CHAINS OF INFINITE ORDER AND THEIR APPLICATION TO LEARNING THEORY

BY

JOHN LAMPERTI AND PATRICK SUPPES

TECHNICAL REPORT NO. 18

OCTOBER 15, 1958

PREPARED UNDER CONTRACT Nonr 225(17)
(NR 171-034)
FOR
OFFICE OF NAVAL RESEARCH

BEHAVIORAL SCIENCES DIVISION
APPLIED MATHEMATICS AND STATISTICS LABORATORY
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

CHAINS OF INFINITE ORDER AND THEIR APPLICATION TO LEARNING THEORY*/

by

John Lamperti and Patrick Suppes

1. Introduction.

The purpose of this paper is to study the asymptotic behavior of a large class of stochastic processes which have been used as models of learning experiments. We will do this by applying a theory of chains of infinite order, or "chaînes à liaisons complètes." Namely, we shall employ certain limit theorems for stochastic processes whose transition probabilities depend on the entire past history of the process, but only slightly on the remote past. Such theorems were given by Doeblin and Fortet [3] in a form close to that we employ; however, in order to accommodate certain cases of learning models we found it necessary to relax somewhat their hypotheses. A self-contained discussion of these and some additional results is the content of Section 2.

The processes which we shall study with these tools are called "linear learning models." From a psychological standpoint these models are very simple. A subject is presented a series of trials, and on each trial he makes a response, which consists of a choice from a finite set of possible actions. This response is followed by a reinforcement (again one of a finite number). The assumption of the model is that the subject's response probabilities on the next trial are linear functions of the probabilities on the present trial,

*/ This research was supported by the Group Psychology and Statistics Branches of the Office of Naval Research under contracts with Stanford University.

where the form of the functions depends upon which reinforcement has occurred. Many results about such models may be found in Bush and Mosteller [2], Estes [4], and Estes and Suppes [6]. We will also study here models constructed along similar lines for experiments involving two or more subjects and a type of interaction between them [6, Section 9] and Atkinson and Suppes [1]. Precise definitions of these processes are given below in Section 3.

The references mentioned above do not, except in very special cases, give a thorough treatment of asymptotic properties. We shall prove that under general conditions linear learning models exhibit "ergodic" behavior; that is, that after much time has passed these processes become approximately stationary and the influence of the initial distributions goes to zero. This is not the case for all models which have been used in experimental work, but it seems as if ergodic behavior can be proved by our method in almost all the cases in which one might expect it. Our theorems to this effect, their proofs and some corollaries are given in Section 4.

The major work so far on limiting behavior of learning models is Karlin [8], who obtains detailed limit theorems for certain classes of models. However, the results and even the techniques of Karlin's paper do not apply to many cases of interest. His starting point is a representation of the linear model as a Markov process whose states are the response probabilities. Two typical situations when such a representation is impractical arise (i) when the probabilities with which the reinforcement is selected depend on two or more previous responses, and (ii) in the many-person situations mentioned above. Both these situations can (and will) be studied using infinite order chains, and ergodic behavior established under mild restrictions. On the other hand, Karlin's work treats interesting non-ergodic cases outside the scope of our approach. For example, consider a T-maze experiment in which the subject (a rat, say) is reinforced (rewarded) on each trial regardless of whether he goes left or right. In the appropriate linear model, the probability of a left turn eventually is either nearly 0 or nearly 1, and which it is depends upon the rat's initial response probabilities. The model of this experiment has been thoroughly studied in [8, Section 2], and these results have been generalized by Kennedy [9].

In conclusion we comment that both more detailed results and other applications seem possible using the ideas of "infinite order chains." We hope to contribute further to this development in the future.

2. Chains of Infinite Order.

In this section we present a theory of non-Markov stochastic processes where the transition probabilities are influenced only slightly by the remote past. The original convergence theorems for this type of process are due to Doeblin and Fortet [3]; they are given here in a generalized form (Theorems 2.1 and 2.2). The weaker hypotheses make the proof of Lemma 2.1 more complicated than it is in [3], but the other proofs are not much affected. T. E. Harris has also studied these chains; we shall not use his results but remark that his paper [7] gives additional references and background on the subject. Finally we point out that the restriction to a finite number of states is not essential, and the theorems can be extended to the denumerable case without much change of methods.

Let I consist of the integers from 1 to N (to represent the states of the chain); we shall use the notation x for a finite sequence $i_o, i_1, \ldots$ of integers from I. The subscript "m" on x_m merely adds the specification that the sequence has m terms; the "sum" $x_m + x'$ will be the combined sequence $i_o, \ldots, i_{m-1}, i'_o, i'_1 \ldots$. The starting point for the theory will be a set of functions $p_i(x)$ defined for all $i \in I$ and all sequences x (including the sequence φ of length zero) and having the properties

$$p_i(x) \geq 0 \ , \quad \sum_i p_i(x) = 1 \ . \tag{2.1}$$

The function $p_i(x)$ will be interpreted as the conditional probability that a path function of the random process will go next to state i, having just occupied state i_o, previously i_1, etc. With this interpretation in mind we define inductively the "higher transition probabilities":

$$p_i^{(n)}(x) = \sum_{j \in I} p_j(x)\, p_i^{(n-1)}(j + x) \ , \tag{2.2}$$

where of course $p_i^{(1)}(x) = p_i(x)$, the given function. It is easy to see that these higher probabilities also satisfy condition (2.1). The functions $p_i^{(n)}(x)$ are the analogues of the terms of the matrix P^n for a Markov chain with transition matrix P ; the theorems we shall give generalize the convergence properties of the matrices P^n .

We shall first impose a positivity condition on the transition probabilities; specifically we assume that for some state j_o , some positive integer n_o , and some $\delta > 0$,

$$(2.3) \qquad p_{j_o}^{(n_o)}(x) \geq \delta \quad \text{for every } x .$$

We also need to make precise the "slight" dependence of these probabilities on the remote past; indeed, this is the crux of the whole theory. Define

$$(2.4) \qquad \epsilon_m = \sup | p_i(x+x') - p_i(x+x'') |$$

where the sup is taken over all states i , all sequences x' and x'' , and all sequences x which contain the state j_o at least m times. We shall use the postulate

$$(2.5) \qquad \sum_{m=0}^{\infty} \epsilon_m < \infty .$$

(In [3], ϵ_m is defined in the same way except that the sup is taken over all x of length at least m . Since this results in larger ϵ_m's , and since it is also assumed there that $\sum \epsilon_m < \infty$, our hypotheses are strictly weaker.) Throughout this section, (2.3) and (2.5) will be assumed.

Lemma 2.1.

$$(2.6) \qquad \lim_{m \to \infty} \sup | p_i^{(n)}(x+x') - p_i^{(n)}(x+x'') | = 0 ,$$

where the sup is the same as in (2.4) (i.e. x contains j_o at least m times); the convergence is uniform in n.

Proof. We define quantities $\epsilon_m^{(k)}$ by using $p_i^{(k)}$ instead of p_i in (2.4); then of course $\epsilon_m^{(1)} = \epsilon_m$, and the conclusion of the lemma is equivalent to $\epsilon_m^{(k)} \to 0$ uniformly in k as $m \to \infty$. Now

$$|p_i^{(k)}(x+x') - p_i^{(k)}(x+x'')| = |\sum_j \{p_i^{(k-1)}(j+x+x')p_j(x+x') - p_i^{(k-1)}(j+x+x'')p_j(x+x'')\}|$$

$$\leq \sum_j p_j(x+x')\,|p_i^{(k-1)}(j+x+x') - p_i^{(k-1)}(j+x+x'')|$$

$$+ \sum_j |p_j(x+x') - p_j(x+x'')|\,p_i^{(k-1)}(j+x+x'') \quad .$$

Suppose that x contains j_o m times. Then the second term of the above estimate is less than $N\epsilon_m$. The absolute value in the first term is less than $\epsilon_m^{(k-1)}$, but if $j = j_o$ this can be improved to $\epsilon_{m+1}^{(k-1)}$. Taking account of (2.3) and assuming that $n_o = 1$, we obtain the estimate

$$\epsilon_m^{(k)} \leq N\epsilon_m + \delta\,\epsilon_{m+1}^{(k-1)} + (1-\delta)\,\epsilon_m^{(k-1)} \quad . \tag{2.7}$$

(In case $n_o > 1$, the same idea can be carried out; the details are more cumbersome and will not be given.)

Now (2.7) can be iterated to obtain an estimate of $\epsilon_m^{(k)}$ in terms of ϵ_m. After some computation the result is

$$\epsilon_m^{(k)} \le N\epsilon_m \sum_{i=0}^{k-1} (1-\delta)^i + N\epsilon_{m+1}\,\delta \sum_{i=0}^{k-2} i(1-\delta)^i$$

$$+ \ldots + N\epsilon_{m+\ell}\,\delta^{\ell} \sum_{i=0}^{k-\ell-1} \binom{i+\ell-1}{i} (1-\delta)^i + \ldots + N\delta^{k-1}\epsilon_{m+k-1} \quad .$$

If the series are extended to infinity, the inequality remains true; calling these (infinite) series $A_o, A_1, \ldots, A_{k-1}$ we have

$$\epsilon_m^{(k)} \le N \sum_{i=0}^{k-1} \epsilon_{m+i}\,\delta^i A_i \quad .$$

But it can be shown without much difficulty that

$$A_{\ell+1} - A_{\ell} = (1-\delta)A_{\ell+1} \;,$$

or $A_{\ell+1} = A_{\ell}/\delta$. Since $A_o = \delta^{-1}$ we obtain $A_{\ell} = \delta^{-(\ell+1)}$, and hence

(2.8) $$\epsilon_m^{(k)} \le N\delta^{-1} \sum_{i=0}^{k-1} \epsilon_{m+i} \quad .$$

Recalling hypothesis (2.5), the uniform convergence of $\epsilon_m^{(k)}$ follows from (2.8).

Lemma 2.2.

(2.9) $$\lim_{n\to\infty} \left| p_i^{(n)}(x') - p_i^{(n)}(x'') \right| = 0$$

and the convergence is uniform in x' and x'' .

Proof. For clarity we shall use probabilistic arguments, although a purely analytic rephrasing is not hard. Consider two stochastic processes operating independently with transition probabilities $p_i(x)$, one with the sequence x' for its past history up to time 0 and the other with x''. In view of Lemma 2.1, for any $\epsilon > 0$ there is an m such that if the two processes have occupied the same states for a period which includes j_o at least m times and ends sometime before time n, then their probabilities of being in state i at time n differ by at most $\epsilon/2$. But it follows from condition (2.3) that with probability one, there will sometime be a period of length m during which both processes remain in state j_o. We can take n large enough so that this simultaneous "run" of state j_o will occur before time n with probability not less than $1 - \epsilon/2$. For this and all greater values of n, therefore, the two processes have probabilities of occupying state i at time n which differ by at most ϵ, and this proves (2.9). It is also easy to see from (2.3) and Lemma 2.1 that n can be chosen uniformly in x' and x''.

With this much preparation we shall now prove the first theorem:

Theorem 2.1. *The quantities*

(2.10) $$\lim_{n \to \infty} p_i^{(n)}(x) = \pi_i$$

exist, are independent of x, *and satisfy* $\sum_I \pi_i = 1$; *the convergence is uniform in* x.

Proof. Applying (2.2) repeatedly, we have

$$p_i^{(n+m)}(x) = \sum_{x_m} p_{i_{m-1}}(x)p_{i_{m-2}}(i_{m-1}+x) \ \ldots \ p_{i_o}(i_1 + \ldots + i_{m-1}+x)p_i^{(n)}(x_m+x) \ ,$$

where $x_m = i_o, i_1, \ldots, i_{m-1}$. Therefore

$$|p_i^{(n+m)}(x)-p_i^{(n)}(x)| \leq \sum_{x_m} p_{i_{m-1}}(x) \ \ldots \ p_{i_o}(i_1 + \ldots + i_{m-1}+x) \, | \, p_i^{(n)}(x_m+x)-p_i^{(n)}(x)|$$

and by Lemma 2.2, for any ϵ there is an n such that each term within absolute value signs on the right is less than ϵ . Since the weights $p_{i_{m-1}}(x) \ \ldots \ p_{i_o}(i_1 + \ldots + i_{m-1} + x)$ sum to one, we have

$$| \, p_i^{(n+m)}(x) - p_i^{(n)}(x) \, | < \epsilon \ ,$$

and so $p_i^{(n)}(x)$ has a (uniform in x) limit π_i . Since there are a finite number of states,

$$\sum_i \pi_i = \sum_i \lim_{n \to \infty} p_i^{(n)}(x) = \lim_{n \to \infty} \sum p_i^{(n)}(x) = 1 \ ,$$

and this completes the proof.

Next we shall define joint probabilities. If x_m is $i_o, i_1, \ldots, i_{m-1}$, let

$$(2.11) \quad p_{x_m}(x') = p_{x_m}^{(1)}(x') = p_{i_{m-1}}(x')p_{i_{m-2}}(i_{m-1}+x') \ldots p_{i_o}(i_1 + \ldots + i_{m-1}+x) \ .$$

This is, of course, the probability of executing the sequence of states x_m starting with past history x'. We can define also the higher joint probabilities:

$$p_{x_m}^{(n)}(x') = \sum_{j \in I} p_j(x')\, p_{x_m}^{(n-1)}(j + x') \quad . \tag{2.12}$$

Analogues of Lemmas 2.1 and 2.2 can be proved for these quantities by the same arguments used already; in this way it is not difficult to prove

Theorem 2.2. _The quantities_

$$\lim_{n \to \infty} p_{x_m}^{(n)}(x') = \pi_{x_m} \tag{2.13}$$

exist, are independent of x' , _and satisfy_ $\sum_{i_o \ldots i_{m-1}} \pi_{x_m} = 1$; _the convergence is uniform in_ x' .

Remark. These two theorems imply the existence of a stationary stochastic process with the $p_i(x)$ for transition probabilities. The idea is that the quantities π_{x_m} can be used to define a probability measure on the "cylinder sets" in the space of infinite sequences of members of I , and this measure can then be extended. This stationary process need not concern us further here.

Finally we will prove convergence theorems for certain "moments" which are useful in studying experimental data. The idea is that if we have a stochastic process with the functions $p_i(x)$ for transition probabilities, the probability $p_i(x_m)$ that the state at time m is i given the past history x_m is itself a random variable, and so it makes sense to study $E(p_i^{\nu}(x_m))$. More formally, define

$$\alpha_i^{\nu}(m,x) = \sum_{i_o,\ldots,i_{m-1}} p_i^{\nu}(x_m+x)\, p_{x_m}(x) \tag{2.14}$$

where $p_{x_m}(x)$ is defined by (2.9). Thus $\alpha_i^1(m,x)$ is the same as $p_i^{(m)}(x)$. Theorem 2.1 states that $\lim_{m\to\infty} \alpha_i^1(m,x) = \pi_i$ exists. We shall now prove

Theorem 2.3. The quantities

$$\lim_{m\to\infty} \alpha_i^{\nu}(m,x) = \alpha_i^{\nu} \tag{2.15}$$

exist for every positive integer ν; convergence is uniform in x and the limit is independent of x.

Proof. We use a simple estimate to show that $\alpha_i^{\nu}(m,x)$ is a Cauchy sequence:

$$\left| \alpha_i^{\nu}(m+k+h, x) - \alpha_i^{\nu}(m+k, x) \right|$$

$$= \left| \sum_{x_{m+k+h}} p_i^{\nu}(x_{m+k+h}+x)p_{x_{m+k+h}}(x) - \sum_{x_{m+k}} p_i^{\nu}(x_{m+k}+x)p_{x_{m+k}}(x) \right|$$

$$\leq \sum_{x_{m+k+h}} \left| p_i^{\nu}(x_{m+k+h}+x) - p_i^{\nu}(x_m+x) \right| p_{x_{m+k+h}}(x)$$

$$+ \sum_{x_{m+k}} | p_i^\nu(x_{m+k}+x) - p_i^\nu(x_m+x) | p_{x_{m+k}}(x)$$

$$+ | \sum_{x_{m+k+h}} p_i^\nu(x_m+x)p_{x_{m+k+h}}(x) - \sum_{x_{m+k}} p_i^\nu(x_m+x)p_{x_{m+k}}(x) | \ .$$

If m is chosen large enough, the first two terms will be arbitrarily small; this involves nothing more than the conditions (resulting from (2.3) and (2.5)) that $\epsilon_m \to 0$, and that a long sequence x contains j_o many times with high probability. The last term may be rewritten by carrying out the summation over all the indices except those in x_m ; this yields

$$| \sum_{x_m} p_i^\nu(x_m+x)(p_{x_m}^{(k+h)}(x) - p_{x_m}^{(k)}(x)) | \le \sum_{x_m} | p_{x_m}^{(k+h)}(x) - p_{x_m}^{(k)}(x) |$$

which is small for all h (and for all x) if k is large enough, by Theorem 2.2. Thus if $n = m + k$, $| \alpha_i^\nu(n+h,x) - \alpha_i^\nu(n,x) |$ is small for all h , and this proves that the limit (2.15) must exist; the limit is uniform in x since $\alpha_i^\nu(m,x)$ is uniformly Cauchy. Another estimate along much the same line can be made to show that for any $\epsilon > 0$,

$$| \alpha_i^\nu(m+k,x) - \alpha_i^\nu(m+k,x') | \le \epsilon$$

provided m and k are large. Since the limit of $\alpha_i^\nu(m+k,x)$ exists as $m+k \to \infty$, we can conclude that the limit is the same for all x .

It is also desirable to consider some additional "cross" moments involving $p_i(x_m)$ for several states at once; accordingly we define

$$(2.16)\quad \alpha_{j_1 j_2 \cdots j_k}^{\nu_1 \nu_2 \cdots \nu_k}(m,x) = \sum_{x_m} p_{j_1}^{\nu_1}(x_m+x)\, p_{j_2}^{\nu_2}(x_m+x) \cdots p_{j_k}^{\nu_k}(x_m+x)\, p_{x_m}(x) \quad .$$

The following theorem is then a generalization of Theorem 2.3, which treats the case $k = 1$:

Theorem 2.4. *The quantities*

$$(2.17)\qquad \lim_{m \to \infty} \alpha_{j_1 \cdots j_k}^{\nu_1 \cdots \nu_k}(m,x) = \alpha_{j_1 \cdots j_k}^{\nu_1 \cdots \nu_k}$$

exist uniformly in x *for all non-negative integers* $\nu_1 \ldots \nu_k$ *and all* $j_1 \ldots j_k \in I$, *and the limits are independent of* x .

The argument used in proving Theorem 2.3 works in this case also with only trivial changes, and need not be repeated. Finally we remark that moments involving several values of n can be considered, and it can be shown that their limits exist also. This provides a generalization of Theorem 2.2.

3. Definition of Linear Learning Models.

The models we consider apply to an experimental situation which consists of a sequence of trials. On each trial the subject of the experiment makes a response, which is followed by a reinforcing event. Thus an experiment may be represented by a sequence $(\underline{A}_1, \underline{E}_1, \underline{A}_2, \underline{E}_2, \ldots \underline{A}_n, \underline{E}_n, \ldots)$ of random variables, where the choice of letters follows conventions established in the literature: the value of the random variable $\underline{A}_n$ is a number j representing the actual response on trial n, and the value of $\underline{E}_n$ is a number k representing the reinforcing event on trial n. The relevant data on each trial may then be represented by an ordered pair (j,k) of integers with $1 \leq j \leq r$, and $0 \leq k \leq t$, that is, we envisage in general r responses and $t+1$ reinforcing events. Any sequence of these pairs of integers is a sequence of values of the random variables and thus represents a possible experimental outcome. The general aim of the theory is to predict the probability distribution of the response random variable when a particular distribution, or class of distributions, is imposed on the reinforcement random variable.

In dealing with the general linear model with r responses and $t+1$ reinforcing events we are following the formulation in Chapter 1 of Bush and Mosteller [2], although our notation is somewhat different, being closer to Estes [4] and Estes and Suppes [6].

The theory is formulated for the probability of a response on trial $n+1$ <u>given</u> the entire preceding sequence of responses and reinforcements. For this preceding sequence we use the notation x_n. Thus

$$x_n = (k_n, j_n, k_{n-1}, j_{n-1}, \ldots, k_1, j_1) \quad .$$

(It is convenient to write these sequences in this order, but note that the numbering here is from past to present, not the reverse as in Section 2.)

Our single axiom is the following linearity assumption:

Axiom L. *If* $E_n = k$ *and* $P(x_n) > 0$ *then*

$$(3.1) \qquad P(A_{n+1} = j \mid x_n) = (1 - \theta_k)P(A_n = j \mid x_{n-1}) + \theta_k \lambda_{jk} ,$$

where $0 \leq \theta_k,\ \lambda_{jk} \leq 1$ *and* $\sum_j \lambda_{jk} = 1$.

We obtain the linear model studied intensively in [6] by setting:

$$(3.2) \qquad \begin{cases} \theta_k = \theta & \text{for } k \neq 0 \\ \theta_k = 0 & \text{for } k = 0 \\ \lambda_{jj} = 1 & \\ \lambda_{jk} = 0 & \text{for } j \neq k \\ t = r . & \end{cases}$$

A linear model satisfying (3.2) we shall term an Estes Model, and for such models (3.1) may be replaced by the simpler condition:

$$(3.3) \quad P(A_{n+1} = j \mid x_n) = \begin{cases} (1-\theta)P(A_n = j \mid x_{n-1}) + \theta & \text{if } E_n = j \\ (1-\theta)P(A_n = j \mid x_{n-1}) & \text{if } E_n = k, k \neq 0, k \neq j \\ P(A_n = j \mid x_{n-1}) & \text{if } E_n = 0 . \end{cases}$$

Axiom L satisfies the combining classes condition of Bush and Mosteller. Upon replacing θ by $1 - \alpha$ in (3.1) essentially their general formulation of the linear model is obtained, although they do not explicitly indicate dependence on the sequence x_n .

We also define here certain moments which are of experimental interest and whose asymptotic properties we investigate subsequently. The moments $\alpha^{\nu}_{j,n}$ of $p_{j,n}(x)$ are:

$$(3.4) \qquad \alpha^{\nu}_{j,n} = \sum_{x_{n-1}} p^{\nu}(\underline{A}_n = j \mid x_{n-1}) P(x_{n-1}) \quad .$$

And if the appropriate limits exist, we define

$$(3.5) \qquad \alpha^{\nu}_{j} = \lim_{n \to \infty} \alpha^{\nu}_{j,n} \quad .$$

The moments (3.4) are formed in an unsymmetrical way; however, they enter in a natural way in the expression of quantities which are easily observed experimentally -- for instance, the joint probability $P(\underline{A}_{n+1} = j, \underline{A}_n = j)$. (For other examples, see [6].)

We are also interested in studying extensions of the linear model to multiperson situations. We may suppose that we have s subjects in a situation such that the probability of a particular reinforcing event for any one subject will depend in general on preceding responses and reinforcements of the other $s-1$ subjects as well as on his own prior responses and reinforcements. The data on each trial may then be represented by an ordered $2s$-tuple $(j_1, k_1, \ldots, j_s, k_s)$ of integers with $1 \leq j_i \leq r_i$, $0 \leq k_i \leq t_i$, for $i = 1, \ldots, s$, and any sequence of such tuples represents a possible experimental outcome. Let $\underline{A}^{(i)}_n$ and $\underline{E}^{(i)}_n$ be the response and reinforcement random variables for the i^{th} subject on trial n . We may then generalize Axiom L to:

Axiom M. For $1 \leq i \leq s$, if $\underline{E}_n^{(i)} = k$ and $P(x_n) > 0$ then

$$(3.6) \qquad P(\underline{A}_{n+1}^{(i)} = j \mid x_n) = (1 - \theta_k^{(i)})P(\underline{A}_n^{(i)} = j \mid x_{n-1}) + \theta_k^{(i)} \lambda_{jk}^{(i)} ,$$

where $0 \leq \theta_k^{(i)}, \ \lambda_{jk}^{(i)} \leq 1$ and $\sum_j \lambda_{jk}^{(i)} = 1$.

Experimental tests of Axiom M for two-person situations are reported in Estes [5] and in Atkinson and Suppes [1]. Let $x_{n-1}^{(i)}$ be just the sequence of first $n-1$ responses and reinforcements of subject i . It is an easy consequence of Axiom M that

$$P(\underline{A}_n^{(i)} = j \mid x_{n-1}^{(i)}) = P(A_n^{(i)} = j \mid x_{n-1}) ,$$

and it is in terms of $x_{n-1}^{(i)}$ that we define moments $\alpha_{\nu,j,n}^{(i)}$ exactly analogous to (3.4). We shall also be interested in the joint moments

$$(3.7) \quad \gamma_{j_1,\ldots,j_s,n}^{\nu} = \sum_{x_{n-1}} P^{\nu}(\underline{A}_n^{(1)} = j_1,\ldots,A_n^{(s)} = j_s \mid x_{n-1})P(x_{n-1}) ,$$

and their asymptotes $\gamma^{\nu} = j_1,\ldots,j_s$ if they exist. To work with these latter moments in terms of Axiom M we need the additional reasonable assumption that when all the $n-1$ preceding responses and reinforcements are given, the s responses on trial n are statistically independent:

Axiom I. If $P(x_{n-1}) > 0$ then

$$P(A_n^{(1)} = j_1,\ldots,\underline{A}_n^{(s)} = j_s \mid x_{n-1}) = \prod_{i=1}^{s} P(A_n^{(i)} = j_i \mid x_{n-1}) .$$

The experimental restriction implied by Axiom I has been satisfied in the multiperson studies employing the linear model.

4. Asymptotic Theorems for Learning Models.

After dealing with some matters of notation, we state general theorems on the existence of asymptotic moments. The hypotheses of the theorems give some broad conditions which guarantee ergodic behavior. We begin with the one-person models satisfying Axiom L.

In this section it will be convenient to use some of the notation of Section 2. Thus we may write $P(\underline{A}_n = j \mid x_m + x')$ in place of $P(\underline{A}_n = j \mid x_{n-1})$ to indicate we are interested in the last m terms of x_{n-1}. The "sum" $x_m + x'$ is just the combined sequence x_{n-1}. We reserve the subscript m for counting back m trials from a given trial n.

To clarify the general theorems it is desirable to define in an exact way the notion of the conditional probability of a reinforcing event depending on only a finite number m of past trial outcomes and independent of the trial number.

Definition. A linear model has a reinforcement schedule with past dependence of length m if, and only if, for all k, n and n' with $n, n' > m$ and all x_m, x' and x''

$$P(\underline{E}_n = k \mid x_m + x') = P(\underline{E}_{n'} = k \mid x_m + x'') \quad . \tag{4.1}$$

(It is understood that x_m includes the response $A_{j,n}$ which precedes $E_{k,n}$ on trial n .) It is to be noticed that the use of n on one side and n' on the other side of (4.1) yields independence of trial number. The term reinforcement schedule has been used because of its frequent occurrence with approximately this meaning in the experimental literature. For the conditional probabilities of (4.1) we shall use the notation

$$\pi_{k,x_m} = P(\underline{E}_n = k \mid x_m + x) \quad . \tag{4.2}$$

We may now state the first general theorem.

Theorem 4.1. Let $\mathcal{L}$ be a linear model such that

(i) *$\mathcal{L}$ has a reinforcement schedule with past dependence of length m^* ,*

(ii) *there is an integer k^* such that*

(a) $\Theta_{k*} \neq 0$

(b) *there is a δ^* and an m_o such that for all sequences x and all integers n*

$$P(\underline{E}_{n+m_o} = k^* \mid x_n) \geq \delta^* > 0 \quad .$$

Then the asymptotic moments α_j^ν of $\mathcal{L}$ all exist and are independent of the initial distribution of responses.

Proof: The central task is to characterize $\mathcal{L}$ as a chain of infinite order and show that satisfaction of the hypotheses of the theorem implies satisfaction of conditions (2.3) and (2.5). With this accomplished the asymptotic theorems of Section 2 may be applied to $\mathcal{L}$. It is most convenient to take as states of the chain the ordered pairs (j,k) , where j is the response on trial n , say, and k is the reinforcement on the preceding trial. Consider now the reinforcement k^* of the hypothesis of the theorem. Let j^* be a response such that $\lambda_{j^*k^*} \neq 0$. (There is at least one such j^* since $\sum_j \lambda_{jk} = 1$; in the Estes model $j^* = k^*$.) With the pair (j^*,k^*) as the state j_o of the infinite order chain, we shall establish (2.3) and (2.5).

To verify (2.3), we use (ii)b of the hypothesis and the following equalities and inequalities, which hold for all x and n :

$$P(\underline{A}_{n+m_o+1}=j^*,\underline{E}_{n+m_o}=k^*|x_n)$$

$$= \sum_{x_{m_o-1}} P(\underline{A}_{n+m_o+1}=j^*|\underline{E}_{n+m_o}=k^*,x_{m_o-1}+x_n)\cdot P(\underline{E}_{n+m_o}|x_{m_o-1}+x_n)P(x_{m_o-1}|x_n) \ .$$

Applying Axiom L, the right-hand side becomes:

$$= \sum_{x_{m_o-1}} [(1-\theta_{k^*})P(\underline{A}_{n+m_o}=j^*|x_{m_o-1}+x_n)+\theta_{k^*}\lambda_{j^*k^*}]P(\underline{E}^*_{n+m_o}=k^*|x_{m_o-1}+x_n)\cdot P(x_{m_o-1}|x_n)$$

$$\geq \theta_{k^*}\lambda_{j^*k^*} \sum_{x_{m_o-1}} P(\underline{E}_{n+m_o}=k^*|x_{m_o-1}+x_n)P(x_{m_o-1}|x_n)$$

$$\geq \theta_{k^*}\lambda_{j^*k^*}P(\underline{E}_{n+m_o}=k^*|x_n)$$

$$\geq \theta_{k^*}\lambda_{j^*k^*}\delta^* \qquad \text{by (ii)b} \ .$$

To establish (2.5), consider the following equalities and inequalities:

$$(4.3)\quad |P(\underline{A}_{n'+1}=j,\ \underline{E}_{n'}=k \mid x+x') - P(\underline{A}_{n''+1}=j,\ \underline{E}_{n''}=k \mid x+x'')|$$

$$= \pi_{k,x_{m*}} |P(\underline{A}_{n'+1}=j \mid \underline{E}_{n'}=k,\ x+x') - P(\underline{A}_{n''+1}=j \mid \underline{E}_{n''}=k,\ x+x'')| ,$$

where x_{m*} means the last m^* terms of x, and where the sequence x contains at least m occurrences of k^*, with $m > m^*$. The equality follows from (i) of the hypothesis, for by virtue of (i)

$$\pi_{k,x_{m*}} = P(\underline{E}_{n'}=k \mid x+x') = P(\underline{E}_{n''}=k \mid x+x'') .$$

Applying Axiom L once to the right-hand side of (4.3) we get, ignoring $\pi_{k,x_{m*}}$:

$$|P(\underline{A}_{n'+1}=j \mid \underline{E}_{n'}=k,\ x+x') - P(\underline{A}_{n''+1}=j \mid \underline{E}_{n''}=k,\ x+x'')|$$

$$= (1-\Theta_k)\,|P(\underline{A}_{n'}=j \mid x+x') - P(\underline{A}_{n''}=j \mid x+x'')| .$$

We do not know that $\Theta_k \neq 0$, but as we apply Axiom L repeatedly, we obtain the factor $(1-\Theta_{k*})$ at least m times, so that

$$(4.4)\quad |P(\underline{A}_{n'+1}=j,\ \underline{E}_{n'}=k \mid x+x') - P(\underline{A}_{n''+1}=j,\ \underline{E}_{n''}=k \mid x+x'')|$$

$$\leq (1-\Theta_{k*})^m\,|P(\underline{A}_{n'-h}=j \mid x') - P(A_{n''-h} \mid x'')| ,$$

where h is the length of x.*/ The difference term on the right of this

*/ If all $\Theta_k \neq 0$, the original condition given in [3] would be satisfied; our weaker condition (2.5) allows inclusion of cases where some of the Θ_k are 0 (i.e. where there can be trials without a reinforcement).

inequality is not more than 1, so that from (4.4) we obtain the estimate for $m > m^*$

$$\epsilon_m \leq (1-\theta_{k^*})^m ,$$

whence

$$\sum_{m=0}^{\infty} \epsilon_m < \infty ,$$

which is (2.5).

On the basis of (2.3) and (2.5) we know from Theorem 2.4 that the asymptotic cross-moments of $\mathcal{L}$ exist and are independent of the initial distribution of responses. But

$$P(\underline{A}_n = j \mid x_{n-1}) = \sum_k P(\underline{A}_n = j, \underline{E}_{n-1} = k \mid x_{n-1}) ,$$

and so the moments $\alpha^{\nu}_{j,n}$ can be expressed as sums of the cross-moments for the infinite order chain $\mathcal{L}$, which insures the existence of the limiting moments (3.5) and that they do not depend upon initial conditions.

There are several remarks to be made about the theorem just proved. First, we observe that a simple sufficient (but not necessary) condition for (ii)b is

$$(4.5) \qquad \min_{x_{m^*}} \pi_{k^*, x_{m^*}} \neq 0 .$$

The interpretation of (4.5) is that the reinforcing event k^* has positive probability on every trial no matter what sequence x_{m^*} of responses and reinforcements preceded. A number of interesting experimental cases of the linear model can be described in terms of (4.5), (i) and (ii)a of Theorem 4.1.

I. Contingent case with lag v. In the Estes model let $P(\underline{E}_n = k \mid \underline{A}_{n-v} = j, x) = \pi_{kj}(v)$, for all x such that $P(\underline{A}_{n-v} = j, x) > 0$. To satisfy (4.5), we need only that for some k, $\pi_{jk}(v) \neq 0$ for all j. Experimental data for $v = 0,1,2$ are given in Estes [5].

II. Double contingent case. Let $P(\underline{E}_n = k \mid \underline{A}_n = j, \underline{A}_{n-1} = j', x) = \pi_{k,jj'}$, for all x such that $P(\underline{A}_n = j, \underline{A}_{n-1} = j', x) > 0$.

Then (i) of Theorem (4.1) is immediately satisfied, and for (ii)a and (4.5) we need a k such that $\Theta_k \neq 0$ and for all j and j', $\pi_{k,jj'} \neq 0$.

An interesting fact about (I) and (II) is that although they are simple to test experimentally and their asymptotic response moments exist on the basis of Theorem 4.1, there is no known constructive method for computing the actual asymptotes. (The Estes [5] test of (I) excludes non-reinforced trials which cause the computational difficulties.) It may also be noted that the convergence theorems in Karlin [8] do not in general apply to (II), and apply to (I) only if $v = 0$.

On the basis of the proof of Theorem 4.1 we may, by virtue of Theorem 2.2, conclude that the asymptotic joint probabilities of successive responses also exist:

Corollary 1. *If the hypothesis of Theorem 4.1 is satisfied, then for every* m *the limit as* $n \to \infty$ *of* $P(\underline{A}_{n+m} = j_m, \underline{A}_{n+m-1} = j_{m-1}, \ldots, \underline{A}_n = j_o)$ *exists.*

We may regard the quantities $P(A_n = j \mid x_{n-1})$, for $1 \leq j \leq r$ as a random probability vector with an arbitrary joint distribution F_1 on trial 1, and distribution F_n on trial n. The following corollary is a consequence of the existence of the moments α_j^ν independent of the initial response probabilities.

Corollary 2. *If the hypothesis of Theorem 4.1 is satisfied, then there is a unique asymptotic distribution* F_∞, *independent of* F_1, *to which the distributions* F_n *converge.*

For the multiperson situations characterized by Axioms I and M, we have a theorem analogous to Theorem 4.1. For use in the hypothesis of this theorem we define the notion of *reinforcement schedule with past dependence of length* m, exactly as we did in (4.1), namely, we have such a schedule if for all k, $1 \leq i \leq s$, all n and n' with $n, n' > m$ and all x_m, x' and x''

$$\pi_{k^{(1)},\ldots,k^{(s)},x_m} = P(\underline{E}_n^{(1)}=k^{(1)},\ldots,\underline{E}_n^{(s)}=k^{(s)} \mid x_m+x') = P(E_n^{(1)}=k^{(1)},\ldots,E_n^{(s)}=k^{(s)} \mid x_m+x'').$$

Theorem 4.2. *Let* $\mathcal{M}$ *be an* s-*person linear model such that*

(i) $\mathcal{M}$ *has a reinforcement schedule with past dependence of length* m^*,

(ii) <u>there are integers</u> $k^{(i)*}$, <u>for</u> $1 \leq i \leq s$, <u>such that</u>

(a) $\theta^{(i)}_{k^{(i)*}} \neq 0$,

(b) <u>there is a</u> δ^* <u>and an</u> m_o <u>such that for all sequences</u> x <u>and all integers</u> n

$$P(\underline{E}^{(1)}_{n+m_o} = k^{(1)*}, \ldots, \underline{E}^{(s)}_{n+m_o} = k^{(s)*} | x_n) \geq \delta^* > 0 \quad .$$

<u>Then the asymptotic moments</u> $\gamma^{\nu}_{j^{(1)},j^{(2)},\ldots,j^{(s)}}$ <u>of</u> $\mathcal{M}$ <u>all exist and are independent of the initial distribution of responses</u>.

<u>Proof</u>: The states of the chain are now defined as 2s-tuples $(j^{(1)},\ldots,j^{(s)}, k^{(1)},\ldots, k^{(s)})$, where $j^{(i)}$ is the response made by the i^{th} subject and $k^{(i)}$ is the reinforcement for that subject on the preceding trial. Using the reinforcements $k^{(i)*}$ of the hypothesis, let $j^{(i)*}$ be such that $\lambda^{(i)}_{j^{(i)*}k^{(i)*}} \neq 0$. We take $(j^{(1)*},\ldots,j^{(s)*}, k^{(1)*},\ldots,k^{(s)*})$ as the state j_o for which we establish (2.3) and (2.5).

To verify (2.3) we proceed exactly as in the proof of Theorem 4.1, applying now Axioms I and M instead of L, and we obtain that

$$P(\underline{A}^{(1)}_{n+m_o+1}=j^{(1)*},\ldots,\underline{A}^{(s)}_{n+m_o+1}=j^{(s)*}, \underline{E}^{(1)}_{n+m_o}=k^{(1)*},\ldots,\underline{E}^{(s)}_{n+m_o}=k^{(s)*}|x_n)$$

$$\geq \prod_{i=1}^{s} \theta^{(i)}_{k^{(i)*}} \lambda^{(i)}_{j^{(i)*}k^{(i)*}} \delta^* \quad .$$

For (2.5), we first observe that by virtue of (i) of the hypothesis and Axiom I

$$|P(\underline{A}_{n'+1}^{(1)} = j^{(1)},\ldots, \underline{A}_{n'+1}^{(s)} = j^{(s)}; \underline{E}_{n'}^{(1)} = k^{(1)},\ldots, \underline{E}_{n'}^{(s)} = k^{(s)} \mid x+x')$$

$$- P(\underline{A}_{n''+1}^{(1)} = j^{(1)},\ldots, \underline{A}_{n''+1}^{(s)} = j^{(s)}; \underline{E}_{n''}^{(1)} = k^{(1)},\ldots, \underline{E}_{n''}^{(s)} = k^{(s)} \mid x+x'')| =$$

$$\pi_{k^{(1)},\ldots,k^{(s)},x_{m*}} \left|\prod_{i=1}^{s} P(\underline{A}_{n'+1}^{(i)}=j^{(i)} \mid \underline{E}_{n'}^{(i)}=k^{(i)},x+x') - \prod_{i=1}^{s} P(\underline{A}_{n''+1}^{(i)}=j^{(i)} \mid \underline{E}_{n''}^{(i)}=k^{(i)},x+x'')\right|.$$

We notice next that the right-hand side is

$$\leq \pi_{k^{(1)},\ldots,k^{(s)},x_{m*}} [P(\underline{A}_{n'+1}^{(1)}=j^{(1)} \mid \underline{E}_{n'}^{(1)}=k^{(1)},x+x') \left|\prod_{i=2}^{s} P(\underline{A}_{n'+1}^{(i)}=j^{(i)} \mid \underline{E}_{n'}^{(i)}=k^{(i)},x+x')\right.$$

$$\left.- \prod_{i=2}^{s} P(\underline{A}_{n''+1}^{(i)}=j^{(i)} \mid \underline{E}_{n''}^{(i)}=k^{(i)},x+x'')\right| + \prod_{i=2}^{s} P(\underline{A}_{n''+1}^{(i)}=j^{(i)} \mid \underline{E}_{n''}^{(i)}=k^{(i)},x+x'') \cdot$$

$$|P(\underline{A}_{n'+1}^{(1)}=j^{(1)} \mid \underline{E}_{n'}^{(1)}=k^{(1)},x+x') - P(\underline{A}_{n''+1}^{(1)}=j^{(1)} \mid \underline{E}_{n''}^{(1)}=k^{(1)},x+x')| \quad .$$

Continuing this same development, we obtain:

$$\leq \sum_{i=1}^{s} | P(\underline{A}_{n'+1}^{(i)}=j^{(i)} \mid \underline{E}_{n'}^{(i)}=k^{(i)},x+x') - P(\underline{A}_{n''+1}^{(i)}=j^{(i)} \mid \underline{E}_{n''}^{(i)}=k^{(i)},x+x'')| \quad .$$

And by the line of reasoning used in the proof of Theorem 4.1, if the sequence x contains state $(j^{(1)*},\ldots,k^{(s)*})$ at least m times the last quantity is

$$\leq \sum_{i=1}^{s} (1 - \theta_{k^{(i)*}}^{(i)})^{m} \quad .$$

Provided $m > m*$ this inequality yields an estimate of ϵ_m from which we conclude that (2.5) holds. The existence of the asymptotic moments then follows from the theory of Section 2 as in the case of Theorem 4.1. Q.E.D.

A pair of corollaries follow from the theorem just proved which are exactly like the two given after Theorem 4.1.

REFERENCES

[1] Atkinson, Richard C. and Suppes, Patrick, "An analysis of two-person game situations in terms of statistical learning theory," _J. of Experimental Psychology_, vol. 55 (1958), pp. 369-378.

[2] Bush, Robert R. and Mosteller, Frederick, _Stochastic Models for Learning_, New York, 1955.

[3] Doeblin, W. and Fortet, R., "Sur des chaînes à liaisons complètes," _Bull. Soc. Math. France_, vol. 65 (1937), pp. 132-148.

[4] Estes, W. K., "Theory of learning with constant, variable, or contingent probabilities of reinforcement," _Psychometrika_, vol. 22 (1957), pp. 113-132.

[5] Estes, W. K., "Of models and men," _Amer. Psychologist_, vol. 12 (1957), pp. 609-617.

[6] Estes, W. K. and Suppes, Patrick, _Foundations of Statistical Learning Theory_, I. _The Linear Model for Simple Learning_, Technical Report No. 16, Contract Nonr 225(17), Applied Mathematics and Statistics Laboratory, Stanford University, 1957.

[7] Harris, T. E., "On chains of infinite order," _Pacific J. of Math._, vol. 5 (1955), pp. 707-724.

[8] Karlin, Samuel, "Some random walks arising in learning models I," _Pacific J. of Math._, vol. 3 (1953), pp. 725-756.

[9] Kennedy, Maurice, "A convergence theorem for a certain class of Markoff processes," _Pacific J. of Math._, vol. 7 (1957), pp. 1107-1124.

STANFORD UNIVERSITY

Technical Reports Distribution List

Contract Nonr 225(17)

(NR 171-034)

Armed Services Technical
Information Agency
Arlington Hall Station
Arlington 12, Virginia 5

Commanding Officer
Office of Naval Research
Branch Office
Navy No. 100, Fleet Post Office
New York, New York 35

Director, Naval Research
Laboratory
Attn: Technical Information
Officer
Washington 25, D. C. 6

Office of Naval Research
Group Psychology Branch
Code 452
Department of the Navy
Washington 25, D. C. 5

Office of Naval Research
Branch Office
346 Broadway
New York 13, New York 1

Office of Naval Research
Branch Office
1000 Geary Street
San Francisco 9, Calif. 1

Office of Naval Research
Branch Office
1030 Green Street
Pasadena 1, California 1

Office of Naval Research
Branch Office
Tenth Floor
The John Crerar Library Building
86 East Randolph Street
Chicago 1, Illinois 1

Office of Naval Research
Logistics Branch, Code 436
Department of the Navy
Washington 25, D. C. 1

Office of Naval Research
Mathematics Division, Code 430
Department of the Navy
Washington 25, D. C. 1

Operations Research Office
6935 Arlington Road
Bethesda 14, Maryland
Attn: The Library 1

Office of Technical Services
Department of Commerce
Washington 25, D. C. 1

The Logistics Research Project
The George Washington University
707 - 22nd Street, N.W.
Washington 7, D. C. 1

The RAND Corporation
1700 Main Street
Santa Monica, California
Attn: Dr. John Kennedy 1

Library
Cowles Foundation for Research
in Economics
Box 2125, Yale Station
New Haven, Connecticut 1

Center for Philosophy of Science
University of Minnesota
Minneapolis 14, Minnesota 1

Institut für Math. Logik
Universitat
Schlossplatz 2
Münster in Westfalen
Germany 1

Professor Ernest Adams
Department of Philosophy
University of California
Berkeley 4, California 1

Professor Maurice Allais
15 Rue des Gates-Ceps
Saint-Cloud (S.-O.)
France 1

Professor Alan Ross Anderson
Department of Philosophy
Yale University
New Haven, Connecticut 1

Professor Norman H. Anderson
Department of Psychology
University of California
Los Angeles 24, California 1

Professor T. W. Anderson
Department of Statistics
Columbia University
New York 27, New York 1

Professor K. J. Arrow
Serra House
Stanford University 1

Professor Richard C. Atkinson
Department of Psychology
University of California
Los Angeles 24, California 1

Dr. R. F. Bales
Department of Social Relations
Harvard University
Cambridge, Massachusetts 1

Professor Alex Bavelas
Department of Psychology
Stanford University 1

Professor Gustav Bergman
Department of Philosophy
State University of Iowa
Iowa City, Iowa 1

Professor E. W. Beth
Bern, Zweerskade 23, I
Amsterdam, Z.,
The Netherlands 1

Professor Max Black
Department of Philosophy
Cornell University
Ithaca, New York 1

Professor David Blackwell
Department of Statistics
University of California
Berkeley 4, California 1

Professor Lyle E. Bourne, Jr.
Department of Psychology
University of Utah
Salt Lake City, Utah 1

Mr. Gordon Bower
Department of Psychology
Yale University
New Haven, Connecticut 1

Professor R. B. Braithwaite
King's College
Cambridge, England 1

Professor C. J. Burke
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor R. R. Bush
106 College Hall
University of Pennsylvania
Philadelphia 4, Pennsylvania 1

Dr. Donald Campbell
Department of Psychology
Northwestern University
Evanston, Illinois 1

Professor Rudolf Carnap
Department of Philosophy
University of California
Los Angeles 24, California 1

Professor Edward C. Carterette
Department of Psychology
University of California
Los Angeles 24, California 1

Professor C. West Churchman
School of Business Administration
University of California
Berkeley 4, California 1

Dr. Clyde H. Coombs
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. Gerard Debreu
Cowles Commission
Box 2125, Yale Station
New Haven, Connecticut 1

Dr. Mort Deutsch
Bell Telephone Laboratories
Murray Hill, New Jersey 1

Professor Robert Dorfman
Department of Economics
Harvard University
Cambridge 38, Massachusetts 1

Dr. Ward Edwards
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Dr. Jean Engler
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Professor W. K. Estes
Department of Psychology
Indiana University
Bloomington, Indiana 1

Professor Robert Fagot
Department of Psychology
University of Oregon
Eugene, Oregon 1

Dr. Leon Festinger
Department of Psychology
Stanford University 1

Professor M. Flood
Willow Run Laboratories
Ypsilanti, Michigan 1

Professor Maurice Fréchet
Institut H. Poincaré
11 Rue P. Curie
Paris 5, France 1

Dr. Milton Friedman
Department of Economics
University of Chicago
Chicago 37, Illinois 1

Dr. Eugene Galanter
Center for Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California 1

Dr. Murray Gerstenhaber
University of Pennsylvania
Philadelphia, Pennsylvania 1

Dr. I. J. Good
25 Scott House
Cheltenham, England 1

Dr. Leo A. Goodman
Statistical Research Center
University of Chicago
Chicago 37, Illinois 1

Professor Nelson Goodman
Department of Philosophy
University of Pennsylvania
Philadelphia, Pennsylvania 1

Professor Harold Gulliksen
Educational Testing Service
20 Nassau Street
Princeton, New Jersey 1

Professor Louis Guttman
Israel Institute of Applied
Social Research
David Hamlech No. 1
Jerusalem, Israel 1

Mr. John Harsanyi
2/165 Victoria Road
Bellevue Hill, N.S.W.
Australia 1

Dr. T. T. ten Have
Social - Paed. Instituut
Singel 453
Amsterdam, Netherlands 1

Professor Carl G. Hempel
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Professor Leonid Hurwicz
Serra House
Stanford University 1

Professor Lyle V. Jones
Department of Psychology
University of North Carolina
Chapel Hill, North Carolina 1

Professor Donald Kalish
Department of Philosophy
University of California
Los Angeles 24, California 1

Dr. Leo Katz
Department of Mathematics
Michigan State College
East Lansing, Michigan 1

Professor John G. Kemeny
Department of Mathematics
Dartmouth College
Hanover, New Hampshire 1

Professor T. C. Koopmans
Cowles Foundation for Research
in Economics
Box 2125, Yale Station
New Haven, Connecticut 1

Professor W. Kruskal
Department of Statistics
Eckart Hall 127
University of Chicago
Chicago 37, Illinois 1

Dr. David La Berge
Department of Psychology
University of Minnesota
Minneapolis 14, Minnesota 1

Professor Douglas Lawrence
Department of Psychology
Stanford University 1

Dr. Duncan Luce
Department of Social Relations
Harvard University
Cambridge 38, Massachusetts 1

Professor Robert McGinnis
Department of Sociology
University of Wisconsin
Madison 6, Wisconsin 1

Dr. W. G. Madow
Engineering Research Division
Stanford Research Institute
Menlo Park, California 1

Professor Jacob Marschak
Box 2125, Yale Station
New Haven, Connecticut 1

Dr. Samuel Messick
Educational Testing Service
Princeton University
Princeton, New Jersey 1

Professor G. A. Miller
Center for Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California 1

Professor Richard C. Montague
Department of Philosophy
University of California
Los Angeles 24, California 1

Dr. O. K. Moore
Department of Sociology
Box 1965, Yale Station
New Haven, Connecticut 1

Professor Sidney Morgenbesser
Department of Philosophy
Columbia University
New York 27, New York 1

Professor Oskar Morgenstern
Department of Economics
and Social Institutions
Princeton University
Princeton, New Jersey 1

Professor Frederick Mosteller
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Professor Ernest Nagel
Department of Philosophy
Columbia University
New York 27, New York 1

Dr. Theodore M. Newcomb
Department of Psychology
University of Michigan
Ann Arbor, Michigan 1

Professor A. G. Papandreou
Department of Economics
University of California
Berkeley 4, California 1

Dr. Juliette Popper
Department of Psychology
University of Kansas
Lawrence, Kansas 1

Dr. Hilary Putnam
Department of Philosophy
Princeton University
Princeton, New Jersey 1

Professor Willard V. Quine
Center for Behavioral Sciences
202 Junipero Serra Boulevard
Stanford, California 1

Professor Roy Radner
Department of Economics
University of California
Berkeley 4, California 1

Professor Howard Raiffa
Department of Statistics
Harvard University
Cambridge 38, Massachusetts 1

Professor Nicholas Rashevsky
University of Chicago
Chicago 37, Illinois 1

Dr. Frank Restle
Department of Psychology
Michigan State University
East Lansing, Michigan 1

Professor David Rosenblatt
American University
Washington 6, D. C. 1

Professor Alan J. Rowe
Management Sciences Research
Project
University of California
Los Angeles 24, California 1

Professor Herman Rubin
Department of Mathematics
University of Oregon
Eugene, Oregon 1

Dr. I. Richard Savage
School of Business
University of Minnesota
Minneapolis, Minnesota 1

Professor L. J. Savage
Committee on Statistics
University of Chicago
Chicago, Illinois 1

Dr. Dana Scott
Eckart Hall
University of Chicago
Chicago 37, Illinois 1

Dr. C. P. Seitz
Special Devices Center
Office of Naval Research
Sands Point
Port Washington
Long Island, New York 1

Dr. Marvin Shaw
School of Industrial Management
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge 39, Massachusetts 1

Mr. R. L. Shuey
General Electric Research Lab.
Schenectady, New York 1

Dr. Sidney Siegel
Department of Psychology
Penn. State University
University Park, Pennsylvania 1

Professor Herbert Simon
Carnegie Institute of Technology
Schenley Park
Pittsburgh, Pennsylvania 1

Dr. Herbert Solomon
Department of Statistics
Stanford University 1

Professor K. W. Spence
Psychology Department
State University of Iowa
Iowa City, Iowa 1

Dr. F. F. Stephan
Box 337
Princeton University
Princeton, New Jersey 1

Mr. Saul Sternberg
Department of Social Relations
Emerson Hall
Harvard University
Cambridge 38, Massachusetts 1

Professor S. Smith Stevens
Memorial Hall
Harvard University
Cambridge 38, Massachusetts 1

Dr. Donald W. Stilson
Department of Psychology
University of Colorado
Boulder, Colorado 1

Dr. Dewey B. Stuit
108 Schadffer Hall
State University of Iowa
Iowa City, Iowa 1

Professor Alfred Tarski
Department of Mathematics
University of California
Berkeley 4, California 1

Professor G. L. Thompson
Department of Mathematics
Ohio Wesleyan
Delaware, Ohio 1

Dr. Robert L. Thorndike
Teachers College
Columbia University
New York, New York 1