

A LIMIT THEOREM OF A FUNCTIONAL OF REGULAR MARKOV PROCESSES AND ITS APPLICATIONS TO LEARNING MODELS

Sugiman, Ikuo
Department of Mathematics, Faculty of Science, Kyushu University

<https://doi.org/10.5109/13340>

出版情報 : Bulletin of informatics and cybernetics. 20 (1/2), pp.115-125, 1982-03. Research
Association of Statistical Sciences

バージョン :

権利関係 :

A LIMIT THEOREM OF A FUNCTIONAL OF REGULAR MARKOV PROCESSES AND ITS APPLICATIONS TO LEARNING MODELS

By

Ikuo SUGIMAN*

(Received December 1, 1981)

Abstract

We introduce a new interpretation of a phenomenon followed by certain subsequent learning experiments such as shift problems in discrimination learning, and show some limit theorems representing the rightness of such experiments on the basis of our interpretation.

In this paper we treat a regular Markov chain $\{X_n\}_{n \geq 0}$ as a mathematical model for subjects' behavior in the first phase and a family of stopping times $\{N_g(d)\}_{d > 0}$ for selecting the time when we change the first phase into the second phase, and show that the asymptotic distribution of $\{f(X_{N_g(d)})\}_{d > 0}$ is the same as the asymptotic distribution of $\{f(X_n)\}_{n \geq 0}$ under some conditions.

1. Introduction.

Since R. R. Bush and F. Mosteller proposed a linear model as a mathematical model for two-choice simple learning experiments in 1951, many psychologists have proposed various Markov chain models and studied them. These models were generalized and analyzed by means of probabilistic methods by M. Iosifescu and R. Theodorescu [4] and M. F. Norman [7]. On the other hand learning experiments have come specific and complex in recent years. In this paper we shall introduce a new interpretation of a *phenomenon* followed by certain subsequent experiments which cause different events in reinforcing each other as the results of a series of the so-called shift problems. And we shall show some limit theorems representing the rightness of such experiments on the basis of our interpretation. Hereafter the word '*an experiment*' is used to represent the whole one of a series of subsequent '*phases*'.

First we describe the above-mentioned experiments more precisely. For the sake of simplicity we treat the experiment composed of two phases. Each phase consists of a sequence of trials. On each trial a subject is given a stimulus in accordance with a schedule settled in advance, and responses are observed. Such trials are repeated until he learns sufficiently. A rule for stopping the phase is to stop it when the probability of his correct response approaches one, although his inner state he

* Department of Mathematics, Faculty of Science, Kyushu University, Fukuoka.

has had at the stopping time is not necessarily near from the state obtained by the perfect learning.

Peculiarities in the two phases are as follows. Main differences are how to give the stimulus to the subject by experimenters and how to respond by the subject. The two phases can be represented by a common model except directions of the change of state. The spaces of inner states of the subject are same between two phases. The subject does not know the time when the second phase starts. His initial state for the second phase coincides with his final state for the first one.

Considering these points we introduce the following interpretation different from the conventional one. Main aims of the experiment are to know the behavior of the subject relative to his learning at the first phase, and to observe a quantity attained by the subject as a consequence of his learning at the first phase. This quantity can not be observed by the experimenter at the first phase. So, we prepare the second phase in order to observe the quantity. Namely, at the second phase we estimate the final state of the first phase the subject obtained and we study how he learned from his observations in the first phase.

In this paper we study the behavior of the subject in the first phase represented by a Markov chain and we give a limit theorem which is satisfied by the observations in the second phase if the subject sufficiently learns in the first phase.

In Section 4, we shall treat the ZHL model and the overlearning reversal effect as examples of 'an experiment' and 'a phenomenon', and apply our results to them.

2. Formulation.

In this paper, we consider a Markov chain $\mathcal{X} = \{X_n\}_{n \geq 0}$ with state space (X, \mathcal{B}) and transition operator U , as a mathematical model for subjects' behavior in the first phase in such experiments. Next let $B(X)$ be the Banach space of bounded \mathcal{B} -measurable real valued functions on X under the supremum norm $|\cdot|$ and let L be a Banach algebra, contained in $B(X)$, under a norm $\|\cdot\|$ such that

$$C := \sup_{f \in L: f \neq 0} \frac{|f|}{\|f\|} < \infty.$$

In addition, the transition operator U is assumed to satisfy one of following conditions.

DEFINITION 1. The operator U is *aperiodic* if there is a bounded linear operator U^∞ on L such that $\|U^n - U^\infty\| \rightarrow 0$ as $n \rightarrow \infty$.

An aperiodic operator U is *regular* if $U^\infty f$ is a constant function on X for each $f \in L$.

Examples of the Markov chain \mathcal{X} and the Banach algebra $(L, \|\cdot\|)$ are described in [7]. We suppose that functions observed in both phases belong to L , then the observation in the second phase is expressed as follows. Let $g(\in L)$ be the function we observe for selecting the time when we change the first phase into the second phase. As stated in the previous section, the time is a stopping rule in the first phase, and it is the time when $g(X_n)$ sufficiently approaches to $U^\infty g$ to which the observed values approach if we don't change the phase and keep on making experiments of the first phase. Let N be this time and let $f(\in L)$ be the function we observe in the second

phase. Then we can express the observed value in such an experiment by $f(X_N)$. So if we use these expressions, we may formulate the problem mentioned in the previous section what value $f(X_N)$ approaches to when the difference between $g(X_N)$ and $U^\infty g$ is very small. In this paper we show that $f(X_N)$ approaches to $U^\infty f$ as well as $f(X_n)$ does and survey conditions for f and g to satisfy.

We cite some results which were proved by M. F. Norman [7].

LEMMA 1. *If U is aperiodic, and $V=U-U^\infty$, then*

$$r(V)=\lim_{n\rightarrow\infty}\|V^n\|^{1/n}<1.$$

It follows from this that for any $r(V)<\alpha<1$ there is a $D=D_\alpha$ such that

$$\|V^n\|\leq D\alpha^n$$

for all $n\geq 0$.

For $n\geq 1$, $u\in Z$ and $f\in L$, let

$$\begin{aligned} S_n(f) &= \sum_{j=0}^{n-1} f(X_j), \\ Z_n(f) &= (S_n(f) - nU^\infty f) / \sqrt{n}, \\ \rho_u(f) &= U^\infty(fV^{|u|}f), \end{aligned}$$

and

$$\sigma^2(f) = \sum_{u=-\infty}^{\infty} \rho_u(f).$$

We use abridged notations S_n , Z_n , ρ_u and σ^2 as far as we don't demand.

THEOREM 1. *If U is regular, then*

- (1) S_n/n converges to $U^\infty f$ almost surely as $n\rightarrow\infty$, and
- (2) the distribution of Z_n converges to the normal distribution with mean 0 and variance σ^2 .

Let $\{t_n\}_{n\geq 1}$ be a sequence of positive numbers such that for all $n\geq 1$, $t_n < n$ and

$$t_n \rightarrow \infty \text{ and } t_n/n \rightarrow 0 \text{ as } n \rightarrow \infty$$

and let

$$\sigma_n = \sigma_n(f) = \frac{1}{n} \sum_{|u|\leq t_n} \sum_{j=0}^{n-1-|u|} \left(f(X_j) - \frac{S_n}{n} \right) \left(f(X_{j+|u|}) - \frac{S_n}{n} \right).$$

Then

- (3) the truncated estimator σ_n converges to σ^2 in quadratic mean as $n\rightarrow\infty$.

We use the well-known Anscombe theorem to introduce stopping rules into these limit theorems.

THEOREM 2. *Let $\{Y_n\}_{n\geq 1}$ be a sequence of random variables. We suppose that there exist a sequence of positive numbers $\{w_n\}_{n\geq 1}$ and a random variable Y_∞ such that Y_n/w_n converges to Y_∞ in distribution as $n\rightarrow\infty$. Let $\{N_r\}_{r\geq 1}$ be a sequence of positive integer-valued random variables defined on the same probability space that $\{Y_n\}_{n\geq 1}$ is defined. We assume that N_r/n_r converges in probability to 1 as $r\rightarrow\infty$ where $\{n_r\}_{r\geq 1}$ is an increasing sequence of positive integers tending to infinity when r tends to infinity. Then the sequence*

Y_{N_r}/w_{n_r} converges to Y_∞ in distribution as $r \rightarrow \infty$ if the following condition is satisfied;
 For every $\epsilon > 0$ and $\eta > 0$ there exist a small positive number c and a natural number n_0 such that for every $n > n_0$

$$P(\max_{m \geq 1; |m-n| < cn} |Y_m - Y_n| > \epsilon) < \eta.$$

The last condition is called the Anscombe condition.

3. Main results.

While we assumed the condition only for the Markov chain \mathcal{X} in Theorem 1, we add an assumption for $f \in L$ which is satisfied in many mathematical models and give some results corresponding to each one in Theorem 1.

THEOREM 3. *If U is aperiodic and $U^\infty f^2 = (U^\infty f)^2$, then*

- (1) $f(X_n)$ converges to $U^\infty f$ almost surely as $n \rightarrow \infty$,
- (2) Z_n converges to 0 almost surely as $n \rightarrow \infty$, and
- (3) σ_n converges to 0 almost surely as $n \rightarrow \infty$.

PROOF. We shall assume that $U^\infty f^2 = U^\infty f = 0$. The general case is obtained by applying this special case to $f' = f - U^\infty f \in L$. For $n \geq 0$, let $\mathcal{F}_n = \sigma(X_i; 0 \leq i \leq n)$ be the smallest Borel field with respect to which X_0, X_1, \dots, X_n are measurable.

Firstly, for $m, j \geq 0$

$$E[f^2(X_{m+j}) | \mathcal{F}_m] = U^j f^2(X_m) = V^j f^2(X_m) \quad \text{a. s.}$$

since $U^\infty f^2 = 0$, so that

$$|E[f^2(X_{m+j}) | \mathcal{F}_m]| \leq |V^j f^2| \leq C \|V^j f^2\| \leq C D \alpha^j \|f^2\|$$

where $r(V) < \alpha < 1$, by Lemma 1. This is expressed

$$E[f^2(X_{m+j}) | \mathcal{F}_m] = O(\alpha^j)$$

in this paper. We use such abbreviated notation when the order of magnitude is uniform over all relevant variables that don't appear on the right. It follows that

$$(3.1) \quad E[f^2(X_j)] = E[E[f^2(X_j) | \mathcal{F}_0]] = O(\alpha^j),$$

hence

$$E[\sum_{j=0}^{\infty} f^2(X_j)] < \infty$$

$$\sum_{j=0}^{\infty} f^2(X_j) < \infty \quad \text{a. s.},$$

and $f(X_j) \rightarrow 0$ a.s. as $j \rightarrow \infty$.

Next it follows from nonnegativity of the bounded linear operator U^∞ on L that for $g, h \in L$

$$U^\infty(gh) \leq (U^\infty g^2)^{1/2} (U^\infty h^2)^{1/2},$$

so that $U^\infty(fg) = 0$ for $g \in L$ since $U^\infty f^2 = 0$. Thus if $0 \leq i \leq j \leq k \leq m$, we can write

$$E[f(X_i)f(X_j)f(X_k)f(X_m)] = E[V^i(fV^{j-i}(fV^{k-j}(fV^{m-k}f)))(X_0)],$$

and get

$$(3.2) \quad |E[f(X_i)f(X_j)f(X_k)f(X_m)]|=0(\alpha^m).$$

It follows that

$$(3.3) \quad E[S_n^4]=0(1),$$

since

$$|E[S_n^4]| \leq \sum_{i,j,k,m=0}^{n-1} |E[f(X_i)f(X_j)f(X_k)f(X_m)]| \leq 4! \sum_{0 \leq i \leq j \leq k \leq m \leq n-1} |E[f(X_i)f(X_j)f(X_k)f(X_m)]|,$$

and

$$\sum_{0 \leq i \leq j \leq k \leq m} \alpha^m = 1/(1-\alpha)^4 < \infty.$$

On division by n^2 , (3.3) yields

$$E[Z_n^4]=0(1/n^2)$$

hence

$$E[\sum_{n=1}^{\infty} Z_n^4]=0(1),$$

$$\sum_{n=1}^{\infty} Z_n^4 < \infty \quad \text{a. s.}$$

and $Z_n \rightarrow 0$ a. s. as $n \rightarrow \infty$.

Clearly

$$\sigma_n = \mathcal{A}_n + \varepsilon_n$$

where

$$\mathcal{A}_n = \frac{1}{n} \sum_{|u| \leq t_n} \sum_{i=0}^{n-1-|u|} f(X_i)f(X_{i+|u|})$$

and

$$\varepsilon_n = \frac{1}{n^2} \sum_{|u| \leq t_n} S_n(\sum_{i=n-|u|}^{n-1} f(X_i) + \sum_{i=|u|}^{n-1} f(X_i)).$$

From (3.2) we get

$$E[f^2(X_i)f^2(X_{i+|u|})]=0(\alpha^{i+|u|})$$

for $i \geq 0$ and $u \in \mathbf{Z}$, so

$$E[f^2(X_i)f^2(X_{i+|u|})]^{1/2} = 0(\beta^{i+|u|})$$

where $\beta = \alpha^{1/2}$. By the Minkowski inequality

$$E[\mathcal{A}_n^2]^{1/2} = \frac{1}{n} \sum_{|u| \leq t_n} \sum_{i=0}^{n-1-|u|} 0(\beta^{i+|u|}) = 0(1/n),$$

and

$$E[\mathcal{A}_n^2]=0(1/n^2).$$

By the estimation similar to show (3.3) we get

$$E[S_n^2(\sum_{i=n-|u|}^{n-1} f(X_i) + \sum_{i=|u|}^{n-1} f(X_i))^2] = o(1),$$

and similarly

$$E[\varepsilon_n^2] = o(1/n^2).$$

Therefore Δ_n and ε_n , then σ_n converges to 0 almost surely. Thus the proof is complete.

Of course, Theorem 3 (2) gives the Anscombe condition for the central limit theorem in Theorem 1 by taking $Y_n = Z_n$ and $w_n = 1$ in Theorem 2. Under the same assumptions as Theorem 3, we also get the Anscombe condition generally used;

For every $\varepsilon > 0$ and $\eta > 0$ there exist a positive number $c (< 1)$ and a natural number n_0 such that for $n \geq n_0$

$$P(\max_{m \geq 1; |m-n| < cn} |S_{m,n}(f)| \geq \varepsilon \sqrt{n}) < \eta$$

where

$$S_{m,n}(f) = \begin{cases} \sum_{i=m}^{n-1} f(X_i) - (n-m)U^\infty f & \text{if } m < n \\ 0 & \text{if } m = n \\ \sum_{i=n}^{m-1} f(X_i) - (m-n)U^\infty f & \text{if } m > n. \end{cases}$$

PROPOSITION. If U is aperiodic and $U^\infty f^2 = (U^\infty f)^2$, then for $\varepsilon > 0$ and $0 < c < 1$

$$P(\max_{m \geq 1; |m-n| < cn} |S_{m,n}(f)| \geq \varepsilon \sqrt{n}) = o(1)$$

as $n \rightarrow \infty$.

PROOF. We can assume that $U^\infty f^2 = 0$ without loss of generality. By the Schwartz inequality and (3.1)

$$\begin{aligned} & E[\max_{m \geq 1; |m-n| < cn} |S_{m,n}(f)|] \\ & \leq E[\max_{m \geq 1; |m-n| < cn} \sqrt{|m-n|} (S_{m,n}(f^2))^{1/2}] \\ & \leq \sqrt{cn} E[(\sum_{i \geq 1; |i-n| < cn} f^2(X_i))^{1/2}] \\ & \leq \sqrt{cn} (\sum_{i \geq 1; |i-n| < cn} E[f^2(X_i)])^{1/2} \\ & = o(\sqrt{n} \cdot \beta^{(1-c)n}) \end{aligned}$$

where $0 < \beta (= \alpha^{1/2}) < 1$. Hence, by the Markov inequality

$$\begin{aligned} & P(\max_{m \geq 1; |m-n| < cn} |S_{m,n}(f)| \geq \varepsilon \sqrt{n}) \\ & \leq E[\max_{m \geq 1; |m-n| < cn} |S_{m,n}(f)|] / \varepsilon \sqrt{n} \\ & = o(\beta^{(1-c)n}) = o(1). \end{aligned}$$

This completes the proof.

The following lemma gives a procedure to constitute the families of stopping times that satisfy the condition for $\{N_r\}_{r \geq 1}$ in Theorem 2.

LEMMA 2. Let $\{Y_n\}_{n \geq 1}$ be a sequence of random variables that converges to a positive random variable \bar{Y} almost surely as $n \rightarrow \infty$, and let $k(n)$ be a sequence of positive constants such that

$$\lim_{n \rightarrow \infty} k(n) = \infty \quad \text{and} \quad \lim_{n \rightarrow \infty} \frac{k(n)}{k(n-1)} = 1.$$

For each $d > 0$, define

$$N_d = \inf \{n \geq 1; 0 < Y_n \leq d^2 k(n)\}.$$

Then the stopping times $\{N_d\}_{d > 0}$ are well-defined and $d^2 k(N_d)$ converges to \bar{Y} in probability as $d \downarrow 0$.

PROOF. As well as the proof of Lemma 1 of Y. S. Chow and H. Robbins [3] we can obtain that $\{N_d\}_{d > 0}$ are well-defined,

$$\lim_{d \downarrow 0} N_d = \infty \text{ a.s. ,}$$

$$\lim_{d \downarrow 0} Y_{N_d} = \lim_{d \downarrow 0} Y_{N_{d-1}} = \bar{Y} \text{ a.s. ,}$$

and

$$Y_{N_d} \leq d^2 k(N_d) < [k(N_d)/k(N_d-1)] Y_{N_{d-1}}$$

on $A_d = \{Y_{N_{d-1}} > d^2 k(N_d-1)\}$. Then for $\epsilon > 0$

$$\begin{aligned} &P(|d^2 k(N_d) - \bar{Y}| \geq \epsilon) \\ &\leq P(A_d \cap \{|d^2 k(N_d) - \bar{Y}| \geq \epsilon\}) + P(A_d^c) \rightarrow 0 \end{aligned}$$

as $d \downarrow 0$. This completes the proof.

Particularly if \bar{Y} is a constant function y , the stopping times $\{N_d\}_{d > 0}$ satisfy the condition of Theorem 2 for $n_d = [y/d^2]$.

In the remainder of this section we get these results into the limit theorems of processes of the form $f(X_N)$. Henceforth let g and f be, respectively, the function observed in the first and second phases as these were in the previous section. We suppose that U is regular and $f(\in L)$ satisfies the Anscombe condition above mentioned. For $d > 0$, let

$$N_1(d) = \inf \{n \geq 1; |S_n(g) - nU^\infty g| \leq n(nd^2 - 1)\}$$

$$N_2(d) = \inf \{n \geq 1; |g(X_n) - U^\infty g| \leq nd^2 - 1\},$$

and

$$T_i(d) = \frac{1}{d} \left(\frac{\sum_{j=0}^{N_i(d)} f(X_j)}{N_i(d)} - U^\infty f \right).$$

Then the distribution of $T_1(d)$ converges to the normal distribution with mean 0 and variance $\sigma^2(f)$, and if $U^\infty g^2 = (U^\infty g)^2$, the distribution of $T_2(d)$ also converges to the same

distribution. And if $\sigma_n(g)$ converges almost surely, let

$$N_3(d) = \inf \{n \geq 1; 0 < \sigma_n(g) \leq nd^2\}$$

$$N_4(d) = \inf \{n \geq 1; |\sigma_n(g)| \leq nd^2 - \varepsilon\} \quad (\varepsilon > 0)$$

for $d > 0$. Then if $\sigma^2(g)$ is positive, the distribution of $T_3(d)$ converges to the normal distribution with mean 0 and variance $\sigma^2(f)/\sigma^2(g)$, and if not, for any $\varepsilon > 0$ the distribution of $T_4(d)$ converges to the normal distribution with mean 0 and variance $\sigma^2(f)/\varepsilon$ as $d \downarrow 0$.

4. Example.

In this section we take up the *ZHL model* which was proposed by D. Zeaman and B. J. House [9] and E. Lovejoy [5] as an example of the mathematical model for discrimination learning stated in introduction and we describe briefly about the phenomenon called the *overlearning reversal effect* and then we show that this example satisfies the framework and the conditions in this paper.

First we describe the typical experiment in discrimination learning. In this experiment stimuli vary along two dimensions such as form and color, position and brightness, etc. Each dimension is represented by two values such as circle and triangle, green and red, right and left, etc. In this section A and B stand for the two dimensions and the values in $A(B)$ is represented by a and \bar{a} (b and \bar{b}). We select the correct value to reinforce (a) and then we say that the dimension A is relevant. On any trial, the subject must choose one of two stimulus objects. On half the trials, the objects are (a, b) and (\bar{a}, \bar{b}) . On the remaining trials, they are (a, \bar{b}) and (\bar{a}, b) . The symbol (a, b) denotes the stimulus with both a and b such as a red circle and a green triangle. And his choice is rewarded only if his chosen stimulus contains the correct value. Such trials are repeated until the subject's performance meets some criterion representing that he learned to choose the rewarded stimuli.

Secondly we describe the *ZHL model* for the above experiment. The state space (X, \mathcal{B}) of the Markov chain $\mathcal{X} = \{X_n\}_{n \geq 0}$ is $[0, 1] \times [0, 1]$ with its Borel σ -field and its element is represented by $x = (v, w)$. The variable v denotes the probability of attending to the relevant dimension and then he choose the correct value with conditional probability w , that is,

$$v = P(A) \quad \text{and} \quad w = P(a|A).$$

We now describe how x change on each trial. This depends on what he attends to and whether his response is correct or incorrect. While the variable v changes on all trials, the variable w does only when he attends to the relevant dimension. And the variables approach to the state which perfectly learned the events reinforced on that trial. These changes are effected by linear transformations. The following table lists the values of changes of v and w corresponding to the combinations of attention and outcomes.

attention	outcome	probability	Δv	Δw
A (relevant)	correct	vw	$\phi(1-v)$	$\theta(1-w)$
	incorrect	$v(1-w)$	$-\phi v$	$\theta(1-w)$
B (irrelevant)	correct	$(1-v)/2$	$-\phi v$	0
	incorrect	$(1-v)/2$	$\phi(1-v)$	0

$$(0 < \phi, \theta < 1)$$

This Markov chain \mathcal{X} is a regular compact Markov chain with single absorbing state $x_\infty=(1, 1)$ [7]. Consequently this satisfies the regularity stated in formulation if $L=L(X)$ or $B(X)$ where $L(X)$ is the set of bounded Lipschitz functions. Besides the condition ' $U^\infty f^2=(U^\infty f)^2$ ' often assumed is satisfied by all functions in L since $U^\infty f$ is represented by $f(x_\infty)$.

Next we describe the overlearning reversal effect (ORE) as an example of the phenomenon ensued from subsequent experiments. There are two phases in this experiment. The experiment in the first phase is done in the same way above mentioned, but in the subsequent phase the correct value changes a into \bar{a} in the relevant dimension. As an interesting question it is whether extra trials before reversal in the first phase facilitates or retards the reverse learning on the second phase. The phenomenon that overtraining facilitates reversal is famous and is called the ORE, but Mackintosh reported the opposite result. In [6] huge results for this experiment are well arranged.

Now we consider the phenomenon on our interpretation described in introduction. The criterion to finish the first phase is that the subject's state $x=(v, w)$ satisfies

$$vw+(1-v)/2=1.$$

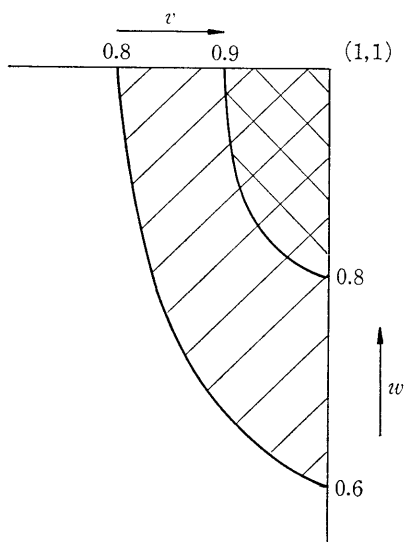


Fig. 1

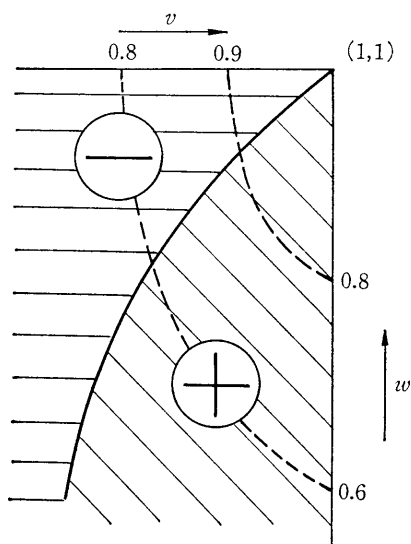


Fig. 2

But we can know no more about the stopping state. Then let f be the expected total number of errors on the second phase and this is the function of the initial state of the second phase, that is, the final state of the first phase. And we analyze the increase or decrease of this function as the result of one additional trial in the first phase. These results are illustrated in the above figures. When his state enters into the dark shading on figure 1, we change the phase and, in addition, if it is in $\oplus(\ominus)$ on figure 2, the function f increases (decreases). It comes to this that if the final state of the first phase is in $\oplus(\ominus)$, the overlearning retards (facilitates) reverse learning.

Acknowledgement

It is a pleasure to acknowledge the encouragement and helpful suggestions of my teachers, especially Professor N. Furukawa of Kyushu University and Mr. E. Isogai of Niigata University.

References

- [1] ANSCOMBE, F.J., *Large sample theory of sequential estimation*. Proc. Cambridge Philos. Soc., **48** (1952), 600-607.
- [2] BUSH, R.R. and MOSTELLER, F., A mathematical model for simple learning. Psychological Review, **58** (1951), 313-323.
- [3] CHOW, Y. S. and ROBBINS, H., *On the asymptotic theory of fixed-width sequential confidence intervals for the mean*. Ann. Math. Statist., **36** (1965), 457-462.
- [4] IOSIFESCU, M. and THEODORESCU, R., *Random processes and learning*. Springer-Verlag, N.Y., 1969.
- [5] LOVEJOY, E., *Analysis of the overtraining reversal effect*. Psychological Review, **73** (1966), 87-103.

- [6] MACKINTOSH, N. J. and SUTHERLAND, N. S., *Mechanisms of animal discrimination learning*. Academic Press, N.Y., 1971.
- [7] NORMAN, M. F., *Markov processes and learning models*. Academic Press, N.Y., 1971.
- [8] NORNAN, M. F., *Effects of overtraining, problem shifts, and probabilistic reinforcement in discrimination learning; predictions of an attentional model*. *Contemporary developments in mathematical psychology*, vol. 1, 185-208, 1974.
- [9] ZEAMAN, D. and HOUSE, B. J., *The role of attention in retardate discrimination learning*. *Handbook of mental deficiency*, 159-223, 1963.

Communicated by N. Furukawa