\begin{aligned} \pi < P\{R(U) = 0 \mid X(t_0) = 1, E_{1,t_0}, \text{Model I}\} < \pi + 0.19, \\ P\{R(U) = 0 \mid X(t_0) = 1, E_{1,t_0}, \text{Model II}\} &= \left(\pi + \frac{1 - \pi}{N}\right) \\ &\quad \times P\{R(U) = 0 \mid X(t_0) = 1, E_{1,t_0}, \text{Model I}\}. \end{aligned}$$

For N large enough ($N \geq 3$), it may be shown that there are values of π for which:

$$\begin{aligned} P\{R(U) = 0 \mid X(t_0) = 1, E_{1,t_0}, \text{Model II}\} \\ < \pi < P\{R(U) = 0, X(t_0) = 1, E_{1,t}, \text{Model I}\}. \end{aligned}$$

For example, $\pi = 0.4$ and $N \geq 3$ would satisfy this inequality.

We now have a test to differentiate between the two models. For a given level of error, one may compute the number of values to be recorded before knowing with the preassigned error on which side of π the previous probability falls, thereby pointing to one model preferentially.

The test presented here does not use the whole available distribution. The reason lies in the difficult formulation of this distribution and the fact that extensive numerical computations are required.

8. PRELIMINARY EXPERIMENTAL TEST OF MODEL I

To give an at least tentative sense of how well the theory fits experimental data, we conclude this paper with a brief report of a pilot experiment. The general experimental conditions under which this experiment was run were not satisfactory enough to warrant a full-scale analysis, or even to apply Model II, which was itself conceived in reflection upon problems of data analysis. This pilot experiment was run under adverse conditions at the Stanford Computation Center before the Computer-based Laboratory for Learning and Teaching was completed.

For a report of much more satisfactory data, although obtained with rats rather than humans as subjects, the reader is referred to Karsh and Suppes (1964). Karsh and Suppes obtained probability matching in continuous time with individual rats and for thousands of reinforcements, verifying thereby the single most important prediction of the theory.

In the experiment to be reported here, adult subjects were required to predict the appearance of a small airplane on a screen. In accordance with the continuous-time theory, the subject was always either in the left-response or right-response state awaiting the arrival of a "target" or signal on a random schedule. His objective was to be on the side a target appeared as often as possible. When he was, the airplane image "exploded" shortly after appearance, in partial simulation of a target "hit."

The experiment was programmed on the PDP-1 computer in the Laboratory. The program generated the reinforcement signals, sent them to a 19-inch cathode-ray tube in front of which the subject was seated, displayed them, recorded both the signals when they were given, and the changes in the response, when they occurred, and at the end of each session, analyzed the data and punched out the results on tape. In the first half-hour session, the only one to be discussed here, 37 subjects participated in the experiment and were given reinforcing signals at a fast rate according to a Poisson process with approximate mean rate of 1 signal per second, and a constant probability $\pi = 0.625$ of reinforcing the right side. Four of the 37 subjects exhibited Bayesian behavior, i.e., once they were aware of the greater frequency of signals on the right, they kept their responses on the right constantly during the remainder of the experiment; data from these 4 subjects were omitted from the analysis.

Figure 1 gives the predicted and observed mean learning curve from minute to minute for the remaining 33 subjects. The net running time was about 22 minutes, as shown in the figure. The mean asymptote value is also shown. Notice that the starting probability of responding A_1 at time 0 is 1: This is so because the starting position was fixed as the same for all subjects. An asymptotic level was reached after 7 or 8

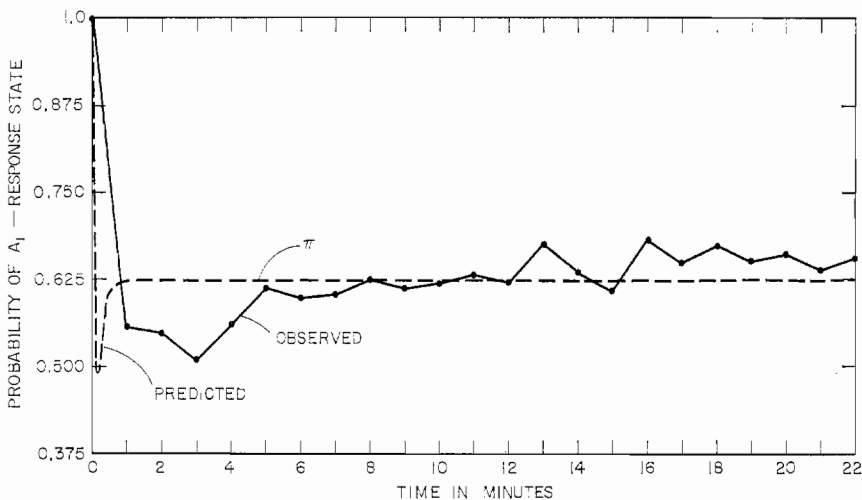


FIG. 1. Mean learning curve A_1 -responses state for Session 1.

minutes. The learning curve indicates that there was a slight tendency to overshoot the asymptotic value of $\pi = 0.625$. However, the asymptotic level observed is close enough to π to warrant deeper experimental investigation of the validity of Models I and II, with particular attention devoted to such questions as the distribution of interarrival times of response changes, and the rather complicated matter of estimating parameters.

REFERENCES

- DONIO, J. Some stochastic learning models with time as a continuous parameter. Unpublished doctoral dissertation. Stanford: Stanford University, 1964.
- ESTES, W. K. Component and pattern models with Markovian interpretations. In R. R. Bush and W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press, 1959. Pp. 9-52.
- KARSH, EILEEN B. AND SUPPES, P. Probability learning of rats in continuous-time experiments. *Psychonomic Science*, 1964, 1, 361-362.
- LEVINE, M. V. *The generalization function in the probability learning experiment*. Tech. Rept. No. 73, Psychology Series, Institute for Mathematical Studies in the Social Sciences, Stanford University, 1965.
- SUPPES, P. Stimulus sampling theory for a continuum of responses. In K. Arrow, S. Karlin, and P. Suppes (Eds.), *Mathematical methods in the social sciences*, 1959. Stanford: Stanford Univer. Press, 1960. Pp. 348-365.
- SUPPES, P. AND ATKINSON, R. C. *Markov learning models for multiperson interactions*. Stanford: Stanford Univer. Press, 1960.

RECEIVED: June 10, 1965