A PROBABILISTIC MODEL FOR SIMPLE LEARNING

Jan Coufal

Abstract

In this paper I shall present an example the experimental extinction of learned response and the basic structure of a mathematical model designed to describe some simple learning situations, with special attention to the acquisition and extinction of behaviour habits in the Graham-Gagné runway and the Skinner box. As measure of behaviour I have chosen probability, that the instrumental response will occur during a specified time. I conceive that probability is increased or decreased amount after each occurrence of the response that the determinants of the amount of change in probability are the environmental events and the work or effort expending in making the response. It is probability model for simple learning, that is examined the first-order linear difference equations with constant coefficients. Equations of mean latent time as a function of trial number are derived for the runway problem. Equations for the mean rate of responding and cumulative numbers of responses versus time are derived for the Skinner box experiments. An example is the particle and wave interpretation of atomic theory.

Key words: linear difference equation, homogeneous difference equation, auxiliary equation, Markov property, probability, simple learning

JEL Code: A12, C02

Introduction

This paper continues and extends the article Coufal (2012). During the Second World War, developments in engineering, mathematical logic and computability theory, computer science and mathematics, and the military need to understand human performance and limitations, brought together experimental psychologists, mathematicians, engineers, physicists, and economists. Out of this mix of different disciplines mathematical psychology arose. Especially the developments in signal processing, information theory, linear systems and filter theory, game theory, stochastic processes and mathematical logic gained a large influence on psychological thinking. (Batchelder, 2002)

Two seminal papers on learning theory in Psychological Review helped to establish the field in a world that was still dominated by behaviorists: A paper by Robert R. Bush and Frederick A. Mosteller (1951) instigated the linear operator approach to learning, and a paper by William K. Estes (1950) that started the stimulus sampling tradition in psychological theorizing. These two papers presented the first detailed formal accounts of data from learning experiments. Mathematical psychology is a sub-field of psychology that started in the 1950s and has continued to grow as an important contributor to formal psychological theory, especially in the cognitive areas of psychology such as learning, memory, classification, choice response time, decision making, attention, and problem solving.

1 The Model

Let us that subject is introduced into following oversimplified learning situation:

a) a stimulus is presented,

b) the subject may or may not react to this stimulus, but

c) if he does respond positively, he is by some means discouraged from repeating this response.

To fix ideas, consider the example in which a rat, previously perfectly conditioned to running a straight runway to find food at its end, is placed in the starting box. In a specified subsequent time interval (long enough to permit completion of the run, but not long enough to allow dawdling along the way) the rat either makes the complete run or does not. If he does, he is disappointed to find that the food reward is not longer present. Let us call the completion of the run in the allotted time a positive response. After this trial run, the rat is once again placed in the starting box and another trial takes place. In this way the rat is subjected to many repeated trial runs, each of which may or may not result in a positive response. If we imagine a large number of rats similarly, but independently, used as subjects in these repeat runway trials, than we can compute the proportion of rats responding positively in trial number 1, in trial number 2, etc. Intuitively, we expect that the response of running to the end of the runway in the fixed allotted time interval will be extinguished owing to the absence of reward and that this “learning” will manifest itself in an ever-decreasing proportion of the rats who respond positively.

Now it actually turns out to be more convenient for the mathematical model to study a subject’s probability of making a positive response, rather than the proportion of positive responses in large group of subjects. Of course, when using this probability model one ordinary takes this empirical proportion as an estimate of the theoretical probability. Since the subject learns as the experimental trials are run, the probability of positive response will change trial to trial.

Suppose we wish to study how behavior changes under certain experimental conditions. We think, in particular, of a sequence of events starting with the perception of stimulus, followed by the performance of a response (pressing bar running maze, etc.), and ending with the occurrence of an environmental event (presentation of food, electric shock, etc.).

Behavior is measured by the probability, p, that the response will occur during some specified time interval after the sequence is initiated. The general idea is that p denotes the subject's level of performance and is increased or decreased after each occurrence of the response according as the environmental factors are reinforcing or inhibiting.

If we imagine an experiment in which a subject is repeatedly exposed to this sequence of events (stimulus-response-environmental event), we may divide the experiment into stages, each stage being a trial during which the subject is run through the sequence. The subject's level of performance is then function of the trial number, denoted by n, and we let p_n be the probability of response (during the specified time interval following the stimulus) in the nth trial run. The number p_0 will be taken as the initial value describing the disposition of the subject toward the response when he is first introduced to the experiment proper. The function p is then defined with domain the set of n-values 0, 1, 2, $\cdots$. By calling p a probability we impose the norming

$$\underset{n=0,1,2,\ldots}{\forall} \left(0 \le p_n \le 1\right), \tag{1}$$

which merely identifies the extremes of no response and certain response with the values 0 and 1 respectively.

We first assume that p_{n+1} depends on p_n only and not on earlier values of the function p. In other words, the subject's performance in trial $n+1$, although dependent on his level of behavior in the preceding trial (as measured by p_n), is independent of the past record of performance leading to trial n. This is referred to as the Markov property (Feller, 1957, p. 369) of the model. Following (Bush & Mosteller, 1951), we make the simplifying assumption that this dependence of p_{n+1} on p_n is a linear one. The slope-intercept form of the equation of this straight line is

$$\underset{n=0,1,2,\ldots}{\forall}\left(p_{n+1}=a+m\ p_n\right), \tag{2}$$

where a is the intercept and m is the slope of the line.

For our purposes, it is more convenient to write this linear relation in the "gain-loss" form. We introduce the parameter b by defining equation

$$m=1-a-b \tag{3}$$

so relation (2) may be written as

$$\underset{n=0,1,2,\ldots}{\forall}\left(p_{n+1}=p_n+a\ \left(1-p_n\right)-b\ p_n\right), \tag{4}$$

If a and b are know, then we may likewise determine m.

If the subject's level of performance at trial number n is given by p_n, then $1-p_n$ is the maximum possible increase in level and $-p_n$ is the maximum possible decrease in moving to trial $n+1$. This follows since 1 and 0 are the largest and smallest values of p_{n+1}. Equation (4) may be translated by saying that the change in performance level, $\Delta\ p_n=p_{n+1}-p_n$ is proportional to the maximum possible gain and maximum possible loss[1]. The constant of proportionality are a and b and we may therefore measure by the parameter a those environmental events which are reinforcing (e.g., presenting a reward) and by b those event which are inhibiting (e.g., punishing the subject).

Restriction on a and b are imposed only order to ensure that no matter what value p_n has, consistent with (1), the following value, p_{n+1}, will also be between 0 and 1 inclusive. If $p_n=0$, then $p_{n+1}=a$, so we require

$$0\le a\le 1. \tag{5}$$

If $p_n=1$, then $p_{n+1}=1-b$, so $(1-b)$ must be between 0 and 1 inclusive. But this means we require

$$0\le b\le 1. \tag{6}$$

We have proved that conditions (5) a (6) are necessary for p_{n+1} to be between 0 and 1 inclusive. It is not hard to show that they are also sufficient conditions. If $0\le a\le 1$ and $0\le b\le 1$, then $p_{n+1}=p_n+a\ \left(1-p_n\right)-b\ p_n\le p_n+1\cdot\left(1-p_n\right)-0=1$ and

[1] It is for this reason that (4) was named the "gain-loss" form.

$p_{n+1} = p_n + a\ (1 - p_n) - b\ p_n \geq p_n + 0 \cdot (1 - p_n) - p_n = 0$.

These restrictions are the only ones imposed on the parameters a and b appearing in the fundamental equation (4). Thus, $a = 0$ describes a situation in which no reward is given after the response occurs, $b = 0$ describes a no-punishment trial, and $a = b$ implies that the measures of reward and punishment are equal.

The apparatus used in this study consists of a runway, at one end of which is a starting box and at the other end a food box. The time taken by the animal to leave the starting box before traversing the runway to food (the latent period) was used as the measure of response. Quoting (Bush & Mosteller, 1951), "we may now describe the progressive change in the probability of response in an experiment such as the Graham-Gagné runway or Skinner box in which the same environmental events follow each occurrence of the response."

2 The Solution

To return to the general case, we note that (4) may be rewritten in the standard form

$$\underset{n=0,1,2,\ldots}{\forall} \left(p_{n+1} - (1 - a - b)\ p_n = a\right). \tag{7}$$

This we recognize as a linear first-order difference equation with constant coefficients. The homogeneous difference equation corresponding to (7) is

$$\underset{n=0,1,2,\ldots}{\forall} \left(p_{n+1} - (1 - a - b)\ p_n = 0\right), \tag{8}$$

and the auxiliary equation to (8) is

$$\lambda - (1 - a - b) = 0.$$

Then the general solution of the homogeneous equation (8) is

$$\overline{p}_n = C\ (1 - a - b)^n,$$

where C is arbitrary real constant.

To find a particular solution of (7) we use a trial solution of the form $p_n^* = k$, a constant. If this to satisfy (7), we must have $k - (1 - a - b)\ k = a$, i.e. $k = \frac{a}{a+b}$, so a particular solution of (7) is given by $p_n^* = \frac{a}{a+b}$.

Thus, the general solution of equation (7) has the form

$$p_n = \overline{p}_n + p_n^* = C\ (1-a-b)^n + \frac{a}{a+b}. \tag{9}$$

The number p_0 is the initial value. We therefore have the solution:

for $n = 0,\ 1,\ 2,\ \cdots$

$$p_n = \begin{cases} (1-a-b)^n \left(p_0 - \frac{a}{a+b}\right) + \frac{a}{a+b} & \text{if } \ a+b \neq 0, \\ p_0 & \text{if } \ a+b = 0. \end{cases} \tag{10}$$

In view (5) a (6) if $a = b = 1$, the sequence (p_n) oscillates between the two values p_0 and $1 - p_0$. But in all other cases the sequence (p_n) converges, to the limit p_0 if $a = b = 0$, and to limit $\frac{a}{a+b}$ otherwise. If $0 < a+b < 1$, then the sequence (p_n) is monotone decreasing to $\frac{a}{a+b}$ if $p_0 > \frac{a}{a+b}$, monotone increasing to $\frac{a}{a+b}$ if $p_0 < \frac{a}{a+b}$. If $1 < a+b < 2$, then the sequence (p_n) is a damped oscillatory sequence with limit $\frac{a}{a+b}$. The special case $a+b=0$ yields a constant sequence with value p_0; $a+b=1$ produces a sequence each of whose elements is $\frac{a}{a+b}$.

Let us conclude with two special cases:

(a) $a = 0$ and

(b) $a = b$.

Case (a) assumes that no reward is given after the response occurs. The difference equation (7) becomes

$$\underset{n=0,1,2,\ldots}{\forall} \left(p_{n+1} - (1-b)\ p_n = 0\right)$$

with solution

$$p_n = (1-b)^n\, p_0 .$$

This is an equation which describes the steady decrease in response probability (as $n \to \infty$) from initial probability p_0.

In case (b), $a = b$, the measures of reward and punishment are equal. If extreme cases $a = b = 0$ and $a = b = 1$ are discounted, then the quantity $(1 - a - b)^n \to 0$ as $n \to \infty$ and solution (10) shows that $p_n \to \frac{a}{a+b}$, which equal to $\frac{1}{2}$ in case (b), $a = b$ That is, ultimately the response tends to occur (in the specified time interval after stimulus is presented) in half the trials. The balancing of reward and punishment forces produces, in the long run, a corresponding symmetry in performance.

Conclusion

A large number of examples of this kind leads to a mathematical analysis involving difference equations. Mathematical modeling has been used to solve problems not only in engineering and physics, but also in biology and psychology. Most mathematics questions are neat and pure and simple. But the real world is often "messy". There are facts that get in the way, the numbers don't always work out nicely, we have to convert answers to different units and so on. Mathematical modeling is a process of representing real world problems in mathematical terms in an attempt to find solutions to the problems. A mathematical model can be considered as a simplification or abstraction of a (complex) real world problem or situation into a mathematical form, thereby converting the real world problem into a mathematical problem. The mathematical problem can then be solved using whatever known techniques to obtain a mathematical solution. This solution is then interpreted and translated into real terms. (Arora & Rogerson, 1991) The crucial problems of justifying the assumptions to be made and of testing usefulness of this particular mode of analysis in the social sciences are not within our purview. These are problems for the social scientist, not the mathematician.

References

Arora, M. S., & Rogerson, A. (1991). Future trends in mathematical modelling and applications. *Teaching of Mathematical Modelling and Applications*, 111-116.

Batchelder, W. H. (2002). Mathematical psychology. In A. E. Kazdin (Ed.), *Encyclopedia of Psychology*. Washington/New York: APA/Oxford University Press.

Bush, R. R., & Mosteller, F. A. (1951). Mathematical model for simple learning. *Psychological Review*, *58*, 313 – 323.

Coufal, J. (2012). A stochastic model for simple learning. *MUNDUS SYMBOLICUS*, *20*, 13-18.

Coufal, J. (2012a). *An application of discrete mathematical model in economics*. In *Mezinárodní statisticko-ekonomické dny na VŠE* (p. 249–256). Praha: Vysoká škola ekonomická v Praze.

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, *57*(2), 94-107. doi: 10.1037/h0058559

Feller, W. (1957). *An introduction to probability theory and its applications*. (2nd ed., Vol. 1). New York: John Wiley.

Contact

Jan Coufal

Vysoká škola ekonomie a managementu (University of Economics and Management), Prague

Nárožní 2600/9a, 158 00 Prague 5, Czech Republic

e-mail: jan.n.coufal@seznam.cz, jan.coufal@vsem.cz