PSYCHOMETRIKA—VOL. 31, NO. 1
MARCH, 1966

PROBABILITY MATCHING*

M. FRANK NORMAN AND JOHN I. YELLOTT, JR.

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
STANFORD, CALIFORNIA

The class of symmetric path-independent models with experimenter-controlled events is considered in conjunction with two-choice probability learning experiments. Various refinements of the notion of probability matching are defined, and the incidence of these properties within this class is studied. It is shown that the linear models are the only models of this class that predict a certain phenomenon that we call stationary probability matching. It is also shown that models within this class that possess an additional property called marginal constancy predict approximate probability matching.

In this paper we are concerned with a class of models for two-choice probability learning (binary prediction) experiments. On each trial in such experiments the subject predicts which of two outcomes will occur. We denote the outcomes O_1 and O_2 and the corresponding predictions A_1 and A_2. Similarly $O_{i,n}$ and $A_{i,n}$ denote, respectively, the occurrence of O_i and A_i on trial n. In the general case of *contingent reinforcement* $O_{i,n}$ follows $A_{i,n}$ with probability π_i determined by the experimenter. The special case of *noncontingent reinforcement*, in which the outcome probabilities do not depend on the subject's response (i.e., $\pi_1 = 1 - \pi_2$), has received most experimental attention. Experiments using noncontingent reinforcement have generally yielded asymptotic proportions of A_1 responses consistent with

$$\lim_{n\to\infty} P(A_{1,n}) = \pi, \tag{1}$$

where π, the common value of π_1 and $1 - \pi_2$, is the probability of $O_{1,n}$. There is also some evidence that the corresponding phenomenon

$$\lim_{n\to\infty} P(A_{1,n}) = \lim_{n\to\infty} P(O_{1,n}) \tag{2}$$

obtains under contingent reinforcement conditions. We will refer to condition (1), and the more general condition (2), as *simple probability matching*. (Estes'

*This research grew out of questions posed by William K. Estes. We are also indebted to Professor Estes for his encouragement and assistance at all stages of this research. During the course of this research J. I. Y. received support from the U. S. Public Health Service (N. I. M. H.). M. F. N.'s present address is the University of Pennsylvania. J. I. Y.'s present address is the University of Minnesota.

recent review [2] provides a discussion of experimental variations that lead to results consistent with (1) and (2).)

Let $\mathbf{p}_{1,n}$ be the random variable that gives the probability of $A_{1,n}$ for a single subject (so that $P(A_{1,n}) = E[\mathbf{p}_{1,n}]$). It is well known that for each $0 < \theta \leq 1$ the linear model

$$\mathbf{p}_{1,n+1} = \begin{cases} \mathbf{p}_{1,n} + \theta(1 - \mathbf{p}_{1,n}) & \text{if } O_{1,n} \\ \mathbf{p}_{1,n} - \theta \mathbf{p}_{1,n} & \text{if } O_{2,n} \end{cases} \tag{3}$$

predicts simple probability matching [1, 3], and in addition these models have proved useful in accounting for fine grain sequential effects in binary prediction experiments [6]. The linear models are members of the general class of "symmetric path-independent learning models with experimenter-controlled events" which will be defined and discussed in the next section. Each model of this class is completely determined by the form of two operators on response probabilities: the operator associated with O_1 and the operator associated with O_2. We will refer to models of this class as "two-operator" models. The point of departure of the present research was an attempt to determine whether the linear models are the *only* two-operator models that predict simple probability matching. We will present some rather trivial examples below to show that they are not. However, by broadening the scope of our inquiry to take account of other predictions of the linear models we have obtained a number of positive results. These give much information about the constraints imposed on the form of a two-operator model by the prediction of various probability matching phenomena. For instance in the case of the linear models the convergence to π represented by (1) is monotonic. In particular, if $P(A_{1,1}) = \pi$ then $P(A_{1,n}) = \pi$ for all n. Theorem 1 below shows that this property, which we call *stationary probability matching*, is possessed by no two-operator models other than the linear models. In Theorem 2 a closely related but somewhat weaker property is shown to characterize linear models within the subclass of two-operator models satisfying some additional restrictions. In the latter part of the paper we show that another property of the linear models, *marginal constancy*, permits an approach to the question of which two-operator models predict approximate probability matching. A learning model is said to predict marginal constancy if, in the *double reward situation* where $\pi_1 = \pi_2 = 1$ (i.e., the outcome always agrees with the response) the marginal response probability $P(A_{1,n})$ does not depend on n. In Theorem 3 an explicit bound is obtained for

$$|\lim_{n\to\infty} P(A_{1,n}) - \lim_{n\to\infty} P(O_{1,n})|$$

for marginally constant two-operator models satisfying a few additional conditions. The form of this bound makes it clear that models of this type that are "close to" a linear model in the form of their operators predict asymptotic response probabilities close to probability matching. Theorem 4 is concerned with families of two-operator models indexed by a "learning rate" parameter like the θ of the linear models. It is shown that as the learning rate becomes small probability matching and marginal constancy become in a certain sense equivalent.

A Class of Models

The class of models discussed here has also been extensively investigated for the double reward situation by Rose [10]. A more thorough discussion of many of our assumptions is given by Sternberg [11].

We now delimit precisely the class of models to which our discussion pertains by making certain general assumptions about the learning process in the binary prediction experiment. We do not regard these assumptions as necessary conditions for models for probability learning, and we will indicate below how certain well-known models that have been proposed for this situation fail to satisfy them. We do believe, however, that each of these assumptions is plausible enough to be entertained, and that the class of models satisfying all of these assumptions is sufficiently broad that our results may be relevant to a variety of theoretical interpretations of the learning process in binary prediction experiments.

Any model for the binary prediction experiment, regardless of what unobservable processes it postulates, determines the (conditional) probabilities

$$\mathbf{p}_{i,n} = P(A_{i,n} \mid O_{i_{n-1},n-1}A_{i_{n-1},n-1} \cdots O_{i_1,1}A_{i_1,1}\mathbf{p}_{i,1}) \tag{4}$$

of response A_i on trial n for a subject with the (observable) experimental history $O_{i_{n-1},n-1} \cdots A_{i_1,1}$ over the preceding trials and with probability $\mathbf{p}_{i,1}$ of A_i on trial 1. These are the basic "response probabilities" in terms of which our assumptions are formulated. This approach is suggested by the treatment of linear models given by Estes and Suppes [3]. In our notation only the bold-face type reminds the reader that $\mathbf{p}_{i,n}$ depends on $O_{i_{n-1},n-1} \cdots A_{i_1,1}$ and $\mathbf{p}_{i,1}$. Since models for binary prediction are typically applied to groups of subjects with different probabilities of $A_{i,1}$, $\mathbf{p}_{i,1}$ must in general be considered a random variable.

Our first assumption is that the learning process is *independent of path* in the sense that $\mathbf{p}_{i,n+1}$ depends on responses and outcomes before trial n and on $\mathbf{p}_{i,1}$ only through $\mathbf{p}_{i,n}$. Thus $\mathbf{p}_{i,n+1}$ is a function only of $\mathbf{p}_{i,n}$ and the response and outcome on trial n. We further assume that this function does not depend on n. Thus

A1 *For each j, k ($j = 1, 2$; $k = 1, 2$) there is a function f_{kj} such that if $A_{j,n}$ and $O_{k,n}$ then $\mathbf{p}_{1,n+1} = f_{kj}(\mathbf{p}_{1,n})$ or, equivalently,*

$$\mathbf{p}_{2,n+1} = 1 - f_{kj}(\mathbf{p}_{1,n}).$$

A great many of the models proposed to date have this property (see [1, 9, 11]), but some rather interesting ones do not. For instance, of the "weighted outcomes" models considered by Feldman and Newell [5] A1 holds only for the linear models, and it seems extremely unlikely that stimulus sampling models in general satisfy A1. (However Estes and Suppes [4] have proved a general approximation theorem which implies that there exists a sequence of N-element stimulus sampling models that satisfies (3) in the limit as $N \to \infty$.)

Next we assume that the occurrence of outcome O_i on trial n has the same effect on the learning process regardless of which response occurs on trial n. More precisely

A2 $$f_{11}(p) = f_{12}(p) \quad \textit{and} \quad f_{22}(p) = f_{21}(p) \quad \textit{for all} \quad 0 \leq p \leq 1.$$

We will use the notation

$$f(p) = f_{11}(p) \quad \text{and} \quad g(p) = f_{21}(p)$$

throughout the rest of the paper. Certainly it is to be expected that the occurrence of O_i on trial n will increase or at least not decrease the probability of A_i on subsequent trials regardless of which outcome the subject predicted on trial n. We know of no completely convincing argument that the effect should be exactly the same in the two cases, but this assumption seems as plausible as any other at this time, and it is quite convenient mathematically as it reduces the number of possible effects of a learning trial to two: the effect of O_i for $i = 1, 2$. In the terminology of Bush and Mosteller [1], A2 is the assumption of *experimenter-controlled events*.

We next suppose that the experimental situation is *symmetric* in the sense that the subject has no bias in his reaction to the two possible trial outcomes. Thus the occurrence of O_1 has the same effect on the A_1 response probability that the occurrence of O_2 has on the A_2 response probability. According to A1 and A2 the effect of O_1 on A_1 response probability is represented by the transformation

$$\mathbf{p}_{1,n} \to f(\mathbf{p}_{1,n})$$

while the effect of O_2 on A_2 response probability is represented by

$$\mathbf{p}_{2,n} \to 1 - g(1 - \mathbf{p}_{2,n}).$$

Thus the symmetry assumption can be expressed as $f(p) = 1 - g(1 - p)$ or

A3 $$g(p) = 1 - f(1 - p) \quad \textit{for} \quad 0 \leq p \leq 1.$$

The function f, which completely determines the subject's response to his experimental environment, is subject to several constraints. Since $f(p)$ is a probability we must have $0 \leq f(p) \leq 1$ for $0 \leq p \leq 1$. And, since f represents the effect of reinforcement we expect $f(p) \geq p$. In addition, however, it seems unreasonable to suppose that there are special values of response probability in the open unit interval such that if a subject happens to have one of these values as his initial A_1 probability his state of learning will be unchanged by an indefinitely long sequence of O_1 outcomes. Consequently we require that

A4 $$f(p) > p \quad for \quad 0 < p < 1.$$

Further it seems reasonable to expect that if two subjects begin a trial with A1 probabilities p and p', and both receive an O_1 outcome on that trial, the new A_1 response probabilities $f(p)$ and $f(p')$ will have the same ordering as p and p'. Consequently we assume that

A5 $$f(p) \text{ is non-decreasing for } 0 \le p \le 1.$$

Finally it is convenient to restrict ourselves to the case in which

A6 $$f(p) \text{ is continuous for } 0 \le p \le 1.$$

For convenience we refer to models satisfying A1 through A6 as *two-operator models* rather than as *symmetric path-independent models with experimenter-controlled events.*

With these preliminaries behind us we can fulfill our promise to provide examples of nonlinear two-operator models that predict simple probability matching. Let $1 > \theta > \frac{1}{2}$ and let $f(p)$ be any continuous non-decreasing function with $1 \ge f(p)$ for all $0 \le p \le 1$,

$$f(p) = p + \theta(1 - p) \qquad 0 \le p \le 1 - \theta, \qquad \theta \le p \le 1,$$

and

$$f(p) > p + \theta(1 - p) \qquad 1 - \theta < p < \theta.$$

Then, for any $0 \le p \le 1$,

$$f(p) \ge \theta \quad \text{and} \quad g(p) = 1 - f(1 - p) \le 1 - \theta.$$

Thus the two-operator model determined by f has the property that $\mathbf{p}_{1,2}$, $\mathbf{p}_{1,3}$, $\mathbf{p}_{1,4}$, $\cdots$ are confined to $[0, 1 - \theta] \cup [\theta, 1]$. But on this set the transitions for this model agree with those of the linear model with parameter θ. Since the latter predicts simple probability matching it follows that the former must also. Examples of this sort do not arise under the assumptions of Theorems 1, 2, and 4. The reason for this is that Theorem 1 puts heavy constraints on the transitions from $\mathbf{p}_{1,1}$ to $\mathbf{p}_{1,2}$, Theorem 2 requires that f be analytic, and Theorem 4 is concerned with small learning rates.

The Implications of Stationary Probability Matching

In the remainder of the paper we will denote by $\mathbf{p}_n$ the A_1 response probability random variable $\mathbf{p}_{1,n}$, and by $\bar{p}_n$ the expectation $E[\mathbf{p}_n]$ of $\mathbf{p}_n$. It follows from (4) that $\bar{p}_n$ is the unconditional probability of $A_{1,n}$, i.e.,

$$\bar{p}_n = P(A_{1,n}).$$

Under noncontingent reinforcement the linear model leads to the well-known equation

$$\bar{p}_n = \pi - (\pi - \bar{p}_1)(1 - \theta)^{n-1}, \qquad n = 1, 2, \cdots, \tag{5}$$

for any π and any distribution of $\mathbf{p}_1$. This implies simple probability matching; moreover, it implies that the convergence of $\bar{p}_n$ is monotonic. In particular, if the initial A_1 response probability random variable $\mathbf{p}_1$ has mean π then $\mathbf{p}_n$ has mean π throughout the experiment. Thus the linear model predicts stationary probability matching in the sense of the following definition.

DEFINITION. *A model is said to predict* stationary probability matching *if, for all* $0 \leq \pi \leq 1$, $P(A_{1,1}) = \pi$ *implies* $P(A_{1,n}) = \pi$ *for all* $n \geq 1$ *under noncontingent reinforcement with* O_1 *probability* π.

Clearly a model might have the property that for each π there is *some* distribution of $\mathbf{p}_1$ which has mean π and such that for this particular distribution of $\mathbf{p}_1$ all of the $\bar{p}_n$ would remain at π when $P(O_1) = \pi$. However the stationary probability matching condition requires that for *every* distribution of $\mathbf{p}_1$ with $\bar{p}_1 = \pi$ we have $\bar{p}_n = \pi$ for all $n \geq 1$ when $P(O_1) = \pi$. Bearing this in mind a simple inductive argument shows that for any model satisfying A1 stationary probability matching is equivalent to the requirement that $\bar{p}_1 = \pi$ and $P(O_1) = \pi$ imply $\bar{p}_2 = \pi$. Theorem 1 below shows that stationary probability matching is predicted by no two-operator models other than the linear models.

THEOREM 1. *A two-operator model predicts stationary probability matching if and only if, for some* $0 < \theta \leq 1$

$$f(p) = p + \theta(1 - p)$$

for all $0 \leq p \leq 1$.

PROOF. In view of (5) we need only show that the form of $f(p)$ indicated in the statement of the theorem is a necessary condition for stationary probability matching. Consequently we assume that $\bar{p}_1 = \pi$ and $P(O_1) = \pi$ imply $\bar{p}_2 = \pi$.

If we define the function $u(p)$ for $0 \leq p \leq 1$ by

$$u(p) = f(p) - p \tag{6}$$

then we can write

$$f(p) = p + u(p) \quad \text{and} \quad g(p) = p - u(1 - p), \tag{7}$$

the latter in view of A3. For an arbitrary distribution of $\mathbf{p}_1$ we have, under noncontingent reinforcement with O_1 probability π,

$$\bar{p}_2 = E[\mathbf{p}_2] = E[E[\mathbf{p}_2 \mid \mathbf{p}_1]] \tag{8}$$
$$= E[\pi f(\mathbf{p}_1) + (1 - \pi)g(\mathbf{p}_1)]$$
$$= \bar{p}_1 + E[\pi u(\mathbf{p}_1) - (1 - \pi)u(1 - \mathbf{p}_1)].$$

One distribution of $\mathbf{p}_1$ for which $\bar{p}_1 = \pi$ (and thus $\bar{p}_2 = \pi$) is the distribution δ_π concentrated at π. For this initial distribution (8) reduces to

$$\pi u(\pi) = (1 - \pi)u(1 - \pi). \tag{9}$$

Thus for all $0 < p \leq 1$

$$u(p) = (1 - p)u(1 - p)/p.$$

Substituting this for $u(\mathbf{p}_1)$ in (8) and assuming that $P(\mathbf{p}_1 = 0) = 0$, we have

$$\bar{p}_2 = \bar{p}_1 + E[(\pi - \mathbf{p}_1)u(1 - \mathbf{p}_1)/\mathbf{p}_1]. \tag{10}$$

Other distributions of $\mathbf{p}_1$ for which $\bar{p}_1 = \pi$ are those with

$$P(\mathbf{p}_1 = \pi + c) = P(\mathbf{p}_1 = \pi - c) = \tfrac{1}{2},$$

where $0 < |c| < \min(\pi, 1 - \pi)$. For a distribution of this form (10) specializes to

$$\frac{u(1 - (\pi + c))}{\pi + c} = \frac{u(1 - (\pi - c))}{\pi - c}. \tag{11}$$

Let any $0 < x, y < 1$ be given, where $x \neq y$. Let $\pi = (x + y)/2$ and $c = (x - y)/2$. Then $0 < |c| < \min(\pi, 1 - \pi)$, (the latter since $1 - \pi = ((1 - x) + (1 - y))/2$, while $c = ((1 - y) - (1 - x))/2$). Thus (11) gives $u(1 - y)/y = u(1 - x)/x$ for all $0 < x, y < 1$, i.e., $u(1 - x)/x = \theta$ for some constant θ and all $0 < x < 1$. Thus $f(p) = p + \theta(1 - p)$ for all $0 < p < 1$. By continuity of f the equality holds at the endpoints also. Clearly $f(0) = \theta \leq 1$, and A4 implies $\theta > 0$. Q.E.D.

Theorem 1 shows that within the class of two-operator models the linear model is characterized by stationary probability matching, which involves arbitrary distributions of $\mathbf{p}_1$. It is natural to ask to what extent the linear model is characterized by a comparable property, *weak stationary probability matching*, that involves only distributions of $\mathbf{p}_1$ concentrated at a single point.

DEFINITION. *A model is said to predict* weak stationary probability matching *if, for all* $0 \leq \pi \leq 1$, $\mathbf{p}_1 \equiv \pi$ *implies* $P(A_{1,n}) = \pi$ *for all* $n \geq 1$ *under noncontingent reinforcement with* O_1 *probability* π.

We use the term "weak" because this property is a consequence of stationary probability matching. The following theorem shows that within a certain important subclass of the class of two-operator models only the linear models predict weak stationary probability matching.

THEOREM 2. *If a two-operator model with $f(p)$ an analytic function for $0 \leq p \leq 1$* and $f(0) > 0$ predicts weak stationary probability matching then for some $0 < \theta \leq 1$*

$$f(p) = p + \theta(1 - p)$$

for all $0 \leq p \leq 1$.

PROOF. Suppose first that there are points $0 \leq p < p' \leq 1$ such that $f(p) = f(p')$. Then by A5, $f(p)$ is constant on the interval $[p, p']$, and thus, by the identity theorem for analytic functions ([8], p. 87) f is constant on [0, 1]. But A4 and A6 imply that $f(1) = 1$. Thus $f(p) = 1$ for $0 \leq p \leq 1$, and the conclusion of the theorem holds with $\theta = 1$. Thus we may assume throughout the rest of the proof that f is strictly increasing on [0, 1] (and thus $f(p) < 1$ for $0 \leq p < 1$).

Whenever $f(p)$ has a finite derivative from the left at $p = 1$, the function $w(p)$ defined by

$$w(p) = \begin{cases} \dfrac{u(1 - p)}{p} & \text{for} \quad 1 \geq p > 0 \\ 1 - f'(1) & \text{for} \quad p = 0 \end{cases} \tag{12}$$

is continuous on [0, 1]. In terms of $w(p)$, (7) can be written

$$\begin{aligned} f(p) &= p + (1 - p)w(1 - p), \\ g(p) &= p - pw(p), \qquad 0 \leq p \leq 1. \end{aligned} \tag{13}$$

In the present case the analyticity of $f(p)$ implies that $w(p)$ is analytic for $0 \leq p \leq 1$ (see [8], p. 90).

Weak stationary probability matching has as a consequence that $\mathbf{p}_1 \equiv \pi$ and $P(O_1) = \pi$ imply $\bar{p}_2 = \pi$ so the reasoning that led to (9) is valid. From (9) we easily conclude that

$$w(p) = w(1 - p) \tag{14}$$

for $0 < p < 1$ and thus, by continuity of w, for $0 \leq p \leq 1$. This means that w is symmetric about $\frac{1}{2}$.

We will now show by induction that, for each $0 < \pi < 1$, $w(p_n)$ has the same value $w_{n,\pi}$ for all A_1 probabilities p_n that can arise on trial n when $\mathbf{p}_1 \equiv \pi$. This is trivially true for $n = 1$. Suppose that it is true for some $n \geq 1$ and all $0 < \pi < 1$. Let some $0 < \pi_0 < 1$ be given and assume that $\mathbf{p}_1 \equiv \pi_0$ and that reinforcement is noncontingent with O_1 probability π_0 in what follows. By the induction hypothesis, for any A_1 probability p_{n+1} which can arise on trial $n + 1$ we have either $w(p_{n+1}) = w_{n,f(\pi_0)}$ (if $p_2 = f(p_1) = f(\pi_0)$), or $w(p_{n+1}) = w_{n,g(\pi_0)}$ (if $p_2 = g(\pi_0)$). Thus the induction will be complete

*That is, f has a convergent power series expansion about every point of the interval [0, 1].

if we can show that $w_{n,f(\pi_0)} = w_{n,g(\pi_0)}$. Now

$$\begin{aligned}\pi_0 &= E[\mathbf{p}_{n+2}] = E[E[\mathbf{p}_{n+2} \mid \mathbf{p}_{n+1}]] \\ &= E[\pi_0 f(\mathbf{p}_{n+1}) + (1 - \pi_0)g(\mathbf{p}_{n+1})] \\ &= \pi_0 + E[(\pi_0 - \mathbf{p}_{n+1})w(\mathbf{p}_{n+1})]\end{aligned}$$

using (14). Thus

$$\begin{aligned}(15) \qquad 0 &= E[(\pi_0 - \mathbf{p}_{n+1})w(\mathbf{p}_{n+1})] \\ &= E[E[(\pi_0 - \mathbf{p}_{n+1})w(\mathbf{p}_{n+1}) \mid \mathbf{p}_2]] \\ &= w_{n,f(\pi_0)}E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2 = f(\pi_0)]\pi_0 \\ &\qquad + w_{n,g(\pi_0)}E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2 = g(\pi_0)](1 - \pi_0)\end{aligned}$$

by the induction hypothesis. But weak stationary probability matching implies that

$$\begin{aligned}0 &= E[\pi_0 - \mathbf{p}_{n+1}] = E[E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2]] \\ &= E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2 = f(\pi_0)]\pi_0 + E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2 = g(\pi_0)](1 - \pi_0)\end{aligned}$$

and, using A4 and the fact that f is strictly increasing

$$E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2 = f(\pi_0)] < E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2 = \pi_0] = E[\pi_0 - \mathbf{p}_n] = 0.$$

Thus $E[\pi_0 - \mathbf{p}_{n+1} \mid \mathbf{p}_2 = f(\pi_0)]\pi_0$ can be cancelled out of (15) to obtain

$$w_{n,f(\pi_0)} = w_{n,g(\pi_0)}.$$

We conclude that, for all $0 < \pi < 1$, and all $n \geq 1$, $w(p_n)$ has the same value for all p_n which can arise on trial n when $\mathbf{p}_1 \equiv \pi$. In particular,

$$(16) \qquad w(f^{(n)}(\pi)) = w(g^{(n)}(\pi))$$

for all $n \geq 0$ and $0 < \pi < 1$, where $f^{(n)}$ and $g^{(n)}$ are the nth iterates of f and g respectively, i.e.,

$$\begin{aligned}f^{(0)}(x) &= x \\ f^{(n+1)}(x) &= f(f^{(n)}(x)), \qquad n \geq 0.\end{aligned}$$

By continuity (16) holds for $\pi=0$ and $\pi=1$. But $g^{(n)}(0)=0$ for all $n \geq 0$, and $0 < f^{(1)}(0) < f^{(2)}(0) < \cdots < f^{(n)}(0) < 1$ by A4. Since

$$f(\lim_{n\to\infty} f^{(n)}(0)) = \lim_{n\to\infty} f^{(n)}(0)$$

we must have

$$\lim_{n\to\infty} f^{(n)}(0) = 1.$$

Thus $w(p)$ has the same value $\theta = w(0)$ on an infinite set of points accumulating at $p = 1$. By the identity theorem for analytic functions

$$w(p) = \theta \tag{17}$$

for $0 \leq p \leq 1$. From A4 it follows that $0 < \theta$, and since $f(0) = \theta$ we must have $\theta < 1$. Q.E.D.

The reader may have noticed that the assumption of analyticity in Theorem 2 is used primarily as a means of "interpolating" the function $w(p)$ between those points which can arise as possible values of $\mathbf{p}_n$ when $\mathbf{p}_1 \equiv 0$. If we were to drop the analyticity assumption, and simply assume that $f(p)$ is strictly increasing we could still conclude that $w(\mathbf{p}_n)$ is a constant $w_{n,x}$ over the set of points that can be values of $\mathbf{p}_n$ when $\mathbf{p}_1 = x$. It might be supposed that as n increases this set, for some x, would become sufficiently large that w could be shown to be constant everywhere in $[0, 1]$. However the remarks at the end of the previous section indicate that this is not feasible. It appears that something like analyticity is required if we are to obtain a complete characterization of w from conditions on the values of $w(\mathbf{p}_n)$.

We will say that a model is *ergodic* if for every $0 < \pi_1, \pi_2 < 1$ there is a distribution $F^{(\pi_1, \pi_2)}(p)$ such that for every distribution of $\mathbf{p}_1$ the distribution of $\mathbf{p}_n$ converges to $F^{(\pi_1, \pi_2)}$ as n approaches infinity under contingent reinforcement with $P(O_{i,n} \mid A_{i,n}) = \pi_i$. The weaker property of *ergodicity under noncontingent reinforcement* is defined similarly. We mention that the argument given by Karlin ([7], sec. 6) need be modified very little to show ergodicity, not only for the linear models which he considers, but for any two-operator model that is *strictly distance diminishing* in the sense that there is some $\lambda < 1$ such that

$$|f(p) - f(p')| \leq \lambda\, |p - p'| \quad \text{for all} \quad 0 \leq p, p' \leq 1.$$

If a model that is ergodic under noncontingent reinforcement predicts weak stationary probability matching then $F^{(\pi, 1-\pi)}(p)$ must have mean π so that simple probability matching under noncontingent reinforcement is also predicted. On the other hand if a model that is ergodic under noncontingent reinforcement predicts simple probability matching on such schedules then if $\mathbf{p}_1$ has the distribution $F^{(\pi, 1-\pi)}$ it follows that $\bar{p}_n = \pi$ for all $n \geq 1$ under noncontingent reinforcement with $P(O_1) = \pi$. Thus for two-operator models that are ergodic under noncontingent reinforcement simple probability matching on this schedule and the two versions of stationary probability matching are very closely related. It is interesting to ask how one might distinguish experimentally between them.

To test for stationary probability matching over and above simple probability matching under noncontingent reinforcement with $P(O_1) = \pi$ it is necessary to find a group of subjects that has a distribution of $\mathbf{p}_1$ with mean π *other than* $F^{(\pi, 1-\pi)}$. Thus we cannot obtain subjects by running a prior noncontingent binary prediction experiment with $P(O_1) = \pi$. However a suitable group of subjects could be obtained by giving various subgroups previous training on different $P(O_1)$'s. For instance, half the subjects might receive previous training under noncontingent reinforcement with $P(O_1) = \pi + c$ and the rest with $P(O_1) = \pi - c$ where $c < \min(\pi, 1 - \pi)$. A binary prediction experiment with a group of subjects thus constituted could easily lead to results that would cast doubt on stationary probability matching.

The Implications of Marginal Constancy

Weak stationary probability matching is not well adapted to direct experimental test; in order to perform an appropriate experiment one would have to be able to evaluate the initial A_1 response probabilities of individual subjects. However for two-operator models this property implies a condition that has direct experimental implications. Consider the contingent reinforcement situation in which the outcome always agrees with the subject's response; i.e., $\pi_1 = \pi_2 = 1$. Under these conditions, which we call double reward contingencies, we have

$$E[\mathbf{p}_{n+1} \mid \mathbf{p}_n = p] \equiv p\, f(p) + (1 - p)g(p),$$

or using the representation given by (7),

$$E[\mathbf{p}_{n+1} \mid \mathbf{p}_n = p] \equiv p + p\, u(p) - (1 - p)u(1 - p). \tag{18}$$

Now if a two-operator model satisfies the functional equation (9), which is a consequence of weak stationary probability matching, (18) simplifies to

$$E[\mathbf{p}_{n+1} \mid \mathbf{p}_n = p] \equiv p. \tag{19}$$

Consequently any two-operator model that satisfies (9) predicts *marginal constancy*.

DEFINITION. *A model is said to predict* marginal constancy *if* $P(A_{1,n}) = P(A_{1,1})$ *for all* $n \geq 1$ *in the double reward situation.*

For the class of two-operator models it is easily seen that (9) is a necessary as well as a sufficient condition for marginal constancy.

Since marginal constancy is a property of the linear model that is subject to direct experimental test it is of interest to examine asymptotic predictions *under contingencies other than double reward* of two-operator models that share this property. In this section we will relate $\lim_{n\to\infty} \bar{p}_n$ for an important subclass of the class of marginally constant two-operator models to the comparable probability matching limits for linear models. Our basic tool will be the characterization (9) of marginally constant two-operator models which does not refer to any particular reinforcement contingency.

We will restrict our attention to two-operator models satisfying the regularity condition

A7 *$f(p)$ has a finite derivative from the left at $p = 1$.*

Under A7 we have at our disposal the representation (13) where $w(p)$ is the continuous function (12). In terms of $w(p)$ the functional equation (9) for marginal constancy assumes the form (14).

We will also assume that

A8 $$f(0) > 0$$

i.e., reinforcement of A_1 is effective, even when the probability of A_1 is 0. In view of (13) this implies that $w(1) > 0$. By (14) it follows that $w(0) > 0$ also under marginal constancy. But A4 and (12) imply that $w(p) > 0$ for $0 < p < 1$. Thus by the continuity of $w(p)$ we have

$$\min_{0 \le p \le 1} w(p) > 0$$

for marginally constant two-operator models satisfying A7 and A8.

Putting together (13) and (14) we obtain the representation

$$\mathbf{p}_{n+1} = \begin{cases} \mathbf{p}_n + w(\mathbf{p}_n)(1 - \mathbf{p}_n) & \text{if} \quad O_{1,n} , \\ \mathbf{p}_n - w(\mathbf{p}_n)\mathbf{p}_n & \text{if} \quad O_{2,n} . \end{cases} \tag{20}$$

Comparing (20) with (3) we see that the class of marginally constant two-operator models satisfying A7 and A8 may be regarded as a generalization of the class of linear models, with the learning rate parameter θ replaced by the continuous positive function $w(p)$. We pause here to show that there exist nonlinear ergodic marginally constant two-operator models satisfying A7 and A8. Let

$$f(p, \gamma, \delta) = p + (1 - p)w(1 - p, \gamma, \delta),$$

where

$$w(p, \gamma, \delta) = \gamma(1 + \delta p(1 - p))$$

and $0 < \gamma < \frac{1}{2}$, $0 \le \delta \le \frac{1}{4}$. Clearly $1 \ge f(p, \gamma, \delta)$ and $w(p, \gamma, \delta) = w(1 - p, \gamma, \delta)$ for $0 \le p \le 1$, and $f(p, \gamma, \delta) > p$ if $0 \le p < 1$. Further

$$\frac{\partial}{\partial p} f(p, \gamma, \delta) = 1 - \gamma + \gamma\delta(1 - p)(1 - 3p).$$

From this it follows that

$$\frac{\partial}{\partial p} f(p, \gamma, \delta) \ge 1 - \gamma - \tfrac{1}{2}\gamma > 0$$

and

$$\frac{\partial}{\partial p} f(p, \gamma, \delta) \le 1 - \gamma + \tfrac{1}{2}\gamma < 1.$$

The fact that $|f'(p)|$ is bounded away from 1 implies that the model determined by $f(p, \gamma, \delta)$ is strictly distance diminishing, hence ergodic.

In the next theorem, (22) shows how near-linearity in the sense of small (max w − min w) implies approximate probability matching in the sense of small $|\bar{p}_\infty - \bar{e}_\infty|$ for ergodic marginally constant two-operator models that satisfy A7 and A8. Equation (21) is a by-product of the proof of (22) which we include in the statement of the theorem because it is interesting in its own right.

THEOREM 3. *For* $0 < \pi_1, \pi_2 < 1$ *ergodic marginally constant two-operator models having properties* A7 *and* A8 *satisfy*

$$\frac{\min w}{\max w}\, l \le \bar{p}_\infty \le \frac{\max w}{\min w}\, l \tag{21}$$

and

$$|\bar{p}_\infty - \bar{e}_\infty| \le \sqrt{(1-\pi_1)(1-\pi_2)}\, \frac{(\max w - \min w)}{\sqrt{\max w \min w}} \tag{22}$$

under contingent reinforcement with

$$P(O_{i,n} \mid A_{i,n}) = \pi_i$$

where

$$\max w = \max_{0 \le p \le 1} w(p), \qquad \min w = \min_{0 \le p \le 1} w(p),$$

$$\bar{p}_\infty = \lim_{n\to\infty} \bar{p}_n, \qquad \bar{e}_\infty = \lim_{n\to\infty} P(O_{1,n}),$$

and

$$l = \frac{1-\pi_2}{2-\pi_1-\pi_2}$$

is the unique value of $\bar{p}_\infty$ *for which probability matching obtains,* i.e., *for which* $\bar{p}_\infty = \bar{e}_\infty$. (The reader should note the extent to which this notation suppresses dependence on π_1 and π_2.)

PROOF. For any function $b(p)$ continuous for $0 \le p \le 1$ we have

$$\begin{aligned} E[b(\mathbf{p}_{n+1})] &= E[E[b(\mathbf{p}_{n+1}) \mid \mathbf{p}_n]] \\ &= E[M(\mathbf{p}_n)b(f(\mathbf{p}_n)) + (1 - M(\mathbf{p}_n))b(g(\mathbf{p}_n))] \end{aligned}$$

where $M(p) = \pi_1 p + (1-\pi_2)(1-p) = P(O_{1,n}|\mathbf{p}_n = p)$. Letting $n \to \infty$ we obtain

$$\int_0^1 b(p)\, dF(p) = \int_0^1 (M(p)b(f(p)) + (1 - M(p))b(g(p))\, dF(p), \tag{23}$$

where $F = F^{(\pi_1, \pi_2)}$ is the asymptotic distribution of $\mathbf{p}_n$ when $P(O_{i,n}|A_{i,n}) = \pi_i$. For $b(p) = p$, (23) yields

$$0 = \int_0^1 [M(p)(1 - p)w(1 - p) - (1 - M(p))pw(p)] \, dF(p). \tag{24}$$

Since $w(p) = w(1 - p)$, (24) can be simplified to

$$\int_0^1 w(p)p \, dF(p) = l \int_0^1 w(p) \, dF(p). \tag{25}$$

It is worth pointing out that in the case of simple probability matching: $\bar{p}_\infty = l$, (25) means that $w(p)$ and p are uncorrelated with respect to the distribution $F(p)$.

Equation (25) implies

$$\min w \int_0^1 p \, dF(p) \leq l \max w$$

and

$$\max w \int_0^1 p \, dF(p) \geq l \min w.$$

Since

$$\bar{p}_\infty = \int_0^1 p \, dF(p),$$

(21) is proved. We next note that

$$P(O_{1,n}) = \pi_1 \bar{p}_n + (1 - \pi_2)(1 - \bar{p}_n).$$

Consequently

$$\bar{p}_\infty - \bar{e}_\infty = (2 - \pi_1 - \pi_2)\bar{p}_\infty - (1 - \pi_2). \tag{26}$$

This means that $\bar{p}_\infty - \bar{e}_\infty$ is small when and only when $\bar{p}_\infty - l$ is small. From (26) and (21) it follows that

$$\frac{\min w - \max w}{\max w}(1 - \pi_2) \leq \bar{p}_\infty - \bar{e}_\infty \leq \frac{\max w - \min w}{\min w}(1 - \pi_2). \tag{27}$$

As a consequence of the symmetry assumption A3 the same argument applied to the probabilities of A_2 and O_2 will yield

$$\frac{\min w - \max w}{\max w}(1 - \pi_1) \leq (1 - \bar{p}_\infty) - (1 - \bar{e}_\infty) \leq \frac{\max w - \min w}{\min w}(1 - \pi_1)$$

or

$$\frac{\min w - \max w}{\min w}(1 - \pi_1) \leq \bar{p}_\infty - \bar{e}_\infty \leq \frac{\max w - \min w}{\max w}(1 - \pi_1). \tag{28}$$

If $\bar{p}_\infty \geq \bar{e}_\infty$ the product of the upper inequalities in (27) and (28) yields (22). If $\bar{p}_\infty - \bar{e}_\infty < 0$ the product of the lower inequalities in (27) and (28) leads to the same conclusion. Q.E.D.

The fact that estimates of the parameter θ of the linear model are usually small in binary prediction experiments leads us to examine two-operator models that predict slow learning. We do this by considering certain one-parameter families of two-operator models depending on a "learning rate" parameter ϵ like the θ of the linear models, and focusing our attention on the predictions of the family as ϵ becomes small. Just as a single two-operator model is determined by a function $f(p)$ of one variable a one-parameter family of two-operator models is determined by a function $f(p, \epsilon)$ of two variables. We first impose the following restriction on the dependence of $f(p, \epsilon)$ on p for fixed ϵ:

F1 *For some $\mu > 0$ and each $0 < \epsilon \leq \mu$, $f(p, \epsilon)$, regarded as a function of p, determines an ergodic two-operator model satisfying* A7, A8, *and*

$$\frac{\partial}{\partial p} f(1, \epsilon) < 1.$$

Under F1, $f(p, \epsilon)$ has the representation

$$f(p, \epsilon) = p + w(1 - p, \epsilon)(1 - p),$$

where $w(p, \epsilon)$ is positive and continuous in p for $0 \leq p \leq 1$, $0 < \epsilon \leq \mu$. Since the magnitude of $w(p, \epsilon)$ is positively related to the rate of learning, our interpretation of ϵ as a learning rate parameter requires that $w(p, \epsilon)$ decrease to 0 as ϵ decreases to 0. To make this precise we first assume that

F2 $$\lim_{\epsilon \to 0} w(p, \epsilon) = 0, \qquad 0 \leq p \leq 1.$$

This suggests that we extend the domain of $w(p, \epsilon)$ by defining $w(p, 0)$ to be 0 for $0 \leq p \leq 1$. We further assume that

F3 *$w(p, \epsilon)$ is continuous and $(\partial/\partial\epsilon)w(p, \epsilon)$ and $(\partial^2/\partial\epsilon^2)w(p, \epsilon)$ exist and are continuous for $0 \leq p \leq 1$, $0 \leq \epsilon \leq \mu$.*

To insure that $w(p, \epsilon)$ decreases as ϵ decreases, at least for sufficiently small ϵ, it is sufficient to assume

F4 $$\frac{\partial}{\partial \epsilon} w(p, 0) > 0 \qquad 0 \leq p \leq 1.$$

Examples of families of models satisfying F1 $-$ F4 are the linear family $w(p, \theta) = \theta$, $0 < \theta \leq 1$, and, more generally, the families determined by the functions $w(p, \gamma) = \gamma(1 + \delta p(1 - p))$, $0 < \gamma \leq \frac{1}{4}$, for any fixed value of δ, $0 \leq \delta \leq \frac{1}{4}$.

We wish to give a definition of approximate probability matching for small learning rates that will be applicable to families of models satisfying F1 $-$ F4. In view of (26) we might consider requiring only that

$$\lim_{\epsilon \to 0} \bar{p}_\infty = l.$$

However it turns out to be useful to require somewhat more.

DEFINITION. *A family of models satisfying* F1 $-$ F4 *is said to predict* approximate probability matching for small learning rates *if and only if*

$$\lim_{\epsilon \to 0} \bar{p}_\infty = l \tag{29}$$

and

$$\lim_{\epsilon \to 0} \lim_{n \to \infty} \text{var } \mathbf{p}_n = 0 \tag{30}$$

for all contingent reinforcement schedules with $0 < \pi_1 , \pi_2 < 1$.

It is easily shown that the condition

$$\lim_{\epsilon \to 0} \int_0^1 (p - l)^2 \, dF_\epsilon(p) = 0 \tag{31}$$

is equivalent to (29) and (30) where $F_\epsilon = F_\epsilon^{(\pi_1, \pi_2)}$ is the asymptotic distribution of $\mathbf{p}_n$ when $P(O_{i,n} | A_{i,n}) = \pi_i$ and the learning rate parameter has the value ϵ. For the linear family under noncontingent reinforcement we have the formula

$$\text{var } \mathbf{p}_\infty = \pi(1 - \pi) \frac{\theta}{2 - \theta}$$

[3] which shows that var $\mathbf{p}_\infty \to 0$ as $\theta \to 0$. The same is true for the linear family under contingent reinforcement as a consequence of Theorem 4 below.

Our previous work suggests that approximate marginal constancy for small learning rates, appropriately defined, will be closely related to approximate probability matching in the above sense. We might think of requiring only that

$$\lim_{\epsilon \to 0} (\bar{p}_{n+1} - \bar{p}_n) = 0$$

for all n under double reward. However it is easily shown that for any family of models satisfying F1 $-$ F4

$$\bar{p}_{n+1} - \bar{p}_n = O(\epsilon)$$

under double reward. Since the term "approximate marginal constancy" suggests that $\bar{p}_{n+1}$ is *especially* close to $\bar{p}_n$ we are led to the following definition.

DEFINITION. *A family of models satisfying* F1 — F4 *is said to predict approximate marginal constancy for small learning rates if and only if*

$$\bar{p}_{n+1} - \bar{p}_n = O(\epsilon^2)$$

for all $n \geq 1$ *under double reward.*

Using the Taylor expansion

$$w(p, \epsilon) = \epsilon \frac{\partial}{\partial \epsilon} w(p, 0) + \frac{\epsilon^2}{2} \frac{\partial^2}{\partial \epsilon^2} w(p, \epsilon^*) \tag{32}$$

where $0 < \epsilon^* < \epsilon$, it is easily shown that the condition

$$\frac{\partial}{\partial \epsilon} w(p, 0) = \frac{\partial}{\partial \epsilon} w(1 - p, 0) \tag{33}$$

for $0 \leq p \leq 1$ is equivalent to approximate marginal constancy for small learning rates, just as (14) is equivalent to marginal constancy for a single model. We can now state and prove the following theorem.

THEOREM 4. *For any family of models satisfying* F1 — F4 *approximate probability matching for small learning rates and approximate marginal constancy for small learning rates are equivalent.*

PROOF. Suppose that (31) is satisfied for some $0 < \pi_1 = \pi = 1 - \pi_2 < 1$. Since $\pi = l$ and $M(p) \equiv \pi$ (24) and (32) yield

$$0 = \int_0^1 \left[\pi(1 - p) \frac{\partial}{\partial \epsilon} w(1 - p, 0) - (1 - \pi)p \frac{\partial}{\partial \epsilon} w(p, 0) \right] dF_\epsilon(p) + O(\epsilon). \tag{34}$$

Shrinking ϵ to 0, (31) implies

$$0 = \pi(1 - \pi) \frac{\partial}{\partial \epsilon} w(1 - \pi, 0) - (1 - \pi)\pi \frac{\partial}{\partial \epsilon} w(\pi, 0)$$

which gives (33) for $p = \pi$. Thus approximate probability matching for small learning rates implies approximate marginal constancy for small learning rates.

To prove the opposite implication we take $b(p) = (p - l)^2$ in (23) to obtain

$$\int_0^1 (p - l)^2 \, dF_\epsilon(p) = \int_0^1 [M(p)((p - l) + w(1 - p, \epsilon)(1 - p))^2 + (1 - M(p))((p - l) - w(p, \epsilon)p)^2] \, dF_\epsilon(p)$$

or

$$0 = 2 \int_0^1 (p - l)[M(p)w(1 - p, \epsilon)(1 - p) - (1 - M(p))w(p, \epsilon)p] \, dF_\epsilon(p) + \int_0^1 [M(p)w^2(1 - p, \epsilon)(1 - p)^2 + (1 - M(p))w^2(p, \epsilon)p^2] \, dF_\epsilon(p).$$

Thus, using (32) and (33) and noting that $M(p) - p = (2 - \pi_1 - \pi_2)(l - p)$ we find that

$$\int_0^1 (p - l)^2 \frac{\partial}{\partial \epsilon} w(p, 0)\, dF_\epsilon(p) = O(\epsilon).$$

Consequently

$$0 \leq \min_{0 \leq p \leq 1} \frac{\partial}{\partial \epsilon} w(p, 0) \int_0^1 (p - l)^2\, dF_\epsilon(p) \leq O(\epsilon).$$

Since

$$\min_{0 \leq p \leq 1} \frac{\partial}{\partial \epsilon} w(p, 0) > 0$$

by F3 and F4, (31) follows. Q.E.D.

REFERENCES

[1] Bush, R. R. and Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.
[2] Estes, W. K. Probability learning. In A. W. Melton (Ed.), *Categories of human learning.* New York: Academic Press, 1964. Pp. 89–128.
[3] Estes, W. K. and Suppes P. Foundations of linear models. In R. R. Bush and W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford: Stanford Univ. Press, 1959. Pp. 137–179.
[4] Estes, W. K. and Suppes, P. *Foundations of statistical learning theory,* II. *The stimulus sampling model.* Tech. Rept. No. 26, Psychol. Ser., Inst. Math. Studies Soc. Sci., Stanford Univer., 1959.
[5] Feldman, J. and Newell, A. A note on a class of probability matching models. *Psychometrika,* 1961, **26**, 333–337.
[6] Friedman, M. P., Burke, C. J., Cole, M., Keller, L., Millward, R. B., and Estes, W. K. Two-choice behavior under extended training with shifting probabilities of reinforcement. In R. C. Atkinson (Ed.), *Studies in mathematical psychology.* Stanford: Stanford Univ. Press, 1964. Pp. 250–316.
[7] Karlin, S. Some random walks arising in learning models. *Pacific J. Math.,* 1953, **3,** 725–756.
[8] Knopp, K. *Theory of functions* (part I). New York: Dover, 1945.
[9] Luce, R. D. *Individual choice behavior.* New York: Wiley, 1959.
[10] Rose, R. M. Models for experiments with two complementary reinforcing events. Unpublished doctoral dissertation, Univ. Pennsylvania, 1964.
[11] Sternberg, S. Stochastic learning theory. In R. D. Luce, R. R. Bush, and E. Galanter (Eds.), *Handbook of mathematical psychology* (vol. II). New York: Wiley, 1963. Pp. 1–120.

Manuscript received 3/2/65
Revised manuscript received 7/26/65