THE MATCHING PRINCIPLE REVISITED

EDWARD J. GREEN and JOHN KEMENY*

University of Tennessee at Chattanooga

Reexamination of anomalous data from an old study sheds new light on statistical learning theory and deterministic assumptions about human behavior.

The generalization of Estes’ original statistical learning theory (Estes, 1950) was published by him and Straughan in 1954 (Estes & Straughan, 1954). It has come to be known as the “matching principle” and is defined by the equation

\[ \bar{p}_n = \Pi - (\Pi - \bar{p}_0) (1 - \theta)^n \]  

where mean response probability approaches the probability of reinforcement, \( \Pi \), asymptotically.

In our model, the constant \( \Pi \), which Estes and Straughan used to represent the probability of reinforcement, or proportion of reinforced trials, was replaced by the probability of response of Subject B for Subject A, and vice versa. The experiment designed to test the model arranged that right-hand responses of one subject would reinforce right-hand responses of the other. The constant, \( \Pi \), replaced by the varying probabilities of reinforcement for each person in the interaction. These variables through their iterative interactions defined a convergence from any two initial response probabilities to a point uniquely determined by the initial probabilities and the two sampling ratios of the interacting individuals. Once convergence was achieved, it was thought that behavior would settle down around the joint asymptotic probabilities in a manner similar to results reported by Estes and Straughan for the verbal learning paradigm.

Formally, our mathematical model assumed that the behavior of individuals A and B may be categorized into two mutually exclusive classes R and R’. Instances of R on the part of individual A reinforce instances of responses R for individual B. Symmetrically, R’ on the part of either member of group reinforces behavior R’ on the part of the other.

Edward J. Green is Guerry Professor Emeritus at the University of Tennessee at Chattanooga.

* John Kemeny, deceased, was formerly Chair of the Department of Mathematics and President of Dartmouth College.
The rate of change in response probability is a function of the size of the sampling ratios, $\theta A$ and $\theta B$. In short, the behavior of each changes as a function of the behavior of the other. The probability that individual A will make an R response, $p(A)$, and $p(B)$ is the probability that individual B will make an R response. Let the changes for individual A be described by the equation

$$
\overline{p_n} A = \overline{p_n} B - (\overline{p_n} B - \overline{p_0} A)(1 - \theta_A)^n
$$

(2)

where individual A is the first to respond in the interaction and by the equation

$$
\overline{p_n} B = \overline{p_{n-1}} A - (\overline{p_{n-1}} A - \overline{p_0} B)(1 - \theta_B)^n
$$

(3)

for individual B. The equations are symmetrical for the instance where B responds first.

Equations 2 and 3 may be rewritten as:

$$
\overline{p_n} A = (1 - \rho^n A)\overline{p_n} B + a \rho^n A
$$

(4)

and

$$
\overline{p_n} B = (1 - \rho^n B)\overline{p_{n-1}} A + b \rho^n A
$$

(5)

where $a = \overline{p_0} A$; $b = \overline{p_0} B$; $A = (1 - \theta A)$; and $B = (1 - \theta B)$. It is seen that these last two equations are of the form

$$
\overline{p_n} A = a + (b - a)f(n)
$$

(6)

and

$$
\overline{p_n} B = b + (a - b)g(n)
$$

(7)

where

$$
f(n) = (1 - \rho^n A)(1 - g(n)), \text{ and } g(n) = (1 - \rho^n B)(1 - f(n - 1)).
$$

The term $f(n)$ may be expanded where $f(0) = 0$ and $g(0) = 1$,

$$
f(n) = 1 - \rho^n A(1 - \rho^n B)f(n - 1) = (1 - \rho^n A)\rho^n B.
$$

(8)

Solving for $n = 3$, one sees

$$
f(3) = (1 - \rho^3 A)\rho^3 B + (1 - \rho^3 A)(1 - \rho^3 B)(1 - \rho^2 A)\rho^2 B + (1 - \rho^3 A)(1 - \rho^3 B)(1 - \rho^2 A) + (1 - \rho^2 B)(1 - \rho A)\rho B,
$$

and where we deal with the simplified case $\rho A = 0$,

$$
g(n) = (1 - \rho B)(1 - \rho^2 B) \ldots (1 - \rho^n B)
$$

(9)
which will be recognized as the Lambert series. No simple formula exists for predicting asymptotes, although for certain special conditions, the model does yield unique asymptotes that can be readily computed from Equations 4 and 5. It is possible to solve for asymptotic constants \( F \) and \( G \) where half the time \( A \) subjects start the interaction and half the time \( B \) subjects start it. This condition is not merely a mathematical curiosity, but also conforms to the requirements of good experimental design. Since

\[
f(n) = (1 - \rho^nA)(1 - g(n))
\]

and

\[
g(n) = 1 - \rho^nB(1 - f(n-1))
\]

where the initial roles of \( A \) and \( B \) are reversed,

\[
f'(n) = (1 - \rho^nA)(1 - g'(n - 1))
\]

and

\[
g'(n) = (1 - \rho^nB)(1 - f'(n))
\]

so

\[
\bar{f}(n) = (1 - \rho^nA)(1 - [(g(n) + g'(n - 1))/2])
\]

and

\[
\bar{g}(n) = (1 - \rho^nA)(1 - [(\bar{f}(n - 1) + f'(n))/2]).
\]

Therefore

\[
F = (1 - \rho A)/[2 - \rho A - \rho B] = \theta A/(\theta A + \theta B) = 1 - G
\]

and it is true that

\[
\bar{p}_\infty A = \bar{p}_\infty B = (Fb = Ga)
\]

which, rewritten in the original notation

\[
\bar{p}_\infty A = \bar{p}_\infty B = [\theta A/(\theta A + \theta B)]\rho_0B + [\theta B/(\theta A + \theta B)]\rho_0A.
\]

Therefore, four parameters uniquely predict the asymptotic convergence for the 2-person group. These are the mean initial probabilities for \( A \) and \( B \), \( \rho_0A \) and \( \rho_0B \), and the sampling ratios, \( \theta A \) and \( \theta B \). It is seldom the case either in practice or in theory that convergence is reached exactly halfway between the initial positions of the individuals participating in the dyadic
interaction. First, whoever initiates the interaction takes the largest step toward the initial position of the other. Second, differing sampling ratios tilt convergence toward the initial probability of the person with the smaller sampling ratio. The convergence value will always be closer to the initial probability of the slower learner; a fact with depressingly unsurprising implications for education and negotiation.

Conditions do exist, however, where these four parameters do not determine convergence. One is where the behavior of one party to the interaction is not reinforced by the behavior of the other. This is the original condition of the matching principle where behavior is controlled or reinforced by adherence to an independent standard. An example would be where one party is supposedly teaching the other. In this case, one does not rely upon taking a vote to verify a correct answer, although that can happen. I once heard of a junior high math teacher who did exactly that when confronted with several answers to a homework problem for which he himself did not know the answer. A second case is where a single person is party to an interaction with several others holding a differing initial position. In that case, the interaction follows the equations described above except that while the exponent over trials for the single individual remains \( n \), the exponent for the group becomes \( n/N \), where \( N \) is the number of persons in the group. Convergence in this instance is more correctly called conformity. The group position is little altered by the individual.

Empirical data gathered by myself and Richard M. Sanders initially confirmed the convergent predictions of the model for a 2-person interaction in that response probabilities of two subjects working against each other did for a time approach a response probability intermediate between their respective initial probabilities at a rate commensurate with their respective sampling ratios.

In Phase 1 of the study, subjects were separately trained according to the matching procedures of Estes and Straughan to differing asymptotic response probabilities. These became initial probabilities for the interaction condition of Phase 2. Respective values of their respective sampling ratios were empirically calculated from the data of Phase 1.

Response probabilities were defined by the proportion of right-hand responses over 10 trial blocks of responses. That there might be such a thing as a 50% response taking place in any individual trial is an absurdity akin to Schrödinger’s Cat. Subjects made either a right-hand or a left-hand response. This definition of probability was necessary to fit empirical results to the mathematical model predicting mean probabilities. An alternative strategy could have been to average left- and right-hand responses trial by trial over a sample of subjects where \( N > 1 \). The experiment was an extraordinarily time intensive pilot study, so such an option was unrealistic.

We favored rate or frequency as the dependent variable to take full advantage of a free operant situation. Response rate is correlated with response probabilities. A response consumes a finite interval of time. Response \( A \) competes with other behaviors for time, so less time is required
to execute that response when competing behaviors do not precede it. Where \( k \) is an empirical constant, \( t \) is time, \( p \) is probability, and \( r \) is rate, the relationship between probability of response, reaction time or latency, and rate is \( p = k/t \). Rates of free operant responding ran at about 120 responses/minute at the limiting probabilities, and at about 30 responses/minute where \( p = 0.5 \). When response probabilities approached either 0 or 1, response rates increased.

Disturbingly, following what appeared to be an initial convergence, positively accelerated oscillations appeared after about 140 interactions. The oscillations increasingly became more extreme until the experimenter terminated the run. The rate of response and frequency of oscillations was quite dramatic. The oscillations were between all right-hand responses in a series alternating with a series of all left-hand responses. Averaging individual responses with those of other subjects might have yielded probabilities more in conformity with prediction. Such averaging was rejected, however, since the oscillation in the responding of individual subjects was too remarkable to be swept under a statistical rug, so to speak. The results of the experiment seemed so anomalous that I not only abandoned the model, but ceased further research in statistical learning theory. The oscillation of response probabilities were not, to my understanding at the time, consequences of any characteristic of the model they were designed to test. That year, I happened to discuss the experiment with Edith Neimark at the annual convention of the American Psychological Association. She noted that she had observed the same oscillation of response probabilities using the standard matching to probability of reinforcement when subjects ran past convergence. We agreed that incorporating ad hoc variables into the model to accommodate these results would compromise the integrity of the theory. In any case, at least for myself, the matter rested for 40 years.

Recently, in reviewing James Gleick’s book on chaos theory (Gleick, 1987), I realized that the anomalous results of Sander’s and my empirical study had also appeared in other fields of empirical research where phenomena follow stochastic processes. Similar results were obtained in May’s studies of nonlinear population changes in the late 1970s. These studies, however, were two decades too late for our experiment. Once I realized that the oscillation is the result of intrinsic variability in stochastic processes I sought to simulate momentary changes although the methods at hand were very time intensive. Ideally a computer simulation could be used, but neither I nor my colleagues had the expertise to program the process, so a mechanical Monte Carlo procedure was adopted. In this, I set up two arrays of “stimuli” consisting of 10 numbered spaces. One of these was used to track changes in the status of individual stimuli for Subject A, and the other for Subject B. If Subject A had responded, say, to Stimuli 2, 4, and 8, they would determine whether he made a right-hand or a left-hand response. The same sampling procedure was followed for individual B. His response then either confirmed the response of A, or it reversed the associations. Both
individuals A and B made a right-hand response if the stimuli which they initially sampled were mostly R stimuli, and they made a left-hand response when the sampled stimuli were mostly L stimuli. If B’s response was the same as A’s had been, no changes occurred in whether the stimuli remained Rs or Ls. If, however, B made a left-hand response to A’s right-hand response, all the R stimuli that A had sampled in his set became L stimuli, in accordance with Estes’ formulation of the conditioning process. In the simulations reported here, either one stimulus was sampled on a trial, or three were sampled, eliminating the possibility of a tie, so resulting response probabilities were unambiguously for either right- or left-hand responses. The results of the various simulations were quite interesting in their own right, apart from the question of how or whether they shed light on behavioral oscillation.

First of all, the problem of asymptotes that Kemeny resolved by postulating that each partner to the interaction had to start half the time is a necessary given in the simulations. Someone has to respond first. It turns out that the person who makes the second response ends by dominating the interaction. The Green-Kemeny model was confirmed with respect to the predictions concerning domination of the interaction by a slower learner. With stimulus sets of 10 and sampling 3 stimuli at a time, the faster learner not only converged toward the slower learner’s initial position, he was soon trapped in it. Using the Monte Carlo technique, convergence was not to an intermediate point, but to the boundary condition itself. With a larger sampling ratio, the same was true of the advantage enjoyed by a person entering second into the interaction. With a smaller sampling ratio, sampling one stimulus at a time, the first starter converges at a point near the initial probability of the second starter without, however, being trapped at the boundary.

Forty or fifty responses into this condition yields data that clearly demonstrate the onset of oscillation. After the initial convergence, whatever it may be, the process settles down essentially into the matching process investigated by Neimark. Oscillation also appears with the matching of response probabilities to a constant probability of reinforcement.

Estes has defined response probability theoretically as the proportion of conditioned stimulus elements in the sample. This number is immediately accessible from the simulation computations, and its relationship to whether a left- or right-hand response results is that it is more sensitive than response probabilities calculated either over a series of trials or as the mean response of a group of subjects. It was always Skinner’s contention that averaged curves obscure more information than they reveal. This certainly is the case in these experiments.

As it happens, the sampling process with all its attendant effects on response probability has the same nonlinear characteristics found in other fields. Long range behavioral predictions are as impossible to achieve as are long range weather forecasts, and for the same reasons. The intrinsic probabilistic nature of behavior qualifies it for membership in the same class as quantum events, and the results of this experiment
empirically validate the inclusion of behavior in this class. The results provide both a powerful confirmation of the appropriateness of a stochastic model as applied to behavior, and a further clarification of the hazards of applying deterministic assumptions to behavior. They empirically support a proposed nondeterministic paradigm on which behavioral analyses must be based (Green, 2002).

Finally, our results may illuminate a question that has long perplexed physicists, namely, where do the laws of classical physics give way to those of quantum mechanics (Wick, 1995)? From the present data, convergence seems to be a classical phenomenon that depends, both in its mathematics and in its phenomenal existence upon an averaging over two or more particles, or holons. In this case, holons are represented by the subjects of the experiment. A convergent asymptote could not have been determined without an averaging of subject probabilities, and, as it turns out, the convergence of the behaviors of individual subjects is overruled by the probabilistic nature of their responses.

Behaviorally, this vindicates B. F. Skinner’s argument against averaging data. He insisted that smooth mean curves are purchased at the unacceptable cost of losing information inherent in the fine structure of individual response curves. His intuitive concern is validated by the shifting of gears between what is essentially a nonlinear probabilistic process of the individual, and an averaged entity that exists in no reality except as some statistical ideal. While the asymptotic convergence does occur for the masses, reliance upon grouped data obliterates the evidence for the most significant underlying properties of the stochastic process.

References


