Offprints From
Handbook of Mathematical Psychology, Volume II
edited by
Luce, Bush and Galanter
Published by John Wiley & Sons, Inc., 1963

9

Stochastic Learning Theory[1]

Saul Sternberg
University of Pennsylvania

1. Preparation of this chapter was supported by Grant G-18630 from the National Science Foundation to the University of Pennsylvania. Doris Aaronson provided valuable help with computations; her work was supported in part by NSF Grant G-14839. I wish to thank Francis W. Irwin for his helpful criticism of the manuscript.

Contents

Stochastic Learning Theory

The process of learning in an animal or a human being can often be analyzed into a series of choices among several alternative responses. Even in simple repetitive experiments performed under highly controlled conditions, the choice sequences are typically erratic, suggesting that probabilities govern the selection of responses. It is thus useful to think of the systematic changes in a choice sequence as reflecting trial-to-trial changes in response probabilities. From this point of view, much of the study of learning is concerned with describing the trial-to-trial probability changes that characterize a stochastic process.

In recent mathematical studies of learning investigators have assumed that there is *some* stochastic process to which the behavior in a simple learning experiment conforms. This is not altogether a new idea (for a sketch of its history, see Bush, 1960b). But two important features appear primarily in the work since 1950 that was initiated by Bush, Estes, and Mosteller. First, the step-by-step nature of the learning process has been an explicit feature of the proposed models. Second, these models have been analyzed and applied in ways that do not camouflage their statistical aspect. Various models have been proposed and studied as possible approximations to the stochastic processes of learning. The purpose of this chapter is to review some of the methods and problems of formulating models, analyzing their properties, and applying them in the analysis of learning data.[2]

The focus of our attention is a simple type of learning experiment. Each of a sequence of trials consists of the selection of a response alternative by the subject followed by an outcome provided by the experimenter. The response alternative may be pressing one of a set of buttons, turning right in a maze, jumping over a barrier before a shock is delivered, or failing to recall a word.[3]

[2] The model enterprise is not and should not be separate from other efforts in the study of learning. It is partly for this reason that I have not attempted a summary of present knowledge vis-à-vis various models. A good proportion of the entire learning literature is directly relevant to many of the questions raised in work with stochastic models. For this reason an adequate survey would be gargantuan and soon outdated.

[3] The reader will note from these examples that the terms "choice" and "response alternative" are used in the abstract sense discussed in Chapter 2 of Volume I. For example, I ignore the question whether a choice represents a conscious decision; classes of responses are defined both spatially (as in the maze) and temporally (as in the shuttlebox); a subject's inability to recall a word is grouped with his "choice" not to say it.

We shall be concerned almost entirely with experiments in which the subject's behavior is partitioned into two mutually exclusive and exhaustive response alternatives. The outcome may be a pellet of food, a shock, or the onset of one of several lights. The outcome may or may not change from trial to trial and its occurrence may or may not depend on the response chosen. When no differential outcome (Irwin, 1961) is provided, as, for example, when the experimenter gives food on every trial or when he does not explicitly provide any reward or punishment, we think of the experiment simply as a sequence of response choices. We do not consider experiments in which the stimulus situation is deliberately altered from trial to trial; for our purposes it can be referred to once and then ignored. Because little mathematical work has been done on bar-pressing or runway experiments, they receive little attention.

The elements of a stochastic learning model correspond to the components of the experiment. The sequence of trials is indexed by $n = 1, 2, \ldots, N$. There is a set of *response alternatives*, $\{A_1, \ldots, A_j, \ldots, A_r\}$, and a set of *outcomes*, $\{O_1, O_2, \ldots, O_s\}$. Each response-outcome pair constitutes a possible *experimental trial event*, E_k. A probability distribution, $\{p_{i,n}(1), \ldots, p_{i,n}(j), \ldots, p_{i,n}(r)\}$, is defined over the set of response alternatives for each subject, $i = 1, 2, \ldots, I$, and each trial, $n = 1, 2, \ldots, N$. The subject subscript is suppressed when we consider a generic sequence. The response probabilities form a *probability vector* with r elements. In most of the examples that have been studied $r = 2$ and $p_n(1) = 1 - p_n(2)$, thus making it possible to reduce the sequence of probability vectors to a sequence of scalars, $p_1, \ldots, p_n, \ldots$.

The crux of a model is its description of response-probability changes from trial to trial. One type of description is in terms of explicit transition rules or *operators*, usually independent of the trial number, that transform the response probabilities of trial n into those of trial $n + 1$. The operator invoked depends on the event that occurs on trial n. A second type of description is in terms of an explicit formula for the dependence of p_n on both n and the sequence of events through trial $n - 1$. The explicit formula approach, although less popular, is somewhat more general, as we shall see, because it can be used for models whose expression in terms of operators of the type mentioned would be cumbersome or impossible.

One or more parameters with unspecified numerical values usually appear in the formulation of a model. These parameters may be initial probability values or they may be quantities that reflect the magnitude of the effects of different trial events. The values of these parameters are usually estimated from the set of data being analyzed. In more stringent tests of a model parameters estimated from one phase of an experiment are used in its application to another phase. A useful, if rough, distinction can be made between a *model type*, consisting of the class of models with all possible values of the parameters, and a *model* in which the numerical values of parameters are specified. Because there is no theory of parameters in this field, investigators have concentrated on determining which model type, if any, can describe a set of data, rather than which model within a type is the appropriate one. Model types themselves can be grouped into different families which reflect basically different conceptions of the learning process, and the choice between families is the fundamental problem. In this chapter, as in much of the published work, the word "model" is used to denote a model type when it is clear from the context what is meant.

The model builder's view of learning differs in its emphasis from that of many experimenters. The idea of trial-to-trial changes in the behavior of individual subjects has been basic in traditional approaches to learning. But, with few exceptions, the changes have been investigated indirectly, often by considering the effects of experimental variations on the gross shape of a curve of mean performance versus trials. Stochastic models have been used increasingly to supplement an interest in the learning curve with analyses of various sequential features of the data, features that reflect more directly the operation of the underlying trial-to-trial changes.

At first glance stochastic models appear to have been remarkably successful in accounting for the data from a variety of learning experiments (Bush & Mosteller, 1955; Bush & Estes, 1959). Recent work, however, suggests that we view the situation with caution. As more model types are investigated, the problem becomes not one of fitting a model to a set of data but of discrediting all but one of the competing models. Apparent agreement between model and data comes easily and can lead to a sense of over-confidence. There are a number of ways of dealing with this problem, such as refining estimation and testing procedures and performing crucial experiments. Criteria have been invoked that involve more than merely the ability of a model to describe a particular set of data. A model is designed to describe some process that occurs in an experiment. But in the present state of the art it is difficult to perform experiments in which no other processes, aside from those described by the model, intrude. We must therefore compromise between making severer demands on the models and acknowledging that at present their descriptions of actual experiments cannot hope to be more than approximations.

1. ANALYSIS OF EXPERIMENTS AND MODEL IDENTIFICATION

In considering what may happen on a trial, we must draw a careful distinction between *experimental events* and *model events*. Not all the events in an experiment are identified with distinct events in the model that is applied to it. A good deal of intuition, with varying degrees of experimental support, leads to assumptions of equivalence and complementarity among experimental events.[4] These assumptions have strong substantive implications. They also fix the number of distinct operators (events) in the model, impose constraints on them, and specify how their application to response probabilities is governed. In general, the more radical the assumptions, the simpler the model, the less the labor in analysis and application, and the greater the chance that the model will fail. A few examples will serve to illustrate some of the relevant considerations.

1.1 Equivalent Events

Each response-outcome pair constitutes an experimental event. Examples of several sets of experimental events are given in Table 1.[5]

At this level of analysis each experiment has four possible events per trial. In one analysis of the prediction experiment (e.g., Estes & Straughan, 1954) the four experimental events are grouped into two equivalence classes, $\{(A_1, O_1), (A_2, O_1)\}$ and $\{(A_1, O_2), (A_2, O_2)\}$, which define the two model events, E_1 and E_2. This amounts to assuming that changes in

[4] Throughout this chapter adjectives such as "equivalent," "complementary," "path-independent," and "commutative" are applied to the term "events." In all cases this is a shorthand way of describing properties of the effects on response probabilities of the occurrence of events. These properties may characterize model events. Whether they also characterize corresponding experimental events is a question to be answered by testing the model.

Occasionally I write as if an event were an active agent, as in "the event transforms the probability . . . ." This is another shorthand form. It stands for "the operator corresponding to the event transforms . . ." in the case of model events with operator representations. It stands for "the occurrence of the event affects the organism so as to change its probability . . ." in the case of experimental events.

The use of the same terms in talking about both kinds of events is intended to emphasize the fact that insofar as a model is successful the properties of its events are also properties of the corresponding experimental events.

[5] Choices have to be made even at the stage of tabulating response and outcome possibilities, as illustrated by the difference between the tables for T-maze and prediction experiments. An alternative analysis of the T-maze experiment, formally identical to the one for the prediction experiment, is illustrated in Table 2. One relevant consideration is the type of experiment to which the analysis may be generalized. Thus, if in the prediction experiment we defined the outcomes to be O_1:correct, O_2:incorrect, then the generalization to an experiment with three buttons and three lights might be inappropriate. For the T-maze experiment the analysis in Table 1 is to be preferred if we wish to include experiments in which on some trials neither or both maze arms are baited. On the other hand, Table 2 provides an analysis that is more easily extended to experiments with a correction procedure.

Table 1 Definition of Experimental Events in Four Experiments

(i) Two-Choice Prediction

Response	Outcome
A_1: Left button press	O_1: Left light onset (correct)
A_2: Right button press	O_1: Left light onset (incorrect)
A_1: Left button press	O_2: Right light onset (incorrect)
A_2: Right button press	O_2: Right light onset (correct)

(ii) T-Maze

Response	Outcome
A_1: Left turn	O_1: Food
A_2: Right turn	O_2: No Food
A_1: Left turn	O_2: No Food
A_2: Right turn	O_1: Food

(iii) Escape-Avoidance Shuttlebox

Response	Outcome
A_1: Jump before US from left to right	O_1: Avoidance of US on left
A_1': Jump before US from right to left	O_1': Avoidance of US on right
A_2: Jump after US from left to right	O_2: Escape of US on left
A_2': Jump after US from right to left	O_2': Escape of US on right

(iv) Continuous Reinforcement in Runway

Response	Outcome
A_1: Run with speed in first quartile	O_1: Food
A_2: Run with speed in second quartile	O_1: Food
A_3: Run with speed in third quartile	O_1: Food
A_4: Run with speed in fourth quartile	O_1: Food

response probability from trial to trial depend only on outcomes and not on responses. Reward of A_1 by O_1 is assumed equivalent to nonreward of A_2 by O_1. This assumption, as we shall see, considerably simplifies models for the experiment. Although a comparable reduction in the number of events in the T-maze (or the analogous "two-armed bandit") experiment is possible, it has, in general, not been made (e.g., Galanter & Bush, 1959). Analysis of the shuttlebox experiment has ignored the alternation in the animal's starting position and used the equivalence classes $\{(A_1, O_1), (A_1', O_1')\}$ and $\{(A_2, O_2), (A_2', O_2')\}$ (Bush & Mosteller, 1955, 1959). Analyses of the runway experiment have grouped all experimental events into one equivalence class, resulting in a single model event (e.g., Bush & Mosteller, 1955). As an alternative, at least one theory about runway behavior proposes that the effect of a trial event depends critically on the running speed of that trial (Logan, 1960).

It should be emphasized that the reduction in the number of events by the definition of equivalence classes entails strong assumptions about what the experimental subjects are indifferent to. Even in forming the lists in Table 1 we have implicitly appealed to the existence of equivalence classes; we have assumed, for example, that all ways of turning left are equivalent.

1.2 Response Symmetry and Complementary Events

When we have determined the set of events $\{E_k\}$ for a model by defining whatever equivalence classes seem reasonable, we can introduce further simplifications by identifying pairs or sets of complementary events. In a two-choice experiment two events, E_1 and E_2, form a complementary pair if, to put it roughly, the effect of E_1 on p is the same as the effect of E_2 on $q = 1 - p$. If the model involves a set of operators, each associated with an event, then E_1 and E_2 will be associated with *complementary operators*.

Let us suppose, for example, that the operators are linear and that E_k transforms p into $Q_k p = \alpha_k p + a_k$, where α_k and a_k are constants. Then the complementarity of E_1 and E_2 requires that when E_2 occurs $q = 1 - p$ will be transformed into $\alpha_1 q + a_1$. This requirement implies that p is transformed into $\alpha_1 p + (1 - \alpha_1 - a_1)$ when E_2 occurs and gives the relations $\alpha_2 = \alpha_1$ and $a_2 = 1 - \alpha_1 - a_1$. The result is that we have one operator and its complement rather than two independent operators.

As in the case of equivalence classes of events, it is the subject's behavior, not the experimenter, that determines whether two events are complementary. In the analysis of prediction experiments it has frequently been assumed that the two equivalence classes are complementary. In analysis of the T-maze experiment the event pairs $\{(A_1, O_1), (A_2, O_1)\}$ and $\{(A_1, O_2), (A_2, O_2)\}$ have been assumed to be complementary. In their treatment of an experiment on imitation Bush and Mosteller (1955) rejected the assumption that rewarding an imitative response was complementary to rewarding a nonimitative response.

It is in dealing with pairs of events in which the same outcome (a food reward, for example) occurs in conjunction with a pair of "symmetric" responses (left turn and right turn, for example) that investigators have been most inclined to assume that the events are complementary. There appears, however, to be no available response theory that would allow us to determine, from the properties of two (or more) responses, whether they are symmetric in the desired sense. Learning model analyses of a variety of experiments would provide one source of information on which such a theory could be based.

At present, therefore, it is primarily intuition that leads us to assume that left and right are symmetric in a sense in which imitation and nonimitation are not. Perhaps more obvious examples of asymmetric responses are alternatives A_1 and A_2 in the shuttlebox experiment and alternatives A_1 and A_4 in the runway (see Table 1).

In the foregoing discussion I have considered the relation between the events determined by two responses for each of which the same outcome is provided, such as "left turn—food" and "right turn—food." A second sense in which response symmetry may be invoked in the design of learning model operators arises when we consider the effects of the same event (such as "left turn—food") on the probabilities of two different responses. In many models the operators that represent the effects of an event are of the same form for all responses; that is, the operators are members of a restricted family, such as multiplicative or linear transformations. These models are, therefore, invariant under a reassignment of labels to responses, so long as the values of one or two parameters are altered. Such invariance represents a second type of response symmetry.

This type of symmetry may be defined only in relation to a specified family of operators; when such symmetry obtains, then the family is *complete* (Luce, 1963) in the sense that it contains the operators appropriate to all the responses.

As an example, let us consider the Bush-Mosteller model for two responses, in which $p = \Pr\{A_1\}$ and the occurrence of E_k transforms p into $\alpha_k p + a_k$. This is, of course, equivalent to the transformation of $q = 1 - p$ into $\alpha_k' q + a_k'$, where $\alpha_k' = \alpha_k$ and $a_k' = 1 - a_k - \alpha_k$. As a result of the occurrence of E_k, the probabilities of the two responses change in the same (linear) manner. The model for changes in $\Pr\{A_1\}$ is of the same form as the model for changes in $\Pr\{A_2\}$.

Not all learning models are characterized by this sort of response-symmetry relative to a simple family of operators. For example, Hull's model (1943, Chapter 18) for changes in the probability of a response, A_1, where the alternative response, A_2, is defined as nonoccurrence of A_1, incorporates a threshold notion in the relationship between the "strength" of A_1 (its ${}_SE_R$) and its probability. No such threshold applies to A_2, and the responses are not symmetric in the sense outlined. (This model is discussed in Secs. 2.5 and 4.1.)

A second model that lacks the symmetry features is Bush and Mosteller's (1959) "late Thurstone model" discussed in Sec. 2.6. In this model the transformations induced by events on the probability of error can be expressed by applying an additive increment to the reciprocal of the probability:

$$\frac{1}{p_{n+1}} = \frac{1}{p_n} + b.$$

The form of the corresponding transformation on the reciprocal of $q_n = 1 - p_n$ is not a member of the family of additive (or even full linear) transformations.

For any two-response model in which the effects of events may be expressed in terms of operators, we can use the response-symmetry condition to impose a restriction on the class of allowed operators. We can do this by translating the condition into the requirement that the operators on $p = \Pr\{A_1\}$ and $q = 1 - p = \Pr\{A_2\}$ that correspond to an event are to be members of the same family of operators. If, for example, we require the family to be expressed by a particular function with two parameters,

$$p_{n+1} = f(p_n; a, b),$$

then symmetry dictates that for all p, $0 \leqslant p \leqslant 1$, and for all allowed values of a and b, f has the property that

$$f(p; a, b) + f(1 - p; c, d) = 1,$$

where $c = c(a, b)$ and $d = d(a, b)$. As indicated by the discussion of the "late Thurstone model," when we require that $f(p)$ be of the form $f(p) = [(a/p) + b]^{-1}$, then its operators do not satisfy the condition. The condition may be generalized in an obvious way to more than two responses.

1.3 Outcome Symmetry

The reader may have questioned the contrast between the treatments of the prediction and the T-maze experiments. In the first experimental events containing different responses are grouped in the same equivalence class, whereas this is not done in the second. The critical difference that guides the definition of equivalence classes here appears to be the extent of outcome symmetry from what is thought to be the subject's viewpoint. From the experimenter's viewpoint the possible outcomes in many T-maze experiments could be symmetrically described by "left-arm baited" and "right-arm baited." With this terminology we can form a new event list that is identical, formally, to the list for the prediction experiment.

Table 2 Alternative Definition of Experimental Events in T-Maze Experiment

Response	Outcome
A_1: Left turn	O_1: Left-arm baited (correct)
A_2: Right turn	O_1: Left-arm baited (incorrect)
A_1: Left turn	O_2: Right-arm baited (incorrect)
A_2: Right turn	O_2: Right-arm baited (correct)

Despite their formal identification with the events in the prediction experiment, many model builders would be loth to assume that the first two events are equivalent in the T-maze experiment. The distinction between the two experiments is more easily seen if they are extended to three alternatives. Food presented after one of three alternatives is unlikely to have the same effect as the absence of food after another of the three. It is perhaps conceivable that the onset of one of three lights would have the same effect independently of the response that precedes it. The important criterion seems to be not whether the outcomes are capable of symmetric description by the experimenter but whether they "appear" symmetric to the subject. The question whether outcomes are symmetric is, of course, finally decided by whether the behavior produced by subjects is described by a model in which symmetry is assumed.

1.4 The Control of Model Events

Experimental events, with assumptions about their equivalence and complementarity, determine a set of model events and thereby give rise to four important classes of models. These classes are defined in terms of how the occurrence of model events is controlled by the sequence of response-outcome pairs in the experiment.

If knowledge of both response and outcome is needed in order to know which model event has occurred on a trial, then the events are *experimenter-subject controlled.* For example, in the analysis of the T-maze experiment given in Table 1 there are four model events, and both the direction of the rat's turn and the schedule of rewards must be known in order to determine which event has occurred. Although the reward schedule may be predetermined, the response is not. Therefore the sequence of probability changes cannot be specified in advance of the experiment. This class of models is relatively intractable mathematically.

The second class of models is illustrated by the Bush and Mosteller (1955) analysis of the shuttlebox experiment. Here there are only two model events (avoid or escape) and the response determines the outcome (no-shock or shock). This is an example of *subject-controlled events* in which the response alone determines the model event. Any experiment in which responses and outcomes are perfectly correlated consists of subject-controlled events. This correlation is produced in the T-maze by baiting the left arm on every trial and never baiting the right arm, for example. Again, the sequence of probability changes cannot be specified in advance and in general will be different for each subject in the experiment. Even if the real subjects correspond to a set of *identical model subjects* (identical in their parameter values and initial response probabilities), they will have a *distribution* of response probabilities on later trials.

The third class of models is illustrated by the Estes and Straughan (1954) analysis of the prediction experiment. This is an example of *experimenter-controlled events* in which the outcome alone (left-light or right-light onset) determines the event. Because the outcome schedule can be predetermined, only the parameter values and initial probabilities are needed in order to specify the trial-to-trial sequence of response probabilities. Identical subjects with identical outcome sequences who behave in accordance with such a model will have the same sequence of response probabilities. Although a subject's successive response probabilities may be different, his successive responses are independent. These features of models with experimenter-controlled events make mathematical analysis relatively simple. Despite the wide use of these models, however, direct experimental evidence that favors the independence assumption has not been forthcoming, and for the prediction experiment there is a certain amount of strong negative evidence, for example, in Hanania (1959) and Nicks (1959). The onset of a light apparently has an effect on the response probability that depends on whether the onset was correctly predicted. It is not known whether there are other experiments for which models with experimenter-controlled events might be appropriate.

Models in the fourth class involve just a single event and are the simplest.

A *single-event model* may be obtained from any model with subject-controlled events by the simplification of grouping all events into a single equivalence class. If the events "left turn—reward" and "right turn—nonreward" in a T-maze experiment with 100:0 reward, for example, are assumed to have equal effects on the probability of the right-turn response, then a single-event model is applicable. A second source for a single-event model is in the application of a model with experimenter-controlled events to the prediction experiment (Sec. 1.1), under the special condition that the same outcome is provided on every trial. Models with a single event are the easiest to study mathematically (see, for example, Bush & Sternberg, 1959) but the assumptions they entail seem seldom to be met in practice (Galanter & Bush, 1959; Sternberg, 1959b).

It appears that the best understood models are poor approximations to the data, and the models more likely to apply are little understood. We probably cannot dispense with subject control of events in learning experiments, and therefore the only choice available to us is whether or not there is experimenter control as well. Insofar as we deal with experiments whose outcomes are perfectly correlated with responses, we simplify matters by eliminating experimenter control. In this chapter I shall consider, in particular, models with subject-controlled events. Although they apply only to a restricted set of experiments, there are points in their favor: they appear to be more realistic than experimenter-controlled models, and more is known about them in terms of both theory and data than about models with experimenter-subject control.

1.5 Contingent Experiments and Contingent Events

It has been emphasized that the analysis of an experiment depends heavily on assumptions made about the subject and that the analysis is not an automatic consequence of the experimental design alone. Some of the current terminology can mislead one into thinking otherwise. Experiments have been categorized as "contingent" or "noncontingent," according to whether the occurrence of an outcome does or does not depend on the subject's response (Bush & Mosteller, 1955). It has been implied that contingent experiments correspond to experimenter-subject (contingent) events and noncontingent experiments to experimenter-controlled events and that this correspondence is unambiguous.

One difficulty with this method of classifying experiments lies in the definition of outcomes. If outcomes in the T-maze experiment are "left-arm baited" and "right-arm baited," then what the experimenter does can be predetermined and is noncontingent, although the relevant model may be a contingent one. This example seems absurd because we have confidence in our intuitions in regard to what constitutes a reinforcing event for a rat: it is surely food versus nonfood rather than left-arm versus right-arm baited. In the analogous two-armed bandit experiment the ambiguity is more obvious, especially if the subject imagines that exactly one of the two responses is correct on each trial.

Even if we reject what may appear to be bizarre definitions of outcomes, the contingent-noncontingent distinction leads to difficulties. Let us suppose that in a T-maze or bandit experiment the subject is rewarded on a preassigned subset of trials regardless of his response. The experiment appears to be noncontingent, but in developing a model few would be willing to assume that the effect of the outcome is independent of the response.

To begin one's analysis of any experiment—contingent or noncontingent—with the *assumption* that events are not subject-controlled would appear to be somewhat rash, unless the assumption is treated as a null hypothesis or an approximating device and is later carefully tested. On the other hand, analysis by means of a model that incorporates subject control should reveal the fact that a model with experimenter control alone (or a single-event model) can represent the behavior, if such is the case.

2. AXIOMATICS AND HEURISTICS OF MODEL CONSTRUCTION

Various considerations, formal and informal, substantive and practical, have been used as guides in constructing models. So far I have discussed the factors that help to determine the number of distinct alternative model events that may occur on a trial and the determinants of their occurrence. There remains the problem of the mathematical form in which to express the effects of events. Suppose that there are two alternative responses, A_1 and A_2, and that $\mathbf{p}_n = \Pr\{A_1 \text{ on trial } n\}$. Let $\mathbf{X}_n$ be a row-vector random variable[6] with t elements corresponding to the t possible events. $\mathbf{X}_n$ can take on the values $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$, $\ldots$, $(0, 0, \ldots, 1)$, which correspond to the occurrence on trial n of $E_1, E_2, \ldots, E_t$. In general, a learning model is a function that gives $\mathbf{p}_n$ in terms of the trial number and the sequence of events through trial $n - 1$,

$$\mathbf{p}_n = F(n, \mathbf{X}_{n-1}, \mathbf{X}_{n-2}, \ldots, \mathbf{X}_1), \tag{1}$$

[6] In this chapter vector random variables are designated by boldface capitals and scalar random variables by boldface lower-case letters. Realizations (particular values) of random variables are designated by the corresponding lightface capital and lower-case letters.

where the initial probability and other parameters are suppressed.[7] This equation makes it clear that $\mathbf{p}_n$, a function of random variables, is itself a random variable. Because it gives $\mathbf{p}_n$ explicitly in terms of the event sequence, we refer to Eq. 1 as the *explicit equation* for $\mathbf{p}_n$. In this section I consider some of the arguments that have been used to restrict the form of F.[8]

2.1 Path-Independent Event Effects

At the start of the nth trial of an experiment, p_n is the subject's response probability, and the sequence $X_1, X_2, \ldots, X_{n-1}$ describes the course of the experiment up to this trial. This sequence, then, specifies the "path" traversed by the subject in attaining the probability p_n. A simplifying assumption which underlies most of the learning models that have been studied is that the event on trial n has an effect that depends on p_n but not on the path. The implication is that insofar as past experience has any influence on the future behavior of the process this influence is mediated entirely by the value of p_n. Another way of saying this is that the subject's state or "memory" is completely specified by his p-value.

The assumption of independence of path leads naturally to a recursive expression for the model and to the definition of a set of operators. The recursive form is given by

$$\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{X}_n), \tag{2}$$

[7] Equation 1, and many of the other equations in this chapter in which response probabilities appear, may be regarded in two different ways. The first alternative, expressed by the notation in Eq. 1, is to regard $\mathbf{p}_n$ as a function of random variables and therefore to consider $\mathbf{p}_n$ itself as a random variable. This alternative is useful in emphasizing one of the important features of modern learning models—the fact that most of them specify a distribution of p-values on every trial after the first. By restricting the event sequence in any way, we determine a new, conditional distribution for the random variable. And we may be interested, for example, in determining the corresponding conditional expectation.

The second alternative is to regard the arguments in a formula such as Eq. 1 as *realizations* of the indicated random variables and the p-values it defines as *conditional* probabilities, conditioned by the particular event sequence. The formula is then more properly written as

$$\begin{aligned} p_n &= \Pr\{A_1 \text{ on trial } n \mid \mathbf{X}_1 = X_1, \mathbf{X}_2 = X_2, \ldots, \mathbf{X}_{n-1} = X_{n-1}\} \\ &= F(n, X_{n-1}, X_{n-2}, \ldots, X_1). \end{aligned}$$

Aside from easing the notation problem by reducing the number of boldface letters required, this alternative is occasionally useful; for example, the likelihood of a sequence of events can be expressed as the product of a sequence of such conditional probabilities. In this chapter, however, I make use of the first alternative.

[8] I omit stimulus sampling considerations, which are discussed in Chapter 10.

which indicates that p_{n+1} depends only on p_n and on the event of the nth trial. Equation 2 is to be contrasted with the explicit form given by Eq. 1. We note that conditional on the value of $\mathbf{p}_n$, $\mathbf{p}_{n+1}$ is not only independent of the particular events that have occurred (the content of the path) but also of their number (the path length). By writing Eq. 2 separately for each possible value of $\mathbf{X}_n$ we arrive at a set of trial-independent operators or transition rules:

$$\mathbf{p}_{n+1} = \begin{cases} Q_1\mathbf{p}_n = f[\mathbf{p}_n; (1, 0, \ldots, 0)] & \text{if} \quad E_1 \text{ on trial } n \\ Q_2\mathbf{p}_n = f[\mathbf{p}_n; (0, 1, \ldots, 0)] & \text{if} \quad E_2 \text{ on trial } n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots & \\ Q_t\mathbf{p}_n = f[\mathbf{p}_n; (0, 0, \ldots, 1)] & \text{if} \quad E_t \text{ on trial } n. \end{cases}$$

A common method for developing a model for an experiment is to begin with a set of plausible operators and rules for their application. If the event probabilities during the course of an experiment are functions of p_n alone, as they usually are, then path independence implies that the learning model is a discrete-time Markov process with an infinite number of states, the states corresponding to p-values.

The assumption is an extremely strong one, as indicated by three of its consequences, each of which is weaker than the assumption itself:

1. The effect of an event on the response probability is completely manifested on the succeeding trial. There can be no "delayed" effects. Examples of direct tests of this consequence are given later in this chapter.
2. When conditioned by the value of $\mathbf{p}_n$ (i.e., for any particular value of $\mathbf{p}_n$), the magnitude of the effect on the response probability of the nth event is independent of the sequence of events that precedes it.
3. When conditioned by the value of $\mathbf{p}_n$, the magnitude of the effect on the response probability of the nth event is independent of the trial number. Operators cannot be functions of the trial number.

Several models that meet conditions 1 and 2 but not 3 have been studied (Audley & Jonckheere, 1956, and Hanania, 1959). These models are quasi-independent of path, involving event effects that are independent of the content of the path but dependent on its length.

2.2 Commutative Events

Events are defined to be commutative if $\mathbf{p}_n$ is invariant with respect to alterations in the order of occurrence of the events in the path. To make this idea more precise, let us define a t-dimensional row vector, $\mathbf{W}_n$, whose kth component gives the cumulative number of occurrences of event E_k on trials $1, 2, \ldots, n-1$. We then have

$$\mathbf{W}_n = \mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_{n-1}.$$

Under conditions of commutativity, the vector $\mathbf{W}_n$ gives sufficient information about the path to determine the value of $\mathbf{p}_n$, and although a recursive expression does not result naturally Eq. 1 is simplified and becomes

$$\mathbf{p}_n = F(\mathbf{W}_n). \tag{4}$$

Several models have been proposed in terms of an explicit equation for $\mathbf{p}_n$ of this simple form, in which $\mathbf{p}_n$ depends only on the components of $\mathbf{W}_n$.

A further simplification arises if we require not only that F be a function of the components of $\mathbf{W}_n$ but also that it be expressible as a continuous function of an argument that is linear in these components. Under these circumstances we have not only commutativity but path independence as well.[9] As we shall see, path-independent models need not be commutative.

Events that commute are, in a certain sense, not subject to being forgotten. In a commutative model the effect of an event on $\mathbf{p}_n$ is the same, whether it occurred on trial 1 or trial $n - 1$. Because the distant past is as significant as the immediate past, models of this kind tend to be relatively unresponsive to changes in the outcome sequence. (An example is given in Sec. 4.) Commutativity leads to considerable simplifications in the analysis of models and in estimation procedures.

2.3 Repeated Occurrence of a Single Event

What is the effect on response probabilities of the repeated occurrence of a particular event? This is an important consideration in formulating a

[9] If F can be written as a continuous function of an argument that is linear in the components of $\mathbf{W}_n$, then events must have path-independent effects. The linearity condition requires that there exist some column vector, A, of coefficients for which $\mathbf{p}_n = F(\mathbf{W}_n) = G(\mathbf{W}_n \cdot A)$. Then

$$\begin{aligned}\mathbf{p}_{n+1} &= G(\mathbf{W}_{n+1} \cdot A) = G[(\mathbf{W}_n + \mathbf{X}_n) \cdot A] = G[\mathbf{W}_n \cdot A + \mathbf{X}_n \cdot A] \\ &= G[G^{-1}(\mathbf{p}_n) + \mathbf{X}_n \cdot A] = f(\mathbf{p}_n \cdot \mathbf{X}_n).\end{aligned}$$

A slight modification of the argument, which uses the continuity of G, is needed if G does not possess a unique inverse. One implication of some recent work by Luce (1963) on the commutativity property is that the converse of this result is also true: if a model is both commutative and path-independent then F can be written as a function of an argument that is linear in the components of $\mathbf{W}_n$.

model. Does **p** converge to any value as E_k is repeated again and again, and, if so, what value? Except for events that have no effect at all on **p**, it has been assumed in many applications of learning models that a particular event has either an incremental or a decremental effect on **p** and has this effect whatever the value of **p** as long as it is in the open (0, 1) interval. In most cases, then, repetition of a particular event causes **p** to converge to zero or one. Although in principle, learning models need not have this convergence property, it seems to be called for by most of the experiments to which they have been applied. This does not imply that the limiting behavior of these models always involves extreme response probabilities; in some instances successive repetition of any one event is improbable.

A related question concerns the sense in which the effect of an event is invariant during the course of an experiment. Neither commutativity nor path independence requires that the effect of an event on **p** be in the same direction—incremental or decremental—throughout an experiment. Commutativity alone, for example, does not require that F in Eq. 4 be monotonic in any one of the components of $\mathbf{W}_n$. Path independence implies that if the direction of effect is to change at all in the course of an experiment the direction can depend at most on p_n. These possibilities and restrictions are relevant to the question whether existing models can handle such phenomena as the development of secondary reinforcement or any changes that may occur in the effects of reward and nonreward.

2.4 Combining-Classes Condition: Bush and Mosteller's Linear-Operator Models

Despite their strong implications, neither path independence nor commutativity is restrictive enough to produce useful models. Further assumptions are needed. Two important families of models have used path independence as a starting point. The first, the family with which most work has been done, comprises Bush and Mosteller's linear-operator models. In this section we shall consider the general characteristics of the linear-operator family and its application to two experiments. The second family, to be discussed in Sec. 2.5, with applications to the same pair of experiments, is Luce's response-strength operator family. In both families the additional assumptions are invariance conditions concerning multiple-response alternatives.

The combining-classes condition is a precise statement of the assumption that the definition of response alternatives is arbitrary and that therefore any set of actions by the subject can be combined and treated as one alternative in a learning model.[10] It might appear that the assumption is untenable because it ignores any natural response organization that may characterize the organism; this issue has yet to be clarified by theoretical and experimental work. The assumption is not inconsistent with current learning theory, however. Concerning the problem of response definition, Logan (1960, p. 117–120) has written:

> The present approach assumes that responses that differ in any way whatsoever are different in the sense that they may be put into different response classes so that separate response tendencies can be calculated for them .... Differences among responses that are not importantly related to response tendency can be suppressed. This conception is consistent with what appears to be the most common basis of response aggregation, namely the conditions of reinforcement. Responses are separated if the environment ... distinguishes between them in administering rewards .... The rules discussed above would permit any aggregation that preserves the reinforcement contingencies in the situation. Thus, if the reward is given independently of how the rat gets into the correct goal box of a T-maze, then the various ways of doing it can be classed together.

The combining-classes condition is concerned with an experiment in which three or more response alternatives, $A_1, A_2, \ldots, A_r$, are initially defined. A subset, $A_{h+1}, A_{h+2}, \ldots, A_r$, of the alternatives is to be combined into A^*, thus producing the reduced set of alternatives $A_1, A_2, \ldots, A_h, A^*$. We consider the probability vector associated with the reduced set of alternatives after a particular event occurs. The combining-classes condition requires this vector to be the same, no matter whether the combining operation is performed before or after the event occurs; this invariance with respect to the combining operation is required to hold for all events and for any subset of the alternatives.

The result of applying the condition is that in a multiple-alternative

[10] As originally stated by Bush, Mosteller, and Thompson (1954) and later by Bush and Mosteller (1955) the actions to be combined must have the same outcome probabilities associated with them. For example, if A_1 and A_2 are the alternatives to be combined, then $\Pr\{O_i \mid A_1\} = \Pr\{O_i \mid A_2\}$ must hold for all outcomes O_i. This was necessary because outcome probabilities conditional on responses were thought of as a part of the model, and without the equality the probability $\Pr\{O_i \mid A_1 \text{ or } A_2\}$ would not be well defined. From the viewpoint of this chapter, the model for a subject's behavior depends on the particular sequence of outcomes, and any probability mechanism that the experimenter uses to generate the sequence is extraneous to the model. What must be well defined is the sequence of outcomes actually applied to any one of the alternative responses, and this sequence is well defined even if the actions to be combined into one alternative are treated differently.

For a formal treatment of the combining classes condition see the references already cited, Mosteller (1955), or Bush (1960b).

experiment the effect of an event on the probability p_n of one of the alternatives can depend on p_n but cannot depend on the relative probabilities of choice among the other alternatives. To see this, let us suppose that the change in p_n does depend on the way in which $1 - p_n$ is distributed among the other alternatives. Then, if alternatives defined initially are combined *before* the event, the model cannot reflect the distribution of $1 - p_n$ among them. Thus, even if we can arrange to have p_{n+1} well defined, its value will not, in general, be the same as if the alternatives were combined *after* the event, in contradiction to the invariance assumption.

Together with the path-independence assumption, the combining-classes condition requires not only that each component of the p_{n+1}-vector depend on X_n and on the corresponding component of the p_n-vector only, but it also requires that this dependence be linear. The effect of a particular event on a set of response probabilities is therefore to transform each of them linearly, allowing us to write the operator Q_k as

$$p_{n+1} = Q_k p_n = \alpha_k p_n + a_k, \tag{5}$$

where the values of α_k and a_k are determined by the event E_k. This result can be proved only when $r \geqslant 3$, where r is the number of response alternatives. However, if we regard $r = 2$ as having arisen from the combination of a larger set of alternatives for which the condition holds, then the form of the operator is that given by Eq. 5.

One aspect of using linear transformations on response probabilities is that the parameters must be constrained to keep the probability within the unit interval. The constraints have usually been determined by the requirement that all possible values of p_n, from 0 to 1, be transformed into probabilities by the operator. A consequence of this requirement is that with r alternatives $-1/(r - 1) \leqslant \alpha \leqslant 1$. It is in the spirit of the combining-classes condition that, in principle, an unlimited number of response classes may be defined. If r becomes arbitrarily large, then we see that one implication of the condition is that negative α's are inadmissible.

In comparing operators Q_k and considering their properties, it is useful to define a new parameter $\lambda_k = a_k/(1 - \alpha_k)$. The general operator can then be rewritten as

$$p_{n+1} = Q_k p_n = \alpha_k p_n + (1 - \alpha_k)\lambda_k, \tag{6}$$

where the constraints are $0 \leqslant \alpha_k \leqslant 1$ and $0 \leqslant \lambda_k \leqslant 1$. The transformed probability $Q_k p_n$ may be thought of as a weighted sum of p_n and λ_k and the operator may be thought of as moving p_n in the direction of λ_k. Because $Q_k\lambda_k = \lambda_k$, the parameter λ_k is the *fixed point* of the operator: the operator does not alter a probability whose value is λ_k. In addition, when $\alpha_k \neq 1$, λ_k is the *limit point* of the operator: repeated occurrence of the event E_k leads to repeated application of Q_k, and this causes the probability to approach λ_k asymptotically. This may be seen by first calculating the effect of m successive applications of the operator Q to p:

$$Q^m p = \alpha^m p + (1 - \alpha^m)\lambda = \lambda - \alpha^m(\lambda - p).$$

If $\alpha < 1$, $\lim_{m\to\infty} \alpha^m = 0$ and therefore $\lim_{m\to\infty} Q^m p = \lambda$.

As noted in Sec. 2.3, values of λ other than the extreme values zero or one have seldom been used in practice. An extreme limit point automatically requires that α be nonnegative if p_{n+1} is to be confined to the unit interval for all values of p_n. In order to justify the assumption that α is nonnegative, we can therefore usually appeal to the required limit point rather than to the extension of the combining-classes condition already mentioned. Because of this and because multiple-choice studies with $r \geqslant 3$ are relatively rare, the combining-classes condition has never been put directly to the test.

The parameter α_k may be thought of as a learning-rate parameter. Its value is a measure of *ineffectiveness* of the event E_k in altering the response probability. When α_k takes on its maximum value of 1, then Q_k is an *identity operator* and event E_k induces no change in the response probability. The smaller the value of α_k, the more p_n is moved in the direction of λ_k by the operator Q_k.

For the two extreme limit points, the operators are either of the form $Q_1 p = \alpha_1 p$ or $Q_2 p = \alpha_2 p + (1 - \alpha_2)$, and these are the two operators that have most commonly been used. In general, pairs of operators do not commute, but under certain conditions (either they have equal limit points or one of them is the identity operator) they do. When all pairs of operators in a model commute, the explicit expression for p_n in terms of the path has a simple form. When the operators in a two-choice experiment are all of the form of Q_2, it is convenient to deal with $q = 1 - p$, the probability of the other response, and to make use of the complementary operators whose form is $\tilde{Q}_2 q = \alpha_2 q$.

To be concrete, we consider examples of the Bush-Mosteller model for two of the experiments discussed in Sec. 1.

ESCAPE-AVOIDANCE SHUTTLEBOX. We interpret this experiment to consist of two subject-controlled events: escape (shock) and avoidance. Both events reduce the probability $\mathbf{p}_n$ of escape in the direction of the limit point $\lambda = 0$. It is convenient to define a binary random variable that represents the event on trial n,

$$\mathbf{x}_n = \begin{cases} 0 & \text{if avoidance } (E_1) \\ 1 & \text{if escape } (E_2), \end{cases} \tag{7}$$

with a probability distribution given by $\Pr\{\mathbf{x}_n = 1\} = \mathbf{p}_n$, $\Pr\{\mathbf{x}_n = 0\} = 1 - \mathbf{p}_n$. The operators and the rule for their application are

$$\mathbf{p}_{n+1} = \begin{cases} Q_1\mathbf{p}_n = \alpha_1\mathbf{p}_n & \text{if} \quad \mathbf{x}_n = 0 \quad \text{(i.e., with probability } 1 - \mathbf{p}_n) \\ Q_2\mathbf{p}_n = \alpha_2\mathbf{p}_n & \text{if} \quad \mathbf{x}_n = 1 \quad \text{(i.e., with probability } \mathbf{p}_n). \end{cases} \tag{8}$$

Because events are subject-controlled, the sequence of operators (events) is not predetermined. The vector $\mathbf{X}_n$ (Sec. 2) is given by $(\mathbf{x}_n, 1 - \mathbf{x}_n)$ and the recursive form of Eq. 2 (Sec. 2.1) is given by

$$\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{X}_n) = \alpha_2^{\mathbf{x}_n}\alpha_1^{(1-\mathbf{x}_n)}\mathbf{p}_n. \tag{9}$$

The operators have equal limit points and therefore commute. This makes possible a simple explicit formula for $\mathbf{p}_n$. Let $\mathbf{W}_n = (\mathbf{s}_n, \mathbf{t}_n)$, where

$$\mathbf{s}_n = \sum_{j=1}^{n-1} \mathbf{x}_j$$

is the number of shocks before trial n and

$$\mathbf{t}_n = \sum_{j=1}^{n-1} (1 - \mathbf{x}_j)$$

is the number of avoidances before trial n. The explicit form of Eq. 4 (Sec. 2.2) is given by

$$\mathbf{p}_n = F(\mathbf{W}_n) = \alpha_2^{\mathbf{s}_n}\alpha_1^{\mathbf{t}_n}p_1. \tag{10}$$

Later I shall make use of the fact that, by redefining the parameters of the model, $\mathbf{p}_n$ may be written as a function of an expression that is linear in the components of $\mathbf{W}_n$. To do this, let $p_1 = e^{-a}$, $\alpha_1 = e^{-b}$, and $\alpha_2 = e^{-c}$. Then Eq. 10 becomes

$$\mathbf{p}_n = \exp\left[-(a + b\mathbf{t}_n + c\mathbf{s}_n)\right]. \tag{11}$$

It should be emphasized that, for this model, $\mathbf{t}_n$ and $\mathbf{s}_n$ and, therefore, $\mathbf{p}_n$ are random variables whose values are unknown until trial $n - 1$. The set of trials is a dependent sequence.

PREDICTION EXPERIMENT. For purposes of illustration we make the customary assumption that this experiment consists of two experimenter-controlled events. Onset of the left light (E_1) increases the probability p_n of a left button press toward the limit point $\lambda = 1$. Onset of the right light decreases p_n toward the limit point $\lambda = 0$.

The events are assumed to be complementary (Sec. 1.2), and therefore the operators have equal rate parameters ($\alpha_1 = \alpha_2 = \alpha$) and complementary limit points ($\lambda_1 = 1 - \lambda_2$). It is convenient to define a binary variable that represents the event on trial n,

$$y_n = \begin{cases} 0 & \text{if} \quad \text{right light } (E_2) \\ 1 & \text{if} \quad \text{left light } (E_1). \end{cases}$$

Because events are experimenter-controlled, the sequence $y_1, y_2, \ldots$ can be predetermined. In some experiments a random device may be used to generate the actual sequence used. For example, the $\{y_n\}$ may be a realization of a sequence $\{\mathbf{y}_n\}$ of independent random variables with $\Pr\{\mathbf{y}_n = 1\} = \pi$. However, insofar as we are interested in the behavior of the subject, the actual sequence, rather than any properties of the random device used to generate it, is of interest. It is shown later how simplifying approximations may be developed by assuming that the subject has experienced the average of all the sequences that the random device generates. For the purpose of such approximations, which, of course, involve loss of information, y_n may be considered a random variable with a probability distribution. The more exact treatment, however, deals with the experiment conditional on the actual outcome sequences that are used.

The operators and the rules for their application are

$$p_{n+1} = \begin{cases} Q_1 p_n = \alpha p_n + 1 - \alpha & \text{if } \; y_n = 1 \\ Q_2 p_n = \alpha p_n & \text{if } \; y_n = 0. \end{cases} \tag{12}$$

The vector X_n (Sec. 2) is given by $(y_n, 1 - y_n)$ and the recursive form of Eq. 2 is given by

$$p_{n+1} = \alpha p_n + (1 - \alpha) y_n. \tag{13}$$

Note that in the exact treatment p_n is not a random variable, unlike the case for an experiment with subject control. The operators do not commute, and therefore the cumulative number of E_1's and E_2's does not determine p_n uniquely. The explicit formula for p_n, in contrast to the shuttlebox example, includes the entire sequence $y_1, y_2, \ldots$ and is given by

$$p_n = F(n, X_1, X_2, \ldots, X_{n-1}) = \alpha^{n-1} p_1 + (1 - \alpha) \sum_{j=1}^{n-1} \alpha^{n-1-j} y_j. \tag{14}$$

Equation 14 shows that (when $\alpha < 1$) a recent event has more effect on p_n than an event in the distant past. By contrast, Eq. 10 indicates that for the shuttlebox experiment there is no "forgetting" in this sense: given that the event sequence has a particular number of avoidances, the effect of an early avoidance on p_n is no different from the effect of a late avoidance. As I mentioned earlier, this absence of forgetting is a characteristic of all experiments with commutative events.

The model for the prediction experiment consists of a sequence of independent binomial trials: the p_n-sequence is determined by the y_n-sequence which is independent of all responses.

2.5 Independence from Irrelevant Alternatives: Luce's Beta Response-Strength Model

Stimulus-response theory has traditionally treated response probability as deriving from a more fundamental response-strength variable. For example, Hull (1943, Chapter 18) conceived of the probability of reaction (where the alternative was nonreaction) as dependent, in a complicated way, on a reaction-potential variable that is more fundamental in his system than the probability itself. The momentary reaction potential was thought to be the sum of an underlying value and an error (behavioral oscillation) that had a truncated normal distribution: the response would occur if the effective reaction potential exceeded a threshold value. The result of these considerations was that reaction probability was related to reaction potential by means of a cumulative normal distribution which was altered to make possible zero probabilities. The alteration implied that a range of reaction potentials could give rise to the same (zero) probability. Such a threshold mechanism has not been explicitly embodied in any of the modern stochastic learning models.

The reaction potential variable was more fundamental partly because it changed in a simple way in response to experimental events and partly because the state of the organism was more completely described by the reaction potential than by the probability. The last observation is clearer if we turn from the single-response situation considered by Hull[11] to an experiment with two symmetric responses. The viewpoint to be considered is that such an experiment places into competition two responses, each of which may vary independently in its strength. Let us suppose that response strengths, symbolized by $v(1)$ and $v(2)$, are associated with each of the two responses and that the probability of a response is given by the ratio of its strength to the sum of the two strengths: $p\{1\} = v(1)/[v(1) + v(2)]$. Then, although the two strengths determine the probability uniquely, knowledge of the probability can tell us only the ratio of the strengths, $v(1)/v(2)$. Multiplying both strengths by the same constant does not alter the response probability, but it might, for example, correspond to the change from an avoidance-avoidance to an approach-approach conflict and might be revealed by response times or amplitudes. The response strengths therefore provide a more basic description of the state of the organism than the response probabilities. This is the sort of thinking that might lead one to favor a learning model whose underlying

[11] For the symmetric two-choice experiment a response-strength analysis that is considerably different from the one discussed here is given by Hull (1952).

variables are response strengths. A critical question regarding this viewpoint is whether there are aspects of the behavior in choice learning experiments that can be accounted for by changes in response strengths but that are not functions of response probabilities alone.

An invariance condition concerning multiple response alternatives, but different from the combining-classes condition, is used by Luce (1959) in arriving at his beta response-strength model. Path independence of the sequence of response strengths is also assumed, and within the model this entails path independence of the sequence of probability vectors. The invariance condition (Luce's Axiom 1) states that, in an experiment in which one of a set of responses is made, the ratio of the probabilities of two alternatives is invariant with respect to changes in the set of remaining alternatives from which the subject can select. As stated, the condition applies to choice situations in which the probabilities of choice from a constant set of alternatives are unchanging. Nonetheless, by assuming that the condition holds during any instant of learning, we can use it to restrict the form of a learning model. The condition implies that a positive response-strength function, $v(j)$, can be defined over the set of alternatives $\{A_j\}$ with the property that

$$\Pr\{A_k\} = \frac{v(k)}{\sum_j v(j)}. \tag{15}$$

The subsequent argument rests significantly on the fact that $v(j)$ is a ratio scale and that the scale values are determined by the choice probabilities only up to multiplication by a positive constant; that is, the unit of the response-strength scale is arbitrary.

The argument begins with the idea that in a learning experiment the effects of an event on the organism can be thought of as a transformation of the response strengths. Two steps in the argument are critical in restricting the form of this transformation. First, it is observed that because the unit of response strength is arbitrary the transformation f must be invariant with respect to changes in this unit: $f[kv(j)] = kf[v(j)]$. Second, it is assumed that the scale of response strength is unbounded and that therefore any real number is a possible scale value. The independence-of-unit condition, together with the unboundedness of the scale, leads to the powerful conclusion that the only admissible transformation is multiplication by a constant.[12] The requirement that response strengths

[12] As suggested by Violet Cane (1960), it is not possible to have both an unbounded response-strength scale and choice probabilities equal to zero or unity ("perfect discrimination"); for, if choice probabilities can take on all values in the closed unit interval, then the v-scale must map onto this closed interval and must therefore itself extend over a closed, and thus bounded, interval. But the unboundedness of the scale

be positive implies that the multiplying constant must be positive. Path independence implies that the constant depends only on the event and not on the trial number or the response strength. This argument completely defines the form of a learning model—called the "beta model"—for experiments with two alternative responses. The model defines a stochastic process on response strengths, which in turn determines a stochastic process on the choice probabilities.

When event E_k occurs, let $v(1)$ and $v(2)$ be transformed into $a_k v(1)$ and $b_k v(2)$. The new probability is then

$$\Pr\{1\} = \frac{a_k v(1)}{a_k v(1) + b_k v(2)} = \frac{(a_k/b_k)[v(1)/v(2)]}{1 + (a_k/b_k)[v(1)/v(2)]}.$$

If we let $v = v(1)/v(2)$ be the ratio of response strengths and $\beta_k = a_k/b_k$ be the ratio of constants, the original probability is $v/(1 + v)$ and the transformed probability is $\beta_k v/(1 + \beta_k v)$. The ratio β_k and the relative response strength v are sufficient to determine p and its transformed value. Because response strengths are important in this chapter only insofar as they govern probabilities, the simplified notation is adequate. We let v_n be the relative response strength $v(1)/v(2)$ on trial n, let β_k be the multiplier of v_n that is associated with the event E_k, and let p_n be $\Pr\{A_1 \text{ on trial } n\}$. Then

$$p_n = \frac{v_n}{1 + v_n} = \frac{1}{1 + v_n^{-1}} \quad \text{and} \quad v_n = \frac{p_n}{1 - p_n}.$$

Moreover, if event E_k occurs on trial n,

$$p_{n+1} = \frac{\beta_k v_n}{1 + \beta_k v_n} = \frac{\beta_k p_n}{(1 - p_n) + \beta_k p_n} = Q_k p_n, \qquad (16)$$

which gives the corresponding nonlinear transformation on response probability.

A number of implications of the model can be seen immediately from the form of the probability operator. If E_k has an incremental (decremental) effect on $\Pr\{A_1\}$, then $\beta_k > 1 (< 1)$. An identity operator results

---

is an important feature of the argument that forces the learning transformation to be multiplicative. It therefore appears that Luce's axiom leads to a multiplicative learning model only when it is combined with the assumption that response probabilities can never be exactly zero or unity. In practice, this assumption is not serious, since a finite number of observations do not allow us to distinguish between a probability that is exactly unity and one that is arbitrarily close to that value. The fact that the additional assumption is needed, however, makes it difficult to disprove the axiom on the basis of a failure of the learning model, since the fault may lie elsewhere.

when $\beta_k = 1$. The only limit points possible are $p = 0$ and $p = 1$, which are obtained, respectively, when $\beta_k < 1$ and $\beta_k > 1$. This follows because

$$Q^m p = \frac{\beta^m v}{1 + \beta^m v} = \frac{v}{v + \beta^{-m}}$$

and, when $\beta \neq 1$, either β^m or β^{-m} approaches zero. These properties imply that the effect of an event on p_n must always be in the same direction; all operators are unidirectional in contrast to operators in the linear model, which may have zero points other than zero and unity. The restriction to extreme limit points appears not to be serious in practice, however; as noted in the Sec. 2.4, most experiments seem to call for unidirectional operators.

Perhaps more important from the viewpoint of applications is the fact that operators in the beta response-strength model must always commute; the model requires that events in learning experiments have commutative effects. That nonlinear probability operators of the form given by Eq. 16 commute can be shown directly, or can be seen more simply by noting the commutativity of the multiplicative transformations of v_n, to which such operators correspond. Whether or not commutativity is realistic, it is a desirable simplifying feature of the model. A final property is that the model cannot produce learning when $p_1 = 0$ because this requires that v_1, hence all v_n, be zero.

Let us consider applications of the beta model to the experiments discussed in Sec. 2.4. When applicable, the same definitions are used as in that section.

ESCAPE-AVOIDANCE SHUTTLEBOX. It is convenient to let p_n be the probability of escape (response A_2, event E_2) and to let v_n be the ratio of escape strength to avoidance strength. Both events reduce the probability of escape and therefore both $\beta_1 < 1$ and $\beta_2 < 1$. The binary random variable $\mathbf{x}_n$ is defined as in Eq. 7. The operators and the rules for their application are

$$\mathbf{p}_{n+1} = \begin{cases} Q_1\mathbf{p}_n = \dfrac{\beta_1\mathbf{p}_n}{(1-\mathbf{p}_n) + \beta_1\mathbf{p}_n} & \text{if } \mathbf{x}_n = 0 \\ & \text{(i.e., with probability } 1 - \mathbf{p}_n) \\ Q_2\mathbf{p}_n = \dfrac{\beta_2 p_n}{(1-\mathbf{p}_n) + \beta_2\mathbf{p}_n} & \text{if } \mathbf{x}_n = 1 \\ & \text{(i.e., with probability } \mathbf{p}_n). \end{cases} \tag{17}$$

The recursive form of Eq. 2 is given by

$$\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{X}_n) = \frac{\beta_2^{\mathbf{x}_n}\beta_1^{(1-\mathbf{x}_n)}\mathbf{p}_n}{(1-\mathbf{p}_n) + \beta_2^{\mathbf{x}_n}\beta_1^{(1-\mathbf{x}_n)}\mathbf{p}_n}. \tag{18}$$

Both expressions are cumbersome. More light is shed by the explicit formula

$$\mathbf{p}_n = F(\mathbf{W}_n) = \frac{1}{1 + \beta_2^{-\mathbf{s}_n}\beta_1^{-\mathbf{t}_n}v_1^{-1}} \cdot \tag{19}$$

Redefining the parameters simplifies Eq. 19. Let $v_1 = e^a$, $\beta_1 = e^b$ and $\beta_2 = e^c$. Then the expression becomes

$$\mathbf{p}_n = \frac{1}{1 + \exp\left[-(a + b\mathbf{t}_n + c\mathbf{s}_n)\right]} \cdot \tag{20}$$

It is instructive to compare this explicit formula with Eq. 11 for the linear model. In the usual experiment a would be positive and b and c would be negative, unlike the coefficients in Eq. 11, all of which are positive. (Recall that as $\mathbf{t}_n$, $\mathbf{s}_n$ increase $\mathbf{p}_n$ decreases. These definitions are awkward, but they will facilitate matters later on.) Again it should be noted that $\mathbf{t}_n$ and $\mathbf{s}_n$ are random variables whose behavior is governed by p-values earlier in the sequence and that the model therefore defines a dependent sequence of trials.

PREDICTION EXPERIMENT. Experimenter-controlled events are assumed here, as in Sec. 2.4. The $\{p_n\}$ and $\{y_n\}$ are defined as in that section. The complementarity of events demands that $\beta_1 = \beta_2^{-1} = \beta > 1$. This can be seen by noting that if E_1 transforms $v(1)/v(2)$ into $\beta v(1)/v(2)$ then for operators to be complementary E_2 must transform $v(2)/v(1)$ into $\beta v(2)/v(1)$. The operators and the rules for their application therefore are

$$p_{n+1} = \begin{cases} Q_1 p_n = \dfrac{\beta p_n}{(1 - p_n) + \beta p_n} & \text{if} \quad y_n = 1 \\[2ex] Q_2 p_n = \dfrac{p_n}{\beta(1 - p_n) + p_n} & \text{if} \quad y_n = 0. \end{cases} \tag{21}$$

The recursive expression is given by

$$p_{n+1} = \frac{\beta^{y_n} p_n}{\beta^{(1-y_n)}(1 - p_n) + \beta^{y_n} p_n} \cdot \tag{22}$$

Because of the universal commutativity of the beta model, the explicit formula is simple in contrast to Eq. 14. We have $W_n = (l_n, r_n)$, where

$$l_n = \sum_{j=1}^{n-1} y_j$$

is the number of left-light outcomes before trial n and

$$r_n = \sum_{j=1}^{n-1} (1 - y_j)$$

is the corresponding number of right-light outcomes. We define $d_n = l_n - r_n$ to be the difference between these numbers. The explicit formula is then

$$p_n = F(W_n) = \frac{1}{1 + \beta^{d_n} v_1^{-1}}. \tag{23}$$

For this model commutativity seems to be of more use than path independence in simplifying formulas. Again we can define new parameters $v_1 = e^a$, $\beta = e^{-b}$ to obtain

$$p_n = \frac{1}{1 + \exp[-(a + bd_n)]}. \tag{24}$$

Equation 24 indicates that the response probability is expressed in terms of d_n by means of the well-known logistic function. Equation 20 is a generalized form of this function. All of the two-alternative beta models have explicit formulas that are (generalized) logistics. Because the logistic function is similar to the cumulative normal distribution, the relation in the beta model between response strength and probability is reminiscent of Hull's treatment of this problem.

2.6 Urn Schemes and Explicit Forms

The treatment of examples in Secs. 2.4 and 2.5 illustrates two of the alternative ways of regarding a stochastic learning model. One approach is to specify the change in $\mathbf{p}_n$ that is induced by the event (represented by $\mathbf{X}_n$) on trial n. This change in probability depends, in general, on $\mathbf{X}_1, \ldots, \mathbf{X}_{n-1}$ as well as on $\mathbf{X}_n$. In the general recursive formula for response probability we therefore express $\mathbf{p}_{n+1}$ as a function of $\mathbf{p}_n$ and the events through trial n:

$$\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{X}_n, \mathbf{X}_{n-1}, \ldots, \mathbf{X}_1).$$

If the model is path-independent, then $\mathbf{p}_{n+1}$ is uniquely specified by $\mathbf{p}_n$ and $\mathbf{X}_n$, and the expression may be simplified to give the recursive formula of Sec. 2.1,

$$\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{X}_n),$$

and its corresponding operator expressions. The second approach is to specify the way in which $\mathbf{p}_n$ depends on the entire sequence of events through trial $n - 1$. This is done by the explicit formula in which $\mathbf{p}_n$ is expressed as a function of the event sequence:

$$\mathbf{p}_n = F(\mathbf{X}_{n-1}, \mathbf{X}_{n-2}, \ldots, \mathbf{X}_1).$$

A model may have a more "natural" representation in one of these forms than in the other. In this section I discuss models for which the explicit form is the more natural.

URN SCHEMES. Among the devices traditionally used in fields other than learning to represent the aftereffects of events on probabilities are games of chance known as urn schemes (Feller, 1957, Chapter V). An urn contains different kinds of balls, each kind representing an event. The occurrence of an event is represented by randomly drawing a ball from the urn. Aftereffects are represented by changes in the urn's composition. Schemes of this kind are among the earliest stochastic models for learning (Thurstone, 1930; Gulliksen, 1934) and are still of interest (Audley & Jonckheere, 1956; Bush & Mosteller, 1959). The stimulus-sampling models discussed in Chapter 10 may be regarded as urn schemes whose balls are interpreted as "elements" of the stimulus situation. In contrast, Thurstone (1930) suggested that the balls in his scheme be interpreted as elements of response classes. In all of these examples two kinds of balls, corresponding to two events, are used. An urn scheme is introduced to help make concrete one's intuitive ideas about the learning process. Except in the case of stimulus-sampling theory, the interpretation of the balls as psychological entities has not been pressed far.

The general scheme discussed by Audley and Jonckheere (1956) encompasses most of the others as special cases. It is designed for experiments with two subject-controlled events. On trial 1 the urn contains w_1 white balls and r_1 red balls. A ball is selected at random. If it is white, event E_1 occurs ($\mathbf{x}_1 = 0$), the ball is replaced, and the contents of the urn are changed so that there are now $w_1 + w$ white balls and $r_1 + r$ red balls. If the chosen ball is red, event E_2 occurs ($\mathbf{x}_1 = 1$), the ball is replaced, and the new numbers are $w_1 + w'$ and $r_1 + r'$. This process is repeated on each trial. The quantities w, w', r and r' have fixed integral values that may be positive, zero, or negative, but, if any of them are negative, then arrangements must be made so that the urn always contains at least one ball and so that the number of balls of either color is never negative.

Let $\mathbf{t}_n = \sum_{j=1}^{n-1} (1 - \mathbf{x}_j)$ be the number of occurrences of E_1 and $\mathbf{s}_n = \sum_{j=1}^{n-1} \mathbf{x}_j$ be the number of E_2 occurrences before trial n. Let $\mathbf{w}_n$ and $\mathbf{r}_n$ be the number of white and red balls in the urn before the nth trial. Then

$$\mathbf{w}_n = w_1 + w\mathbf{t}_n + w'\mathbf{s}_n, \qquad \mathbf{r}_n = r_1 + r\mathbf{t}_n + r'\mathbf{s}_n$$

and

$$\begin{aligned}\mathbf{p}_n &= \Pr\{\text{event } E_2 \text{ on trial } n\} \\ &= \frac{\mathbf{r}_n}{\mathbf{r}_n + \mathbf{w}_n} \\ &= \frac{r_1 + r\mathbf{t}_n + r'\mathbf{s}_n}{(r_1 + w_1) + (r + w)\mathbf{t}_n + (r' + w')\mathbf{s}_n}.\end{aligned} \tag{25}$$

Equation 25 gives the explicit formula for $\mathbf{p}_n$. It demonstrates the most important property of these models—their commutativity.

If $\mathbf{w}_n$ and $\mathbf{r}_n$ are interpreted as response strengths, the model can be regarded as a description of additive (rather than multiplicative) transformations of these strengths.[13]

The recursive formula and corresponding operators are unwieldy and are not given. Suffice it to say that the operators are nonlinear and depend on the trial number (that is, on the path length) but not on the particular sequence of preceding events (the content of the path). The model, therefore, is only quasi-independent of path (Sec. 2.1). This is the case because the change induced by an event in the proportion of red balls depends on p_n, on the numbers of reds and whites added (which depend only on the event), and on the total number of balls in the urn before the event occurred (which can in general be inferred only from knowledge of both p_n and n).

Two special cases of the urn scheme that are exceptions to the foregoing statement and produce path-independent models are (1) those for which $r + w = r' + w' = 0$, so that the total number of balls is constant, and (2) those for which either $r = r' = 0$ or $w = w' = 0$, so that the number of balls of one color is constant. The first condition is met by Estes' model, which, however, departs in another respect from the general scheme: its additive increments vary with the changing composition of the urn instead of being constant. (This modification sacrifices commutativity, but it is necessary if the probability operators are to be linear. The modification follows from the identification of balls with stimulus elements, and so is less artificial than it sounds.)

The second condition is assumed in the urn scheme that Bush and Mosteller (1959) apply to the shuttlebox experiment. They assume that $r = r' = 0$, so that only white balls are added to the urn; neither escape nor avoidance alters the "strength" of the escape response. The model is modified so that w and w' are continuous parameters rather than discrete numbers, as they would have to be in a strict interpretation as numbers of balls. The result may be expressed in simple form by defining $a = (r_1 + w_1)/r_1$. To be consistent with Eqs. 11 and 20, we let $b = w/r_1$ and $c = w'/r_1$ and obtain

$$\mathbf{p}_n = \frac{1}{a + b\mathbf{t}_n + c\mathbf{s}_n}. \tag{26}$$

[13] The model is also appropriate if it is thought that strength $v(j)$ is transformed multiplicatively but that response probability depends on logarithms of strengths: $p(1) = \log v(1)/\log [v(1)v(2)]$. In such a case $\mathbf{w}_n$ and $\mathbf{r}_n$ are interpreted as logarithms of response strengths.

The operators are given by

$$\mathbf{p}_{n+1} = \begin{cases} Q_1\mathbf{p}_n = \dfrac{\mathbf{p}_n}{1 + b\mathbf{p}_n} & \text{if} \quad \mathbf{x}_n = 0 \quad \text{(i.e., with probability } 1 - \mathbf{p}_n) \\ Q_2\mathbf{p}_n = \dfrac{\mathbf{p}_n}{1 + c\mathbf{p}_n} & \text{if} \quad \mathbf{x}_n = 1 \quad \text{(i.e., with probability } \mathbf{p}_n). \end{cases} \tag{27}$$

LINEAR MODELS FOR SEQUENTIAL DEPENDENCE. It has been indicated earlier that the responses produced by learning models consist of stochastically dependent sequences, except for the case of experimenter-controlled events. Moreover, insofar as experimenter control is present, the sequence of responses will be dependent on the sequence of outcomes. The autocorrelation of responses and the correlation of responses with outcomes are interesting in themselves, whether in learning experiments or, for example, in trial-by-trial psychophysical experiments in which there is no over-all trend in response probability. Several models have arisen directly from hypotheses about repetition or alternation tendencies that perturb the learning process and produce a degree or kind of response-response or response-outcome dependence that is unexpected on the basis of other learning models. The example to be mentioned is neither path-independent nor commutative.

The one trial perseveration model (Sternberg, 1959a,b) is suggested by the following observation: in certain two-choice experiments with symmetric responses the probability of a particular response is greater on a trial after it occurs and is rewarded than on a trial after the alternative response occurs and is not rewarded. There are several possible explanations. One is that reward has an immediate and lasting effect on $\mathbf{p}_n$ that is greater than the effect of nonreward. This hypothesis attributes the observed effect to a differential influence of outcomes in the cumulative learning process. One of the models already discussed could be used to describe this mechanism: for example, the model given by Eq. 8 (Sec. 2.4) with $\alpha_1 < \alpha_2$.

A second hypothesis is that the two outcomes are equally effective (i.e., they are symmetric) but that there is a short-term one-trial tendency to repeat the response just made. This hypothesis, when applied to an experiment with 100:0 reward[14] leads to the one-trial perseveration model.

Without the repetition tendency, the assumption of outcome symmetry leads to a model with experimenter-controlled events of the kind that was applied in Sec. 2.4 to the prediction experiment. The 100:0 reward

[14] The term "π_1:π_2 reward" describes a two-choice experiment in which one choice is rewarded with probability π_1 and the other with probability π_2.

schedule implies that $y_n = 0$ on all trials and therefore that the same operator, Q_2 in Eq. 12, is applied on every trial. Equation 14 shows the explicit formula to be

$$p_n = \alpha^{n-1} p_1. \tag{28}$$

This single-operator model was discussed by Bush and Sternberg (1959). It may also be regarded as a special case of the subject-controlled model used for the shuttlebox (Eq. 8) with $\alpha_1 = \alpha_2 = \alpha$.

In developing the perseveration model, the single-operator model is taken to represent the "underlying" learning process. Define $\mathbf{x}_n$ so that $\Pr\{\mathbf{x}_n = 1\} = \mathbf{p}_n$ and $\Pr\{\mathbf{x}_n = 0\} = 1 - \mathbf{p}_n$. We note that the strongest possible tendency to repeat the previous response can be described by the model $\mathbf{p}_n = \mathbf{x}_{n-1}$. This is the effect that perturbs the learning process.

To combine the underlying and perturbing processes, we take a weighted combination of the two, with nonnegative weights $1 - \beta$ and β. This gives the explicit formula[15] for the subject-controlled model:

$$\mathbf{p}_n = F(n, \mathbf{x}_{n-1}) = (1 - \beta)\alpha^{n-1} p_1 + \beta \mathbf{x}_{n-1}, \qquad (n \geqslant 2). \tag{29}$$

Knowledge of the trial number and of only the last response is needed to determine the value of $\mathbf{p}_n$. The two possible values that $\mathbf{p}_n$ can have on a particular trial differ by the (constant) value of β; $\mathbf{p}_n$ takes on the higher of the two values when $\mathbf{x}_{n-1} = 1$ and the lower when $\mathbf{x}_{n-1} = 0$. The extent to which the learning process is perturbed by the repetition tendency is greater with larger β.

That the model is path-dependent is shown by the form of its recursive expression:

$$\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{x}_n, \mathbf{x}_{n-1}) = \alpha \mathbf{p}_n + \beta \mathbf{x}_n - \alpha\beta \mathbf{x}_{n-1}, \qquad (n \geqslant 2). \tag{30}$$

Knowledge of the values of $\mathbf{p}_n$ and $\mathbf{x}_n$ alone is insufficient to specify the value of $\mathbf{p}_{n+1}$. For this model, in contrast to most others, more past history is needed in order to specify $\mathbf{p}_n$ by the recursive form than by the explicit form. Data from a two-armed bandit experiment have been fruitfully analyzed with the perseveration model.

The development of the perseveration model illustrates a technique that is of general applicability and is occasionally of interest. A tendency to alternate responses may be represented by a similar device. Linear equations may also be used to represent positive or negative correlation between outcome and subsequent response—for example, a tendency in a

[15] A trivial modification in this expression is made by Sternberg (1959a) based on considerations about starting values.

prediction experiment to avoid predicting the event that most recently occurred.

LOGISTIC MODELS. The most common approach to the construction of models begins with an expression for trial-to-trial probability changes, an expression that seems plausible and that may be buttressed by more general assumptions. An alternative approach is to consider what features of the entire event sequence might affect p_n and to postulate a plausible expression for this dependence in terms of an explicit formula. The second approach is exemplified by the perseveration model and also by a suggestion by Cox based on his work on the regression analysis of binary sequences (1958).

In many of the models we have considered the problem arises of containing p_n in the unit interval, and it is solved by restrictions on the parameter values, restrictions that are occasionally complicated and interdependent. The problem is that although p_n lies in the unit interval the variables on which it may depend, such as total errors or the difference between the number of left-light and right-light onsets, may assume arbitrarily large positive or negative values. The probability itself, therefore, cannot depend linearly on these variables. If a linear relationship is desired, then what is needed is a transformation of p_n that maps the unit interval into the real line.[16]

Such a transformation is given by logit $p = \log\,[p/(1-p)]$. Suppose that this quantity depends linearly on a variable, x, so that logit $p = a + bx$. Then the function that relates p to x is the logistic function that we have already encountered in Eq. 24 and is represented by

$$p = \frac{1}{1 + \exp\,[-(a + bx)]}. \tag{31}$$

As was mentioned in Sec. 2.5, the logistic function is similar in form to the normal ogive and therefore it closely resembles Hull's relation between probability and reaction potential. One advantage of the logistic transformation is that no constraints on the parameters are necessary. A second advantage, to be discussed later, is that good estimates of the parameter values are readily obtained.

Cox (1958) has observed that many studies utilizing stochastic learning models,

> . . . have led to formidable statistical problems of fitting and testing. When these studies aim at linking the observations to a neurophysiological mechanism, it is reasonable to take the best model practicable and to wrestle as vigorously

[16] When the dependent variables are nonnegative, the unit interval needs to be mapped only into the positive reals. This can be achieved, for example, by arranging that the transformations p^{-1} or $\log\,(p^{-1})$ depend linearly on the variables, as illustrated by Eq. 11 and Eq. 25.

as possible with the resulting statistical complications. If, however, the object is primarily the reduction of data to a manageable and revealing form, it seems fair to take for the probability of a success . . . as simple an expression as possible that seems to be the right general shape and which is flexible enough to represent the various possible dependencies that one wants to examine. For this the logistic seems a good thing to consider.

The desirable features of the logistic function carry over into its generalized form, in which logit p is a linear function of several variables. When these variables are given by the components of $\mathbf{W}_n$ (the cumulative number of times each of the events $E_1, E_2, \ldots, E_t$ has occurred in the first $n - 1$ trials), then the logistic function is exactly equivalent to Luce's beta model, so that the same model is obtained from quite different considerations. An example of the logistic function generalized to two dependent variables is given by Eq. 20 for the shuttlebox experiment.

A second example of a generalized logistic function, one that does not follow from Luce's axioms, is given by the analogue of the one-trial perseveration model (Eq. 29) in which

$$\text{logit } \mathbf{p}_n = a + bn + c\mathbf{x}_{n-1}$$

or

$$\mathbf{p}_n = \frac{1}{1 + \exp\left[-(a + bn + c\mathbf{x}_{n-1})\right]}. \tag{32}$$

2.7 Event Effects and Their Invariance

The magnitude of the effect of an event is usually represented by the value of a parameter that, ideally, depends only on constant features of the organism and of the apparatus and therefore does not change during the course of an experiment. There are some experiments or phases of experiments in which the ideal is clearly not achieved, at least not in the context of the available models and the customary definitions of events. Magazine training, detailed instructions, practice trials, and other types of pretraining are some of the devices used to overcome this difficulty. Probably few investigators believe that the condition is ever exactly met in actual experiments, but the principle of parameter invariance within an experiment is accepted as a working rule with the hope that it will at least approximate the truth. (It is also desirable that event effects be invariant from experiment to experiment; this principle provides one test of a model.)

A careful distinction should be drawn between invariance of parameter values and equality of an event's effects in the course of an experiment.

All of the models that have thus far been mentioned imply that the probability change induced by an event *varies* systematically in the course of an experiment; different models specify different forms for this variation. It is therefore only in the context of a particular model that the question of parameter invariance makes sense. Insofar as changes in event effects are in accord with the model, parameters will appear invariant, and we would be inclined to favor the model.

In most models, event effects, defined as probability differences, change because $p_{n+1} - p_n$ depends on at least the value of p_n. This dependence arises in part from the already mentioned need to avoid using transition rules that may take p_n outside the unit interval. But the simple fact that most learning curves (of probability, time, or speed versus trials) are not linear first gave rise to the idea that event effects change.

Gulliksen (1934) reviewed the early mathematical work on the form of the learning curve, and he showed that most of it was based on one of two assumptions about changes in the effect of a trial event. Let t represent time or trials and let y represent a performance measure. Models of Type A begin with the assumption that the improvement in performance induced by an event is proportional to the amount of improvement still possible. The chemical analogy was the monomolecular reaction. This assumption led to a differential equation approximation, $dy/dt = a(b - y)$ whose solution is the exponential growth function $y = b - c \exp(-at)$. Models of Type B begin with the assumption that the improvement in performance induced by an event is proportional to the product of the improvement still possible and the amount already achieved. The chemical analogy was the monomolecular autocatalytic reaction. The assumption led to a differential equation approximation, $dy/dt = ay(b - y)$, whose solution is a logistic function $y = b/[1 + c \exp(-At)]$.

The modern versions of these two models are, of course, Bush and Mosteller's linear-operator models and Luce's beta models. In the linear models the quantity $p_{n+1} - p_n$ is proportional to $\lambda_k - p_n$, the magnitude of the change still possible on repeated occurrences of the event E_k. If all events change p_n in the same direction, let us say toward $p = 0$, then their effects are greatest when p_n is large. In contrast, the effect of an event in the beta-model is smallest when p_n is near zero and unity and greatest when p_n is near 0.5; these statements are true whether the event tends to increase or decrease p_n. The sobering fact is that in more than forty years of study of learning curves and learning a decision has not been reached between these two fundamentally different conceptions.

There is one exception to the rule that no model has the property that the increment or decrement induced in p by an event is a constant. In the middle range of probabilities the effects vary only slightly in many models, and Mosteller (1955) has suggested an additive-increment model to serve as an approximation for this range. The transitions are of the form $\mathbf{p}_{n+1} = Q_k\mathbf{p}_n = \mathbf{p}_n + \delta_k$, where δ_k is a small quantity that may be either positive or negative. This model is not only path-independent and commutative, it is also "p_n-independent."

The foregoing discussion is restricted to path-independent models. In other models the magnitude of the effect of an event depends on other variables in addition to the p-value.

2.8 Simplicity

In model construction appeal is occasionally made to a criterion of simplicity. Because this criterion is always ambiguous and sometimes misleading, it must be viewed with caution: simplicity in one respect may carry with it complexity in another. The relevant attributes of the model are the form of its expressions and the number of variables they contain. Linear forms are thought to be simpler than nonlinear forms (and are approximations to them), which suggests that models with linear operators are simpler than those whose operators are nonlinear. Path-independent models have recursive expressions containing fewer variables than those in path-dependent models, and so they may be thought to be simpler. Classification becomes difficult, however, when other aspects of the models are considered, as a few examples will show.

If we consider the explicit formula, our perspective changes. Commutativity is more fruitful of simplicity than path independence. A conflict arises when we find in the context of an urn (or additive response-strength) model that we can have one only at the price of losing the other (see Sec. 2.6). Also, in such a model even the path-independence criterion taken alone is somewhat ambiguous: one must choose between path independence of the numbers of balls added and path independence of probability changes. Among models with more than one event, the greatest reduction of the number of variables in the explicit formula is achieved by sacrificing both commutativity and path independence, as illustrated by the one-trial perseveration model (Sec. 2.6).

To avoid complicated constraints on the parameters, it appears that nonlinear operators are needed. On the other hand, by using the complicated logistic function, we are assured of the existence of simple *sufficient statistics* for the parameters.

These complications in the simplicity argument are unfortunate: they suggest that simplicity may be an elusive criterion by which to judge models.

3. DETERMINISTIC AND CONTINUOUS APPROXIMATIONS

The models we have been dealing with are, in a sense, doubly stochastic. Knowledge of starting conditions and parameters is not only insufficient to allow one to predict the future response sequence exactly, but, in general, it does not allow exact prediction of the future behavior of the underlying probability or response-strength variable. Even if all subjects, identical in initial probability and other parameter values, behave exactly in accordance with the model, the population is characterized by a distribution of p-values on every trial after the first. For a single subject both the sequence of responses and the sequence of p-values are governed by probability laws.

The variability of behavior in most learning experiments is undeniable, and probably few investigators have ever hoped to develop a mathematical representation that would describe response sequences exactly. Early students of the learning curve, such as Thurstone (1930), acknowledged behavioral variability in the stochastic basis of their models. This basis is obscured by the deterministic learning curve equations which they derived, but these investigators realized that the curves could apply only to the average behavior of a large number of subjects. Stimulus-response theorists, such as Hull, have dealt somewhat differently with the variability problem. In such theories the course of change of the underlying response-strength variable (effective reaction potential) is governed deterministically by the starting values and parameters. Variability is introduced through a randomly fluctuating error term which, in combination with the underlying variable, governs behavior.

Although the stochastic aspect of the learning process has therefore usually been acknowledged, it is only in the developments of the last decade or so that its full implications have been investigated and that probability laws have been thought to apply to the aftereffects of a trial as well as to the performed response.[17]

A second distinguishing feature of recent work is its exact treatment of the discrete character of many learning experiments. This renders the models consistent with the trial-to-trial changes of which learning experiments consist. In early work the discrete trials variable was replaced by a

[17] This change parallels comparable developments in the mathematical study of epidemics and population growth. For discussions of deterministic and stochastic treatments of these phenomena, see Bailey (1955) on epidemics and Kendall (1949) on population growth.

continuous time variable, and the change from one trial to the next was averaged over a unit change in time. The difference equations, representing a discrete process, were thereby approximated by differential equations. The differential equation approximations mentioned in Sec. 2.7 are examples.

Roughly speaking, then, a good deal of early work can be thought of as dealing in an approximate way with processes that have been treated more exactly in recent years. Usually the exact treatment is more difficult, and modern investigators are sometimes forced to make continuous or deterministic approximations of a discrete stochastic process. Occasionally these approximations lead to expressions for the average learning curve, for example, that agree exactly with the stochastic process mean obtained by more difficult methods, but sometimes the approximations are considerably in error. In general, the stochastic treatment of a model allows a greater richness of implications to be drawn from it.

It is probably a mistake to think of deterministic and stochastic treatments of a stochastic model as dichotomous. Deterministic approximations can be made at various stages in the analysis of a model by assuming that the probability distribution of some quantity is concentrated at its mean. A few examples will illustrate ways in which approximations can be made and may also help to clarify the stochastic-deterministic distinction.

3.1 Approximations for an Urn Model

In Sec. 2.6 I considered a special case of the general urn scheme, one that has been applied to the shuttlebox experiment. A few approximations will be demonstrated that are in the spirit of Thurstone's (1930) work with urn schemes. Red balls, whose number, r, is constant, are associated with escape; white balls, whose number, $\mathbf{w}_n$, increases, are associated with avoidance. Pr {escape on trial n} $= \mathbf{p}_n = r/(r + \mathbf{w}_n)$. An avoidance trial increases $\mathbf{w}_n$ by an amount b; an escape trial results in an increase of c balls. Therefore, if D_k represents an operator that acts on $\mathbf{w}_n$,

$$\mathbf{w}_{n+1} = \begin{cases} D_1\mathbf{w}_n = \mathbf{w}_n + b & \text{with probability } 1 - \mathbf{p}_n = \dfrac{\mathbf{w}_n}{r + \mathbf{w}_n} \\[2ex] D_2\mathbf{w}_n = \mathbf{w}_n + c & \text{with probability } \mathbf{p}_n = \dfrac{r}{r + \mathbf{w}_n}. \end{cases} \tag{33}$$

Consider a large population of organisms that behave in accordance with the model and have common values of r, w_1, b, and c. On the first trial all subjects have the same probability $\mathbf{p}_1 = p_1$ of escape. Some will escape and the rest will avoid. If $b \neq c$, there will be two subsets on the second trial, one for which $\mathbf{w}_2 = w_1 + b$ and another for which $\mathbf{w}_2 = w_1 + c$. Each of these subsets will divide again on the second trial, but because of commutativity there will be three, not four, distinct values of $\mathbf{w}_3$: $w_1 + 2b$, $w_1 + b + c$, and $w_1 + 2c$. Each distinct value of $\mathbf{w}_n$ corresponds to a distinct p-value. On every trial after the first there is a distribution of p-values.

With two events there are, in general, 2^{n-1} distinct p-values on the nth trial, each corresponding to a distinct sequence of events on the preceding $n - 1$ trials. If the events commute, as in this case, then the number of distinct p-values is reduced to n, the trial number.

Our problem for the urn model is to determine the mean probability of escape on the nth trial, the average being taken over the population.

We let $1 \leqslant \nu \leqslant n$ be the index for the n subsets with distinct p-values on trial n. Let $P_{\nu n}$ be the proportion of subjects in the νth subset on trial n and let $p_{\nu n}$ be the p-value for this subset. Then the mth raw moment of the distribution on trial n is defined by

$$V_{m,n} = E(\mathbf{p}_n^m) = \sum_{\nu} p_{\nu n}^m P_{\nu n}. \tag{34}$$

We use this definition later.

Because $\mathbf{w}_n$, but not $\mathbf{p}_n$, is transformed linearly, it is convenient to determine $E(\mathbf{w}_n) = \bar{w}_n$ first. The increment in numbers of white balls, $\Delta\mathbf{w}_n = \mathbf{w}_{n+1} - \mathbf{w}_n$, is either b or c, and its conditional expectation, conditional on the value of $\mathbf{w}_n$, is given by

$$E_b(\Delta\mathbf{w}_n \mid \mathbf{w}_n) = b\left(\frac{\mathbf{w}_n}{r + \mathbf{w}_n}\right) + c\left(\frac{r}{r + \mathbf{w}_n}\right) = \frac{b\mathbf{w}_n + cr}{\mathbf{w}_n + r}, \tag{35}$$

where E_b denotes the operation of averaging over the binomial distribution of the increment. The unconditional expectation of the increment is obtained by averaging Eq. 35 over the distribution of $\mathbf{w}_n$-values. Using the expectation operator E_w to represent this averaging process, we have

$$E(\Delta\mathbf{w}_n) = E_w E_b(\Delta\mathbf{w}_n \mid \mathbf{w}_n) = E_w\left(\frac{b\mathbf{w}_n + cr}{\mathbf{w}_n + r}\right). \tag{36}$$

Note that the right-hand member of this expression is not in general expressible as a simple function of $\bar{w}_n$.

Now we perform two steps of deterministic approximation. First, we replace $\Delta\mathbf{w}_n$, which has a binomial distribution, by its average value. From Eq. 35 the increment in $\mathbf{w}_n$ (conditional on the value of $\mathbf{w}_n$) can then be written

$$\Delta\mathbf{w}_n \simeq \bar{\Delta}\mathbf{w}_n = \frac{b\mathbf{w}_n + cr}{\mathbf{w}_n + r}. \tag{37}$$

Second, we act as if the distribution of $\mathbf{w}_n$ is entirely concentrated at its mean value $\bar{w}_n$. The expectation of the ratio in Eq. 36 is then the ratio itself, and we have

$$\Delta \mathbf{w}_n \simeq \bar{\Delta} \bar{w}_n = \frac{b\bar{w}_n + cr}{\bar{w}_n + r}. \tag{38}$$

These two steps accomplish what Bush and Mosteller (1955) call the *expected-operator approximation.* In this method, the change in the distribution of p-values (or w-values) on a trial is represented as the mean p-value (or w-value) acted on by an "average" operator (that is, subject to an average change). Two approximations are involved: the first replaces a distribution of quantities by its mean value and the second replaces a distribution of changes by the mean change. The average operator $\bar{D}$ is revealed in this example if we rewrite Eq. 38 as

$$\bar{w}_{n+1} \simeq \bar{w}_n + \bar{\Delta}\bar{w}_n \equiv \bar{D}\bar{w}_n = \bar{w}_n + b\left(\frac{\bar{w}_n}{\bar{w}_n + r}\right) + c\left(\frac{r}{\bar{w}_n + r}\right)$$

and compare it to Eq. 33. The increments b and c are weighted by their approximate probabilities of being applied. In general,

$$V_{1,n} = E_w\left(\frac{r}{r + \mathbf{w}_n}\right) \neq \left(\frac{r}{r + E_w(\mathbf{w}_n)}\right) = \frac{r}{r + \bar{w}_n}.$$

But notice that our approximation, which assumes that every w_n-value is equal to $\bar{w}_n$, leads to this simple relationship.

The discrete stochastic process given by the urn scheme has surely been transformed by virtue of the approximations—but transformed into what? There are at least two interpretations. The first is that the approximate process defines a determined sequence of approximate probabilities for a subject. Like the original process the approximation is stochastic, but on only one "level": the response sequence is governed by probability laws but the response probability sequence is not. According to this interpretation, the approximate model is not deterministic, but it is "more deterministic" than the original urn scheme.

The second interpretation is that the approximate process defines a determined sequence of proportions of white balls, $\bar{w}_n$, for a population of subjects, and thereby defines a determined sequence of proportions of correct responses, that is, the mean learning curve. According to this interpretation the approximate model is deterministic and it applies only to groups of subjects.

However we think of the approximate model, it is defined by means of a nonlinear difference equation for w_n (Eq. 38). Solution of such equations is difficult and a continuous approximation is helpful. We assume that the trials variable n is continuous and that the growth of $\bar{w}_n$ is gradual rather than step-by-step. The approximate difference equation given by Eq. 38 can thus itself be approximated by a differential equation:

$$\frac{dw}{dn} = \frac{bw + cr}{w + r} . \tag{39}$$

Integration gives a relation between w and n and therefore between $V_{1,n}$ and n. For the special case of $b = c$ and $w_1 = 0$ the relation has the simple form $w/c = n - 1$, giving

$$V_{1,n} = \frac{r}{r + (n - 1)c} . \tag{40}$$

Equation 40 is an example of an approximation that is also an exact result. In this example it occurs for an uninteresting reason: equating the values of b and c transforms the urn scheme into a single-event model in which the approximating assumption, namely, that all p-values are concentrated at their mean, is correct.

3.2 More on the Expected-Operator Approximation

The expected-operator approximation is important because results obtained by this method are, unfortunately, the only ones known for certain models. Because the approximation also generates a more deterministic model, it is discussed here rather than in Sec. 5 on methods for the analysis of models.

Suppose that a model is characterized by a set $\{Q_k\}$ of operators on response probabilities, where Q_k is applied on trial n with probability $P_k = P_k(p_n)$. This discussion is confined to path-independent models, and therefore P_k can be thought of as a function of at most the p-value on the trial in question and fixed parameters. Because the p-value on a trial may have a distribution over subjects, the probability P_k may also have a probability distribution. The expected operator $\bar{Q}$ is defined by the conditional expectation

$$\bar{Q}\mathbf{p}_n = E_k(Q_k\mathbf{p}_n \mid \mathbf{p}_n) = \sum_k P_k(\mathbf{p}_n)Q_k\mathbf{p}_n. \tag{41}$$

The expectation operator E_k represents the operation of averaging over values of k. The first deterministic approximation in the expected-operator method is the assumption that the same operator—the expected operator—is applied on every trial. Therefore $\mathbf{p}_{n+1} \simeq \bar{Q}\mathbf{p}_n$ for all n.

What is of interest is the average probability on the $(n + 1)$st trial. This can be obtained by removing the condition on the expectation in Eq. 41 by averaging again, this time over the distribution of p_n-values.

Symbolizing this averaging process by E_p, we have $V_{1,n+1} \simeq E_p(\bar{Q}\mathbf{p}_n)$. The second approximation is to replace $E_p(\bar{Q}\mathbf{p}_n)$ by $\bar{Q}[E_p(\mathbf{p}_n)] = \bar{Q}(V_{1,n})$. This approximation is equivalent to the (deterministic) assumption that the $\mathbf{p}_n$-distribution is concentrated at its mean. In cases in which $\bar{Q}\mathbf{p}$ is linear in $\mathbf{p}$, however, $E_p[\bar{Q}\mathbf{p}] = \bar{Q}[E_p(\mathbf{p})]$ is exact and therefore no assumption is needed. (In nonlinear models the method does not seem to give exact results. Of course, it is just for these models that exact methods are difficult to apply.) Applying the second approximation to Eq. 41, we get

$$V_{1,n+1} \simeq \bar{Q}V_{1,n} = \sum_k P_k(V_{1,n})Q_k V_{1,n} \tag{42}$$

as an approximate recursive formula for the mean of the p-value distribution on the nth trial.

EXPECTED OPERATOR FOR TWO EXPERIMENTER-CONTROLLED EVENTS. Consider the model in Sec. 2.4 for the prediction experiment. The operators and the rules for their application are given by Eq. 12:

$$p_{n+1} = \begin{cases} Q_1 p_n = \alpha p_n + 1 - \alpha & \text{if } \ y_n = 1 \\ Q_2 p_n = \alpha p_n & \text{if } \ y_n = 0. \end{cases}$$

Recall that the y_n-sequence, and thus the sequence of operators, can be predetermined. Therefore the probability p_n is known exactly from Eq. 14. Often the event sequence is generated by a random device, and, as mentioned in Sec. 2.4, an approximation for p_n can be developed by assuming that the subject experiences the average of all the sequences that the random device generates.

Because the subject has, in fact, experienced one particular event sequence, the approximation may be a poor one. An alternative interpretation of the approximation is that, like many deterministic models, it applies to the average behavior of a large group of subjects. This interpretation is reasonable only if the event sequences are independently generated for each of a large number of subjects. In many experiments to which the approximation has been applied this proviso has unfortunately not been met.

Suppose that the $\{\mathbf{y}_n\}$ are independent and that $\Pr\{\mathbf{y}_n = 1\} = \pi$ and $\Pr\{\mathbf{y}_n = 0\} = 1 - \pi$. Then $P_1(p) = \pi$ and $P_2(p) = 1 - \pi$ are independent of the p-value, as is always the case with experimenter control. Equation 42 gives the recursive relation

$$V_{1,n+1} = \alpha V_{1,n} + (1 - \alpha)\pi. \tag{43}$$

Unlike Eq. 38 for the urn model, Eq. 43 is easily solved and a continuous approximation is not necessary. The solution has already been given in

Sec. 2.4 for the repeated application of the same linear operator. The approximate learning curve equation is

$$V_{1,n} = \alpha^{n-1}V_{1,1} + (1 - \alpha^{n-1})\pi = \pi - \alpha^{n-1}(\pi - V_{1,1}). \tag{44}$$

In this example the result of the approximation is exact in a certain sense: if we average the explicit equation for the model, Eq. 14, over the binomial event distributions, then the result obtained for $V_{1,n}$ will be the same as that given by Eq. 44.

EXPECTED OPERATOR AND THE ASYMPTOTIC PROBABILITY FOR EXPERIMENTER-SUBJECT EVENTS. If one is reluctant to assume for the prediction experiment that reward and nonreward have identical effects, then changes in response probability may depend on the response performed, and the model events are under experimenter-subject control. Let us assume that the outcomes are independent and that $\Pr\{\mathbf{O}_j = O_1\} = \pi$, $\Pr\{\mathbf{O}_j = O_2\} = 1 - \pi$. If we assume response-symmetry and outcome-symmetry and use the symbols given in Table 1, the appropriate Bush-Mosteller model is given as follows:

Event	Operator, Q_k	$P_k(\mathbf{p}_n)$
A_1, O_1	$Q_1\mathbf{p}_n = \alpha_1\mathbf{p}_n + 1 - \alpha_1$	$\mathbf{p}_n\pi$
A_2, O_1	$Q_2\mathbf{p}_n = \alpha_2\mathbf{p}_n + 1 - \alpha_2$	$(1 - \mathbf{p}_n)\pi$
A_1, O_2	$Q_3\mathbf{p}_n = \alpha_2\mathbf{p}_n$	$\mathbf{p}_n(1 - \pi)$
A_2, O_2	$Q_4\mathbf{p}_n = \alpha_1\mathbf{p}_n$	$(1 - \mathbf{p}_n)(1 - \pi)$

(45)

The expected operator approximation (Eq. 42) gives

$$V_{1,n+1} \simeq (1 - \alpha_2)\pi + [\pi + (\alpha_1 - \alpha_2)(1 - 2\pi)]V_{1,n} + (\alpha_1 - \alpha_2)(1 - 2\pi)V_{1,n}^2. \tag{46}$$

This quadratic difference equation is difficult to solve, and Bush and Mosteller approximate it by a differential equation which can then be integrated to give an approximate value for $V_{1,n}$.

The continuous approximation is not necessary if we confine our attention to the asymptotic behavior of the process. At the asymptote the moments of the p-value distribution are no longer subject to change, and therefore $V_{1,n+1} = V_{1,n} = V_{1,\infty}$. Using these substitutions in Eq. 46, we get a quadratic equation whose solution is

$$V_{1,\infty} \simeq \frac{2\pi(1 - \gamma) - 1 + \sqrt{2(\pi - 1)^2 + 4\pi(1 - \pi)\gamma^2}}{2(2\pi - 1)(1 - \gamma)}, \tag{47}$$

where $\gamma = (1 - \alpha_2)/(1 - \alpha_1) \neq 1$ (Bush & Mosteller, 1955, p. 289). For the model defined by Eq. 45 no expression for the asymptotic proportion of A_1 responses is known other than the approximation of Eq. 47. There is little evidence concerning its accuracy. Our ignorance about this model is especially unfortunate because of the considerable recent interest in asymptotic behavior in the prediction experiment.

Several conclusions about the use of the expected-operator approximation are illustrated in Figs. 1 and 2. Each figure shows average proportions of A_1 responses for 20 artificial subjects behaving in accordance with the model of Eq. 45. All 20 subjects in each group experienced the same event sequence. For both sets of subjects, $\alpha_1 = 0.90$, $\alpha_2 = 0.95$, and $p_1 = V_{1,1} = 0.50$. Reward, then, had twice the effect of nonreward. For the subjects of Fig. 1, $\pi = 0.9$; for those of Fig. 2, $\pi = 0.6$. In both examples the expected operator estimate for the asymptote (Eq. 47) seems too high. Also shown in each figure are exact and approximate (Eq. 44)

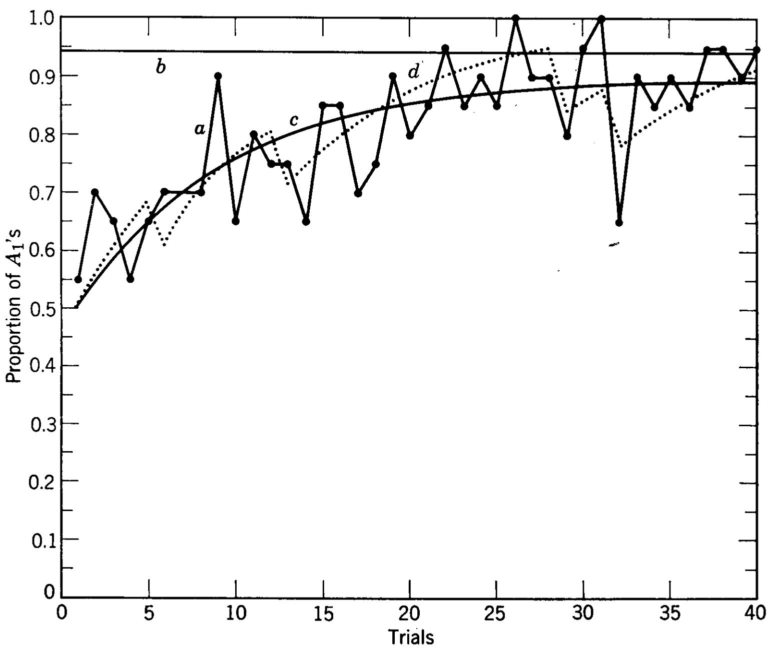

Fig. 1. *a*. The jagged solid line gives the mean proportion of A_1 responses of 20 stat-organisms behaving in accordance with the four-event model with experimenter-subject control (Eq. 45) with $\alpha_1 = 0.90$, $\alpha_2 = 0.95$, $p_1 = 0.50$ and $\pi = 0.90$. *b*. The horizontal line gives the expected-operator approximation of the four-event model asymptote (Eq. 47). *c*. The smooth curve gives the approximate learning curve (Eq. 44) for the two-event model with experimenter control (Eq. 12) with $\hat{\alpha} = 0.8915$ estimated from stat-organism "data," and $p_1 = 0.50$. *d*. The dotted line gives the exact learning curve (Eq. 14) for the two-event model.

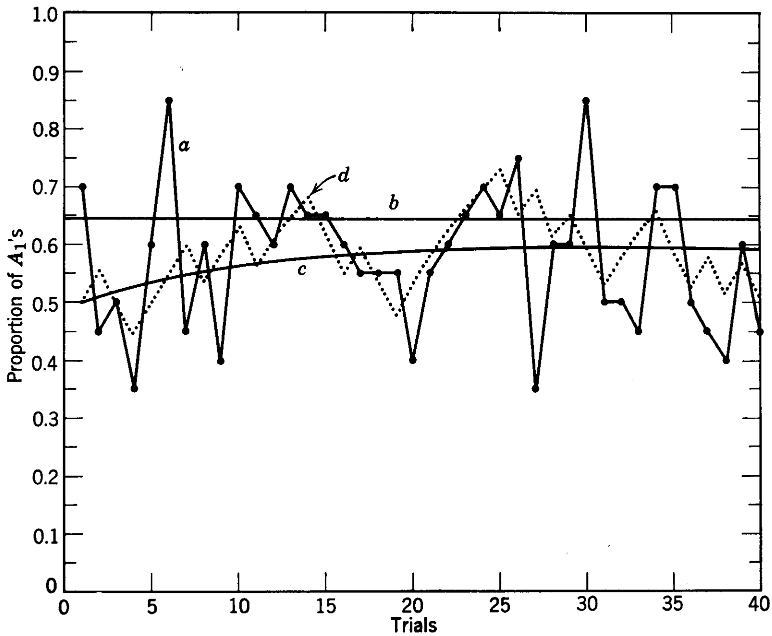

Fig. 2. *a.* The jagged solid line gives the mean proportion of A_1 responses of 20 stat-organisms behaving in accordance with the four-event model with experimenter-subject control (Eq. 45) with $\alpha_1 = 0.90$, $\alpha_2 = 0.95$, $p_1 = 0.50$, and $\pi = 0.60$. *b.* The horizontal line gives the expected-operator approximation of the four-event model asymptote (Eq. 47). *c.* The smooth curve gives the approximate learning curve (Eq. 44) for the two-event model with experimenter-control (Eq. 12) with $\hat{\alpha} = 0.8960$ estimated from stat-organism "data," and $p_1 = 0.50$. *d.* The dotted line gives the exact learning curve (Eq. 14) for the two-event model.

learning curves of the model with two experimenter-controlled events, which has been fitted to the data. The superiority of the exact curve is evident. I shall discuss later the interesting fact that even though the data were generated by a model (Eq. 45) in which reward had more effect than nonreward a model that assumes equal effects (Eq. 12) produces learning curves that are in "good" agreement with the data.

3.3 Deterministic Approximations for a Model of Operant Conditioning

Examples of approximations that transform a discrete stochastic model into a continuous and completely deterministic model are to be found in treatments of operant conditioning (Estes, 1950, 1959; Bush & Mosteller, 1951). To demonstrate the flavor of these treatments and the approximations used, a sketch of a model along the lines of Estes' is given. In applying a choice-experiment analysis to a free-operant situation, each interresponse period is thought of as a sequence of short intervals of constant length h. It is these intervals that are identified as "trials." During each interval the subject chooses either to press (A_1) or not to press (A_2) the lever; pressing occurs with some probability and is rewarded. The probability is assumed to be unchanged by trials (intervals) on which A_2 occurs and increased by trials (intervals) on which A_1 occurs. The problem is to describe the resulting sequence of interresponse times. To define the model completely it is necessary to consider the way in which $\Pr\{A_1\}$ is increased by reward. We do this in the context of an urn scheme.

An urn contains $\mathbf{x}$ white balls (which correspond to A_1) and $b - \mathbf{x}$ red balls (which correspond to A_2). At the beginning of a trial a sample of balls is randomly selected from the urn. Each ball has the same fixed probability of being included in the sample, which is of size $\mathbf{s}$. The proportion of white balls in the sample defines the probability $\mathbf{p} = \Pr\{A_1\}$ for the trial in question. At the end of an interval in which A_2 occurs the sample of balls is retained and used to define $\mathbf{p}$ for the interval that follows. At the end of an interval in which A_1 occurs all the red balls in the sample [there are $\mathbf{s}(1 - \mathbf{p})$ of them] are replaced by white balls and the sample is returned to the urn. The number of trials (intervals of length h) from one press to the next, including the one on which the lever is pressed, is $\mathbf{m}$. The interresponse time thus defined is $\tau = \mathbf{m}h$.

The deterministic approximations are as follows:

1. The sample size $\mathbf{s}$ is binomially distributed. It is replaced by its mean $\bar{s}$.
2. Conditional on the value of $\mathbf{s}$, the proportion of white balls $\mathbf{p}$ is binomially distributed. It is replaced by its mean $\mathbf{x}/b$.
3. Conditional on the value of $\mathbf{p}$, the number of intervals $\mathbf{m}$ is distributed geometrically, with $\Pr\{\mathbf{m} = m\} = \mathbf{p}(1 - \mathbf{p})^{m-1}$. It is replaced by its mean, $1/\mathbf{p}$. By combining the other approximations with this one, we can approximate the number of intervals in the interresponse period by $\mathbf{m} \simeq b/\mathbf{x}$ and therefore the interresponse time is approximated by $\tau \simeq hb/\mathbf{x}$.
4. Finally, $\mathbf{s}(1 - \mathbf{p})$, which is the increase in $\mathbf{x}$ (the number of white balls in the urn) that results from reward, is replaced by the product of the means of $\mathbf{s}$ and $1 - \mathbf{p}$, and becomes $\bar{s}(b - \mathbf{x})/b$.

The result of this series of approximations is a deterministic process. Given a starting value of x/b and a value for the mean sample size $\bar{s}$, the approximate model generates a determined sequence of increasing values of x and of decreasing latencies.

The final approximation is a continuous one. The discrete variables x and τ are considered to be continuous functions of time, $x(t)$ and $\tau(t)$, and $n = n(t)$ is the cumulative number of lever presses. A first integration of an approximate differential equation gives the rate of lever pressing as a (continuous) function of time; integration of this rate gives $n(t)$.

Little work has been done on this model or its variants in a stochastic form. We therefore have little knowledge as to which features of the stochastic process are obscured by the deterministic approximations. One feature that definitely is obscured depends on sampling fluctuations of the proportion of white balls $\mathbf{p}$. When $\mathbf{p}$ has a high value, then the interresponse time will tend to be short and the increment in $\mathbf{x}$ small; when the value of $\mathbf{p}$ happens to be low, then the interresponse time will tend to be above its mean and the increment in $\mathbf{x}$ large. One consequence is that interresponse times constitute a dependent sequence such that the variance of the cumulated interresponse time will be less than the sum of the variances of its components.

4. CLASSIFICATION AND THEORETICAL COMPARISON OF MODELS

A good deal of work has been devoted to the mathematical analysis of various model types, but less attention has been paid to the development of systematic criteria by which to characterize or compare models. What are the important features that distinguish one model from another? More pertinent, in what aspects or statistics of the data do we expect these features to be reflected? The need to answer these questions arises primarily in comparative and "baseline" applications of models to data.

Comparative studies, in which we seek to determine which of several models is most appropriate for a set of data, require us to discover *discriminating statistics* of the data: these are statistics that are sensitive to the important differences among the models and that should therefore help us to select one of several models as best.

Once a model is selected as superior, the statistician may be satisfied but the psychologist is not; the data presumably have certain properties that are responsible for the model's superiority, properties that the psychologist wants to know about.

Finally, a model is occasionally used as a baseline against which data are compared in order to discover where the discrepancies lie. Again, a study is incomplete if it leads simply to a list of agreeing and disagreeing statistics;

what is needed as well is an interpretation of these results that suggests which of the model's features seem to characterize the data and which do not.

Analysis of the distinctive features of model types and how they are reflected in properties of the data is useful in the discovery of discriminating statistics, in the interpretation of a model's superiority to others, and in the interpretation of points of agreement and disagreement between a model and data. Some of the important features of several models were indicated in passing as the models were introduced in Sec. 2. A few examples of more systematic methods of comparison are given in this section. Where possible, they are illustrated by reference to one of the comparative studies that have been performed on the Solomon-Wynne shuttlebox data (Bush & Mosteller, 1959; Bush, Galanter, & Luce, 1959), on the Goodnow two-armed bandit data (Sternberg, 1959b), and on some T-maze data (Galanter & Bush, 1959; Bush, Galanter, & Luce, 1959).

4.1 Comparison by Transformation of the Explicit Formula

A comparable form of expression can be used for all of the path-independent commutative-operator models that were introduced in Sec. 2. By suitably defining new parameters in terms of the old, we can write the explicit formula for $\mathbf{p}_n$ as a function of an expression that is linear in the components of $\mathbf{W}_n$. (Recall that $\mathbf{W}_n$ is the vector whose t components give the cumulative number of occurrences of events $E_1, \ldots, E_t$ prior to the nth trial). Suppose $\mathbf{W}_n = (\mathbf{t}_n, \mathbf{s}_n)$, as in the shuttlebox experiment, where $\mathbf{t}_n$ is the total number of avoidances and $\mathbf{s}_n$ the total number of shocks before the nth trial. Let $\mathbf{p}_n$ be the probability of error (nonavoidance) which decreases as $\mathbf{s}_n$ and $\mathbf{t}_n$ increase. Then for the linear-operator model (Eq. 11) we have

$$\mathbf{p}_n = \exp\left[-(a + b\mathbf{t}_n + c\mathbf{s}_n)\right],$$

and thus

$$\log \mathbf{p}_n = -(a + b\mathbf{t}_n + c\mathbf{s}_n). \tag{48}$$

For Luce's beta model (Eq. 20)

$$\mathbf{p}_n = \frac{1}{1 + \exp(a + b\mathbf{t}_n + c\mathbf{s}_n)},$$

and thus

$$\operatorname{logit} \mathbf{p}_n = -(a + b\mathbf{t}_n + c\mathbf{s}_n). \tag{49}$$

For the special case of the urn scheme (Eq. 26)

$$\mathbf{p}_n = \frac{1}{a + b\mathbf{t}_n + c\mathbf{s}_n},$$

and thus

$$\frac{1}{\mathbf{p}_n} = a + b\mathbf{t}_n + c\mathbf{s}_n. \tag{50}$$

Finally, for Mosteller's additive-increment-approximation model (Sec. 2.7), applied to the two-event experiment,

$$\mathbf{p}_n = -(a + b\mathbf{t}_n + c\mathbf{s}_n), \tag{51}$$

For each model some transformation, $g(p)$, of the response probability is a linear function of $\mathbf{t}_n$ and $\mathbf{s}_n$. The models differ only in the transformations they specify.

The behavior of models for which this type of expression is possible can be described by a simple nomogram. Examples for the first three models above are given in Fig. 3, in which the transformation $x = dg(p) + e$ is plotted for each model (d and e are constants). The units on the abscissa are arbitrary. To facilitate comparisons, the coefficients d and e are

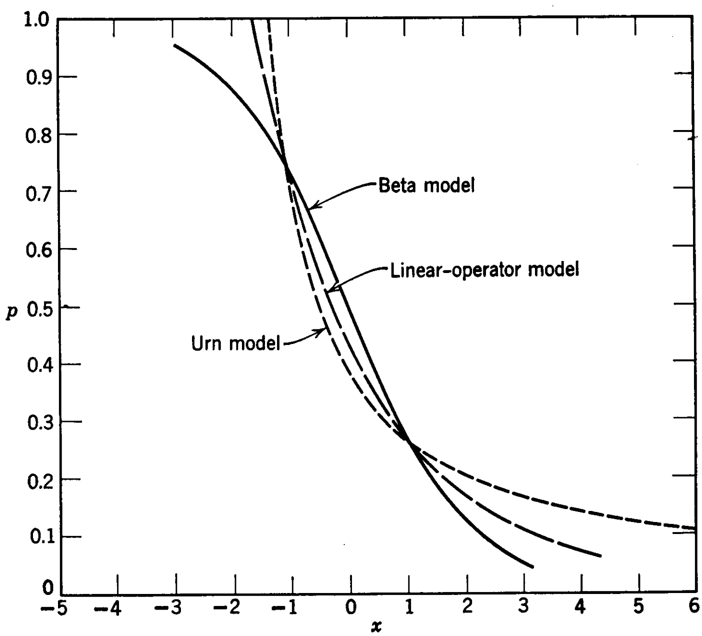

Fig. 3. Nomograms for three models. The linear-operator model (Eq. 48) is represented by $\log_e p = -0.500x + 0.837$. The beta model (Eq. 49) is represented by logit $p = x$. The urn scheme (Eq. 50) is represented by $p^{-1} = 1.214x + 2.667$. Constants were so chosen that curves would agree at $p = 0.25$ and $p = 0.75$. The units on the abscissa are arbitrary.

chosen so that values of x agree at $p = 0.25$ and $p = 0.75$. The additive-increment model is represented by a straight line passing through the two common points.

The nomogram is interpreted as follows: a subject's state (the value of the linear expression $a + b\mathbf{t}_n + c\mathbf{s}_n$) is represented by a point on the abscissa, and his error probability is given by the corresponding point on the p-axis. The occurrence of an event corresponds to a displacement along the abscissa whose magnitude depends only on the event and not on the starting position. For this example all displacements are to the right and correspond to reductions in the error probability. An avoidance corresponds to a displacement b units to the right, and a shock to a displacement c units to the right. If events have equal effects, then we have a single-event model, and each trial corresponds to the same displacement.

Although a displacement along the abscissa is a constant for a given event, the corresponding displacement on the probability axis depends on the slope of the curve at that point. Because the slope depends on the p-value (except for the additive-increment model), the probability change corresponding to an event depends on that value. Because the probability change induced by an event depends only on the p-value, this type of nomogram is limited to path-independent models. Its use is also limited to models in which the operators commute. For the additive-increment, path-independent urn, and beta models it can be used when there are more than two events. For these models events that increase p_n correspond to displacements to the left.

A number of significant features of the models can be seen immediately from Fig. 3. First, the figure indicates that in the range of probabilities from 0.2 to 0.8 the additive-increment model approximates each of the other three models fairly well. This supports Mosteller's (1955) suggestion that estimates for the additive model based on data from this range be used to answer simple questions such as which of two events has the bigger effect. Caution should be exercised, however, in applying the model to data from a subsequence of trials that begins after trial 1. Even if we had reason to believe that all subjects have the same p-value on trial 1, we would probably be unwilling to assume that the probabilities on the first trial of the subsequence are equal from subject to subject. Therefore the estimation method used should not require us to make this assumption.

A second feature disclosed by Fig. 3 concerns the rate of learning (rate of change of $\mathbf{p}_n$) at the early stages of learning. When $\mathbf{p}_n$ is near unity, events in the urn and linear operator models have their maximum effects. In contrast, the beta model requires that $\mathbf{p}_n$ change slowly when it is near unity. If the error probability is to be reduced from its initial value to, let us say, 0.75 in a given number of early trials, then for the beta model to accomplish this reduction it must start at a lower initial value than the other models or its early events must correspond to larger displacements along the x-axis than they do in the other models. Early events tend predominantly to be errors, and therefore the second alternative corresponds to a large error-effect.

A third feature has to do with the behavior of the models at low values of $\mathbf{p}_n$. The models differ in the rates at which the error probability approaches zero. Especially notable is the urn model which, after a translation, is of the form $p = x^{-1}$ $(x \geqslant 1)$. Unlike the situation in the other models, the area under the curve of p versus x diverges, so that in an unlimited sequence of trials we expect an unlimited number of errors. (The expected number of trials between successive errors increases as the trial number increases but not rapidly enough to maintain the total number of errors at a finite value). The urn model, then, tends to produce more errors than the other models at late stages in learning.

This analysis of the three models helps us to understand some of the results that have been obtained in applications to the Solomon-Wynne avoidance-learning data. The analyses have assumed common values over subjects for the initial probability and other parameters. The relevant results are as follows:

1. A "model-free" analysis, which makes only assumptions that are common to the three models, shows that an avoidance response leads to a greater reduction in the escape probability than an escape response. This analysis is described in Sec. 6.7.

2. The best available estimate of p_1 in this experiment is 0.997 and is based on 331 trials, mainly pretest trials (Bush & Mosteller, 1955).

3. The linear-operator model is in good agreement with the data in every way that they have been compared (Bush & Mosteller, 1959). The estimates are $\hat{p}_1 = 1.00$, $\hat{\alpha}_1$ (avoidance parameter) $= 0.80$, and $\hat{\alpha}_2$ (escape parameter) $= 0.92$. According to the parameter values, for which approximately the same estimates are obtained by several methods (Bush & Mosteller, 1955), escape is less effective than avoidance in reducing the escape probability.

4. There is one large discrepancy between the urn model and the data. Twenty-five trials were examined for each of 30 subjects. The last escape response occurred, on the average, on trial 12. The comparable figure for the urn model is trial 20. As in the linear-operator model, the estimates suggest that escape is less potent than avoidance (Bush & Mosteller, 1959).

5. One set of estimates for the beta model is given by $\hat{p}_1 = 0.94$, $\hat{\beta}_1$ (avoidance parameter) $= 0.83$, and $\hat{\beta}_2$ (escape parameter) $= 0.59$ (Bush, Galanter, & Luce, 1959). With these estimates, the model differs from the

data in several respects, notably producing underestimates of intersubject variances of several quantities, such as total number of escapes. As discussed later, this probably occurs because the relative effects of avoidance and escape trials are incorrectly represented by the model.

6. The approximate maximum-likelihood estimates for the beta model are given by $\hat{p}_1 = 0.86$, $\hat{\beta}_1 = 0.74$, and $\hat{\beta}_2 = 0.81$. In contrast to the inference made by Bush, Galanter, and Luce, these estimates imply that escape is less effective than avoidance. The estimate of the initial probability of escape is lower than theirs, however.

These results strongly favor the linear-operator model. The results of analysis of the data with the other models are intelligible in the light of our study of the nomogram. The first set of estimates for the beta model gives an absurdly low value of p_1. In addition, the relative magnitudes of escape and avoidance effects are reversed. The second set of estimates, which avoids attributing to error trials an undue share of the learning, underestimates p_1 by an even greater amount. Apparently, if the beta model is required to account for other features of the data as well, it cannot describe the rapidity of learning on the early trials of the experiment. The major discrepancy between the urn model and data is in accord with the exceptional behavior of that model at low p-values; the average trial of the last error is a discriminating statistic when the urn model is compared to others.

Before we leave this type of analysis, it is instructive to consider Hull's model (1943) in the same way. This model was discussed briefly in Sec. 2.5. It is intended to describe the change in probability of reactions of the all-or-none type, such as conditioned eyelid responses and barpressing in a discrete-trial experiment. If we assume that incentive and drive conditions are constant from trial to trial, then the model involves the assumptions that (1) reaction potential (${}_SE_R$) is a growth function of the number of reinforcements and (2) reaction probability (q) is a (normal) ogival function of the difference between the reaction potential and its threshold (${}_SL_R$) when this difference is positive; otherwise the probability is zero.

The assumptions can be stated formally as follows:

$$1. \qquad {}_SE_R = M(1 - e^{-AN}), \qquad (A, M > 0). \tag{52}$$

The quantity N is defined to be the number of reinforcements, not the total number of trials. Unrewarded trials correspond to the application of an identity operator to ${}_SE_R$.

$$2. \qquad \operatorname{logit} q = \begin{cases} B + C({}_SE_R - {}_SL_R), & ({}_SE_R > {}_SL_R) \\ -\infty, & ({}_SE_R \leqslant {}_SL_R). \end{cases} \tag{53}$$

For uniformity and ease of expression the logistic function is substituted for the normal ogive. When these equations are combined, we obtain for the probability p of not performing the response

$$\log(\text{logit}\, p + c) = \begin{cases} -(a + bN), & (N > k) \\ -\infty, & (N \leqslant k). \end{cases} \tag{54}$$

This is to be compared to Eq. 48 and Eq. 49.

From Hull's figures a rough estimate of $c = 5$ is obtained. This allows us to construct a nomogram, again choosing constants so that the curve agrees with the other models at $p = 0.25$ and $p = 0.75$. The result is displayed in Fig. 4, with nomograms for the linear and beta models. The curve for Hull's model falls in between those of the other two. The existence of a threshold makes the model difficult to handle mathematically, but, in contrast to the beta model, it allows learning to occur in a finite number of trials even with an initial error-probability of unity.

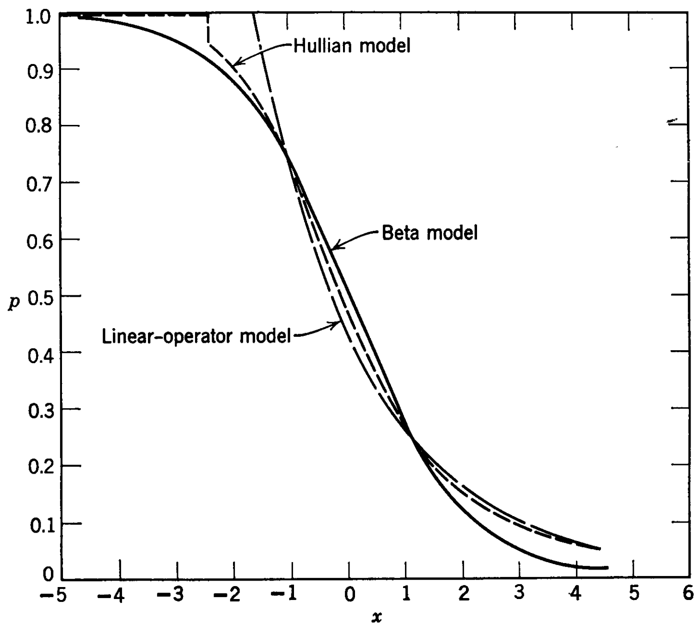

Fig. 4. Nomogram for Hullian model (Eq. 48) and two other models. The Hullian model is represented by $\log_e(\text{logit}\, p + 5) = -0.204x + 1.585$. The nomograms for linear-operator and beta models are those presented in Fig. 3.

4.2 Note on the Classification of Operators and Recursive Formulas

The classification of operators and recursive formulas is probably more relevant to the choice of mathematical methods for the analysis of models than it is to the direct appreciation of their properties. (Two exceptions considered later are the implications of commutativity and of the relative magnitudes of the effects of rewarded and nonrewarded trials.) The classification described here is based on the arguments that appear in the recursive formula. Let us consider models with two subject-controlled events, in which $\mathbf{x}_n = 1$ if E_2 occurs on trial n, $\mathbf{x}_n = 0$ if E_1 occurs on trial n, and $\mathbf{p}_n = \Pr\{\mathbf{x}_n = 1\}$. A rough classification is given by the following list:

1. $p_{n+1} = f(p_n)$. Response-independent, path-independent. *Example:* single-operator model (Eq. 28).
2. $p_{n+1} = f(n, p_n)$. Response-independent, quasi-independent of path. *Example:* urn scheme (Eq. 25) with equal event-effects.

Classes 1 and 2 produce sequences of independent trials and, if there are no individual differences in initial probabilities and other parameters, they do not lead to distributions of p-values.

3. $\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{x}_n)$. Response-dependent, path-independent. *Example:* linear commutative operator model (Eq. 8).
4. $\mathbf{p}_{n+1} = f(n, \mathbf{p}_n; \mathbf{x}_n)$. Response-dependent, quasi-independent of path. *Example:* general urn scheme (Eq. 25).
5. $\mathbf{p}_{n+1} = f(\mathbf{p}_n; \mathbf{x}_n, \mathbf{x}_{n-1})$. Path-dependent. *Example:* one-trial perseveration model (Eq. 30).

4.3 Implications of Commutativity for Responsiveness and Asymptotic Behavior

In Sec. 2.2 I pointed out that in a model with commutative events there is no "forgetting": the effect of an event on $\mathbf{p}_n$ is the same whether it occurred on trial 1 or on trial $n - 1$. The result is that models with commutative events tend to respond sluggishly to changes in the experiment.

As an example to illustrate this phenomenon we take the prediction experiment and, for the moment, consider it as a case of experimenter-controlled events. We use the linear model (Eq. 12) to illustrate non-commutative events and the beta model (Eq. 21) to illustrate commutative events. The explicit formulas are revealing. The quantity d_n is defined as before as the number of left-light outcomes less the number of right-light outcomes, cumulated through trial $n - 1$. The beta model is then represented by Eq. 23 which is reproduced here:

$$p_n = \frac{1}{1 + \beta^{d_n} v_1^{-1}}.$$

In this model all trials with equal d_n-values also have equal p_n-values. The response probability can be returned to its initial value simply by introducing a sequence of trials that brings d_n back to its initial value of zero. The response of the model to successive reversals is illustrated in Fig. 5 with the event sequence $E_1E_1E_1E_1E_2E_2E_2E_2E_1E_1E_1$. Despite the fact that on the ninth trial, on which d_9 is zero, the most recent outcomes have been right-light onsets, the probability of predicting the left light is no lower than it was initially.

The behavior of the commutative model is in contrast to that of the linear model, whose explicit formula (Eq. 14) is reproduced here:

$$p_n = \alpha^{n-1}p_1 + (1 - \alpha)\sum_{j=1}^{n-1} \alpha^{n-1-j}y_j.$$

The formula shows that when $\alpha < 1$ more recent events are weighted more heavily and that equal d_n-values do not in general imply equal probabilities. The response of this model to successive reversals is also illustrated in Fig. 5. Parameters were chosen so that the two models would agree on the first and fifth trials. This model is more responsive to the reversal than the beta model; not only does the curve of probability versus trials change more rapidly, but its direction of curvature is also altered by the reversal.

At first glance the implications of commutativity for responsiveness of a model seem to suggest crucial experiments or discriminating statistics that would allow an easy selection to be made among models. The question is more complicated, however. The contrast shown in Fig. 5 is clear-cut only if we are willing to make the dubious assumption that events in the prediction experiment are experimenter-controlled. Matters become complicated if we allow reward and nonreward to have different effects. The relative effectiveness of reward and nonreward trials is then another factor that determines the responsiveness of a model.

To show this, we shift attention to models with experimenter-subject control of events. To make the conditions extreme, we compare equal-parameter models (experimenter-control) with models in which the

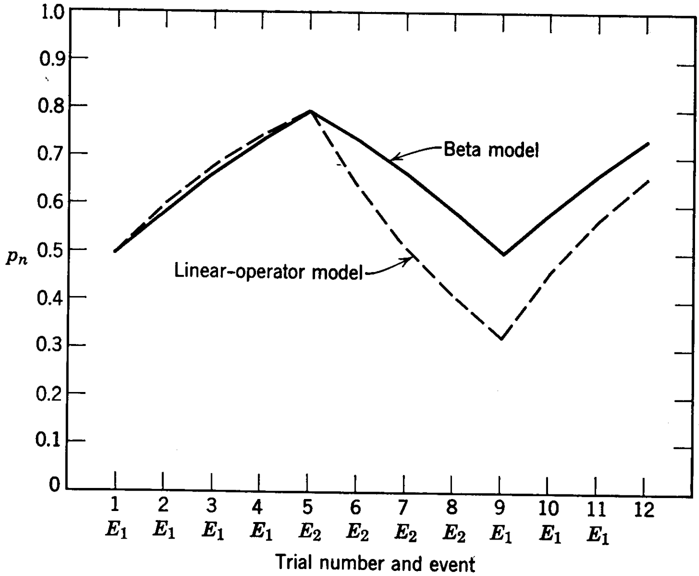

Fig. 5. Comparison of models with commutative and noncommutative experimenter-controlled events. The broken curve represents p_n for the linear-operator model (Eq. 12) with $\alpha = 0.8$ and $p_1 = 0.5$. The continuous curve represents p_n for the beta model with $\beta = 0.713$ and $p_1 = 0.5$. Parameter values were selected so that curves would coincide at trials 1 and 5.

identity operator is associated with either reward or nonreward. The results are shown in Figs. 6 and 7. Parameter values are chosen so that the models agree approximately on the value of $V_{1,5}$. In Fig. 6 the "equal alpha" linear model (Eq. 12) with $\alpha = 0.76$ is compared with the two models defined in (55):

Response	Outcome	Model with Identity Operator for Nonreward ($\alpha = 0.60$)	Model with Identity Operator for Reward ($\alpha = 0.40$)
A_1	O_1	$\mathbf{p}_{n+1} = \mathbf{p}_n + 1 - \alpha$	$\mathbf{p}_{n+1} = \mathbf{p}_n$
A_2	O_1	$\mathbf{p}_{n+1} = \mathbf{p}_n$	$\mathbf{p}_{n+1} = \alpha\mathbf{p}_n + 1 - \alpha$
A_1	O_2	$\mathbf{p}_{n+1} = \mathbf{p}_n$	$\mathbf{p}_{n+1} = \alpha\mathbf{p}_n$
A_2	O_2	$\mathbf{p}_{n+1} = \alpha\mathbf{p}_n$	$\mathbf{p}_{n+1} = \mathbf{p}_n$

(55)

In Fig. 7 the "equal beta" model (Eq. 21) with $\beta = 0.68$ is compared with the two models defined in (56):

Response	Outcome	Model with Identity Operator for Nonreward ($\beta = 0.50$)	Model with Identity Operator for Reward ($\beta = 0.30$)
A_1	O_1	$\mathbf{p}_{n+1} = \dfrac{\beta \mathbf{p}_n}{(1 - \mathbf{p}_n) + \beta \mathbf{p}_n}$	$\mathbf{p}_{n+1} = \mathbf{p}_n$
A_2	O_1	$\mathbf{p}_{n+1} = \mathbf{p}_n$	$\mathbf{p}_{n+1} = \dfrac{\beta \mathbf{p}_n}{(1 - \mathbf{p}_n) + \beta \mathbf{p}_n}$
A_1	O_2	$\mathbf{p}_{n+1} = \mathbf{p}_n$	$\mathbf{p}_{n+1} = \dfrac{\mathbf{p}_n}{\beta(1 - \mathbf{p}_n) + \mathbf{p}_n}$
A_2	O_2	$\mathbf{p}_{n+1} = \dfrac{\mathbf{p}_n}{\beta(1 - \mathbf{p}_n) + \mathbf{p}_n}$	$\mathbf{p}_{n+1} = \mathbf{p}_n$

(56)

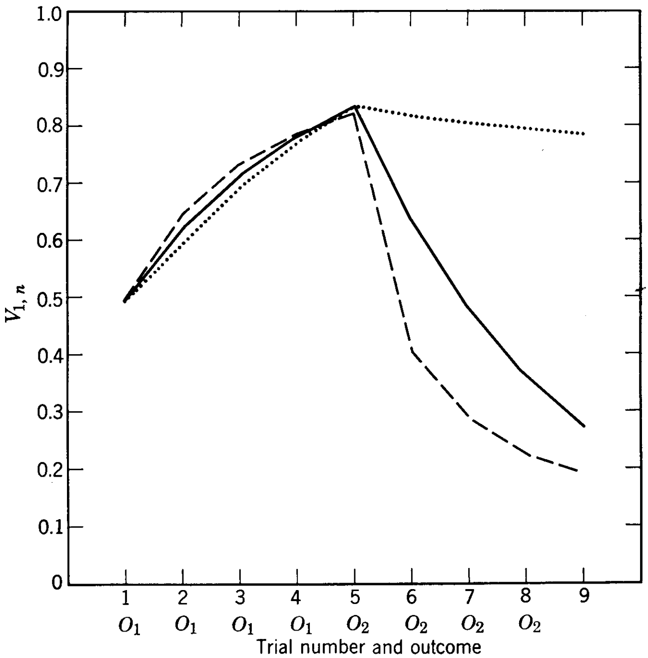

Fig. 6. Responsiveness of the linear-operator model (Eq. 55) depends on the relative effectiveness of reward and nonreward. The solid curve represents $V_{1,n}$ for the linear-operator model with equal reward and nonreward parameters ($\alpha_1 = \alpha_2 = 0.757$, $p_1 = 0.5$). The broken curve represents $V_{1,n}$ for the linear-operator model with an identity operator for reward ($\alpha_1 = 1.0$, $\alpha_2 = 0.4, p_1 = 0.5$). The dotted curve represents $V_{1,n}$ for the linear-operator model with an identity operator for nonreward ($\alpha_1 = 0.6$, $\alpha_2 = 1.0$, $p_1 = 0.5$). Parameter values were selected so that the models would agree approximately on the values of $V_{1,1}$ and $V_{1,5}$.

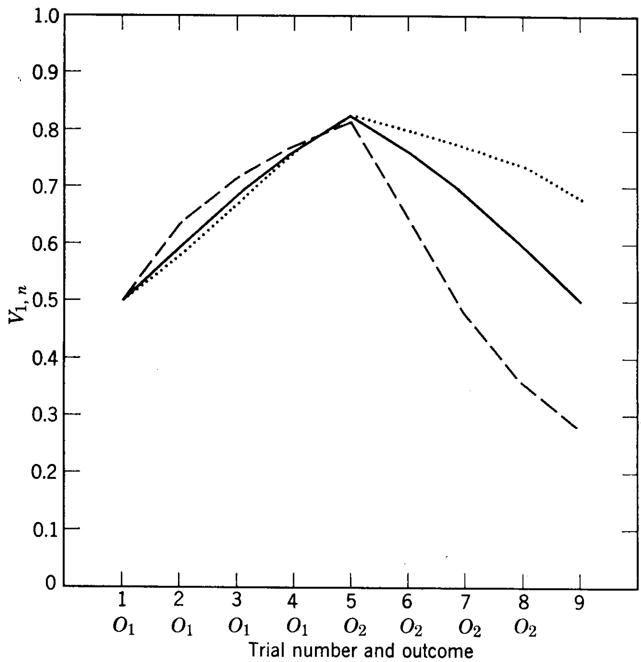

Fig. 7. Responsiveness of the beta model (Eq. 56) depends on the relative effectiveness of reward and nonreward. The solid curve represents $V_{1,n}$ for the beta model with equal reward and nonreward parameters ($\beta_1 = \beta_2 = 0.68, p_1 = 0.5$). The broken curve represents $V_{1,n}$ for the beta model with an identity operator for reward ($\beta_1 = 1.0, \beta_2 = 0.3, p_1 = 0.5$). The dotted curve represents $V_{1,n}$ for the beta model with an identity operator for nonreward ($\beta_1 = 0.5$, $\beta_2 = 1.0$, $p_1 = 0.5$). Parameter values were selected so that the models would agree approximately on the values of $V_{1,1}$ and $V_{1,5}$.

Roughly the same pattern appears for both models. When nonreward is less effective than reward, the response to a change in the outcome sequence that leads to a higher probability of nonreward is sluggish. When reward is less effective than nonreward, the response is rapid. From these examples it appears that the influence on responsiveness of changing the relative effects of reward and nonreward is less marked in the commutative-operator beta model than in the linear model.

Responsiveness alone, then, is not useful in helping us to choose between the models. We must use it in conjunction with knowledge about the relative effects of reward and nonreward. This situation is typical in working with models: observation of a single aspect of the data is often insufficient to lead to a decision. If, in this example, we observe that subjects' behavior is highly responsive, this might imply that a model with commutative operators is inappropriate, but, alternatively, it might mean that the effect of nonreward is relatively great. Also, if we examined such data under the hypothesis that events in the prediction experiment are experimenter-controlled, then the increased rate of change of $V_{1,n}$ after the reversal would probably lead us to conclude, perhaps in error, that a change in the value of a learning rate parameter had occurred. This example indicates how delicate are the conclusions one draws regarding event invariance (Sec. 2.7) and illustrates how the apparent failure of event invariance may signify that the wrong model has been applied.

In a number of studies the prediction experiment has been analyzed by the experimenter-controlled event model of Eq. 12 (Estes & Straughan, 1954; Bush & Mosteller, 1955). This model also arises from Estes' stimulus sampling theory. One of the findings that has troubled model builders is that estimates of the learning-rate parameter α tend to vary systematically from experiment to experiment as a function of the outcome probabilities. It is not known why this occurs, but the phenomenon has occasionally been interpreted as indicating that event effects are not invariant as desired. Another interpretation, which has not been investigated, is that because reward and nonreward have different (but possibly invariant) effects the estimate of a single learning-rate parameter is, in effect, a weighted average of reward and nonreward parameters. Variation of outcome probabilities alters the relative number of reward-trials and thus influences the weights given to reward and nonreward effects in the over-all estimate. The estimation method typically used depends on the responsiveness of the model, which, as we have seen, depends on the extent to which rewarded trials predominate.

4.4 Commutativity and the Asymptote in Prediction Experiments

One result of two-choice prediction experiments that has interested many investigators is that when $\Pr\{\mathbf{y}_n = 1\} = \pi$ and $\Pr\{\mathbf{y}_n = 0\} = 1 - \pi$ then for some experimental conditions the asymptotic mean probability $V_{1,\infty}$ with which human subjects predict $\mathbf{y}_n = 1$ appears to "match"

$\Pr\{\mathbf{y}_n = 1\}$, that is, $V_{1,\infty} \simeq \pi$.[18] (The artificial data in Figs. 1 and 2 illustrate this phenomenon.) The phenomenon raises the question of which model types or model families are capable of mimicking it. No answer even approaching completeness seems to have been proposed, but a little is known. Certain linear-operator models are included in the class, and we shall see that models with commutative events can be excluded, at least when the events are assumed to be experimenter-controlled and symmetric. [Feldman and Newell (1961) have defined a family of models more general than the Bush-Mosteller model that displays the matching phenomenon.]

Figures 1 and 2 suggest that a linear-operator model with experimenter-subject control can approximate the matching effect. As already mentioned, an exact expression for the asymptotic mean of this model is not known. It is easy to demonstrate that the linear model with experimenter-controlled events can produce the effect exactly; indeed, a number of investigators have derived confidence in the adequacy of this particular model from the "probability matching" phenomenon (e.g., Estes, 1959; Bush & Mosteller, 1955, Chapter 13). The value of $V_{1,\infty}$ for the experimenter-controlled model when the $\{\mathbf{y}_n\}$ are independent binomial random variables can easily be determined from its explicit formula (Eq. 14), which is reproduced here:

$$\mathbf{p}_n = \alpha^{n-1}p_1 + (1-\alpha)\sum_{j=1}^{n-1}\alpha^{n-1-j}\mathbf{y}_j.$$

We take the expectation of both sides of this equation with respect to the independent binomial distributions of the $\{\mathbf{y}_j\}$, making use of the fact that $E(\mathbf{y}_j) = \pi$. Performing the summation, we obtain

$$V_{1,n} = \alpha^{n-1}p_1 + (1-\alpha^{n-1})\pi. \tag{57}$$

Note that this is the same result given by the expected-operator approximation in Eq. 44. The final result, $V_{1,\infty} = \pi$, is obtained by letting $n \to \infty$.

As an example of a model with experimenter-controlled events that cannot produce the effect, we consider Luce's model (Eq. 23), with

[18] The validity of this finding and the particular conditions that lead to it have been the subjects of considerable controversy. Partial bibliographies may be found in Edwards (1956, 1961), Estes (1962), and Feldman & Newell (1961). The reader should also consult Restle (1961, Chapter 6) and, for work with several nonhuman species, the papers of Bush & Wilson (1956) and of Bitterman and his colleagues (e.g., Behrend & Bitterman, 1961). The general conclusions to be drawn are that the phenomenon does not occur under all conditions or for all species, that when it seems to occur the response probability may deviate slightly but systematically from the outcome probability, that matching may characterize a group average although it occurs for only a few of the individuals within the group, and that an asymptote may not have been reached in many experiments.

$0 < \beta < 1$. We restrict our attention to experiments in which $\pi \neq \frac{1}{2}$; without loss of generality we can restrict it further to $\pi > \frac{1}{2}$. Because $\mathbf{p}_n$ is governed entirely by the value of $\mathbf{d}_n$, we must concern ourselves with the behavior of

$$\mathbf{d}_n = \sum_{j=1}^{n-1} (2\mathbf{y}_j - 1).$$

Roughly speaking, because the number of left-light outcomes (E_1) increases faster than the number of right-light outcomes (E_2), the difference between their numbers increases, and with an unlimited number of trials this difference $\mathbf{d}_n$ becomes indefinitely large. More precisely, we note that

$$E(\mathbf{d}_n) = \sum_{j=1}^{n-1} (2\pi - 1)$$

and that therefore $E(\mathbf{d}_n) \to \infty$ as $n \to \infty$. From the law of large numbers (Feller, 1957, Chapter X) we conclude that with probability one $\mathbf{d}_n \to \infty$ as $n \to \infty$. Using Eq. 23, it follows that for this model $\mathbf{p}_n \to 1$ when $\pi > 0.50$.

The asymptotic properties of other examples of the beta model for the prediction experiment have been studied by Luce (1959) and Lamperti and Suppes (1960). They find that there are special conditions, determined by the values of π and model parameters, under which $V_{1,n} = E(\mathbf{p}_n)$ does not approach a limiting value of either zero or unity. Therefore it should not be inferred from the foregoing example that the beta-model is incapable of producing probability matching. (In view of the fact that the phenomenon does not occur regularly or in all species, one might consider a model that invariably produces it to be more suspect than one in which its occurrence depends on conditions or parameter values.)

As an illustration of the present state of knowledge, we consider the beta model with experimenter-subject control for the prediction experiment. In this model, absorption at a limiting probability of zero or unity does not always occur. The outcomes are $O_1 : \mathbf{y} = 1$ and $O_2 : \mathbf{y} = 0$. The responses are A_1 : predict O_1, and A_2 : predict O_2. We assume that the pairs of events, $\{A_1O_1, A_2O_2\}$ and $\{A_1O_2, A_2O_1\}$ are complementary. The transformations of the response-strength ratio $v = v(1)/v(2)$ are therefore as follows:

Event	Transformation
A_1O_1	$v \to \beta v$
A_2O_1	$v \to \beta' v$
A_1O_2	$v \to \frac{1}{\beta'} v$
A_2O_2	$v \to \frac{1}{\beta} v$

The parameters β and β' correspond to reward and nonreward, respectively; both are greater than one.

For this model the results of Lamperti and Suppes (1960, Theorem 3) imply that the asymptotic value of $\mathbf{p}_n = \Pr\{A_1 \text{ on trial } n\}$ is either zero or unity *except* when the following inequality is satisfied:

$$\frac{\log \beta}{\log \beta'} < \frac{\pi}{1 - \pi} < \frac{\log \beta'}{\log \beta}.$$

Luce has shown (1959, Chapter 4, Theorem 17) that when the inequality is satisfied the asymptotic value of $V_{1,n}$ is given by

$$V_{1,\infty} = \pi + \frac{2\pi - 1}{(\log \beta')/(\log \beta) - 1}.$$

(It is interesting to note that the value of $V_{1,\infty}$ for the corresponding linear-operator model, given by Eq. 47, is known only approximately.)

From these results several conclusions may be drawn. First, if $\beta > \beta'$, then this model *always* produces asymptotic absorption at zero or one; only if nonreward is more potent than reward ($\beta' > \beta$) is a limiting average probability other than zero or one possible. Second, for a fixed pair of parameter values, $\beta' > \beta \geqslant 1$, absorption at zero or one can be avoided, but only for a limited range of π-values. Third, when $V_{1,\infty}$ is between zero and one, it is equal to π only if $\pi = \frac{1}{2}$; otherwise the asymptote is further from $\frac{1}{2}$ than π is and in the same direction, with the magnitude of the "overshoot" or "undershoot" increasing linearly with $|\pi - \frac{1}{2}|$.

In the experimenter-controlled-events example, which was first discussed, it is the commutativity of the beta model that is responsible for its asymptotic behavior. An informal argument shows that the same asymptotic behavior characterizes any model with two events that are complementary, commutative, and experimenter-controlled and in which repeated occurrence of a particular event leads to a limiting probability of zero or unity. In any such model the response probability returns to its initial value on any trial on which $\mathbf{d}_n = 0$. Moreover, the probability on any trial is invariant under changes in the order of the events that precede that trial. Therefore the probability $\mathbf{p}_n$ after a mixed sequence composed of m E_2's and $(n - m - 1)$ E_1's is the same as the value of $\mathbf{p}_n$ after a block of $(n - m - 1)$ E_1's preceded by a block of m E_2's. Now let $\mathbf{m}(n)$ be an integral random variable whose value is the number of E_2-events in n trials. Note that $E[\mathbf{m}(n)] = n(1 - \pi)$. For $\pi > \frac{1}{2}$ we have already seen that as n increases we have $n - \mathbf{m}(n) - 1 > \mathbf{m}(n)$ with probability one. Consider what happens when the order of the events is rearranged so that a block of all the E_2's precedes a block containing all the E_1's. On the

mth trial of the second block the probability returns to its initial value. After this trial there are $[n - 2\mathbf{m}(n) - 1]$ E_1-trials; but

$$E[n - 2\mathbf{m}(n) - 1] = n(2\pi - 1) - 1,$$

which becomes indefinitely large. The behavior of the model is the same as if, starting at the initial probability, an indefinitely long sequence of E_1-trials occurred. The limiting value of $\mathbf{p}_n$ is therefore unity. A similar result applies when the event-effects have different magnitudes. Without further calculations we know that the urn scheme of Eq. 25 cannot mimic the matching effect.

I have discussed the asymptotic behavior of these models partly because it is of interest in itself but mainly to emphasize the strong implications of the commutativity property. As a final example of the absence of "forgetting," let us consider an experiment in which first a block of E_1's occurs and then a series in which E_1's and E_2's occur independently with probability $\pi = \frac{1}{2}$. In both the beta and linear models for complementary experimenter-controlled events the initial block of E_1-events will increase the probability to some value, say p'. In the linear-operator model the mixed event series will reduce the probability from p' toward $p = \frac{1}{2}$. In the beta model, on the other hand, the mixed event series will cause the probability to fluctuate indefinitely about p' with, on the average, no decrement. The last statement is true for any model whose events are complementary, experimenter-controlled, and commutative.

4.5 Analysis of the Explicit Formula[19]

In this section I consider some of the important features of explicit formulas for models with two subject-controlled events. These models are meant to apply to experiments such as the escape-avoidance shuttlebox and 100:0 prediction, bandit, and T-maze experiments. The event (response) on trial n is represented by the value of $\mathbf{x}_n$, where $\mathbf{x}_n = 0$ if the rewarded response is made and $\mathbf{x}_n = 1$ if the nonrewarded response (an "error") is made. The probability $\mathbf{p}_n = \Pr\{\mathbf{x}_n = 1\}$ decreases over the sequence of trials toward a limiting value of $p = 0$.

EXAMPLES USED. The following models are used as examples:

Model I $\quad p_n = F(n) = \alpha^{n-1}p_1, \qquad (0 < \alpha < 1);$ (58)

Model II $\quad \mathbf{p}_n = F(n, \mathbf{x}_{n-1})$

$$= \alpha^{n-1}p_1(1 - \beta) + \beta\mathbf{x}_{n-1}, \qquad (0 < \beta < 1, n \geqslant 2); \quad (59)$$

[19] Much of this discussion is drawn from Sternberg (1959b).

Model III $\quad \mathbf{p}_n = F(n, \mathbf{x}_{n-1}, \mathbf{x}_{n-2}, \ldots, \mathbf{x}_1)$

$$= \alpha^{n-1}p_1 + \beta \sum_{j=1}^{n-1} \alpha^{n-1-j}\mathbf{x}_j, \qquad (0 < \alpha, \beta < 1); \tag{60}$$

Model IV $\quad \mathbf{p}_n = F(n, \mathbf{s}_n) = \exp\left[-(a + bn + c\mathbf{s}_n)\right], \qquad (0 < a, b). \quad (61)$

We have seen Models I, II, and IV before. Model I is the single-operator model of Eq. 28 (Bush & Sternberg, 1959) and is an example of the family of single-event models. Model II is the one-trial perseveration model of Eq. 29. (An analogous nonlinear model is given by the generalized logistic in Eq. 32.) Model IV is the Bush-Mosteller model of Eqs. 10 and 11, rewritten by using the fact that $\mathbf{t}_n = n - 1 - \mathbf{s}_n$. The quantity $\mathbf{s}_n = \sum_{j=1}^{n-1} \mathbf{x}_j$ is the number of errors before trial n, and $c = -\log(\alpha_2/\alpha_1)$. If the effect of reward is greater than the effect of nonreward ($\alpha_1 < \alpha_2$), then $c < 0$ and more errors (larger $\mathbf{s}_n$) imply a higher probability of error (larger $\mathbf{p}_n$); if $\alpha_1 > \alpha_2$, then $c > 0$ and the converse holds. This model has been studied by Tatsuoka and Mosteller (1959). (Analogous beta and urn models are given by Eqs. 19 and 26.)

Almost all the models that have been applied to data involve either identity operators or operators with limit points of zero or unity. One exception is Model III, whose operators are given by

$$\mathbf{p}_{n+1} = \begin{cases} \alpha\mathbf{p}_n & \text{if} \quad \mathbf{x}_n = 0 \\ \alpha\mathbf{p}_n + \beta & \text{if} \quad \mathbf{x}_n = 1. \end{cases}$$

Referred to as the "many-trial perseveration model," this model has been applied to two-armed bandit data by Sternberg (1959b). The explicit formula is similar in form to Eq. 14 for the linear model for two experimenter-controlled events; more recent events are weighted more heavily.

DIRECT RESPONSE EFFECTS. Consider first the *direct effect* of a response, $\mathbf{x}_j$, on the probability $\mathbf{p}_n$. By "direct effect" is meant the influence of $\mathbf{x}_j$ on the magnitude of $\mathbf{p}_n$ when intervening responses $\mathbf{x}_{j+1}, \ldots, \mathbf{x}_{n-1}$ are held fixed. Response $\mathbf{x}_j$ has a direct effect on $\mathbf{p}_n$ if it appears as an argument of the explicit formula F. The effect is *positive* if $\mathbf{x}_j = 1$ results in a larger value of $\mathbf{p}_n$ than does $\mathbf{x}_j = 0$; otherwise the effect of $\mathbf{x}_j$ is *negative*. Models II and III show positive response effects, achieved by associating an additive constant with $\mathbf{x}_j$ in the explicit formula. In Model IV the direct effects can be positive or negative, depending on the sign of c. The effect is achieved by adding a constant to $\log \mathbf{p}_n$ when $\mathbf{x}_j = 1$; this is equivalent to applying a multiplicative constant to $\mathbf{p}_n$. When response effects occur in one of these models, they are all of the same sign; the direction of the effect of an event does not depend on when the event occurred. Let us confine our discussion to this type of model.

If none of the $\mathbf{x}_j$ appears in F, then there are no response effects and the model is response-independent. This is a characteristic of all single-event models. Model I is an example.

If any of the $\mathbf{x}_j$ appear in F, there are direct response effects. If only $\mathbf{x}_{n-1}$ appears, then $\mathbf{p}_n$ is directly affected only by the immediately preceding response, as in Model II. Because the $\mathbf{p}_m$ for $m > n$ are not affected by $\mathbf{x}_{n-1}$ we say that the direct effect is *erased* as the process advances. If several, say k, of the $\mathbf{x}_j$ appear in F, then the direct effect of a response continues for $k - 1$ trials and is then erased. [Audley and Jonckheere (1956) have considered a special case of their urn scheme that has this property.] If all the $\mathbf{x}_j$ ($j = n - 1, n - 2, \ldots, 1$) appear in F, the effect of a response is never erased and continues indefinitely. This last condition must hold for any response-dependent model that is also path-independent. Models III and IV are examples.

When more than one $\mathbf{x}_j$ appears in F, we can ask two further questions concerned with the way in which the arguments $\mathbf{x}_j$ appear in F. The first is whether there is *damping* of the continuing effects. We define the *magnitude* of the effect of $\mathbf{x}_j$ on $\mathbf{p}_n$ to be the change in the value of $\mathbf{p}_n$ when the value of $\mathbf{x}_j$ in $F(n, 0, 0, 0, \ldots)$ is increased from 0 to 1. When the magnitude of the effect of $\mathbf{x}_j$ is smaller for earlier $\mathbf{x}_j$, then we say that direct response effects are damped. If the magnitudes are equal, then the effects are *undamped.* (Direct effects might also be augmented with trials; this could occur in a model in which the full effect of a response took more than one trial to appear. No such models have been studied, however. In what follows we assume that effects are either damped or undamped.)

The second question we can ask, when two or more of the $\mathbf{x}_j$ appear in F, is whether their effects *accumulate.* If so, then the effect on $\mathbf{p}_n$ when two of the $\mathbf{x}_j$ are errors is greater than the effect when either one of them alone is an error. In all of the models mentioned in this chapter for which effects continue they also accumulate.

If a model exhibits damped response effects, the cumulative number of errors alone is not sufficient to tell us the value of $\mathbf{p}_n$; we must also know on which trials the errors occurred. Therefore, the events in a model with damped effects cannot commute; and, conversely, if a model with commutative events shows response effects, then these effects cannot be damped. Models III and IV provide examples of the foregoing statement and its converse. In Model III the response effects are damped and events do not commute; in Model IV, a commutative event model, response effects are undamped. These two models are analogous to the linear and beta models for experimenter-controlled events (Sec. 4.3). In the linear model outcome effects are damped (there is "forgetting") and events do not commute; in the beta model there is no damping and we have commutativity.

By means of these ideas models can be roughly ordered in terms of the extent to which direct response effects occur. First is the response-independent, single-event model in which there is no effect at all (Model I). Then we have a model in which an effect occurs but is erased (Model II). Next is a model in which the effect continues but is damped (Model III); and finally we have a model with an undamped, continuing effect (Model IV).

INDIRECT RESPONSE EFFECTS. One of the most important properties of models with subject-control of events is the fact that the responses in a sequence are not independent. This is the property that causes subjects' response probabilities to differ even when they have common parameter values and are run under identical reinforcement schedules. One result is that, in contrast to models in which only the experimenter controls events, we must deal with distributions rather than single values of the response probabilities. A second implication is that events (responses) have indirect as well as direct effects on future responses, effects that are transmitted by the intervening trials. In contrast, experimenter-controlled events have only direct effects.

Until now we have been considering only the direct effect of a response $\mathbf{x}_j$ on $\mathbf{p}_n$. If $j < n - 1$, so that trials intervene between responses $\mathbf{x}_j$ and $\mathbf{x}_n$, there also may be *indirect effects* mediated by the intervening responses. For example, whether or not $\mathbf{x}_{n-2}$ has a direct effect on $\mathbf{p}_n$, it may have an indirect effect, mediated through its direct effect on $\mathbf{p}_{n-1}$ and the relation of $\mathbf{p}_{n-1}$ to the value of $\mathbf{x}_{n-1}$. Therefore, even if the direct effect of $\mathbf{x}_j$ on $\mathbf{p}_n$ is erased, the response may influence the probability. Model II provides an example. In this model the p-value on a trial is determined uniquely by the trial number and the preceding response, so that, conditional on the value of $\mathbf{x}_{n-1}$, $\mathbf{x}_n$ is independent of all the $\mathbf{x}_m$, $m < n - 1$. On the other hand, if the value of $\mathbf{x}_{n-1}$ is not specified, then $\mathbf{x}_n$ depends on any one of the $\mathbf{x}_m$, $m < n - 1$, that may be selected. Put another way, the conditional probability $\Pr\{\mathbf{x}_n = 1 \mid x_{n-1}\}$ is uniquely determined, whatever the $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n-2}$ sequence is. But, given any trial at all before the nth, the unconditional ("absolute") probability $\Pr\{\mathbf{x}_n = 1\}$ depends on the response that is made on that trial. A more familiar example is the one-step Markov chain, in which the higher-order conditional probabilities are not the same as the corresponding "absolute probabilities" (Feller, 1957), despite the fact that the direct effects extend only over a single trial. Because $\mathbf{x}_{n-1}$ has no effect at all on $\mathbf{p}_n$ in a response-independent model, it cannot have any indirect effect on $\mathbf{p}_m$ $(m > n)$.

The *total effect* of a response $\mathbf{x}_j$ on the probability $\mathbf{p}_n$ can be represented by the difference between two conditional probabilities:[20]

$$\Pr\{\mathbf{x}_n = 1 \mid \mathbf{x}_j = 1\} - \Pr\{\mathbf{x}_n = 1 \mid \mathbf{x}_j = 0\}.$$

When direct effects are positive, the total effect of $\mathbf{x}_j$ on $\mathbf{p}_n$ cannot be less than its direct effect alone. The extent to which the total effect is greater depends in part on whether there is accumulation of the direct effects of $\mathbf{x}_j$ and the intervening responses and in part on whether and how effects are damped. When direct effects are negative, the situation is more complicated, and the relation between total and direct effects depends on whether the number of intervening trials is even or odd as well as on accumulation and damping.

SUBJECT-CONTROLLED EVENTS AS A PROCESS WITH FEEDBACK. In most of the foregoing discussion we have been considering the effects of responses on probabilities. The altered probabilities influence their associated responses, and these responses in turn have effects on future probabilities. Thus the effects we have been considering "feed back" the "output" of the p_n-sequence so as to influence that sequence. Insofar as two response sequences have different p_n-values on some trial, the nature of the response effects determines whether this probability difference will be enhanced, maintained, reduced, or reversed in sign on the next trial.

Each of the p_n-sequences produced by a model is an *individual learning curve*, and the "area" under this curve represents the expected number of errors associated with that sequence. In a large population of response sequences (subjects) the model specifies a proportion of the population that will be characterized by each of the possible individual learning curves. The mean learning curve is the average of these individual curves. If there are no response effects, there is, of course, only one individual curve.

When response effects exist and are positive, we may speak of a *positive feedback of probability differences* and determine measures of its magnitude. With more positive feedback of p_n-differences, individual learning curves have a greater tendency to deviate from their mean curve as n increases. The *negative feedback* of p_n-differences, which may occur if response effects exist and are negative, may cause the opposite result: p_n-differences that arise among sequences may be neutralized or reversed in sign. Thus an individual curve that deviated from the mean curve would tend to return to it or to cross it and therefore to compensate for the deviation. A rough idea of the magnitude of the feedback can be obtained by comparing an assumed p_n-difference of Δp_n on trial n with the associated expected difference of $\overline{\Delta p}_{n+1}$ on the next trial.

[20] For a binary-event sequence that is generated by a stationary stochastic process this expression gives the autocorrelation function with lag $n - j$.

Also relevant to the feedback question is the range of the p_n-values that a model can produce on a given trial. For example, in a model with positive response effects the maximum possible value of $\mathbf{p}_n$ is attained when all the responses have been errors and the minimum is attained when they have all been successes. For a model with negative effects the reverse holds. Therefore the *p_n-range* is given by the absolute value of

$$F(n, 1, 1, 1, \ldots) - F(n, 0, 0, 0, \ldots).$$

Whatever the sign or magnitude of any feedback of probability differences that may occur, it cannot lead to p_n-differences larger than the p_n-range. The p_n-values corresponding to the extremes of the p_n-range produce the pair of individual learning curves that differ maximally in area. In general, the p_n-range imposes a limit on all the response effects discussed in this section.

For Model I the p_n-range is zero. For Model II it is a constant. For Models III and IV the range increases with n.

DISCRIMINATING STATISTICS: SEQUENTIAL PROPERTIES. The analysis of response effects presented above is useful, first in suggesting statistics of the data that may discriminate among Models I to IV and second in helping us to interpret the results of applications of these models. To illustrate these uses, let us consider results of Sternberg's (1959b) application of these models to data collected by Goodnow in a two-armed bandit experiment with 100:0 reward.

The analysis tells us that fundamental differences among the four models lie in the extent to which response effects occur and are erased or damped. This suggests that the models differ in their sequential properties and that it is among the sequential features of the data that we should find discriminating statistics. This suggestion is confirmed by the following results that were obtained in application of the models to the Goodnow data:

1. Parameter values can be chosen for all four models so that they produce mean learning curves in good agreement with the observed curve of trial-by-trial proportions of errors. The observed and fitted curves are shown in Fig. 8. Despite the differences among the models, the mean learning curve does not discriminate one from another.

2. Now we begin to examine sequential properties. First we consider the mean number of runs of errors. The parameters in Model I cannot be adjusted so that it will retain its good agreement with the learning curve and at the same time produce few enough runs of errors; this model can be immediately disqualified. In contrast, parameters in Models II, III, and IV can be chosen so that these models will agree with both the learning curve and the number of error runs. (This difference is not altogether

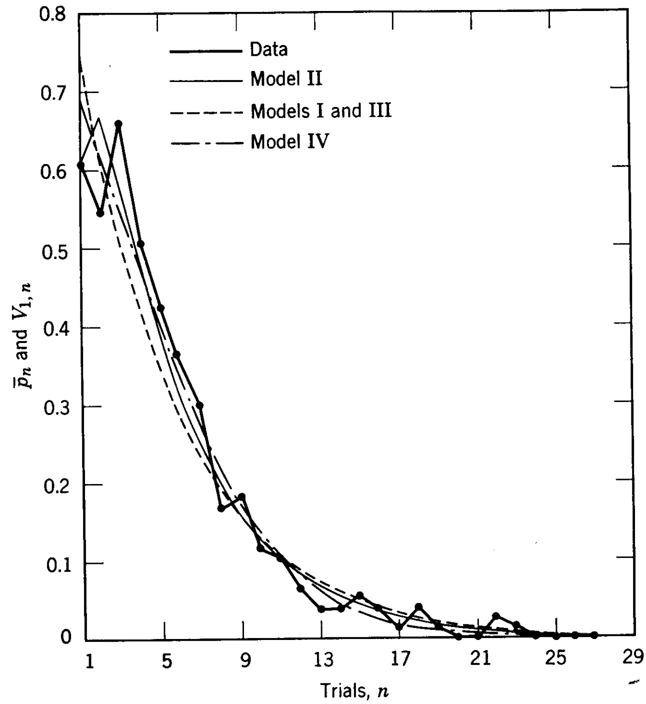

Fig. 8. Observed trial-by-trial proportions $\bar{p}_n$ of errors in the Goodnow experiment and theoretical means $V_{1,n}$ for the four models of Eqs. 58 to 61.

surprising because Model I has one less free parameter than the others.) A finer analysis of error runs, considering the average number of runs of each length, j, $1 \leqslant j \leqslant 7$, does not help; Models II, III, and IV produce equally good agreement with the distribution of error-run lengths. This is shown by Fig. 9.

3. Finally, we examine the serial autocovariance of errors at lags 1 to 10. This statistic is defined to be the mean value of

$$\mathbf{c}_k = \sum_n \mathbf{x}_n \mathbf{x}_{n+k},$$

where the lag is given by the value of k. (Models II, III, and IV all agree with the observed value of $\bar{c}_1$, but this tells us nothing new, since their agreement follows automatically from agreement with the learning curve and the number of runs.) What is of interest is the behavior of $\bar{c}_k$ as k increases: its observed value falls rapidly, and only Model II is able to produce so rapid a decrease. The slowest decrease is produced by Model IV, with Model III a close second. The results are illustrated in Fig. 10.

Our analysis of the four models in terms of response effects aids the interpretation of these results. Figure 8 suggests, as did Figs. 1 and 2, that interesting and important differences among models and between models and data may be totally obscured if we restrict our attention to the learning curve. The results regarding runs of errors reflect the fundamental difference between Model I, which is response-independent, and the others. Responses in this experiment were clearly not independent of each other; when errors occurred they tended to occur in clusters, suggesting a positive response effect. This suggestion is confirmed by the values of the estimated parameters for the other models. The behavior of $\bar{c}_k$ is sensitive to the extent of erasing or damping the positive response effect. Its value drops most rapidly with k in Model II, in which the effect is erased after a single

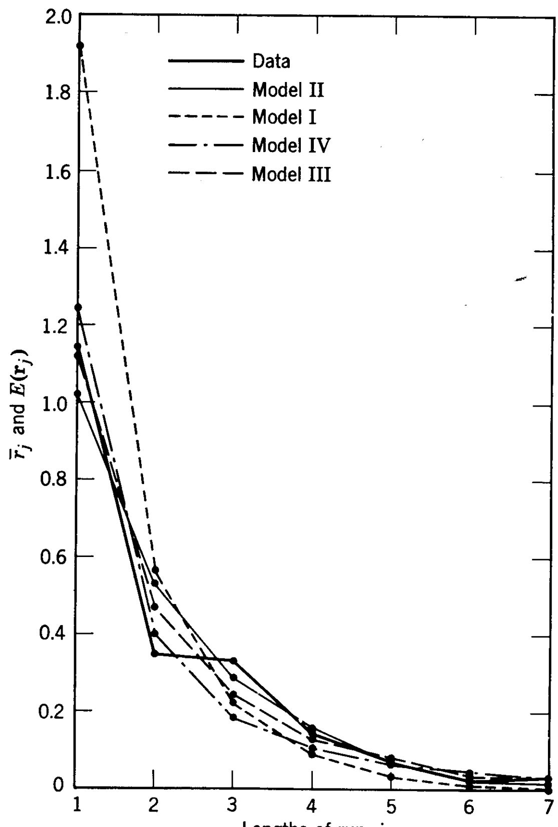

Fig. 9. Observed mean number of error runs $\bar{r}_j$ of length j in the Goodnow experiment and theoretical values $E(\mathbf{r}_j)$ for the four models of Eqs. 58 to 61.

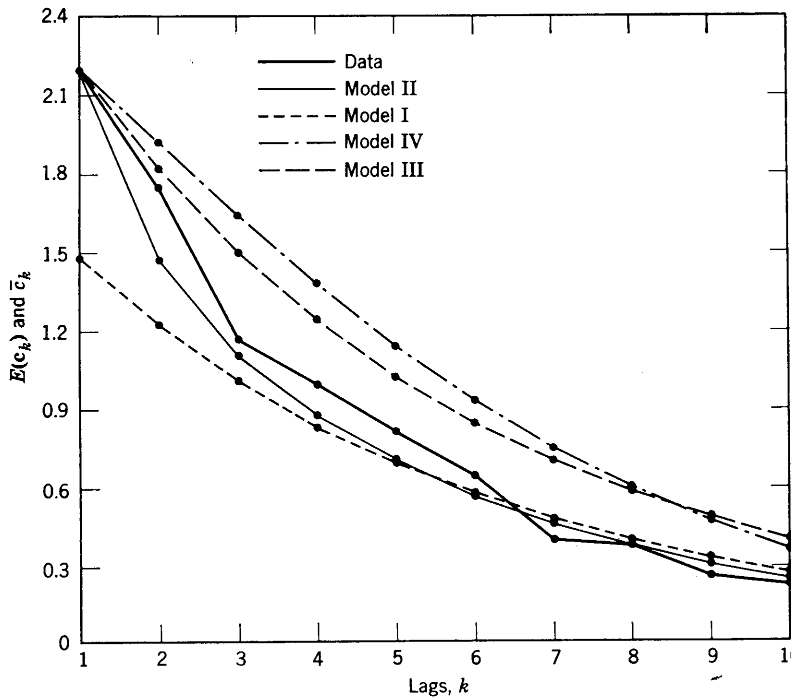

Fig. 10. Observed mean values $\bar{c}_k$ of the serial autocovariance of errors in the Goodnow experiment at lags 1 to 10 and theoretical values $E(\mathbf{c}_k)$ for the four models of Eqs. 58 to 61.

trial. Its value drops most slowly in Model IV, in which the effect continues and is undamped. Model III, with a continuing but damped effect, is intermediate.

The interpretation of these results, based on our analysis of response effects, leads us to conclude that Goodnow's data exhibit short-lived, positive response effects. If a model for these data is to reproduce the learning curve, a degree of positive sequential dependence is required for it to match the number of error runs as well. But this response effect must be radically damped or erased if the model is to describe the autocovariance of errors.

DISCRIMINATING STATISTICS: REWARD AND NONREWARD AND THE VARIANCE OF TOTAL ERRORS. One of the questions that has interested investigators concerns the relative magnitudes of the effects of reward and nonreward in situations in which their effects have the same sign. For example, one plausible interpretation of the events in a T-maze with 100:0 reward is that both rewarded and nonrewarded trials increase the probability of making the rewarded response. Similarly, in the shuttle-box, both the escape and avoidance responses might be thought to increase the probability of avoidance. How would differences in the relative magnitudes of these effects manifest themselves in the data? One answer is suggested if we translate the question into the language of positive and negative response effects. If the effect of reward is the greater, then the error probability after an error is higher than the error probability after a success, and we have a positive response effect. If the effect of nonreward is the greater, a negative response effect exists.

We have already considered the relation between response effects and the feedback of probability differences. When response effects are positive, feedback is positive, individual learning curves tend to diverge from the mean learning curve, and subjects with more early errors tend to have more errors late in learning. When response effects are negative, feedback is negative, differences between individual learning curves tend to be neutralized or reversed, and subjects with more early errors tend to have fewer later errors. Areas under individual learning curves tend to be more variable with positive effects than they are with negative effects.

Because the area under an individual learning curve represents the total number of errors for that individual, this informal argument suggests that there should be a relation between response effects and the variance of the total number of errors. Roughly speaking, if two experiments produce the same average number of total errors, then the experiment in which reward is the more effective should produce a greater variance of total errors than the experiment in which nonreward is the more effective. Moreover, there should be a positive correlation beween the extent of positive response effects (magnitude, continuation, damping) and the variance of total errors.

Both conclusions are borne out by studies of experiments and models. One experiment that provides clear evidence of a negative response effect is a study of reversal after overlearning in a T-maze (Galanter & Bush, 1959; Bush, Galanter, & Luce, 1959). (A "model-free" demonstration that the effect is negative is discussed in Sec. 6.7.) In Table 3 this experiment is compared with five others in which analyses have suggested that there is a positive response effect. The coefficient of variation behaves appropriately. The second conclusion is supported by the theoretical values of the variance of total errors for Models I to IV when their parameters are selected to produce the same learning curve and, when possible, the same number of error runs as observed in the Goodnow data. The figures for the variance correspond roughly to the extent of positive response effects: Model I: 2.53; Model II: 4.04; Model III: 10.78; Model IV: 10.42. (The corresponding figure for the data is 5.17.)

Table 3 Positive and Negative Response Effects and the Variance of Total Errors in Several Experiments

Experiment	Response Effect	Mean Total Errors, $\bar{U}_1$	S. D. of Total Errors, $S(U_1)$	Coefficient of Variation $C = S(U_1)/\bar{U}_1$
T-Maze reversal after overlearning (rats) (Galanter & Bush, 1959)	Negative	24.68	2.24	0.09
T-Maze reversals (rats) (Galanter & Bush, 1959)				
Period 2	Positive	14.10	3.50	0.24
Period 3	Positive	9.50	3.04	0.32
Period 4	Positive	12.70	3.85	0.30
Solomon-Wynne Shuttlebox (dogs) (Bush & Mosteller, 1959)	Positive	7.80	2.52	0.32
Goodnow two-armed bandit (Sternberg, 1959b)	Positive	4.32	2.27	0.53

5. MATHEMATICAL METHODS FOR THE ANALYSIS OF MODELS

In the preceding sections I have considered informal and approximate methods of analyzing and comparing models. A good many of the conclusions have been qualitative, and, although we should not belittle their usefulness in guiding research and in aiding the interpretation of results, it is in the quantitative properties of learning models that the core of our knowledge lies. No unified methods of analysis exist, however. Various devices have been used, of which only a few samples are illustrated here. Other examples are contained in numerous references, including Bush & Mosteller (1955), Bush & Estes (1959), Karlin (1953), Lamperti & Suppes (1960), Bush (1960), and Kanal (1962a,b).

In principle, only the investigator's imagination limits the number of different statistics of response sequences, short of individual responses, that his model can be coaxed to describe. Examples, some of which we have already come across, are the mean learning curve, the number of trials before the kth success, the number of runs of errors of a particular length, the autocovariance of errors, the number of occurrences of a particular outcome-response pair, and the trial on which the last error occurs. The entire distribution for a statistic or, if desired, just its mean and variance can be described by a model. In practice, analytic methods are limited, often severely so, and a good many of these statistics must be estimated from Monte Carlo sampling experiments.

Usually the expectation of a statistic produced by a model is a function of parameter values, and therefore the problem of estimating these values cannot be bypassed in the analysis of data. On the other hand, as we shall see later, the dependence of statistics on parameter values has been exploited a good deal in estimation procedures. Occasionally a model makes a parameter-free prediction; examples are the asymptotes of the beta and linear models that were discussed in Sec. 4.3. When this occurs, we can often dismiss or get favorable evidence for a model type without bothering to narrow it down to a particular model by estimating its parameters.

In addition to their use in estimation, model statistics are, of course, used in evaluating goodness of fit. This is often done by enumerating the statistics thought to be pertinent and comparing their expected values with those observed. The general point of view has been that the observed response sequences constitute a sample from a population of sequences, and the question is which model type describes the population.

5.1 The Monte Carlo Method

The Monte Carlo method is the generation of an artificial realization of a stochastic process by a sampling procedure that satisfies the same probability laws. A random device, usually a table of random numbers, is substituted for the behaving organism and is used to select responses with the appropriate probabilities. Once we have selected values for the initial probability and other parameters of a model we can generate as large a sample of artificial response sequences as we wish. Any feature of these artificial data or "stat-organisms" can be compared with the corresponding feature of the real response sequences. The method is therefore extremely versatile for testing models whose parameters we are able to estimate.[21]

If the outcome sequence in the real experiment is generated by a probability mechanism, as it often is in the prediction experiment, for example, there is a choice between generating new sequences for the Monte Carlo experiment or generating the artificial data conditional on the actual

[21] The Monte Carlo method is discussed in Bush & Mosteller (1955, Chapter 6), and in Chapters 7 and 8 of Vol. I of this *Handbook*. Examples and references are given by Barucha-Reid (1960, Appendix C).

sequences used in the experiment. In this instance, most workers agree on using conditional Monte Carlo calculations.

If the model involves subject-control of events, then a similar choice is available between letting p_n in the Monte Carlo experiment be determined by the preceding sequence of artificial responses and letting it be determined by the real response sequences. In other words, the value of $\mathbf{p}_n$ is specified by an explicit formula whose arguments consist of parameters and variables that represent the responses on trials 1 through $n - 1$. The responses used can be those in the real data, thus conditioning the Monte Carlo experiment by those data, or artificial responses can be used. Most workers have used artificial data that are not conditioned by the observed responses, but the relative merits of these methods for learning-model research have not yet been assessed.

Sampling experiments are extremely inefficient for handling the estimation problem, and here analytic methods have a considerable practical advantage as well as their usual greater elegance. We turn now to a few examples of analytic methods.

5.2 Indicator Random Variables

We have already used random variables whose values indicate which of two responses or which of two outcomes occurs on a trial. By appropriately choosing the two possible values, generally as zero and unity, many of the statistics in which we are interested can be represented easily by products and sums of these random variables. This type of representation facilitates the calculation of expectations and variances.

Let $\mathbf{x}_n = 1$ if the response on trial n is an error and $\mathbf{x}_n = 0$ if it is a success. Then the number of errors in a sequence of N trials is defined by

$$\mathbf{u}_{1,N} = \sum_{n=1}^{N} \mathbf{x}_n. \tag{62}$$

When we consider an infinite sequence of trials, we drop the subscript N; for example, $\mathbf{u}_1$ denotes the number of errors in an infinite sequence. (This number may or may not be finite.) One approximation often used, when the probability of error approaches zero as n increases, replaces error statistics for finite sequences by their infinite counterparts.

The number $\mathbf{r}_{T,N}$ of runs of errors during N trials is expressed in terms of the $\{\mathbf{x}_j\}$ by noting that every error run, except the one that terminates a sequence, is followed by a success. Therefore,

$$\mathbf{r}_{T,N} = \sum_{n=1}^{N-1} \mathbf{x}_n(1 - \mathbf{x}_{n+1}) + \mathbf{x}_N = \sum_{n=1}^{N} \mathbf{x}_n - \sum_{n=1}^{N-1} \mathbf{x}_n\mathbf{x}_{n+1}. \tag{63}$$

For an infinite sequence the upper limit of the summations becomes infinite and we use the symbol $\mathbf{r}_T$.

If we define $\mathbf{u}_j$, the number of "j-tuples" of errors in an infinite sequence, by

$$\mathbf{u}_j = \sum_{n=1}^{\infty} \mathbf{x}_n \mathbf{x}_{n+1} \cdots \mathbf{x}_{n+j-1}, \tag{64}$$

then $\mathbf{r}_k$, the number of runs of errors of a particular length k, can be expressed in terms of the $\{\mathbf{u}_j\}$ by

$$\mathbf{r}_k = \mathbf{u}_k - 2\mathbf{u}_{k+1} + \mathbf{u}_{k+2}$$

(Bush, 1959).

In experiments in which both outcomes and responses may vary, similar expressions can be used for various response-outcome patterns. On trial n for subject i ($i = 1, 2, \ldots, I$), let $\mathbf{x}_{i,n} = 1$ if response A_1 occurs, $\mathbf{x}_{i,n} = 0$ if response A_2 occurs, and let $\mathbf{y}_{i,n} = 1$ if outcome O_1 follows, $\mathbf{y}_{i,n} = 0$ if outcome O_2 follows. The number of subjects for which O_1 occurred on trial n and A_1 on trial $n + 1$ (a measure of the correlation of responses with prior outcomes) is given by

$$\sum_{i=1}^{I} \mathbf{y}_{i,n} \mathbf{x}_{i,n+1}. \tag{65}$$

5.3 Conditional Expectations

Partly because of the "doubly stochastic" nature of most learning models (Sec. 3) in which both the responses and the p-values have probability distributions, it is often convenient when finding the expectation of a statistic to determine first the expectation conditional on, say, the p-value and then to average the result over the distribution of p-values. A few examples will illustrate this use of conditional expectations. We let $\mathbf{p}_n = \Pr\{\mathbf{x}_n = 1\}$. It will be convenient to let $V_{1,k}(p)$ denote the first moment of the p-value distribution of a process that started k trials ago at probability p; that is,

$$V_{1,k}(p) = \Pr\{\mathbf{x}_{n+k} = 1 \mid \mathbf{p}_{n+1} = p\} = E(\mathbf{x}_{n+k} \mid \mathbf{p}_{n+1} = p). \tag{66}$$

It is also useful to let E_x denote an average over the binomial distribution of responses, E_y an average over a binomial distribution of outcomes, and E_p an average over a p-value distribution. Recall that when the response probability is considered to be a random variable it is written $\mathbf{p}_n$; a particular value is p_n.

Suppose we wish to evaluate $E(\mathbf{r}_T)$ for an experiment with subject-controlled events.

$$E(\mathbf{r}_T) = \sum_{n=1}^{\infty} E(\mathbf{x}_n) - \sum_{n=1}^{\infty} E(\mathbf{x}_n \mathbf{x}_{n+1}).$$

$$E(\mathbf{x}_n) = E_p E_x(\mathbf{x}_n \mid \mathbf{p}_n) = E_p(\mathbf{p}_n) = V_{1,n}.$$

$$E(\mathbf{x}_n \mathbf{x}_{n+1}) = E_p E_x(\mathbf{x}_n \mathbf{x}_{n+1} \mid \mathbf{p}_n) = E_p[\mathbf{p}_n \Pr\{\mathbf{x}_{n+1} = 1 \mid \mathbf{x}_n = 1, \mathbf{p}_n\}].$$

To evaluate the last expression further requires us to specify the model. Consider the commutative linear-operator model that we discussed in connection with the shuttlebox experiment (Eq. 8). Then $\Pr\{\mathbf{x}_{n+1} = 1 \mid \mathbf{x}_n = 1, \mathbf{p}_n\} = \alpha_2 \mathbf{p}_n$ and therefore

$$E(\mathbf{x}_n \mathbf{x}_{n+1}) = E_p(\alpha_2 \mathbf{p}_n^{\,2}) = \alpha_2 V_{2,n},$$

where $V_{2,n}$ is the second (raw) moment of the p-value distribution on trial n. We thus have

$$E(\mathbf{r}_T) = \sum_{n=1}^{\infty} V_{1,n} - \alpha_2 \sum_{n=1}^{\infty} V_{2,n}, \tag{67}$$

and the sums can be evaluated in terms of the model parameters (Bush 1959). For the single-operator model ($\alpha_1 = \alpha_2 = \alpha$), $V_{1,n} = \alpha^{n-1} p_1$ and $V_{2,n} = \alpha^{2(n-1)} p_1^{\,2}$, and Eq. 67 gives

$$E(\mathbf{r}_T) = \frac{p_1}{1 - \alpha} - \frac{\alpha p_1^{\,2}}{1 - \alpha^2},$$

which function is illustrated, for $p_1 = 1$, in Fig. 12, p. 91.

As a second example, suppose we wish to evaluate the expectation of the statistic

$$\mathbf{c}_k = \sum_{n=1}^{\infty} \mathbf{x}_n \mathbf{x}_{n+k},$$

which we encountered in Sec. 4.5. We have

$$\begin{aligned} E(\mathbf{x}_n \mathbf{x}_{n+k}) &= E_p E_x(\mathbf{x}_n \mathbf{x}_{n+k} \mid \mathbf{p}_n) \\ &= E_p[\mathbf{p}_n \Pr\{\mathbf{x}_{n+k} = 1 \mid \mathbf{x}_n = 1, \mathbf{p}_n\}] \\ &= E_p(\mathbf{p}_n V_{1,k}[\Pr\{\mathbf{x}_{n+1} = 1 \mid \mathbf{x}_n = 1, \mathbf{p}_n\}]). \end{aligned}$$

Again we use the commutative linear-operator model as our example. The conditional probability is $\alpha_2 \mathbf{p}_n$ and therefore

$$E(\mathbf{c}_k) = \sum_{n=1}^{\infty} E_p[\mathbf{p}_n V_{1,k}(\alpha_2 \mathbf{p}_n)].$$

Turning to experiments in which outcomes may vary from trial to trial, let us consider the evaluation of the expectation of

$$\mathbf{t} = \frac{1}{NI} \sum_{n=m}^{m+N-1} \sum_{i=1}^{I} \mathbf{y}_{i,n}\mathbf{x}_{i,n+1},$$

which is the proportion of outcome-response pairs in the indicated block of trials for which A_1 on trial $n + 1$ follows O_1 on trial n. Statistics of this type were considered by Anderson (1959) and are examples of aspects of the data that are of interest even after the average response probability has stabilized. We assume a linear operator model with experimenter control and let $\Pr\{\mathbf{y}_{i,n} = 1\} = \pi$.

First let us consider $E(\mathbf{t})$ conditional on the particular $\{y_{i,n}\}$ sequences used. Let E_n denote an average taken over trials and E_i an average over subjects.

$$\begin{aligned} E(\mathbf{t} \mid \{y_{i,n}\}) &= E_n E_i E_x(y_{i,n}\mathbf{x}_{i,n+1} \mid p_{i,n}) \\ &= E_n E_i(\alpha_1 y_{i,n} p_{i,n} + a_1 y_{i,n}) \\ &= \alpha_1 E_n E_i(y_{i,n} p_{i,n}) + a_1 \bar{y}, \end{aligned} \tag{68}$$

where $\bar{y}$ is the average value of $y_{i,n}$ for the sequences used. The corresponding equation in terms of statistics of the data is

$$\frac{\sum_n \sum_i y_{i,n} x_{i,n+1}}{NI} = \alpha_1 \frac{\sum_n \sum_i y_{i,n} x_{i,n}}{\sum_n \sum_i y_{i,n}} + a_1 \bar{y}. \tag{69}$$

The expectation in Eq. 68 can be evaluated if parameters are known, or Eq. 69 can be used for estimation or testing.

By using the fact that the $y_{i,n}$ are generated by a probability mechanism, we can arrive at an approximation that is easier to work with. We do this by evaluating $E(\mathbf{t})$ for the "average" $y_{i,n}$-sequence produced by the probability mechanism rather than for the particular sequences used in the experiment. The approximation is obtained by applying to Eq. 68 the expectation operator E_y, which averages over the binomial outcome distribution of $\mathbf{y}_{i,n}$. Let $V_1 = E_n(V_{1,n})$, where the expectation is taken over the indicated trial block. Then,

$$\begin{aligned} E_y E(\mathbf{t} \mid \{\mathbf{y}_{i,n}\}) &= \alpha_1 E_n E_i E_y(\mathbf{y}_{i,n}\mathbf{p}_{i,n}) + a_1 E_y(\bar{\mathbf{y}}) \\ &= \alpha_1 \pi E_n E_i(\mathbf{p}_{i,n}) + a_1 \pi, \end{aligned}$$

and so

$$E_y E(\mathbf{t}) = \alpha_1 \pi V_1 + a_1 \pi. \tag{70}$$

The corresponding equation in terms of statistics of the data is

$$\frac{\sum_n \sum_i y_{i,n} x_{i,n+1}}{NI} = \alpha_1 \pi \frac{\sum_n \sum_i x_{i,n}}{NI} + a_1 \pi, \tag{71}$$

which is to be compared to the more exact Eq. 69.

5.4 Conditional Expectations and the Development of Functional Equations

Conditional expectations are useful also in establishing functional equations for interesting model properties. Let $G_1, G_2, \ldots, G_k, \ldots$ be a set of mutually exclusive and exhaustive events, and let $\mathbf{h}$ be a statistic whose expectation is desired. Then the property used is

$$E(\mathbf{h}) = \sum_k \Pr\{G_k\} E(\mathbf{h} \mid G_k), \tag{72}$$

and we consider two examples of its application to path-independent models with two subject-controlled events. Let $\mathbf{x}_n = 1$ if there is an error on trial n and $\mathbf{x}_n = 0$ if there is a success; let the operator for error be Q_2 and for success, Q_1. As our first example we let $\mathbf{h} = \mathbf{u}_1$, the total number of errors in an infinite sequence. The conditioning events are the possible responses on trial 1, so that G_1 corresponds to $\mathbf{x}_1 = 0$ and G_2 corresponds to $\mathbf{x}_1 = 1$. Equation 72 becomes

$$E(\mathbf{u}_1 \mid p_1 = p) = \Pr\{\mathbf{x}_1 = 0\} E(\mathbf{u}_1 \mid \mathbf{x}_1 = 0, p_1 = p) + \Pr\{\mathbf{x}_1 = 1\} E(\mathbf{u}_1 \mid \mathbf{x}_1 = 1, p_1 = p).$$

Now we note that $E(\mathbf{u}_1 \mid \mathbf{x}_1 = 0, p_1 = p) = E(\mathbf{u}_1 \mid p_1 = Q_1 p)$; that is, we can consider the process as if it began on the second trial with a different initial probability. Similarly, $E(\mathbf{u}_1 \mid \mathbf{x}_1 = 1, p_1 = p) = 1 + E(\mathbf{u}_1 \mid p_1 = Q_2 p)$; in this case we consider the process as beginning on the second trial but we add the error that has already occurred. The result is

$$E(\mathbf{u}_1 \mid p_1 = p) = (1 - p) E(\mathbf{u}_1 \mid p_1 = Q_1 p) + p[1 + E(\mathbf{u}_1 \mid p_1 = Q_2 p)]. \tag{73}$$

For a particular model, the expectation $E(\mathbf{u}_1 \mid p_1 = p)$ depends on the parameters and the value of p; we can suppress the parameters and write it simply as $f(p)$, a function of p. This function is unknown, but Eq. 73 tells us that it has the property that

$$f(p) = (1 - p) f(Q_1 p) + p[1 + f(Q_2 p)].$$

If $Q_1 p = \alpha_1 p$ and $Q_2 p = \alpha_2 p$, then

$$f(p) = (1 - p)f(\alpha_1 p) + p[1 + f(\alpha_2 p)], \tag{74}$$

with the boundary condition $f(0) = 0$. Equation 74 is an example of a *functional equation*, which defines some property of an unknown function that we seek to specify explicitly. It has been studied by Tatsuoka and Mosteller (1959).

In the preceding example a relation is given among the values of the function at an infinite set of triples of the values of its argument; that is, the set defined $\{p, \alpha_1 p, \alpha_2 p \mid 0 \leqslant p \leqslant 1\}$. A more familiar example of a functional equation is a difference equation; the values of the argument differ only by multiples of some constant. An example of such a set of arguments is $\{p, p + h, p + 2h \mid p = 0, h, 2h, \ldots, Nh\}$. Without loss of generality, a difference equation of this kind can be converted into one in which the arguments of the function are a subset of successive integers. A second familiar example of a functional equation is any differential equation. For both of these special types of functional equations, there is a much wider variety of methods of solution—methods of specifying the unknown function—than for the more general equations.

As a second example of the use of Eq. 72 in developing a functional equation let us consider a model with two subject-controlled events in which $\mathbf{x}_n = 1$ results in an increase in $\Pr\{\mathbf{x}_n = 1\} = \mathbf{p}_n$ toward $\mathbf{p}_n = 1$ and $\mathbf{x}_n = 0$ results in a decrease in $\mathbf{p}_n$ toward $\mathbf{p}_n = 0$. In such a model, after a sufficient number of trials, any response sequence will consist of either all "errors" or all "successes"; there are two asymptotically absorbing barriers, at $p = 1$ and $p = 0$.

An example of a linear-operator model of this kind is

$$\mathbf{p}_{n+1} = \begin{cases} Q_1\mathbf{p}_n = \alpha_1\mathbf{p}_n + (1 - \alpha_1), & \text{with probability } \mathbf{p}_n, \\ Q_2\mathbf{p}_n = \alpha_2\mathbf{p}_n, & \text{with probability } 1 - \mathbf{p}_n. \end{cases}$$

One of the interesting questions about such a model is to determine the probability of asymptotic absorption at $\mathbf{p}_\infty = 1$. Bush and Mosteller (1955, p. 155) show by an elementary argument that the distribution of $\mathbf{p}_\infty$ in this model is entirely concentrated at the two absorbing barriers. Therefore $\Pr\{\mathbf{p}_\infty = 1\} = E(\mathbf{p}_\infty)$, and it is fruitful to identify the $\mathbf{h}$ in Eq. 72 with $\mathbf{p}_\infty$. As before, we let G_1 correspond to $\mathbf{x}_1 = 0$ and G_2 correspond to $\mathbf{x}_1 = 1$, and we consider the expectation as a function of the starting probability. Equation 72 thus becomes

$$\begin{aligned}\Pr\{\mathbf{p}_\infty = 1 \mid p_1 = p\} &= E(\mathbf{p}_\infty \mid p_1 = p) \\ &= \Pr\{\mathbf{x}_1 = 0\}\, E(\mathbf{p}_\infty \mid \mathbf{x}_1 = 0, p_1 = p) \\ &\quad + \Pr\{\mathbf{x}_1 = 1\}\, E(\mathbf{p}_\infty \mid \mathbf{x}_1 = 1, p_1 = p).\end{aligned}$$

We note that $E(\mathbf{p}_\infty \mid \mathbf{x}_1 = 0, p_1 = p) = E(\mathbf{p}_\infty \mid p_1 = Q_1 p)$ and similarly that $E(\mathbf{p}_\infty \mid x_1 = 1, p_1 = p) = E(p_\infty \mid p_1 = Q_2 p)$. We then have

$$E(\mathbf{p}_\infty \mid p_1 = p) = pE(\mathbf{p}_\infty \mid p_1 = Q_1 p) + (1 - p)E(\mathbf{p}_\infty \mid p_1 = Q_2 p).$$

Letting $g(p)$ represent the expectation as a function of the starting probability, we arrive at

$$g(p) = pg(Q_1 p) + (1 - p)\, g(Q_2 p) \tag{75}$$

as a functional equation for the probability of absorption at $p = 1$. Boundary conditions are given by $g(1) = 1$ and $g(0) = 0$. The function $g(p)$ is understood to depend on parameters of the model in addition to p_1. Mosteller and Tatsuoka (1960) and others[22] have studied this functional equation for the foregoing linear-operator model. In general, no simple closed solution seems to be available. For the symmetric case of $\alpha_1 = \alpha_2 = \alpha < 1$ the solution is $g(p) = p$.

A similar functional equation can be developed for the beta model with two absorbing barriers and symmetric events. It is convenient to work with logit $\mathbf{p}_n$ instead of $\mathbf{p}_n$ itself. Following Eq. 49, we have, for this model

$$\text{logit } \mathbf{p}_n = -(a + b\mathbf{t}_n - b\mathbf{s}_n),$$

and the corresponding operator expression is

$$\text{logit } \mathbf{p}_{n+1} = \begin{cases} \text{logit } \mathbf{p}_n + b & \text{if } \mathbf{x}_n = 1 \quad \text{(i.e., with probability } \mathbf{p}_n), \\ \text{logit } \mathbf{p}_n - b & \text{if } \mathbf{x}_n = 0 \quad \text{(i.e., with probability } 1 - \mathbf{p}_n). \end{cases}$$

Let $\mathbf{L}_n$ = logit $\mathbf{p}_n$ and let $g(L)$ be the probability of absorption at $\mathbf{L} = \infty$ (which corresponds to $\mathbf{p}_\infty = 1$) for a process that starts at $L_1 = L$. Then Eq. 75 becomes the linear difference equation

$$g(L) = pg(L + b) + (1 - p)\, g(L - b), \tag{76}$$

where p = antilogit $L = 1/(1 + e^{-L})$. The boundary conditions are $g(-\infty) = 0$ and $g(+\infty) = 1$. This equation has been studied by Bush (1960) and Kanal (1962b).

5.5 Difference Equations

The discreteness of learning models makes difference equations ubiquitous in their exact analysis. The recursive equation for $\mathbf{p}_n$ is a difference equation whose solution is given by the explicit equation for $\mathbf{p}_n$; the argument of the difference equation is in this case the trial number n.

[22] See Shapiro and Bellman, cited by Bush & Mosteller, 1955.

A simple example is the recursive equation for the single linear operator model $p_{n+1} = \alpha p_n$, whose solution is $p_n = \alpha^{n-1}p_1$. A more interesting case is Eq. 13 for the prediction experiment,

$$p_{n+1} = \alpha p_n + (1 - \alpha)y_n,$$

with the solution

$$p_{n+1} = \alpha^{n-1}p_1 + (1 - \alpha)\sum_{j=1}^{n-1}\alpha^{n-1-j}y_j.$$

There are systematic methods of solution for many linear difference equations such as these (see, for example, Goldberg, 1958). Often a little manipulation yields a conjectured solution whose validity can be proved by mathematical induction.

Partial difference equations occasionally arise in learning-model analysis; they are more difficult to solve. We have seen in Sec. 5.2 that it is often necessary to know the moments $\{V_{m,n}\} = \{E(\mathbf{p}_n{}^m)\}$ of the p-value distributions generated by a model; properties of a model are often expressed in terms of these moments. The transition rules of linear-operator models lead to linear difference equations or other recurrence formulas for the moments. Occasionally these equations are "ordinary": one of the subscripts of $V_{m,n}$ is constant throughout the equation. An example is the equation for $V_{1,n}$ in the experimenter-controlled events model above, when we consider the $\{y_j\}$ to be random variables and $\Pr\{\mathbf{y}_n = 1\} = \pi$; it is given by the ordinary difference equation (Eq. 43)

$$V_{1,n+1} = \alpha V_{1,n} + (1 - \alpha)\pi, \tag{77}$$

whose solution is easily obtained (Eq. 44).

It is more usual for neither m nor n to be constant in the recurrence formula for $V_{m,n}$, and the formula is then a partial difference equation. In this case we cannot ignore the fact that $V_{m,n}$ is a function of a bivariate argument, and methods of solution are correspondingly more difficult. As an example we consider the linear-operator model with two commutative events and subject control, for which

$$\mathbf{p}_{n+1} = \begin{cases} \alpha_1\mathbf{p}_n & \text{with probability } 1 - \mathbf{p}_n \\ \alpha_2\mathbf{p}_n & \text{with probability } \mathbf{p}_n. \end{cases}$$

First let us consider how the partial difference equation for $V_{m,n}$ is derived. We assume a population of subjects with common values of p_1, α_1, and α_2. On trial n, after $n - 1$ applications of the operators, the population consists of n distinct subgroups defined by the number of times α_1 has been applied. Let $1 \leqslant \nu \leqslant n$ be the index for these subgroups, let $p_{\nu,n}$ be the p-value for the νth subgroup on trial n, and let $P_{\nu,n}$ be the size of this subgroup, expressed as a proportion of the population. Now let us consider the fate of the νth subgroup on trial n. A proportion,

$p_{v,n}$, of the subgroup makes an error on that trial, and its p-value becomes $\alpha_2 p_{v,n}$. The remaining proportion of the subgroup, $1 - p_{v,n}$, performs a correct response, and its p-value becomes $\alpha_1 p_{v,n}$. The result is expressed in the following table:

New p-Values	New Proportions
$\alpha_2 p_{v,n}$	$p_{v,n} P_{v,n}$
$\alpha_1 p_{v,n}$	$(1 - p_{v,n}) P_{v,n}$.

(78)

Therefore,

$$V_{m,n+1} = \sum_{v=1}^{n+1} p_{v,n+1}^{m} P_{v,n+1}$$

$$= \sum_{v=1}^{n} (\alpha_2 p_{v,n})^m p_{v,n} P_{v,n} + \sum_{v=1}^{n} (\alpha_1 p_{v,n})^m (1 - p_{v,n}) P_{v,n}$$

$$= (\alpha_2{}^m - \alpha_1{}^m) \sum_{v=1}^{n} p_{v,n}^{m+1} P_{v,n} + \alpha_1{}^m \sum_{v=1}^{n} p_{v,n}^{m} P_{v,n}$$

$$V_{m,n+1} = (\alpha_2{}^m - \alpha_1{}^m) V_{m+1,n} + \alpha_1{}^m V_{m,n}. \tag{79}$$

One feature of this equation, which is generally true of models with subject control, is that $V_{m,n}$ is expressed in terms of moments higher than the mth moment of the p-value distribution on preceding trials. With experimenter control this complicating feature is absent, as illustrated by Eq. 77.

Equation 79 has been solved by conjecture and inductive proof rather than by any direct method. To illustrate how cumbersome some of the results become in this field, I reproduce the solution here:

$$V_{m,n} = \alpha_1^{m(n-1)} p_1{}^m + \sum_{k=2}^{n} \alpha_1^{m(n-k)} p_1^{j+m-1} \prod_{j=m}^{k+m-2} \frac{(\alpha_2{}^j - \alpha_1{}^j)(1 - \alpha_1^{n-j+m-1})}{1 - \alpha_1^{j-m+1}}$$

$$(m \geqslant 1, n \geqslant 1), \tag{80}$$

where the sum is defined to be zero for $n = 1$.

For more examples of the development of recursive formulas for moments, see Bush & Mosteller (1955, Chapter 4), and for some examples of their use see Bush (1959), Estes & Suppes (1959, Sec. 8), and Sternberg (1959b).

5.6 Solution of Functional Equations[23]

Two methods by which functional equations have been studied are illustrated here; the first is a power-series expansion and the second is a differential equation approximation.

[23] See Kanal (1962a,b) for the formulation of some functional equations arising in the analysis of the linear and beta models and for methods of solution and approximation.

Tatsuoka and Mosteller (1959) solved Eq. 74 by using a power-series expansion. Assume that $f(p)$ is expressible as a power series in p:

$$f^*(p) = \sum_{k=0}^{\infty} c_k p^k. \tag{81}$$

The boundary condition $f^*(0) = 0$ implies that $c_0 = 0$. By substituting the series expansion into the functional equation and equating coefficients of like powers of p we find

$$c_k = \frac{\prod_{j=1}^{k-1} (\alpha_2^{\,j} - \alpha_1^{\,j})}{\prod_{j=1}^{k} (1 - \alpha_1^{\,j})}, \quad k \geqslant 1. \tag{82}$$

For certain special cases this expression can be simplified. For example, with $\alpha_2 = 1$ (identity operator for "error") and $0 \leqslant \alpha_1 < 1$, $c_k = 1/(1 - \alpha_1^{\,k})$ and therefore

$$f^*(p) = \sum_{k=1}^{\infty} \frac{p^k}{1 - \alpha_1^{\,k}}. \tag{83}$$

A closed form for this expression has not been found, but there are tables (Bush, 1959) and approximations (Tatsuoka & Mosteller, 1959).

The only solution of Eq. 74 that satisfies the boundary condition *and* has an expansion in powers of p is $f^*(p)$. To prove this, we assume that there is a second power-series solution, $f^{**}(p) = \sum_{k=1}^{\infty} d_k p^k$ and replace $f(p)$ in Eq. 74 first by $f^*(p)$ and second by $f^{**}(p)$. By subtracting the second resulting equation from the first we obtain an equation for

$$f^*(p) - f^{**}(p) = \sum_{k=1}^{\infty} (c_k - d_k) p^k,$$

whose solution requires that $c_k - d_k = 0$ for $k \geqslant 1$. In general, however, a functional equation may possess solutions for which a power-series expansion is not possible. For this reason it is necessary either to provide a general proof of the uniqueness of $f^*(p)$ or to show that we are interested only in solutions of Eq. 74 with power-series expansions.

Kanal (1960, 1962a) has shown that $f^*(p)$ is the only solution of Eq. 74 that is continuous at $p = 0$ but no *general* proof of uniqueness is available at present. Fortunately, we can use Eq. 80 to show that for the model in question $E(\mathbf{u}_1)$ has a series expansion in powers of p and that power-series solutions of Eq. 74 are therefore the only ones of interest. To do this, we note that

$$E(\mathbf{u}_1) = E\left(\sum_{n=1}^{\infty} \mathbf{x}_n\right),$$

and that

$$E\left(\sum_{n=1}^{\infty} \mathbf{x}_n\right) = \sum_{n=1}^{\infty} E(\mathbf{x}_n) \tag{84}$$

if the right-hand series converges. Equation 80 provides an expression for $E(\mathbf{x}_n) = V_{1,n}$; it is a polynomial in $p_1 = p$. If $\alpha_1 < 1$ and either $\alpha_2 < 1$ or $p_1 < 1$, $\sum_{n=1}^{\infty} V_{1,n}$ converges. We therefore know, incidentally, that under these conditions $E(\mathbf{u}_1)$ exists. Moreover, because it is the sum of a convergent infinite series of polynomials, it must have a power-series expansion.

For some functional equations the power series that is obtained may not converge, and we cannot apply the foregoing method. As an example of a second method we consider Bush's (1960) solution of Eq. 76. First the equation is written in terms of first differences and $(1 - p)/p$ is replaced by e^{-L}:

$$g(L + b) - g(L) = e^{-L}[g(L) - g(L - b)]. \tag{85}$$

In order to convert (85) into a linear equation, the logarithmic transformation is applied to both sides, and the logarithm of the first difference is defined as a new function, $h(L) \equiv \log\,[g(L) - g(L - b)]$, to give

$$h(L + b) = h(L) - L. \tag{86}$$

Equation 86 is the difference equation to be solved.

In this case a solution is sought, not by a power series expansion but by a differential equation approximation of the difference equation. We write Eq. 86 in a form symmetric about L:

$$h\left(L + \frac{b}{2}\right) - h\left(L - \frac{b}{2}\right) = -\left(L - \frac{b}{2}\right),$$

divide by b,

$$\frac{\Delta h}{\Delta L} = -\frac{L}{b} + \frac{1}{2},$$

and treat the result as a derivative,

$$\frac{dh}{dL} = -\frac{L}{b} + \frac{1}{2}.$$

Integration gives

$$h(L) = -\frac{L^2}{2b} + \frac{L}{2} + C \tag{87}$$

as a conjectured solution. The result satisfies Eq. 86, and therefore a particular solution of the complete equation is given by Eq. 87 with $C = 0$.

The homogeneous equation $h(L + b) = h(L)$ has as its general solution

$\tilde{P}(L)$, an arbitrary periodic function of L with period b. For the general solution of the complete equation we then have

$$h(L) = \frac{L}{2} - \frac{L^2}{2b} + \tilde{P}(L). \tag{88}$$

To recover g, we use

$$g(L) - g(L - b) = \exp\,[h(L)] = \exp\left[\frac{L}{2} - \frac{L^2}{2b} + \tilde{P}(L)\right],$$

and completing the square gives us

$$g(L) - g(L - b) = P(L) \exp\left[-\frac{1}{2b}\left(L - \frac{b}{2}\right)^2\right], \tag{89}$$

where $P(L)$ is some other periodic function of L, with period b. This new difference equation is simpler than the original (Eq. 85) because it contains one difference instead of two, and therefore routine procedures can be used. We first note the boundary conditions $g(-\infty) = 0$ and $g(\infty) = 1$. Then we replace L by $L - b$, $L - 2b$, and so on, to obtain the semi-infinite system

$$\begin{aligned} g(L) - g(L - b) &= P(L) \exp\left[-\frac{1}{2b}\left(L - \frac{b}{2}\right)^2\right] \\ g(L - b) - g(L - 2b) &= P(L) \exp\left[-\frac{1}{2b}\left(L - b - \frac{b}{2}\right)^2\right] \\ g(L - 2b) - g(L - 3b) &= P(L) \exp\left[-\frac{1}{2b}\left(L - 2b - \frac{b}{2}\right)^2\right] \\ &\cdots\cdots\cdots\cdots \end{aligned} \tag{90}$$

Addition of these equations and use of the first boundary condition gives

$$g(L) = P(L) \sum_{k=0}^{\infty} \exp\left[-\frac{1}{2b}\left(L - kb - \frac{b}{2}\right)^2\right].$$

The sum may be approximated by a normal integral. The periodic function is still arbitrary; to specify it, we write the full infinite system corresponding to Eqs. 90, sum, and use both boundary conditions; the final result is

$$g(L) = \frac{\displaystyle\sum_{k=0}^{\infty} \exp\left[-\frac{1}{2b}\left(L - kb - \frac{b}{2}\right)^2\right]}{\displaystyle\sum_{k=-\infty}^{\infty} \exp\left[-\frac{1}{2b}\left(L - kb - \frac{b}{2}\right)^2\right]}, \tag{91}$$

which is roughly of the form of a normal integral. In most practical cases $P(L)$ may be approximated by a constant. When antilogits are taken, the absorption probability as a function of p is no longer a normal integral, but it is similar in character; the resulting S-shaped curve is to be compared to the result for the symmetric linear-operator model for which the absorption probability is equal to the initial probability.

6. SOME ASPECTS OF THE APPLICATION AND TESTING OF LEARNING MODELS

6.1 Model Properties: A Model Type as a Subspace

In the last three sections I have mentioned examples of many of the model properties that have been studied and compared with data. Linear models are the best known: they were studied by Bush and Mosteller (1955) and recent progress, a good deal of which is represented in Bush & Estes (1959), has been considerable. Even so, our knowledge tends to be spotty, concentrated at certain special examples of linear models. Models with extreme or equal limit points and those with equal learning-rate parameters are better understood than the others. The single-operator model and models with experimenter control are the most thoroughly studied (Bush & Sternberg, 1959; Estes & Suppes, 1959); analytic expressions are available for the expectations of a good many statistics of these models. On the other hand, except for some of its asymptotic properties, we have less information about Luce's beta model (Bush, 1960; Lamperti & Suppes, 1960; Kanal, 1962a,b). For this model, and for any more general logistic models, there is a compensating advantage: standard methods of estimation can easily be applied. Once estimates are obtained, Monte Carlo calculations can be used for detailed comparisons. It can be argued, moreover, that the advantages of optimal (maximum-likelihood) estimation methods outweigh the convenience of having analytic expressions for model properties.

To arrive at one view of the properties of a model—a view that is helpful in considering the problems of fitting and testing the model—we begin by considering the m-dimensional "property-space" consisting of all values of the vector $(s_1, s_2, \ldots, s_m)$, where s_j denotes a property (the expectation or variance of a statistic) of the model. The corresponding statistic for some observed data sequences is denoted by $\bar{s}_j$. In general, the properties depend on parameter values, and therefore $s_j = s_j(\Theta)$, where Θ is a vector of parameters corresponding to a point in the parameter space. As the point moves through the entire parameter space, the s_j take on all the combinations of values allowed by the model type. For certain purposes we can now ignore the parameters and consider only these allowed combinations, which define a subspace of the property-space. If there were no

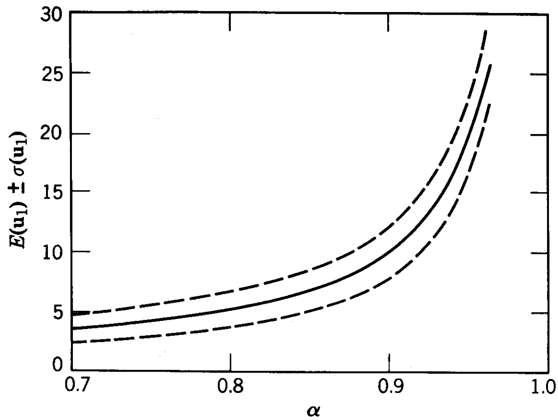

Fig. 11. The solid curve represents the expected number of errors in an infinite sequence of trials, $E(\mathbf{u}_1)$, as a function of the learning-rate parameter α for the single-operator model (Eq. 28) with $p_1 = 1$. The distance between the solid curve and a broken curve represents the standard deviation of the number of errors.

sampling variability, the problem of testing a model type would reduce to the question whether the observed $(\bar{s}_1, \bar{s}_2, \ldots, \bar{s}_m)$ is a point in the subspace. The existence of sampling fluctuations means that the question must be modified so that a certain degree of discrepancy is tolerated.

A simple example with a one-dimensional parameter space illustrates this viewpoint. We consider the single-operator model given by Eq. 28 and discussed in Sec. 4.4. A good many properties of this model are known (Bush & Sternberg, 1959), and in Figs. 11, 12, and 13 three are illustrated graphically. The total number of errors (in an infinite sequence of trials) is symbolized by $\mathbf{u}_1$, the total number of runs of errors by $\mathbf{r}_T$, and the number of trials before the first success by $\mathbf{f}$. These three properties suffice for our purpose, and the property-space we consider is therefore three-dimensional. In order for the model to have only a single free parameter we assume that p_1, the initial probability of error, is known to be unity. The dependence of $E(\mathbf{f})$, $E(\mathbf{u}_1)$, and $E(\mathbf{r}_T)$ on the value of the learning-rate parameter α is shown by the figures and also by the following equations.[24]

$$E(\mathbf{u}_1) = \frac{1}{1-\alpha}, \qquad E(\mathbf{r}_T) = \frac{1}{1-\alpha^2}, \qquad E(\mathbf{f}) = \sum_{k=0}^{\infty} \alpha^{k(k+1)/2}. \qquad (92)$$

[24] The infinite sum for $E(\mathbf{f})$ is tabulated in Bush & Mosteller (1955, Table A) and approximated in Galanter & Bush (1959).

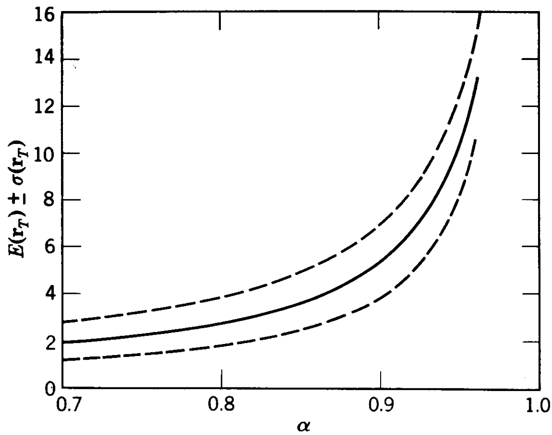

Fig. 12. The solid curve represents the expected number of runs of errors in an infinite sequence of trials, $E(\mathbf{r}_T)$, as a function of the learning-rate parameter α for the single-operator model (Eq. 28) with $p_1 = 1$. The distance between the solid curve and a broken curve represents the standard deviation of the number of runs of errors.

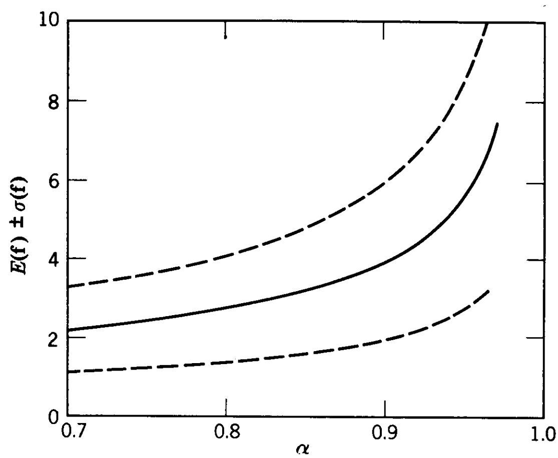

Fig. 13. The solid curve represents the expected number of trials before the first success $E(\mathbf{f})$ as a function of the learning-rate parameter α for the single-operator model (Eq. 28) with $p_1 = 1$. The distance between the solid curve and a broken curve represents the standard deviation of the number of trials before the first success.

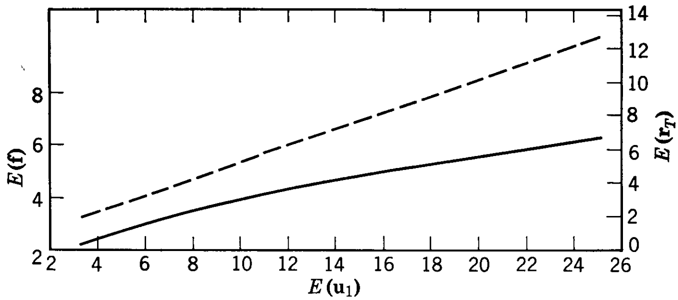

Fig. 14. The solid curve (left-hand ordinate) describes $E(\mathbf{f})$ as a function of $E(\mathbf{u}_1)$ for the single-operator model (Eq. 28) with $p_1 = 0$. The broken curve (right-hand ordinate) describes $E(\mathbf{r}_T)$ as a function of $E(\mathbf{u}_1)$ for this model. Units are chosen so that scales are comparable in standard deviation units.

Also shown for each statistic is the interval defined by twice its standard deviation, for example, $\pm\sqrt{\text{Var}(\mathbf{u}_1)}$. The parameter α can be eliminated from any pair of the functions $s_j(\alpha)$, to define a relation between the two expectations. Two such relations are shown in Fig. 14, in which $E(\mathbf{f})$ and $E(\mathbf{r}_T)$ are plotted against $E(\mathbf{u}_1)$. These two relations are projections of a curve in the three-dimensional property-space with dimensions $E(\mathbf{f})$, $E(\mathbf{r}_T)$, and $E(\mathbf{u}_1)$; this curve is the subspace of the property-space to which the model corresponds. The units are so chosen that the scales are roughly comparable in standard deviation units, that is, a 1-cm discrepancy on the $E(\mathbf{u}_1)$ scale is as serious as a 1-cm discrepancy on either of the other scales.

In Table 4 the average values of the three statistics are given for three

Table 4 Observed Values of u_1, f, r_T in Three Experiments

Experiment	$\bar{u}_1$	$\bar{f}$	$\bar{r}_T$
T-Maze reversal after overlearning (Galanter & Bush, 1959)	24.68	13.32	6.11
T-Maze reversal (Galanter & Bush, 1959, Period 2)	14.10	5.30	6.60
Solomon-Wynne shuttlebox (Bush & Mosteller, 1959)	7.80	4.50	3.24

experiments in which the assumption that $p_1 = 1$ is tenable. It is instructive to examine these values in conjunction with the four graphs for the single-operator model. In none of the experiments does any of the pairs of statistics satisfy the relations for this simple model. Put another way, in no case does the observed point $(\bar{u}_1, \bar{f}, \bar{r}_T)$ fall within the allowed subspace. How large a discrepancy is tolerable is a statistical problem that is touched on later.

A good deal of the work that has been done on fitting and testing models can be thought of as being analogous to the process exemplified above: mathematical study of a model yields several functions, $s_j(\Theta)$, of properties in terms of parameter values, and the question is asked whether there is a choice of parameter values (a point in the parameter space) for which the observed $\bar{s}_j$ are close to their theoretical values. The process is usually conducted in two stages: first, estimation, in which parameter values $\hat{\Theta}$ are selected so that a subset of the $\bar{s}_j$ agrees exactly with the theoretical values, and, second, testing, in which the remaining $\bar{s}_j$ are compared to their corresponding $s_j(\hat{\Theta})$. In the second stage some or all of the theoretical values may be estimated from Monte Carlo calculations. These stages correspond in Fig. 14, for example, to first letting $\bar{u}_1$ determine a point on the abscissa and, second, comparing the corresponding ordinate values to $\bar{f}$ and $\bar{r}_T$.

Conclusions from this method are conditional on the choice of properties used in each of the two stages. To assert that "model X cannot describe the observed distribution of error-run lengths" is stronger than is usually warranted. More often the appropriate statement is of the form "when parameters for model X are chosen so that it describes $\bar{s}_1$ and $\bar{s}_2$ exactly then it cannot describe $\bar{s}_3$." Occasionally there are exceptions in which some property of a model is independent of its parameter values. For example, we can assert unconditionally that "an S-shaped curve of probability versus trials cannot be described by the single-operator model" or that "the linear-operator model with complementary experimenter-controlled events for the prediction experiment must (if learning occurs at all) produce asymptotic probability-matching for all values of π." Such parameter-free properties of a model are worthy of energetic search.

6.2 The Estimation Problem

Most model types have one or more free parameters whose values must be estimated from data. Estimates that satisfy over-all optimal criteria, such as maximum likelihood or minimum chi-square, cannot usually be obtained explicitly in terms of statistics of the data. Because the iterative or other numerical methods that are needed in order to obtain such estimates are inconvenient, they have seldom been used in research with learning models. The more common method has been briefly touched on in Sec. 6.1: parameter values are chosen to equate several of the observed statistics of the data with their expectations as given by the model. The estimates $\hat{\Theta}$ are therefore produced by the solution of a set of equations of the form $s_j(\Theta) = \bar{s}_j$.

Because the properties of estimates so obtained are not well understood, this method may lead to serious errors, as is illustrated by an example. Let us suppose that a learning process behaves in accordance with the single-operator model with $\alpha = 0.90$ and that we do not know this but wish to test the model as a possible description of the data. Suppose that we have a single sequence of responses and that we have reason to assume that $p_1 = 1$. Suppose, further, that because of sampling variability the number of errors at the beginning of the sample sequence is accidentally too large and that the observed values of u_1 and f are inflated, each by an amount equal to its theoretical standard deviation. Figures 12 and 14 then provide us with the following values:

	u_1	f
True (population) value	10.00	3.91
Sample value	12.18	5.91

We have two properties s_j, namely u_1 and f, and a single parameter, α, to estimate. One property is needed for estimation and the remaining one is available for testing. The choice of which property to use for which purpose involves only two alternatives; but it has the essential character of the more complicated choice usually available to the investigator. One procedure has been to use "gross" features of the data, such as the total number of errors, for estimation, and "fine-grain" features, such as the distribution of error-run lengths, for testing. (For examples, see Bush & Mosteller, 1959, and Sternberg, 1959b). Occasionally the investigator cannot choose; whatever few statistics he is lucky enough to have analytic expressions for are automatically elected for use in estimation, and for testing he must resort to Monte Carlo calculations. (For an example, see Bush, Galanter, & Luce, 1959.) Let us examine, in the case of the two statistics tabulated above, how the choice that is made affects our inference about goodness of fit of the model.

First, suppose that we use f for estimation, choosing α so that $E(\mathbf{f} \mid \alpha) = \bar{f}$. Entering Fig. 13 with the observed value of $\bar{f} = 5.9$, we find the corresponding estimate to be $\hat{\alpha} = 0.957$. Now to test the model we refer to Fig. 11. Corresponding to $\alpha = 0.957$ are the values of $E(\mathbf{u}_1) = 23.3$ and $\sigma(\mathbf{u}_1) = 3.35$. The difference between $E(\mathbf{u}_1)$ and its observed value of

$\bar{u}_1 = 12.18$ is more than three times its theoretical standard deviation, a sizable discrepancy. On these grounds we would be inclined to discard the model. But first let us consider the result if we take the second option. We use u_1 for estimation, and Fig. 11 gives the value of $\hat{\alpha} = 0.917$ for $E(\mathbf{u}_1) = 12.18$. To test the model, we enter Fig. 13 with $\alpha = 0.917$; the corresponding theoretical values are $E(\mathbf{f}) = 4.3$, $\sigma(f) = 2.15$. The observed value $\bar{f} = 5.91$ is therefore within one standard deviation of the theoretical value. The second option inclines us to accept the model. It is worth noting that, if anything, this example is conservative: had the number of errors been accidentally large, but toward the end of the sequence instead of the beginning, $\bar{f}$ would have been close to its theoretical value, $\bar{u}_1$ would have been inflated, and the two results would have been still more discrepant.

The reason for the disagreement may be clarified by Fig. 14. Here it can be seen that for this particular model an error in $E(\mathbf{f})$ corresponds to a much larger error in $E(\mathbf{u}_1)$, in terms of standard deviation units. The total-errors statistic is the most "sensitive" of the three; therefore, to give the model the best chance, it is the one that should be used for estimation.

The question of the choice of an estimating statistic is therefore a delicate one. In the lucky instance in which a model approximates the data well in many respects it is unimportant how the estimation is carried out. Such an instance is the Bush-Mosteller (1955) analysis of the Solomon-Wynne data, in which several methods gave estimates in very good agreement; indeed, this fact in itself is strong evidence in favor of the model. But this is a rare case, and more often estimates are in conflict.

The question becomes especially important when several models are to be compared in their ability to describe a set of data. It is crucial that the estimation methods be equally "fair" to the models, and the standard procedures do not ensure this. For example, we might be comparing with the single-operator model of Fig. 14 another hypothetical model for which the curve of $E(\mathbf{f})$ versus $E(\mathbf{u}_1)$ had a slope greater rather than less than unity. If we then used u_1 in estimation for both models, we would be prejudicing the comparison in favor of the single-operator model. For different models different sets of statistics may be the best estimators: we do not ensure equal fairness by using the same estimating statistics for all the models to be compared. This observation, for which I am indebted to A. R. Jonckheere,[25] casts doubt on the results of certain comparative studies, such as those of Bush and Mosteller (1959), Bush, Galanter, and Luce (1959), and Sternberg (1959b).

One possibility for retrieving the situation is to search for aspects of the data that one or more of the competing models are incapable of describing,

[25] Personal communication, 1960.

regardless of the values of its parameters. An example arises in the analysis of reversal after overlearning in a T-maze, one of the experiments included by Bush, Galanter, and Luce (1959) in their comparison of the linear and beta models. The observed curve of proportion of successes versus trials starts at zero and is markedly S-shaped, rising steeply in its middle portion. Parameters can be chosen for the beta model so that its curve agrees well with the one observed. But, as Galanter and Bush (1959) show, although the linear model of Eq. 8 is capable of producing an S-shaped curve that starts at zero ($p_1 = 1$), no choice of α_1 and α_2 permits its curve to rise both slowly enough at the beginning and end of learning and steeply enough in its middle portion. As the analysis stands, then, the beta model is to be preferred for these data. The problem is that if we search long enough we may be able to find a property of the data that this model cannot describe and that the linear model can. To choose between the models, we would then have to decide which of the two properties is the more "important," and the problem of being equally fair to the competing models would again face us.

A solution to the problem lies in the use of maximum likelihood (or other "best") estimates (Wald, 1948), despite their frequent inconvenience, and in the comparison of the maximized likelihoods and the use of likelihood-ratio tests to assess relative goodness of fit. Bush and Mosteller (1955) discuss several over-all measures of goodness of fit. The use of such over-all tests has occasionally been objected to on grounds that they may be sensitive to uninteresting differences among models or between models and data and that they may not reveal the particular respects in which a model is deficient. Our example of the $\mathbf{f}$ and $\mathbf{u}_1$ statistics shows that the first objection applies to the more usual methods as well. In answer to the second objection, there is no reason why detailed comparison of particular statistics cannot be used as a supplement to the over-all test.

One of the desirable features of the beta model and of more general logistic models is that a simple set of sufficient statistics exists for the parameters and that the standard iterative method (Berkson, 1957) for obtaining the maximum-likelihood estimates is easily generalized for more than two parameters, converges rapidly, and is facilitated by existing tables. Cox (1958) suggests that, in applications to learning, initial estimates be obtained by the minimum-logit χ^2 method (Berkson, 1955; Anscombe, 1956), which does not require iteration.

Examples of the results of these methods applied to the Solomon-Wynne data (Sec. 4.1) are given in Table 5. Both the maximum-likelihood and minimum-logit χ^2 methods can be thought of as ways of fitting the linear regression equation given by Eq. 49:

$$\text{logit } \mathbf{p}_n = -(a + b\mathbf{t}_n + c\mathbf{s}_n).$$

The random variables $\mathbf{t}_n$ and $\mathbf{s}_n$ are considered to be the independent variables, and the logit of the escape probability is the dependent variable. The equation defines a plane; the observed points to which the plane is fitted are the logits of the proportions of escapes at given values of $(\mathbf{t}_n, \mathbf{s}_n)$. A difficulty arises with the minimum-logit χ^2 method when the observed proportion for a $(\mathbf{t}_n, \mathbf{s}_n)$ pair is zero, as happens often when $\mathbf{s}_n + \mathbf{t}_n$ is large in the later trials. Most of these zero observations were omitted in obtaining the values in the second row of Table 5, so that this row of values depends principally on early trials. Relations between the values of $\hat{p}_1$, $\hat{\beta}_1$, and $\hat{\beta}_2$ and the estimates of a, b, and c are given in Sec. 2.5.

Table 5 Results of Four Procedures for Estimating Parameters of the Beta Model (Eqs. 17, 19) from the Solomon-Wynne Data

Method	$\hat{p}_1$ (initial escape probability)	$\hat{\beta}_1$ (avoidance)	$\hat{\beta}_2$ (escape)
Bush-Galanter-Luce (1959)	0.94	0.59	0.83
Minimum logit χ^2	0.864	0.760	0.778
One maximum-likelihood iteration	0.857	0.805	0.718
Two maximum-likelihood iterations	0.857	0.811	0.735

When there are only two parameters, as in the case of the beta model for two symmetric experimenter-controlled events (Eq. 24),

$$\operatorname{logit} p_n = -(a + bd_n),$$

a simple graphical method (Hodges, 1958) provides close approximations to the maximum-likelihood estimates; it is probably preferable to minimum-logit χ^2 for obtaining starting values for maximum-likelihood iteration. The minimum-logit χ^2 method should be used with caution; it may occasionally be misleading, perhaps because of the difficulty with zero entries already mentioned. Consider, as an example, the logistic one-trial perseveration model (Eq. 32): $\operatorname{logit} \mathbf{p}_n = a + b(n-2) + c\mathbf{x}_{n-1}$, $(n \geqslant 2)$. A simple graphical method is the visual fitting of a pair of parallel lines to the proportions that estimate $\Pr(\mathbf{x}_n = 1 \mid \mathbf{x}_{n-1} = 0)$ and $\Pr(\mathbf{x}_n = 1 \mid \mathbf{x}_{n-1} = 1)$ when they are plotted against $n \geqslant 2$ on logistic (or normal probability) paper. These lines then represent $\operatorname{logit} p_n = a + b(n-2)$ and $\operatorname{logit} p_n = (a + c) + b(n-2)$ and provide estimates of the three parameters. Values obtained for the Goodnow data (Sec. 4.5), using the graphical method and then applying one cycle of maximum-likelihood iteration to its results, are presented in Table 6. For these data, the minimum-logit χ^2 method gave values that departed more from the maximum-likelihood values than the simple graphical procedure.

The advantages of maximum-likelihood estimates are that their variances are known, at least asymptotically, and that their values tend to represent much of the information in the data. When the maximum-likelihood method is not used, alternative methods that have these properties are to be preferred. As an example, let us consider estimation for the linear-operator model with experimenter control (Eq. 12) that has been used for the prediction experiment with $\Pr\{\mathbf{y}_n = 1\} = \pi$ and $\Pr\{\mathbf{y}_n = 0\} = 1 - \pi$. If

Table 6 Estimates for the Logistic Perseveration Model (Eq. 32) from the Goodnow Data

Method	$\hat{a}$	$\hat{b}$	$\hat{c}$
Visual fit of parallel lines on logistic paper	0	−0.24	0.94
One maximum-likelihood iteration	0.035	−0.236	0.927

$\mathbf{t}_i$ is the total number of A_1 responses by the ith subject during the first N trials, then for this model

$$E(\mathbf{t}_i) = N\pi - (\pi - V_{1,1})\left(\frac{1-\alpha^N}{1-\alpha}\right), \tag{93}$$

and, having determined $\hat{V}_{1,1}$, we can estimate α by setting $E(\mathbf{t}_i)$ equal to the value of t for a group of subjects. This is the method used by Bush and Mosteller (1955), Estes and Straughan (1954), and others. Equation 93 is obtained by adding both sides of the approximate equation for the learning curve (Eq. 44) over trials, $1 \leqslant n \leqslant N$.

One of the observed features of estimates obtained by this method is that $\hat{\alpha}$ varies with the value of π: the higher π (the "easier the discrimination"), the more potent the learning-rate parameter. Psychological mechanisms have been proposed to explain this effect (e.g., Estes, 1959), but very little is known about the method of estimation itself. For example, is $\hat{\alpha}$ unbiased? If not, how does the bias depend on π? The estimate depends entirely on preasymptotic data. (This can be seen from the fact that its value is indeterminate if $V_{1,1} = \pi$.) For experiments in which $V_{1,1} \simeq 0.5$, therefore, the higher the value of π, the more data are used in the estimate of α, hence the more reliable the estimate. Other information about this estimation procedure that is not known but that is vital for the interpretation of the findings is the extent to which perturbations in the process affect the estimate. Examples of such perturbations are intersubject differences in initial probabilities and in values of α and the difference between the effects of nonreward and reward discussed in Sec. 4.3.

The curious fact that no information about α seems to be available from asymptotic data is a result of averaging over the distribution of outcomes and using the resulting approximate equation (Eq. 44). Clearly there is more information in the data than is used for the estimate. This sequential information has been exploited by Anderson (1959) in a variety of procedures for estimation and testing. A simple improvement on the average learning curve procedure arises from the idea that the extent to which responses are correlated with the immediately preceding outcomes depends on the value of α and leads to the use of Eq. 71, or its exact version, Eq. 69, with $a_1 = 1 - \alpha_1$, to estimate α. By this method, estimates can be obtained even from trend-free portions of the data. Such an improvement is in the direction of the use of sufficient statistics, to which the maximum-likelihood method often leads. But it is still inferior to what is possible for the comparable beta model.

6.3 Individual Differences

In most of the discussion in this chapter, and in most applications of learning models, it is assumed that the same values of the initial probability and other parameters characterize all the subjects in an experimental group. When events are subject-controlled, differences in p-values arise on trials after the first, but under this homogeneity assumption these are due entirely to differences between event sequences.

It must be kept in mind, when this assumption is made in the application of a model type, that what is tested by comparisons between data and model is the conjunction of the assumption and the model type and not the model type alone. It is convenience, not theory, that leads to the homogeneity assumption. The question of primary interest is whether each individual subject, with his own parameter values, can be said to behave in accordance with the model type. It is usually thought that if the assumption is not entirely justified then the discrepancy will cause the model to underestimate the intersubject variances of response-sequence statistics. It is hoped (but not known) that the discrepancy will have no other adverse effects. We therefore expect the variances given by a model to be on the small side, and we are not perturbed when this occurs, as it often does (Bush & Mosteller, 1959; Sternberg, 1959b).

On the other hand, unless we are interested specifically in testing the homogeneity assumption, it is probably unwise to use an observed variance as a statistic for estimation, and this is seldom done. One difficulty with the customary procedure, in which the assumption of homogeneity is made, is that estimation and testing methods for different models may be differentially sensitive to deviations from homogeneity. For comparing models, therefore, it is probably preferable to estimate parameters separately for each subject. Audley and Jonckheere (1956) argued for the desirability of this procedure, and Audley (1957) carried it out for a model that describes both choice and choice time.

Estimates for an individual subject that are based on few observations may be unstable. One way of avoiding both the instability of individual estimates and the assumption of homogeneity is to study long sequences of responses from individual subjects. Anderson (1959) favors this method and gives estimation procedures. However, it clearly cannot be applied to experiments in which a single response is perfectly learned in a small number of trials.

Certain types of inference, based on between-subject comparisons, may be misleading if the homogeneity assumption is not met. For example, we might observe a positive correlation between the number of errors before and after some arbitrary trial. One possible cause of the correlation is a positive response effect. A second is the existence of individual differences. Even if a response-independent, single-event model describes the process for each subject, differences in initial probabilities or learning rates will produce a positive correlation of this kind. If there is a negative response effect as well as individual differences, statistics such as the variance of total errors or the correlation of early and late errors might lead us to infer no response effect at all if we assume homogeneity. If, on the other hand, we observe a negative correlation between the number of early and late errors, despite the possible interference of individual differences, we are on sure ground when we infer a negative response effect. By the same token, when homogeneity is assumed and the theoretical variances are too large, we have especially strong evidence against the model type. This last effect has occurred in the analysis of data from both humans (Sternberg, 1959b) and rats (Galanter & Bush, 1959; Bush, Galanter, & Luce, 1959).

When we look at the over-all picture of results from the application of learning models, it is remarkable how weak the evidence against the homogeneity assumption usually appears to be. One is reluctant to believe that individuals are so alike, but the only alternative seems to be that our testing methods are insensitive. N. H. Anderson[26] has argued

[26] Personal communication, 1962.

that this alternative gains support from the fact that analyses of variance of repeated measurements almost always yield significant individual differences.

Direct tests of the homogeneity assumption have occasionally been performed. Data from a single trial alone are, of course, useless as evidence of any more than the first moment of the p-value distribution. But if the p-value for each subject is approximately constant during a block of m trials, then raw moments from the first to the mth can be estimated from the block. As an example, suppose we use a block containing trials one and two. The method[27] depends on the two relations

$$E(\mathbf{x}_1 + \mathbf{x}_2) = E_p E_x(\mathbf{x}_1 + \mathbf{x}_2) = E_p(2\mathbf{p}) = 2V_1$$

$$E[(\mathbf{x}_1 + \mathbf{x}_2)^2] = E_p E_x[(\mathbf{x}_1 + \mathbf{x}_2)^2] = E_p(2\mathbf{p} + 2\mathbf{p}^2) = 2V_1 + 2V_2,$$

where p is the (approximately constant) probability on the two trials and V_1 and V_2 are the (approximate) first and second moments of its distribution. The expectations on the left are replaced by the averages of $x_1 + x_2$ and $(x_1 + x_2)^2$ over subjects, and then the equations are solved for $\hat{V}_1$ and $\hat{V}_2$.

The homogeneity assumption requires that on the first trial $V_2 = V_1^2$. In his analysis of the Goodnow two-armed bandit data Bush estimated these two quantities by using three-trial blocks, drawing a smooth curve through the estimates and extrapolating back to the first trial. The result was $\hat{V}_{2,1} = 0.13$ and $\hat{V}_{1,1}^2 = 0.11$, making the homogeneity assumption tenable for initial probability.

In another test of the assumption Bush and Wilson (1956) examined the number of A_1 responses in the first 10 trials of a two-choice experiment for each of 49 paradise fish. The distribution of number of choices had more spread than could be accounted for by a common probability for all subjects. The assumption was therefore rejected and a distribution of initial probabilities was used. Instead of one initial probability parameter, two were then needed, one giving the mean and the other giving the variance of the distribution of initial probabilities, whose form was assumed to be that of a beta distribution.

Even less work has been done in which variation in the learning-rate parameters is allowed. One example appears in Bush and Mosteller's (1959) analysis of the Solomon-Wynne data: the linear single-operator model was used with a distribution of α-values. In certain respects this generalization improved the agreement between model and data.

[27] This "block-moment" method was developed by Bush, as a general estimation scheme, in an unpublished manuscript, 1955.

6.4 Testing a Single Model Type

In a good deal of the work with learning models a single model is used in the analysis of a set of data. Estimates are obtained, and then several properties of the model are compared with their counterparts in the data. There is little agreement as to which properties should be examined or how many. Informal comparisons, sometimes aided by the theoretical variances of the statistics considered, are used in order to decide on the model's adequacy. Values of the parameter estimates may be used as descriptive statistics of the data.

As with any theory, a stochastic learning model can be more readily discredited than it can be accepted. Two reasons, however, lead investigators to expect and allow some degree of discrepancy between model and data. One reason is the view, held by some, that a model is intended only as an approximation to the process of interest. A second is the fact that today's experimental techniques probably do not prevent processes other than the one described by the model from affecting the data. The matter is one of degree: how good an approximation to the data do we desire and to which of their properties? And how deeply must we probe for discrepancies before we can be reasonably confident that there are no important ones? Recent work that reveals how difficult it may be to select among models (e.g., Bush, Galanter, & Luce, 1959; Sternberg, 1959b) suggests that some of our testing methods for a single model may lack power with respect to alternatives of interest to us and that we may be accepting models in error.

One finding is that the learning curve is often a poor discriminator among models. Two examples have already been illustrated in this chapter. In Figs. 1 and 2 a model with experimenter-controlled events provides an excellent description of learning curves generated by a process with a high degree of subject control. In Fig. 9 four models that differ fundamentally in the nature of their response effects produce equally good agreement with an observed learning curve.

It would be an error to conclude from these examples that the learning curve can never discriminate between models; this is far from true. Occasionally it provides us with strong negative evidence. We have already seen (Sec. 4.4) that the beta model with experimenter control cannot account for the asymptote of the learning curve in prediction experiments with $\pi \neq \frac{1}{2}$, in those experiments in which probability matching occurs. On the other hand, the linear experimenter-controlled event model can be eliminated for a T-maze experiment (Galanter & Bush, 1959) in which its theoretical asymptote is exceeded. The shape of the preasymptotic learning curve may also occasionally discriminate between models; for example, as mentioned in Sec. 6.2, the linear-operator model cannot produce a learning curve that is steep enough to describe the T-maze data of Galanter and Bush (1959) on reversal after overlearning, whereas the beta model provides a curve that is in reasonable agreement with these data. The important point, however, is that agreement between an observed learning curve and a curve produced by a model cannot, alone, give us a great deal of confidence in the model.

More surprising, perhaps, is that the distribution of error-run lengths also seems to be insensitive. In Fig. 9 it can be seen that three distinctly different models can be made to agree equally well with an observed distribution. As another example, let us consider the fourth period of a T-maze reversal experiment of Galanter and Bush (1959, Experiment III). In this experiment three trials were run each day, and by the fourth period there appeared a marked daily "recovery" effect: on the first trial of each day there was a large proportion of errors. Needless to say, this effect was not a property of the path-independent model used for the analysis. Despite the oscillating feature of the learning curve, a feature that one might think would have grave consequences for the sequential aspects of the data, the agreement between model and data, as regards the run-length distribution, was judged to be satisfactory. Again, as for the learning curve, there are examples in which the run-length distribution can discriminate. In Fig. 9 it can be seen that one of the four models cannot be forced into agreement with it. And Bush and Mosteller (1959) show that a Markov model and an "insight" model, when fitted to the Solomon-Wynne data, produce significantly fewer runs than the other models studied.

We do not wish to be limited to negative statements about the agreement between models and data, yet we have evidence that some of the usual tests are insensitive, and we have no rules to tell us when to stop testing. In comparative studies of models this situation is somewhat ameliorated: we continue making comparisons until all but one of the competing models is discredited. Another possible solution is to use over-all tests of goodness of fit. As already mentioned, these tests suffer from being powerful with respect to uninteresting alternatives: such a test, for example, might lead us to discard a model type under conditions in which only the homogeneity assumption is at fault. In contrast, the usual methods seem to suffer from low power with respect to alternatives that may be important.

One role proposed for estimates of the parameters of a model is that they can serve as descriptive statistics of the data. Such descriptive statistics are useful only if the model approximates the data well and if the values are not strongly dependent on the particular method used for their estimation. I have already discussed how an apparent lack of parameter invariance from one experiment to another may be an artifact of applying the wrong model. This has been recognized in recent suggestions that the invariance of parameter estimates from experiment to experiment be used as an additional criterion by which to test a model type.

6.5 Comparative Testing of Models

As I have already suggested, the comparative testing of several models improves in some ways on the process of testing models singly. The investigator is forced to use comparisons sensitive enough so that all but one of the models under consideration can be discredited. Attention is thereby drawn to the features that distinguish the models used, and this allows a more detailed characterization of the data than might otherwise be possible.

As an example, let us take the Bush-Mosteller (1959) comparison of eight models in the analysis of the Solomon-Wynne data. Examination of several properties eliminates all but two of the models. The remaining two are the linear-operator model (Eq. 8) and a model of the kind developed by Restle (1955). A theoretical comparison of the two models is necessary to discover a potential discriminating statistic. The "Restle model" is an example of a single-event model; under the homogeneity assumption all subjects have the same value of p_n. The linear-operator model, on the other hand, involves two subject-controlled events, and parameter estimates suggest a positive response effect. This difference should be revealed in the magnitude of the correlation between number of early and late errors (shocks). The linear-operator model calls for a positive correlation; Restle's model (together with homogeneity) calls for a zero correlation. The observed correlation is positive, and the linear-operator model is selected as the better. (This inference exemplifies the type discussed in Sec. 6.3 that depends critically on the validity of the homogeneity assumption.)

As a second example of a comparative study let us take the Bush-Wilson study (1956) of the two-choice behavior of paradise fish. On each trial the fish swam to one of two goalboxes. On 75% of the trials food was presented in one (the "favorable" goalbox); on the remaining 25% the food was presented in the other. For one group of subjects food in one goalbox was visible from the other. Two models were compared, each of which expressed a different theory about the effects of nonfood trials. The first theory suggests that on these trials the performed response is weakened, giving a model that has commonly been applied to the prediction experiment with humans:

Information Model

Event	$\mathbf{p}_{n+1}$
Favorable goalbox, food	$\alpha\mathbf{p}_n + 1 - \alpha$
Favorable goalbox, no food	$\alpha\mathbf{p}_n$
Unfavorable goalbox, food	$\alpha\mathbf{p}_n$
Unfavorable goalbox, no food	$\alpha\mathbf{p}_n + 1 - \alpha$

The second theory suggests that on nonfood trials the performed response is strengthened:

Secondary Reinforcement Model

Event	$\mathbf{p}_{n+1}$
Favorable goalbox, food	$\alpha_1\mathbf{p}_n + 1 - \alpha_1$
Favorable goalbox, no food	$\alpha_2\mathbf{p}_n + 1 - \alpha_2$
Unfavorable goalbox, food	$\alpha_1\mathbf{p}_n$
Unfavorable goalbox, no food	$\alpha_2\mathbf{p}_n$

We have already seen that the information model produces "probability-matching behavior": if this model applied, each fish would tend to divide its choices in the ratio 75:25, and no individual would consistently make one choice. In regard to the proportion of "favorable" choices averaged over subjects, probability-matching can also be produced by the secondary reinforcement model with the appropriate choice of parameters. (Again we have a case in which the average learning curve does not help us to discriminate between models.) There is a clear difference, however, if we examine properties, other than the mean, of the distribution over subjects of asymptotic choice proportions. The secondary reinforcement model implies that a particular fish will stabilize on either one choice or the other in the long run and that some will consistently choose the favorable goalbox, others the unfavorable. (If the proportion of fish that stabilized on the favorable side were 0.75, then the average learning curve would suggest probability matching.)

The observed distribution of the proportion of choices to the favorable side on the last 49 trials of the experiment is U-shaped, with most fish either at very low values or very high values, giving support to the secondary reinforcement model. One merit of this study that should be mentioned is that the decision between the two models can be made without having to estimate values for the parameters.

A third example of a comparative study that makes use of data from a two-armed bandit experiment was discussed in Sec. 4.5.

Occasionally we wish to compare models that contain different numbers of free parameters. When this occurs, a new problem is added to that of equal "fairness" of the estimation and testing procedures discussed in Sec. 6.2: the model with fewer degrees of freedom will be at a disadvantage. This difficulty can be overcome if one model is a special case of another. If so, we can apply the usual likelihood-ratio test procedure, which takes into account the difference in the number of free parameters.

Suppose, for example, that we wish to decide between the single-event beta model and the beta model with two subject-controlled events (Eq. 20) in application to an experiment such as the escape-avoidance shuttlebox. The second model is the more general and can be written

$$\text{logit}\,\mathbf{p}_n = -(a + b\mathbf{t}_n + c\mathbf{s}_n). \tag{94}$$

The first model is given by the same equation, but with $c = b$, so that

$$\text{logit}\,p_n = -[a + b(\mathbf{t}_n + \mathbf{s}_n)] = -(a + bn). \tag{95}$$

The test is equivalent to the question: are the magnitudes of the two event effects equal? It is performed by obtaining maximum-likelihood estimates of a, b, and c in Eq. 94 and of a and b in Eq. 95 and calculating the maximized likelihood for each model. These steps are straightforward for logistic models. Because Eq. 94 has an additional degree of freedom, the likelihood associated with it will generally be greater than the likelihood associated with the first model. The question of how much greater it must be in order for us to reject the first model and decide that the two events have unequal effects can be answered by making use of the (large sample) distribution of the likelihood ratio λ under the hypothesis that the equal event model holds (Wilks, 1962).[28]

In an alternative procedure a statistic that behaves monotonically with the likelihood ratio is used, and its (small sample) distribution under the hypothesis of equal event effects can be obtained analytically or, if this is difficult, from a Monte Carlo experiment based on Eq. 95. A comparable test for a generalized linear-operator model has been developed and applied by Hanania (1959) in the analysis of data from a prediction experiment. She concludes for those data that the effect of reward is significantly greater than the effect of nonreward.

6.6 Models as Baselines and Aids to Inference

Most of our discussion to this point has been oriented towards the question whether a model can be said to describe a set of data. Usually a model entails several assumptions about the learning process, and

[28] For this example $-2 \log \lambda$ is distributed as chi-square with one degree of freedom.

therefore in asking about the adequacy of a model we are testing the set of assumptions taken together. Models are also occasionally useful when a particular assumption is at stake or when a particular feature of the data is of interest. Occasionally, as with the "null hypothesis" in other problems, it is the discrepancy between model and data that is of interest, and the analysis of discrepancies may reveal effects that a model-free analysis might not disclose. A few examples may make these ideas clear.

In Sec. 6.5 the choice between the two models of Eqs. 94 and 95 was equivalent to the question whether the effects of reward and nonreward are different. We might, on the other hand, start with this question and perform the same analysis, not being concerned with whether either of the models fitted especially well but simply with whether one (Eq. 94) was significantly preferable to the other (Eq. 95).

With the same kind of question in mind we could estimate reward and nonreward parameters for some model and compare their values. One difficulty with this procedure is that unless we know the sampling properties of the estimator it is difficult to interpret such results. A second difficulty is that if a model is not in accord with the data the estimates may depend strongly on the method used to obtain them. An example of this dependence is shown in Table 5; different estimates lead to contradictory answers to the question which of the two events has the larger effect. That the parameters in a model properly represent the event effects in a set of data to which the model is fitted may be conditional on the validity of many of the assumptions embodied in the model. What is needed for this type of question is a model-free test—one that makes as few assumptions as possible. Applications of such tests are illustrated in Sec. 6.7.

If a model agrees with a number of the properties of a set of data, then the discrepancies that do appear may be instructive and may be useful guides in refining the model. One example is provided by the analysis of free-recall verbal learning by a model developed by Miller and McGill (1952). The model is intended to apply to an experiment in which, on each trial, a randomly arranged list of words is presented and the subject is asked to recall as many of the words as he can. The model assumes that the process that governs whether or not a particular word is recalled on a trial is independent of the process for any of the other words. The model is remarkably successful, but one discrepancy appears when the estimated recall probability after v recalls, $\hat{p}_v$, is examined as a function of v and compared to the theoretical mean curve (Bush & Mosteller, 1955, p. 234). The observed proportions oscillate about the mean more than the model allows. This suggests the hypothesis that a subject learns words in clusters rather than independently and that either all or none of the words in a cluster tend to be recalled on a trial.

A second example of the baseline use of a model is provided by Sternberg's (1959b) analysis of the Goodnow data by means of the linear one-trial perseveration model (Eq. 29). Several properties of the data are described adequately by the model, but in at least one respect it is deficient. The model implies that for all trials $n \geqslant 2$, $\Pr\{\mathbf{x}_n = 1 \mid \mathbf{x}_{n-1} = 1\} - \Pr\{\mathbf{x}_n = 1 \mid \mathbf{x}_{n-1} = 0\} = \beta$, a constant. What is observed is that the difference between the estimates of these conditional probabilities decreases somewhat during the course of learning. It was inferred from this finding and from other properties of the data that the tendency to perseverate may change as a function of experience. This example may serve to caution us, however, and to indicate that the hypotheses suggested by a baseline analysis may be only tentative. The inference mentioned depends strongly on the use of a linear model. If the logistic perseveration model (Eq. 32) is used instead, the observed decrease in the difference between conditional probabilities is produced automatically, without requiring changes in parameter values during the course of learning.

A final use of models as aids to inference is in the study of methods of data analysis. In an effort to reveal some feature of data the investigator may define a statistic whose value is thought to reflect the feature. Because the relations between the behavior of the statistic and the feature of interest may not be known, errors of inference may occur. These errors are sometimes known as *artifacts*: the critical property of the statistic arises from its definition rather than from the feature of interest in the data.

A simple example of an artifact has already been mentioned in Sec. 6.3. If individuals differ in their p_1-values and if the existence of response effects is assessed by means of the correlation over subjects between early and late errors, then a positive response effect may be inferred when there is none.

A second example is the evaluation of response-repetition tendencies. If subjects differ in their p-values on trial $n - 1$ and if the existence of a perseverative tendency is assessed by comparing the proportions that correspond to $\Pr\{\mathbf{x}_n = 1 \mid \mathbf{x}_{n-1} = 1\}$ and $\Pr\{\mathbf{x}_n = 1 \mid \mathbf{x}_{n-1} = 0\}$, then a perseverative tendency may be inferred where none is present. This occurs because the samples used to determine the two proportions are not randomly selected: they are chosen on the basis of the response on trial $n - 1$. The subjects used to determine $\Pr\{\mathbf{x}_n = 1 \mid \mathbf{x}_{n-1} = 1\}$ tend to have higher p-values on trial $n - 1$ (and consequently on trial n) than the subjects in the other sample. Selective sampling effects of this kind have been discussed by Anderson (1960, p. 85) and Anderson and Grant (1958).

The question in both of these examples is how extreme a value of the statistic must be observed in order to make the inference valid. In order to answer this question, the behavior of the statistic under the hypothesis of no effect must be known. Such behavior can be studied by applying the method of analysis to data for which the underlying process is known, namely Monte Carlo sequences generated by a model.

A more complicated problem to which this type of study has been applied is the criterion-reference learning curve (Hayes & Pereboom, 1959). This is a method of data analysis in which an average learning curve is constructed by pooling scores for different subjects on trials that are specified by a performance criterion rather than by their ordinal position. Underwood (1957) developed a method of this kind in an effort to detect a cyclical component in serial verbal learning. The result of the analysis was a learning curve with a distinct cyclical component. Whether the inference of cyclical changes in p_n is warranted depends on the result of the analysis when it is applied to a process in which p_n increases monotonically. Hayes and Pereboom apply the method to Monte Carlo sequences generated by such a process and obtain a cyclical curve. We conclude that to infer cyclical changes in p_n the magnitude of the cycles in the criterion-reference curve must be carefully examined; the existence of cycles is not sufficient.

6.7 Testing Model Assumptions in Isolation

A particular learning model can be thought of as the embodiment of several assumptions about the learning process. Testing the model, then, is equivalent to testing all of these assumptions jointly. If the model fails, we are still left with the question of the assumptions that are at fault. Light may be shed on this question by the comparative method and the detailed analysis of discrepancies discussed in the last two sections. But a preferable technique is to test particular assumptions in as much isolation from the others as possible. Examples of several techniques are described in this section.

To illustrate the equivalence of a model to several assumptions, each of which might be tested separately, let us consider the analysis of the data from an escape-avoidance shuttlebox experiment by means of the linear-operator model (Eqs. 7 to 11 in Sec. 2.4). Suppose that an estimation procedure yields an avoidance operator that is more potent than the escape operator ($\alpha_1 < \alpha_2$). Some of the assumptions involved in this model are listed:

1. The effect of an event on $\mathbf{p}_n = \Pr\{\text{escape}\}$ is manifested completely on the next trial.
2. Conditional on the value of $\mathbf{p}_n$, the effect of an event is independent of the previous sequence of events.

3. Conditional on the value of $\mathbf{p}_n$, the effect of avoidance (reward) on $\mathbf{p}_n$ is greater than the effect of escape (nonreward).
4. The reduction in $\mathbf{p}_n$ caused by an event is proportional to the value of $\mathbf{p}_n$.
5. The proportional reduction is constant throughout learning; therefore the change in $\mathbf{p}_n$ induced by avoidance, for example, decreases during the course of learning.
6. The value of $\mathbf{p}_n$ for a subject depends only on the number of avoidances and escapes on the first $n - 1$ trials and not on the order in which they occurred.
7. All subjects have the same values of p_1, α_1, and α_2.

Let us consider the third assumption listed: avoidance has more effect than escape. It has already been indicated that to test this assumption simply by examining the parameter estimates for a model may be misleading. For the Solomon-Wynne data, Bush and Mosteller (1955) have found the values $\hat{\alpha}_1 = 0.80$, $\hat{\alpha}_2 = 0.92$, and $p_1 = 1.00$ for the linear model, confirming the third assumption. In Table 5 it is shown that estimates for the beta model may or may not confirm the assumption, depending on the method of estimation used.

We wish to test this assumption without the encumbrance of all the others. One step in this direction is to apply a more general model, in which at least some of the constraints are discarded. Hanania's work (1959) provides an example of this approach, in which she uses a linear, commutative model, but with trial-dependent operators, so that it is quasi-independent of path:

$$\mathbf{p}_{n+1} = \begin{cases} w_{n+1}\mathbf{p}_n & \text{if} \quad \mathbf{x}_n = 1 \qquad \text{(nonreward)} \\ \theta w_{n+1}\mathbf{p}_n & \text{if} \quad \mathbf{x}_n = 0 \qquad \text{(reward).} \end{cases} \tag{96}$$

Although θ is assumed to be constant, w_n may take on a different value for each trial number. The explicit formula, after a redefinition of the parameters, is

$$\mathbf{p}_n = \theta^{\mathbf{s}_n} u_n, \tag{97}$$

where

$$\mathbf{s}_n = \sum_{j=1}^{n-1} \mathbf{x}_j.$$

When the w_n (or the u_n) are all equal, this formulation reduces to the Bush-Mosteller model. The relative effect of reward and nonreward is reflected by the value of θ, for which Hanania develops statistical tests.

This method is an improvement, but a good many assumptions are still needed. A more direct, if less powerful, method that requires fewer

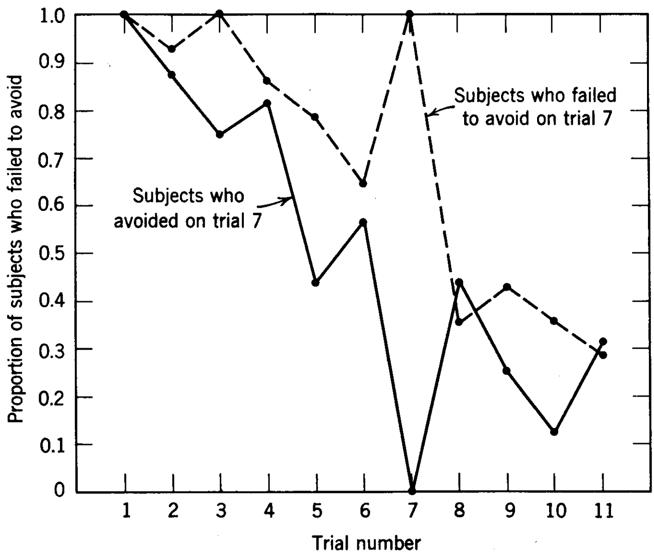

Fig. 15. Performance of two groups of dogs in the Solomon-Wynne avoidance-learning experiment, selected on the basis of their response on trial 7. The broken curve represents the performance of the 14 dogs who failed to avoid on trial 7. The solid curve represents the performance of the 16 dogs who avoided on trial 7.

assumptions is illustrated in Figs. 15 and 16. A trial is selected on which each response is performed by about half the subjects. The subjects are divided into two groups, according to the response they perform, and learning curves are plotted separately for each group. In Fig. 15 this analysis is performed on the Solomon-Wynne shuttlebox data. Subjects are selected on the basis of their seventh response ($\mathbf{x}_7$). It can be seen that animals who escape on trial 7 have a relatively high escape probability on the preceding trials. There is a positive correlation between escape on trial 7 and the number of escapes on earlier trials.

One assumption is needed in order to make the desired inference: the absence of individual differences in parameter values. Individual differences alone, with no positive response effect, could produce a result of this kind; slower learners would tend to fall into the escape group on trial seven. On the other hand, if we assume no individual differences, the result strongly suggests that there is a positive response effect, confirming the third assumption. This result also casts doubt on the validity of the beta model for these data. As shown by Table 5, estimates for that model are either in conflict with the third assumption or require an absurdly low value for the initial escape probability.

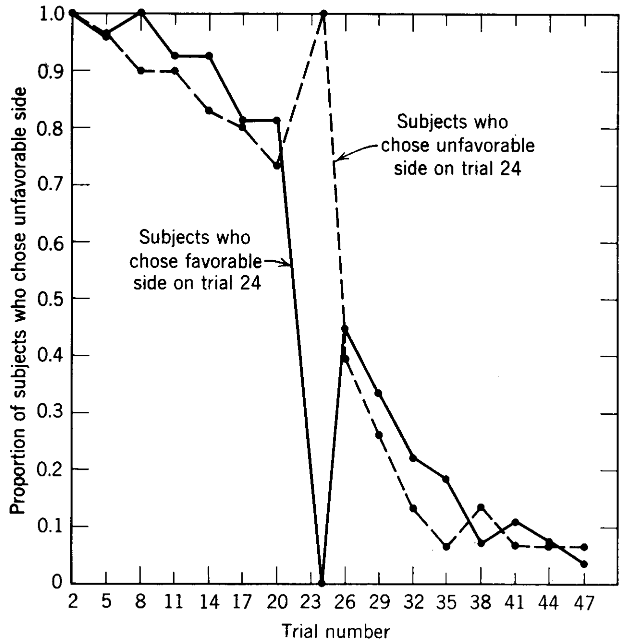

Fig. 16. Performance of two groups of rats in the Galanter-Bush experiment on reversal after overlearning, selected on the basis of their response on trial 24. The broken curve represents the performance of the 10 rats that chose the unfavorable side on trial 24. The solid curve represents the performance of the nine rats that chose the favorable side on trial 24. Except for the point at trial 24, points represent average proportions for blocks of three trials.

Figure 15 also illustrates the errors in sampling that can occur if a subject-controlled event is used as a criterion for selection. This type of error was touched on briefly in Sec. 6.6. It is exemplified by an alternative method that we might have used for assessing the response effect. In this method we would compare the performance of the two subgroups on trials *after* the seventh. Escape on trial 7 is associated with a high escape probability on future trials. It would be an error, however, to infer from this fact alone that the seventh response caused the observed differences; the subgroups are far from identical on trials before the seventh. The method of selecting subjects on the basis of an outcome and examining their future behavior to assess the effect of the outcome must be used with caution. It can be applied when either the occurrence of the outcome is controlled by the experimenter and is independent of the state of the subject or when it can be demonstrated that the subgroups do not differ before the outcome. For an example of this type of demonstration in a model-free analysis of the effects of subject-controlled events, see Sheffield (1948).

Figure 16 gives the result of dividing subjects on the basis of the twenty-fourth trial of an experiment on reversal after overlearning (Galanter & Bush, 1959). These data are mentioned in Sec. 4.5 (Table 3), in which the coefficient of variation of $\mathbf{u}_1$ is shown to be unusually small, and in Sec. 6.4, in which the inability of a linear-operator model to fit the learning curve is discussed. The results of Fig. 16 are the reverse of those in Fig. 15: animals that make errors on trial 24 tend to make *fewer* errors on preceding and following trials than those that give the correct response on trial 24. This negative relationship cannot be attributed to failure of the homogeneity assumption; individual differences in parameter values would tend to produce the opposite effect. Therefore we can conclude, without having to call on even the homogeneity assumption, that in this experiment there is a negative response effect.

This result gives additional information to help choose between the linear (Galanter & Bush, 1959) and beta (Bush, Galanter, & Luce, 1959) models for the data on reversal after overlearning. Estimation for the linear model suggested a positive response effect, which had to be increased if the learning curve was to be even roughly approximated. Because of its positive response effect, the model produced a value for Var ($\mathbf{u}_1$) that was far too large. For the beta model, on the other hand, estimation produces results in agreement with the analysis of Fig. 16 and the value for Var ($\mathbf{u}_1$) is slightly too small, a result consistent with the existence of small individual differences that are not incorporated in the model. The conclusion seems clear that, of the two, the beta model is to be preferred.

We are in the embarrassing position of having discredited both the linear and beta models, each in a different experiment. Unfortunately we cannot conclude that one applies to rats and the other to dogs; evidence similar to that presented clearly supports the linear model for a T-maze experiment on reversal without overlearning (Galanter & Bush, 1959, Experiment III, Period 2).

It has not been mentioned that when outcomes are independent of the state of the subject a model-free analysis of their effects can be performed. As an example, let us consider a T-maze experiment with a 75:25 reward schedule. The favorable (75%) side is baited with probability 0.75, independent of the rat's response. Suppose that we wish to examine the effect of reward on response probability. Rats that choose the favorable side on a selected trial are divided into those that are rewarded on that

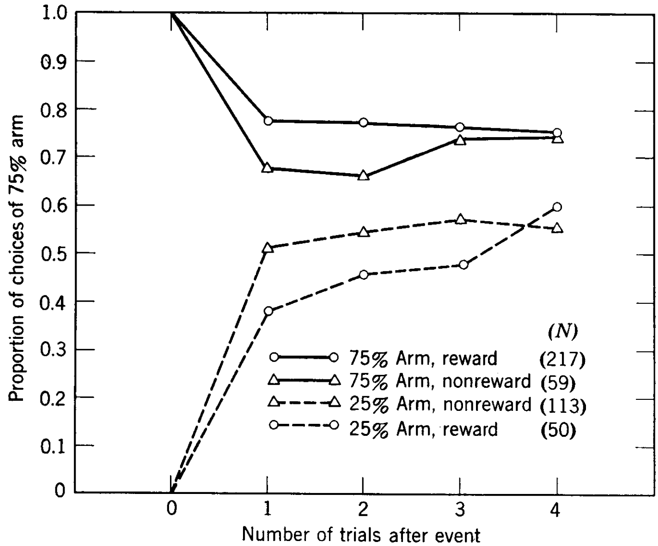

Fig. 17. Performance after four different events in the Weinstock 75:25 T-maze experiment, as a function of the number of trials after the event. The solid (broken) curves represent performance after choice of the 75% (25%) arm. Circles (triangles) represent performance after reward (nonreward). The number of observations used for each curve is indicated.

trial and those that are not. The behavior of the subgroups can be compared on future trials, and any differences can be attributed to the reward effect. To enlarge the sample size, averaging procedures can be employed. The same comparison can be made among the rats that choose the unfavorable side on the selected trial. Results of an analysis of this kind are shown in Fig. 17. The data are the first 20 trials of a 75:25 T-maze experiment conducted by Weinstock (1955). On each trial, n, the rats were divided into four subgroups on the basis of the response and outcome, and the number of choices during each of the next four trials was tabulated for each subgroup. These choice frequencies for all values of n, $1 \leqslant n \leqslant 20$, were added and the proportions given in the figure were obtained.

The results are in comforting agreement with the assumptions in several of our models. After reward, the performed response has a higher probability of occurrence than after nonreward, and this is true for both responses. In keeping with the first assumption mentioned in this section, there is no evidence of a delayed effect of the reinforcing event: the hypothesis that its full effect is manifested on the next trial cannot be rejected. On the contrary, there is a tendency for the effect to be reduced as trials proceed; this last finding would be expected if the effect of reward were less than that of nonreward.

A method that has been used to study a "negative-recency effect" in the binary prediction experiment (e.g., Jarvik, 1951; Nicks, 1959; Edwards, 1961) provides us with a final example of the testing of model assumptions in isolation. The assumption in question is that the direction in which $\mathbf{p}_n$ is changed by an event is the same, regardless of the value of $\mathbf{p}_n$ and the sequence of prior events.[29]

Consider the prediction experiment in terms of four experimenter-subject controlled events, as presented in Table 1. On a trial on which O_1 occurs, either A_1 or A_2 may occur, so that two events are possible. Moreover, which of these two events occurs on a trial depends on the state of the subject. Separate analysis of their effects is therefore difficult, as explained earlier in this section. Fortunately, we are willing to assume that both events have effects that are in the same direction: if either (A_1O_1) or (A_2O_1) occurs on trial n, then $\mathbf{p}_{n+1} > \mathbf{p}_n$. This assumption allows us to perform the test by averaging over all subjects that experienced O_1 on trial n, regardless of their response, and thus examining the average effect of O_1 on the average p-value. This is equivalent to examining an average of the effects of the two events (A_1O_1) and (A_2O_1), and therefore the test lacks power: if only one of the events violates the assumption in question, the test can fail to detect the violation.[30]

The variable of interest is the length of the immediately preceding tuple of O_1's. Does the direction of the effect of O_1 on $\Pr\{A_1\}$ depend on this length? The method involves averaging the proportion of A_1 responses on trials after all j-tuples of O_1's in the outcome sequence for various values of j and considering the average proportion as a function of j. The results of such an analysis, for O_2's as well as O_1's (Nicks, 1959) are given in Fig. 18. The data are from a 380-trial, 67:33 prediction experiment, and the analysis is performed separately for each quarter of the outcome sequence. Under the assumption in question, and in the light of the fact that a 1-tuple of O_1's (an O_1 that is preceded by an O_2) markedly increases $\Pr\{A_1\}$, all curves in the figure should have slopes that are uniformly nonnegative. Such is not the case, and the results lead us to reject the assumption.[31] If we assume in this experiment that the effects

[29] As mentioned in Sec. 2.3, simple models exist in which the direction of the effect of an event depends on the value of $\mathbf{p}_n$. Such models have seldom been applied, however.

[30] If we assume that the experiment consists of two experimenter-controlled events, then this criticism does not apply; however, this is precisely the sort of additional assumption that we do not wish to make.

[31] Using this type of analysis of performance in a prediction experiment after 520 trials of practice, Edwards (1961) obtained results favorable to the assumption.

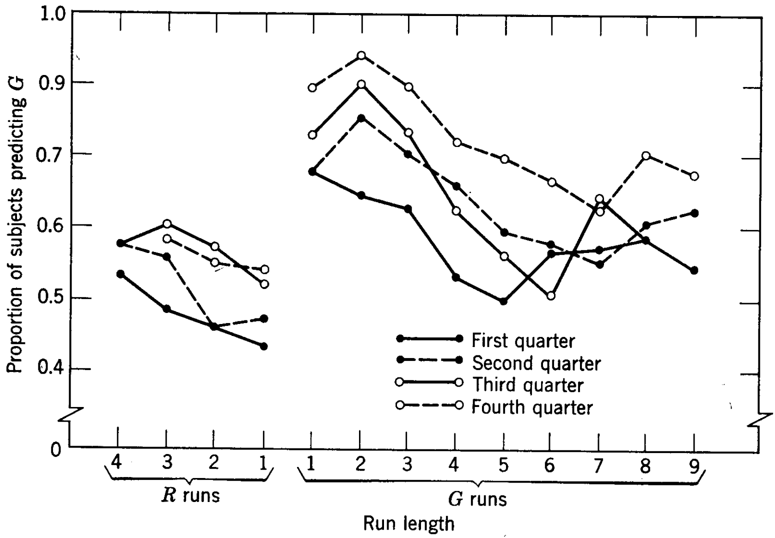

Fig. 18. Proportion of subjects predicting "green" immediately after outcome runs of various lengths of green (G) and red (R) lights in Nick's 67:33 binary prediction experiment. Separate curves are presented for the four quarters of the 380-trial sequence. After Nicks, 1959, Fig. 3.

of events sharing the same outcome are in the same direction, then the direction of the effect of an event appears to depend in a complicated way on the prior sequence of events.

7. CONCLUSION

Implicit in these pages are two alternative views of the place of stochastic models in the study of learning. The first view is that a model furnishes a sophisticated statistical method for drawing inferences about the effects of individual trial events in a learning experiment or for providing descriptive statistics of the data. The method has to be sophisticated because the problem is difficult: the time-series in question is usually nonstationary[32] and involves only a small number of observations per subject; if the observations are discrete rather than continuous, each one gives us little information. The use of a model, then, can be thought of as a method of

[32] In an experiment in which there is no over-all trend (i.e., $V_{1,n}$ and $V_{2,n}$ are constant), statistical methods for the analysis of stationary time series can be used (see, e.g., Hannan, 1960, and Cox, 1958). Interesting models for such experiments have been developed and applied by Cane (1959, 1961).

combining data—of averaging—in a process with trend. The model is not expected to fit exactly even the most refined experiment; it is simply a tool.

The second view is that a model, or a family of models, is a mathematical representation of a theory about the learning process. In this case our focus shifts from features of a particular set of data to the extent to which a model describes those data and to the variety of experiments the model can describe. We become more concerned with the assumptions that give rise to the model and with crucial experiments or discriminating statistics to use in its evaluation. We attempt to refine our experiments so that the process purportedly described by the model can be observed in its pure form.

Whether we are concerned more with describing particular aspects of the process or more with evaluating an over-all theory, many of the fundamental questions that arise about learning can be answered only by the use of explicit models. The use of models, however, does not automatically produce easy answers.

References

Anderson, N. H. An analysis of sequential dependencies. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 248–264.

Anderson, N. H. Effect of first-order conditional probability in a two-choice learning situation. *J. exp. Psychol.*, 1960, **59,** 73–93.

Anderson, N. H., & Grant, D. A. Correction and reanalysis. *J. exp. Psychol.*, 1958, **56,** 453–454.

Anscombe, F. J. On estimating binomial response relations. *Biometrika*, 1956, **43,** 461–464.

Audley, R. J. A stochastic description of the learning behavior of an individual subject. *Quart. J. exp. Psychol.*, 1957, **9,** 12–20.

Audley, R. J., & Jonckheere, A. R. The statistical analysis of the learning process. *Brit. J. Stat. Psychol.*, 1956, **9,** 87–94.

Bailey, N. T. J. Some problems in the statistical analysis of epidemic data. *J. Roy. Stat. Soc. (B)*, 1955, **17,** 35–57.

Barucha-Reid, A. T. *Elements of the theory of Markov processes and their applications.* New York: McGraw-Hill, 1960.

Behrend, E. R., & Bitterman, M. E. Probability-matching in the fish. *Amer. J. Psychol.*, 1961, **74,** 542–551.

Berkson, J. Maximum likelihood and minimum χ^2 estimates of the logistic function. *J. Amer. Stat. Assoc.*, 1955, **50,** 130–162.

Berkson, J. Tables for the maximum-likelihood estimation of the logistic function. *Biometrics*, 1957, **13,** 28–34.

Bush, R. R. *A block-moment method of estimating parameters in learning models.* Unpublished manuscript, 1955.

Bush, R. R. Sequential properties of linear models. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 215–227.

Bush, R. R. Some properties of Luce's beta model for learning. In K. J. Arrow, S. Karlin, & P. Suppes (Eds.), *Mathematical methods in the social sciences, 1959.* Stanford: Stanford Univer. Press, 1960. Pp. 254–264. (a)

Bush, R. R. A survey of mathematical learning theory. In R. D. Luce (Ed.), *Developments in mathematical psychology.* Glencoe, Ill. Free Press, 1960. Pp. 123–170. (b)

Bush, R. R., & Estes, W. K. (Eds.). *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959.

Bush, R. R., Galanter, E., & Luce, R. D. Tests of the "beta model." In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 381–399.

Bush, R. R., & Mosteller, F. A mathematical model for simple learning. *Psychol. Rev.*, 1951, **58**, 313–323.

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.

Bush, R. R., & Mosteller, F. A comparison of eight models. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 293–307.

Bush, R. R., Mosteller, F., & Thompson, G. L. A formal structure for multiple-choice situations. In R. M. Thrall, C. H. Coombs, & R. L. Davis (Eds.), *Decision processes.* New York: Wiley, 1954. Pp. 99–126.

Bush, R. R., & Sternberg, S. H. A single-operator model. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 204–214.

Bush, R. R., & Wilson, T. R. Two-choice behavior of paradise fish. *J. exp. Psychol.*, 1956, **51**, 315–322.

Cane, Violet R. Behaviour sequences as semi-Markov chains. *J. Roy. Stat. Soc. (B)*, 1959, **21**, 36–58.

Cane, Violet R. Review of R. D. Luce, Individual Choice Behavior. *J. Roy. Stat. Soc. (A)*, 1960, **22**, 486–488.

Cane, Violet R. Some ways of describing behavior. In W. H. Thorpe & O. L. Zangwill (Eds.), *Current problems in animal behavior.* Cambridge: Cambridge Univer. Press, 1961. Pp. 361–388.

Cox, D. R. The regression analysis of binary sequences. *J. Roy. Stat. Soc. (B)*, 1958, **20**, 215–232.

Edwards, W. Reward probability, amount, and information as determiners of sequential two-alternative decisions. *J. exp. Psychol.*, 1956, **52**, 177–188.

Edwards, W. Probability learning in 1000 trials. *J. exp. Psychol.*, 1961, **62**, 385–394.

Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.

Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: a study of a science. Vol. II. General systematic formulations, learning, and special processes.* New York: McGraw-Hill, 1959. Pp. 380–491.

Estes, W. K. Learning theory. *Ann. Rev. Psychol.*, 1962, **13**, 107–144.

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

Estes, W. K., & Suppes, P. Foundations of linear models. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 137–179.

Feldman, J., & Newell, A. A note on a class of probability matching models. *Psychometrika*, 1961, **26,** 333–337.

Feller, W. *An introduction to probability theory and its applications*, second edition. New York: Wiley, 1957.

Galanter, E., & Bush, R. R. Some T-maze experiments. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 265–289.

Goldberg, S. *Introduction to difference equations.* New York: Wiley, 1958.

Gulliksen, H. A rational equation of the learning curve based on Thorndike's law of effect. *J. gen. Psychol.*, 1934, **11,** 395–434.

Hanania, Mary I. A generalization of the Bush-Mosteller model with some significance tests. *Psychometrika*, 1959, **24,** 53–68.

Hannan, E. J. *Time series analysis.* London: Methuen, 1960.

Hayes, K. J., & Pereboom, A. C. Artifacts in criterion-reference learning curves. *Psychol. Rev.*, 1959, **66,** 23–26.

Hodges, J. L., Jr. Fitting the logistic by maximum likelihood. *Biometrics*, 1958, **14,** 453–461.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century Crofts, 1943.

Hull, C. L. *A behavior system.* New Haven: Yale Univer. Press, 1952.

Irwin, F. W. On desire, aversion, and the affective zero. *Psychol. Rev.*, 1961, **68,** 293–300.

Jarvik, M. E. Probability learning and a negative recency effect in the serial anticipation of alternating symbols. *J. exp. Psychol.*, 1951, **41,** 291–297.

Kanal, L. Analysis of some stochastic processes arising from a learning model. Unpublished doctoral thesis, Univer. Penn., 1960.

Kanal, L. A functional equation analysis of two learning models. *Psychometrika*, 1962, **27,** 89–104. (a)

Kanal, L. The asymptotic distribution for the two-absorbing-barrier beta model. *Psychometrika*, 1962, **27,** 105–109. (b)

Karlin, S. Some random walks arising in learning models. *Pacific J. Math.*, 1953, **3,** 725–756.

Kendall, D. G. Stochastic processes and population growth. *J. Roy. Stat. Soc. (B)*, 1949, **11,** 230–264.

Lamperti, J., & Suppes, P. Some asymptotic properties of Luce's beta learning model. *Psychometrika*, 1960, **25,** 233–241.

Logan, F. A. *Incentive.* New Haven: Yale Univer. Press, 1960.

Luce, R. D. *Individual choice behavior.* New York: Wiley, 1959.

Luce, R. D. Some one-parameter families of commutative learning operators. In R. C. Atkinson (Ed.), *Studies in mathematical psychology, 1963.* Stanford: Stanford Univer. Press, 1963, in press.

Miller, G. A., & McGill, W. J. A statistical description of verbal learning. *Psychometrika*, 1952, **17,** 369–396.

Mosteller, F. Stochastic learning models. In *Proc. Third Berkeley Symp. Math. Stat. Probability*, 1955, **5.** Pp. 151–167.

Mosteller, F., & Tatsuoka, M. Ultimate choice between two attractive goals: predictions from a model. *Psychometrika*, 1960, **25,** 1–17.

Nicks, D. C. Prediction of sequential two-choice decisions from event runs. *J. exp. Psychol.*, 1959, **57,** 105–114.

Restle, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, **62,** 11–19.

Restle, F. *Psychology of judgment and choice.* New York: Wiley, 1961.

Sheffield, F. D. Avoidance training and the contiguity principle. *J. comp. physiol. Psychol.*, 1948, **41**, 165–177.

Sternberg, S. H. A path-dependent linear model. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press, 1959. Pp. 308–339. (a)

Sternberg, S. H. Application of four models to sequential dependence in human learning. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press: 1959. Pp. 340–380. (b)

Tatsuoka, M., & Mosteller, F. A commuting-operator model. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press, 1959. Pp. 228–247.

Thurstone, L. L. The learning function. *J. gen. Psychol.*, 1930, **3**, 469–491.

Underwood, B. J. A graphical description of rote learning. *Psychol. Rev.*, 1957, **64**, 119–122.

Wald, A. Asymptotic properties of the maximum-likelihood estimate of an unknown parameter of a discrete stochastic process. *Ann. math. Stat.*, 1948, **19**, 40–46.

Weinstock, S. Unpublished experiment, 1955.

Wilks, S. S. *Mathematical statistics*. New York: Wiley, 1962.