不確実な手がかりのもとでの選択：
予備的考察——確率学習について——

小　野　　　茂

不確実な手がかりのもとでの選択：
予備的考察——確率学習について——

人が判断を下したり，選択を行なうとき通常は何らかの手がかりを用いる。いかなる手がかりもなしに判断とか選択を行なうことは，まず無いと云ってよかろう。

しかし，その手がかりは妥当なものであるとは限らない。与えられた手がかりが，選択の結果を予想する上で客観的関係を反映したものであることもあろうが，多くの場合，手がかりというものは，多かれ少かれ不確実なものである。

むろん，手がかりの妥当性は，完全に妥当か，完全に非妥当であるかという悉無的なものではない。通常は，手がかりは，完全に妥当と完全に非妥当の中間，すなわち，ある程度妥当であると共に，ある程度非妥当なものであろう。程度に差はあるにせよ，人は，このような不確実な手がかりを用いて，判断とか選択を行っている。

しかも，その手がかりは，一般には複数個ある。ただ一つの手がかりにより判断したり選択することもあるかもしれないが，それはむしろ特殊な場合であって，普通は同時にいくつかの手がかりを用いる。

このように複数個の手がかりがあり，それらの手がかりは，いずれも程度は異にするが不確実なものである。この論文では，こうした複数個の不確実な手がかりを用いた選択行動の分析を目的としているが，本稿ではとくに〈確率学習〉probability learning を取上げて分析を試みることにしたい。

1. 多重手がかりの確率学習

複数個の手がかり，すなわち n 次元の手がかりが与えられるような確率学習は〈多重手がかりの確率学習〉multiple cue probability learning (以下，MCPL と略す) と呼ばれている (Castellan, N. J. 1977)。

確率学習の実験においては，ある選択を行なったとき，選択によって得られるアウトカム outcome (結果) が不確定的，確率的である。しかし，選択を行なったあと，確率的にせよアウトカムが与えられると，これはアウトカムについての〈情報のフィードバック〉として作用する。すなわち，〈結果の知識〉knowledge of result (KR と略される) が与えられたことになり，これは，次の機会での選択に影響を及ぼす。こうして，選択行動を反復している

と，結果の知識に基づいて行動は改善，あるいは改善ではないにせよ変容する。このような不確定的なアウトカムによる行動の変容が，確率学習である。とくに MCPL では，不確実な多重手がかりの下で選択が行われ，選択後には KR が与えられ，これの反復により選択行動が変容してゆく。

さて，用語の慣用によれば，単に〈確率学習〉と云ったときには，狭義の確率学習を指すことが多い。すなわち，〈単純確率学習〉か，あるいは，ただ一つの手がかりしか含まない〈弁別的確率学習〉のことを指すのである。この論文でも，とくに断らない限りは，〈確率学習〉という言葉を，このような狭義の意味で用いることにし，広義の確率学習，すなわち MCPL とは区別することにしよう。

しかし，当然のことながら，MCPL は狭義の確率学習と深い関係がある。上述のように，確率学習には，単純確率学習と弁別的確率学習が含まれる。ここで弁別的確率学習は単純確率学習に手がかりを導入したものであって，その意味で，弁別的確率学習は単純確率学習を拡張したものと云ってよかろう。MCPL は，この弁別的確率学習において，手がかりの次元数を 1 次元から n 次元に拡張したものである。すなわち，別の表現をすれば，MCPL の特殊な場合が弁別的確率学習であり，さらに，弁別的確率学習の特殊な場合が単純確率学習である。

このような関係があることから，MCPL の理解のためには，確率学習，すなわち単純確率学習と弁別的確率学習の両者を分析することが必要なことが分かる。本稿では，MCPL の特殊形態である確率学習について考察する。

MCPL は，いま述べたように確率学習を拡張したものであるが，それと同時に，いわゆる〈概念学習〉concept learning（概念学習については，例えば小野，1966参照）を拡張したものとも考えられる。MCPL は，一方では，手がかりの次元数を 1 次元から n 次元にすることにより弁別的確率学習を拡張したものであった。しかし，他方では，概念学習における手がかりの妥当性を完全なものから不完全なものへと変えた拡張でもある。通常の概念学習では手がかりの妥当性は完全であるけれども，MCPL では不完全である。（因みに，手がかりの妥当性は，不完全である方が一般的であって，完全であるのは不完全なときの特殊な場合である。）

MCPL を分析するためには，その特殊形態である確率学習を吟味するだけでなく，概念学習の面からも考察する必要がある。しかし，概念学習の面からの吟味は，別に稿を改めて考察することにし，以下，本稿では，MCPL に関する予備的考察として確率学習だけを取上げよう。

2．単純確率学習

まず，弁別的手がかりを含まない単純確率学習から見てゆくことにしよう。単純確率学習の実験手つづきはきわめて簡単である。毎試行，2つの事象 E_1，E_2 のいずれかが生じる。〈事象〉は，正確には〈結果の事象〉outcome event と云うべきであるが，慣例的には簡単に事象と呼ばれているので，ここでもそれにしたがっておく。被験者は，事象が生起する前に，どちらの事象が生じるかを予想する。そして，例えば，手許にある2つのボタンのいずれかを押す。（被験者の予想は2つの選択肢の一方を選ぶことであるから，一種の選択である。）このような被験者の予想は A_1，A_2 という記号で表わすことにする。ここで，A_1 は事象 E_1 が生じるという予想であり，A_2 は事象 E_2 が生じるという予想である。被験者が予想を行ったあと，実際に事象が生じて，被験者は自分の予想が的中したかどうかを知る。すなわち，情報のフィードバックが行われ，KR が与えられる。確率学習の実験では，被験者の予想 $A=A_1$ or A_2 がなされ，そのあと事象 $E=E_1$ or E_2 が生起する。1回の“AE”が1試行であって，このような試行を何回も（通常，数百試行）反復する。

確率学習については多くの実験が行われ，いろいろの事実が見いだされている（例えば小野，1966；Jones, M. R. 1971）。ここでは，その主なものを一，二あげるにとどめる。

まず注目されるのは，確率の〈一致〉matching と〈極大化〉maximizing である。確率学習の実験では，事象 E_1 と E_2 はそれぞれ，ある確率でもってランダムの順序で提示される。いま，事象 E_1 の生起確率 $p(E_1)$ が $\pi(0\leqq\pi\leqq1)$ であったとしよう。すなわち

$$p(E_1)=\pi$$

とする。（π は例えば.75。）そのとき，$p(E_2)=1-\pi$ である。また，被験者は，確率 p でもって，E_1 が生じると予想するものとしよう。すなわち

$$p(A_1)=p$$

とする。そのとき $p(A_2)=1-p$ となる。

被験者の予想確率 p は，試行を重ねるにつれて変化する。そこで，予想確率 p について学習曲線を描いてみる。そのとき，試行の初期では，被験者はやみくもに予想するので $p\fallingdotseq1/2$ である。しかし，試行を重ねるにつれて

$$p\rightarrow\pi$$

という傾向が生じることがある。このような現象，すなわち，多数回の試行を反復したとき，p の漸近値が π になるという現象を，確率の一致と呼んでいる。

ここで，確率学習の実験事態で，予想確率 p の最適値（適中率を最大にする値）を求めてみる。すなわち，多数回試行を行ったとき，適中率を最大にするには予想確率 p をどのような値にすればよいかという問題を考えてみる。詳細は省略するが，簡単な計算により，p の最適値は π の値に依存し

$$p=\begin{cases}1, & \pi>1/2\\ 0, & \pi<1/2\end{cases}$$

であることが分かる。（$\pi=1/2$ のときには，p の値のいかんによらず，適中率は1/2になる。）

実際，実験条件によっては（例えば，適中したときに賞を与える），予想確率 p（漸近値）は，最適化の方向にずれることがある。これを，確率の極大化と呼んでいる。

単純確率学習について，もう一つよく知られている現象は〈系列依存関係〉sequential dependency である。確率学習の実験では，被験者が予想 A を行い，その後，事象 E が生起し，このような試行が多数回くりかえされる。すなわち，$AEAE\cdots\cdots AE$ という予想・事象系列の後に，現試行の予想 A が行われる。現試行の予想 A は，過去に生じた予想・事象系列と関係があり，とりわけ前試行での予想と事象と関係が深い。

前試行で A_i（$i=1, 2$）と予想し，その後，事象 E_j（$j=1, 2$）が生じたときに，現試行で A_1 と予想する確率を $p(A_1|E_jA_i)$ と表わすことにしよう。$i, j=1, 2$ であるから，4通りの $p(A_1|E_jA_i)$ がある。これら4通りの確率の大きさについて考えてみよう。

事象 E_j（$j=1, 2$）を固定し，前試行で A_1 と予想したときと，A_2 と予想したときを比較する。前試行で A_1 と予想したことは，A_1 が生じやすいということであるから，次の不等式が成立するであろう。

$$p(A_1|E_jA_1)>p(A_1|E_jA_2)$$

次に，前試行での予想 A_i（$i=1, 2$）を固定して，E_1 が生じたときと E_2 が生じたときを比較する。そのとき，E_1 が生じれば，E_2 が生じたときよりも，次の試行では A_1 が生じやすいであろうと考えられる。そこで

$$p(A_1|E_1A_i)>p(A_1|E_2A_i)$$

上記の2つの不等式をまとめれば，

$$p(A_1|E_1A_1)>\{p(A_1|E_2A_1),\ p(A_1|E_1A_2)\}>p(A_1|E_2A_2)$$

と表わされる。すなわち，4通りの $p(A_1|E_jA_i)$ のうち，$p(A_1|E_1A_1)$ が最大，$p(A_1|E_2A_2)$ が最小であって，残りの2つが両者の中間の値になる。このような予測が成立つが，実際，実験の結果もほぼこのようになることが知られている（例えば，Suppes, P. & Atkinson,

R. C. 1960)。なお，$p(A_1|E_jA_i)$ について実験結果と照合するときには，多数回試行を重ねたあとの漸近状態（安定状態）において吟味するのが普通である。

3．学習の線形モデル

単純確率学習では，〈確率一致現象〉と〈系列依存関係〉とが実験的事実としてよく知られている。これらの現象に関連して，ここで，学習の〈線形モデル〉に触れておこう（Bush, R. R. & Mosteller, F. 1955)。

いま，予想 A_1 を行なう確率を P，予想 A_2 を行なう確率を $1-P$ とする。（ここで考えている予想確率は，過去の事象系列によって異なる値をとる確率変数であるので，大文字の P でもって表わす。前に示した小文字の $p(=p(A_1))$ は，A_1 の平均反応確率，すなわち，事象系列についての P の期待値である。）

学習の線形モデルでは，予想確率 P について，次のようなきわめて単純な仮定を設ける。ある試行で E_1 が生起すれば，次の試行で予想 A_1 がなされる確率 P は増加する。したがって，E_1 のみが反復生起すれば P は1に近づく。また，ある試行で E_2 が生起すれば，次の試行で P は減少し，E_2 のみが反復生起すれば P は0に近づく。

数式化に際しては，E_1 または E_2 の生起による P の変化は，P と1または0の加重平均（1次結合）によって示される。第 n 試行の P を，添字 n を付して P_n と表わし，また，第 n 試行の E_i $(i=1, 2)$ を $E_{i,n}$ と表わそう。そのとき，第 n 試行から第 $n+1$ 試行への P の変化は次式によって示される。

$$P_{n+1}=\begin{cases}\alpha P_n+(1-\alpha), & E_{1,n}\\ \alpha P_n & , E_{2,n}\end{cases}$$

ここで，α のとる値の範囲は $0\leqq\alpha<1$ であって，α の値が小さいほど P の変化は著しい。上記のようなモデルを〈学習の線形モデル〉と呼んでいる。（正確に云えば，線形モデルの一つの特殊な場合である。）

このモデルから次のような帰結が得られる。まず第一には，確率一致現象についてである。第 n 試行において予想 A_1 がなされる確率を p_n としよう。すなわち $p_n=p(A_{1,n})$ とする。また，前と同じく $\pi=p(E_1)$ とする。そのとき，上のモデルから

$$p_n=\pi-(\pi-p_1)\alpha^{n-1}$$

となる。したがって，$n\to\infty$ のとき

$$p_n\to\pi$$

であり，これは，試行を重ねるにつれて確率一致現象が生じることを示している。

上記の線形モデルから系列依存関係も誘導できる。系列依存関係は，漸近状態で吟味するのが普通であるから，$n\to\infty$のときの $p(A_1|E_jA_i)$ を取上げて，

$$\lim_{n\to\infty}p(A_{1,n+1}|E_{j,n},\ \ A_{i,n})=p(A_1|E_jA_i)$$

と表わすことにしよう（i，$j=1$，2）。上記のモデルから $p(A_1|E_jA_i)$ を計算してみる。事象 E についてのある簡単な条件のもとで（事象 E は予想 A と独立，また E の系列は独立過程），次式が成立つ。

$$p(A_1|E_1A_1)=(1-\alpha)(1+k)+2k\alpha\pi$$
$$p(A_1|E_1A_2)=1-\alpha+2k\alpha\pi$$
$$p(A_1|E_2A_1)=(1-\alpha)k+2k\alpha\pi$$
$$p(A_1|E_2A_2)=2k\alpha\pi$$

ここで，$k=\alpha/(1+\alpha)$ であり，したがって $0\leqq k<1/2$ である。このことから，上の4つの $p(A_1|E_jA_i)$ の値は，上に書いた順序で小さくなっていることが分かる。そして，この順序（不等関係）は，前節で述べたものと矛盾しない。

線形モデルは，確率一致現象とか系列依存関係という確率学習における基本的事実を説明できる点で優れたモデルであると云える。しかし，問題点もないわけではない。その主要なものは，過度の単純化ということ，また，被験者の頭の中で生じる内面過程を考慮していない点であろう。

単純化については，あらゆるモデルにおいて，事実の単純化が行われていると云えるのであり，単純化それ自体は決してわるいことではない。モデル構成の意義は，枝葉末節を切り捨てて対象の本質を掴むことにあるとするならば，単純化は，非難されるどころか，むしろ賞揚されるべきことである。

問題は，過度の単純化である。対象の本質までも切り捨てる単純化は，明らかにゆきすぎであり，過度と云われても致し方あるまい。ただし，何をもって本質を見失っていると見なすかについては，多かれ少かれ批判者の依って立つ観点に左右されるであろう。ここで取上げている線形モデルについても，上述のように，確率一致現象とか系列依存関係は説明できるのであるから，十分に本質を捉えたモデルであるとなす評価もあるかもしれない。

しかし，このモデルでは，確率学習の事態における被験者の内面過程については，ほとんど何も語っていないのである。この点で，線形モデルは，何がしかの物足りなさを感じさせる。すなわち，内面過程に全く触れないで，確率学習事態の本質を把えられるか否かについては，多少とも疑問が残るのである。そこで，次に，この点について少しばかり考察してお

こう。

4．課題解決事態としての確率学習

確率学習という実験場面は，一種の課題解決事態であると見なすことができる。被験者に与えられる課題は，２つの事象 E_1，E_2 のいずれが生じるかを予想し，その予想をできるかぎり多く適中させることである。

単純確率学習の場面では，顕在的には予想の手がかりになるものは何も与えられていない。しかし，上述のような課題を与えられた被験者は，何かを手がかりにして予想を行ない，課題を解決しようとするであろう。その際，被験者が予想の手がかりとして用いることができるのは，その試行までに与えられた〈事象系列〉のみである。

事象系列を手がかりにして予想を行なうとすれば，与えられた事象系列の規則性を発見する必要がある。事象系列の規則性とは，例えば次のようなものである。(１)E_1 は E_2 よりも高頻度で生じる。(２)E_1 が生じたとき，次の試行でも E_1 が生じやすい（同一事象の反復傾向)。(３)あるいは，反対に，E_1 が生じたときには，次の試行では E_2 が生じやすい（交替傾向)。(４)E_1 が何回もつづけて生じたあとでは E_2 が起りやすい（これは〈賭博者の錯誤〉と呼ぶことがある)。(５)場合によっては，もっと複雑な規則性もあって，例えば，E_1 が２回つづけて生じ，そのあと E_2 が生じる。すなわち，$E_1E_1E_2$ という系列が生じる。

元来，事象系列は，実験条件としては多少ともランダムな系列であるから，例外があることを許すような規則性でしかあり得ない。その意味で，きわめてゆるい規則性ではあるが，被験者は予想課題の解決のために，事象系列のうちに何らかの規則性を発見し，それを予想の手がかりに用いようとするであろう。

ここでは，課題解決事態と見た確率学習を，いわゆる〈2段階モデル〉(例えば，Brehmer, B. 1974) に基づいて考えることにしたい。つまり，まず第１段階として，規則を発見し，つづいて第２段階として，発見された規則を使用するという２つの段階を経て，課題が解決される。事象系列において見いだされた規則性は，その規則を実際に用いることにより，予想の手がかりとして現実に機能する。

規則の使用に当ってまず留意すべきは〈記憶〉の役割である。事象系列の規則性を手がかりにして予想するとき，先行数試行の事象を記憶しておくことが必要であろう。例えば $E_1E_1E_2$ という系列がしばしば生じることに気がついたとする。その場合は，先行２試行で E_1 が生じたとき，次の試行では E_2 が生じるであろうと予想する。このとき，先行２試行で生じた事象を記憶しておかねばならない。

前に系列依存関係について述べた。つまり，現試行での予想は，先行する予想・事象系列

に依存している。ここでは，とくに先行事象への依存が問題であるが，このような依存関係が生じる一つの理由は，上述のように，先行事象系列の記憶に基づいてこれを手がかりにして予想が行われるからであろう。

一般に，先行事象の記憶は，過去へ遡るにつれてうすれてゆくであろう（短期記憶の特性)。そこで，現在の予想は，直前の試行の事象に対して最も大きく依存し，現試行から遠く離れるにしたがって依存性は小さくなってゆくであろう。

規則の使用に当ってもう一つ留意すべき点は，ランダムネス（確率性）に対する対処の仕方，つまり〈方略〉である。単純確率学習における予想は，事象系列の何らかの規則性を手がかりにして行われる。しかし，その事象系列はデターミニスティック（確定的）なものではなくて，確率的なランダム系列である。すなわち，大ざっぱな規則性はあっても，その規則性には多かれ少かれ例外がある。確定的な規則にしたがって事象が生起するわけではない。

確率学習において与えられる課題には，厳密な意味での正解は存在せず，その意味では解決不能の課題状況であるとも云える。ランダムネスが含まれているために完全な解決が不可能な状況では，人は，そのランダムネスに対し，どのような方略でもって対処しようとするのであろうか。

その一つの方略は，〈例外の無視〉である。例えば，90% E_1 が生じるとき，実際には10% E_2 が生じるのであるが，それを例外として無視し，E_1 が生じると予想する。天気予報において90%雨であれば，傘を持って出かけるのと同様である。小さい確率でしか起きない事柄については，起きないものときめてかかるのである。

確率学習の実験では，実験開始時に，事象系列の性質に関してある教示を与えるのが普通であるが，この教示に2通りのものがある。一つは，事象系列にはある規則性があることを強調するような教示である。完全な規則性ではないにせよ，事象系列によく注意していれば，かなり適中率を向上させることができるような規則性があることを示唆する。標準的状況での確率学習の実験では，このような教示が与えられている。

もう一つの教示の仕方は，事象系列がランダム系列であることを告げるものである。このような2通りの教示の仕方があり，教示の仕方により実験の結果も多少異なる。すなわち，前者の教示（規則性を強調）によれば，確率一致現象が現われ，後者の教示（ランダムネスを強調）を与えたときには，確率の極大化が生じる傾向がある。

事象系列がランダムであることが告げられると，事象系列の系列特性に関する規則性には注意せず，もっぱら E_1 と E_2 の頻度に注意し，しかも例外を無視して，より高頻度で生じる事象を予想するのではなかろうか。もしそうだとすれば，その結果として極大化の現象が生じることになる。もっとも，実際には完全な極大化 ($p=0$ or 1) が生じることは珍しく，確率の一致と極大化の中間，すなわち，一致よりは極大化の方向にずれたような予想がなさ

れることが多い。これは，例外を無視する傾向はあっても，だからと云って完全に例外を無視してしまうことも困難であることにより生じるものであろう。

5．弁別的確率学習

次に，単純確率学習を拡張した形態である〈弁別的確率学習〉について考察しよう。弁別的確率学習は，MCPL と比較すれば，手がかりの次元数が1次元であるために，より特殊な形態の確率学習である。しかし，単純確率学習と比較すれば，明示的な手がかりを含んでいる点で，より一般的である。すなわち，弁別的確率学習は，今まで述べてきた単純確率学習において，〈弁別的手がかり〉を導入したものである。例えば，2種類の音 C_1 または C_2 を手がかりとして提示し，そのあと事象 E_1, E_2 のいずれが生じるかを予想させる（予想 A_1 または A_2 を行わせる)。予想がすんだあと，実際に E_1 または E_2 が生じて，情報のフィードバック，すなわち KR が与えられる。単純確率学習では，予想を行うとき，先行事象系列しか手がかりがなかったのであるが，弁別的確率学習では，毎試行，明示的に手がかり C_1 または C_2 が提示される。この点が，単純確率学習と異なる点である。

1）手がかりの関連性と妥当性

いま，手がかり C_1, C_2 をそれぞれ確率 β, $1-\beta$ でもって，ランダムの順序で提示するものとしよう。（β の値としては，例えば $\beta=.5$。）すなわち

$$p(C_1)=\beta$$

とする。（そのとき $p(C_2)=1-\beta$。）

手がかり C は，事象 E の生起と関連があって，例えば C_1 が提示されたときには E_1 が起りやすく，C_2 が提示されたときには E_2 が起りやすい，といった具合である。このように，手がかり C の生起は，事象 E の生起と関連があるので，どの事象が生起するかを予想するとき，C は予想の手がかりになる。

いま，手がかり C_i $(i=1,\ 2)$ が提示されたときに事象 E_1 が生じる確率を，π_i でもって表わすことにしよう。すなわち

$$p(E_1|C_1)=\pi_1$$
$$p(E_1|C_2)=\pi_2$$

とする。この π_i $(i=1,\ 2)$ は，手がかり C_i の事象 E_1 に対する〈関連性〉relevance を示していて，手がかりの〈妥当性〉validity と深い関係がある。

しかし，π_i の数値がそのまま手がかりの妥当性の度合を示しているのではない。π_i の値が意味するものは，$p(E_1)$ の値との関係で考える必要がある。

手がかり C_i の妥当性は，π_i が $p(E_1)$ から離れるほど大きくなる。例えば，$p(E_1)=.3$ であったとする。そのとき $\pi_1=.8$ であれば，手がかりが与えられないとき，E_1 は30%の割合で生じることが予想されるが，手がかり C_1 が与えられれば，E_1 は80%の割合で生じることが予想される。つまり，手がかりが提示されることにより，予想の適中率は増加する。

これに対し，もし $\pi_1=\pi_2=p(E_1)$ であれば，手がかりと事象とは独立になり，手がかりが提示されても事象の予想の改善には役だたない。このような場合には，手がかりの妥当性は0である。このようなわけで，π_i の値がそのまま手がかりの妥当性を示しているのではないが，妥当性と深く関係していることは確かである。

手がかりの妥当性は，手がかり C と事象 E の相関の程度により示される。C と E の相関の程度が大になるほど C の妥当性は大きくなり，反対に C と E が無相関（独立）になるほど，C の妥当性は小さくなる。

相関の程度を見るとき，出発点になるのは2変量間の関連表（2次元の度数分布あるいは確率分布）である。そこで，手がかり C と事象 E の2次元確率分布 $p(C,\ E)$ を求めると表1のようになる。ここで $p(C_i,\ E_j)$，$i, j=1, 2$ は同時確率であり，$p(C_i)$，$p(E_j)$ は周

表1　手がかり C と事象 E の2次元確率分布 $p(C,\ E)$

C ＼ E	E_1	E_2	計
C_1	$p(C_1,\ E_1)$ 〔$\pi_1\beta$〕	$p(C_1,\ E_2)$ 〔$(1-\pi_1)\beta$〕	$p(C_1)$ 〔β〕
C_2	$p(C_2,\ E_1)$ 〔$\pi_2(1-\beta)$〕	$p(C_2,\ E_2)$ 〔$(1-\pi_2)(1-\beta)$〕	$p(C_2)$ 〔$1-\beta$〕
計	$p(E_1)$	$p(E_2)$	1

辺確率である。同時確率 $p(C_i,\ E_j)$ については

$$p(C_i,\ E_j)=p(E_j|C_i)p(C_i)$$

という関係がある。そこで，同時確率は β, π_1, π_2 によって表わすことができる。表中の下段の〔 〕内は，β, π_1, π_2 を用いて表わした値である。なお，この表において自由度は3であるから，β, π_1, π_2 の3つのパラメーターがきまれば，確率分布 $p(C,\ E)$ は定まってしまう。

表1に示した手がかり C と事象 E の関連表に基づいて，C と E の相関を求めれば，この相関が，手がかり C の事象 E を予想する上での妥当性を示す指数になる。いまの場合，変量 C，E は カテゴリカルな変量であって，数量的な変量ではない。そこで Pearson の積率相関を用いることはできない。いわゆる属性相関を用いる必要があるが，種々の属性相関のうち，MCPL の研究でよく用いられるのは〈ϕ 係数〉である。

ϕ 係数は，実測値が，2変量が独立であると仮定したときの期待値とどれだけずれているか，そのずれの量によって相関の程度を示そうとするものである。したがって，基本的には χ^2 $(=\Sigma(O-E)^2/E$，O は実測値，E は期待値）と同じものである。実際，ϕ 係数は $\phi^2=\chi^2/n$（n は実測総数）と定義されている。

通常，ϕ 係数の計算式は，度数分布を用いた形式で表わされるが，これを表1のような確率分布を用いた形式でもって示すと，手がかり C と事象 E の間の相関を示す ϕ 係数は次のようになる。

$$\phi^2=\frac{\{p(C_1,E_1)p(C_2,E_2)-p(C_1,E_2)p(C_2,E_1)\}^2}{p(C_1)p(C_2)p(E_1)p(E_2)}$$

表1の関連表は自由度3であるから，$\beta=p(C_1)$，$\pi_1=p(E_1|C_1)$，$\pi=p(E_1)$ の3つのパラメーターで $p(C, E)$ を表わすことにすれば，ϕ^2 は次のようになる。

$$\phi^2=\frac{(\pi_1-\pi)^2}{\pi(1-\pi)}\ \frac{\beta}{1-\beta}$$

実験条件から ϕ 係数を計算するときには，この形の式が使いやすい。

ϕ 係数の定義から分かるように，手がかり C と事象 E が独立であれば，$\phi=0$ となる。ϕ 係数はカテゴリカルな相関であるから，負相関と正相関の区別はなく，常に非負，すなわち $0\leqq\phi$ である。ϕ は非独立であるほど（相関の程度が大であるほど）大きい値をとり，その上限は，関連表が $K\times L$ のとき，$\phi^2\leqq\min(K, L)-1$ である。いまの場合，関連表は表1のように 2×2 であるから $\phi^2\leqq1$ になる。

2）手がかりの予想確率に対する影響

弁別的確率学習で重要なのは，手がかり C_1 または C_2 を提示したとき，被験者はどのような予想を行なうかということである。いま手がかり C_1，C_2 を提示したとき，事象 E_1 が生じるという予想 A_1 を行なう確率をそれぞれ p_1，p_2 と表わそう。すなわち

$$p(A_1|C_1)=p_1$$
$$p(A_1|C_2)=p_2$$

とする。

予想確率 p_i ($i=1$, 2) は，いろいろの条件の影響をうけるが，ここでは手がかりに関するものとして，次の3つの条件を取上げる。(1)手がかりの事象に対する関連性，π_i ($i=1$, 2)。(2)手がかりC_1とC_2の間の類似度。(3)手がかりの提示確率，β。

単純確率学習における予想確率 p がそうであったように，弁別的確率学習における p_i もまた，試行を重ねるにつれて変化するが，ここでは多数回の試行のあとの漸近状態について考えることにしよう。

まず第一に，手がかりの事象に対する関連性を取上げよう。手がかりに関する条件を分析するとき，次のような実験手つづきが，標準的方法としてよく用いられる。$\pi_1(=p(E_1|C_1))$, $\pi_2(=p(E_1|C_2))$ のうち一方（例えば π_1）を固定し，他方（例えば π_2）を動かして，予想確率 p_1, p_2 がどのようになるかを見る，というやり方である。π_i ($i=1$, 2) の効果を吟味するときにも，このような手つづきが用いられる。

このような実験事態（π_1 を固定し，π_2 を動かす）において，p_1, p_2（とくに，試行を重ねたときの漸近値）は，π_2 の関数としてどのようになるであろうか。

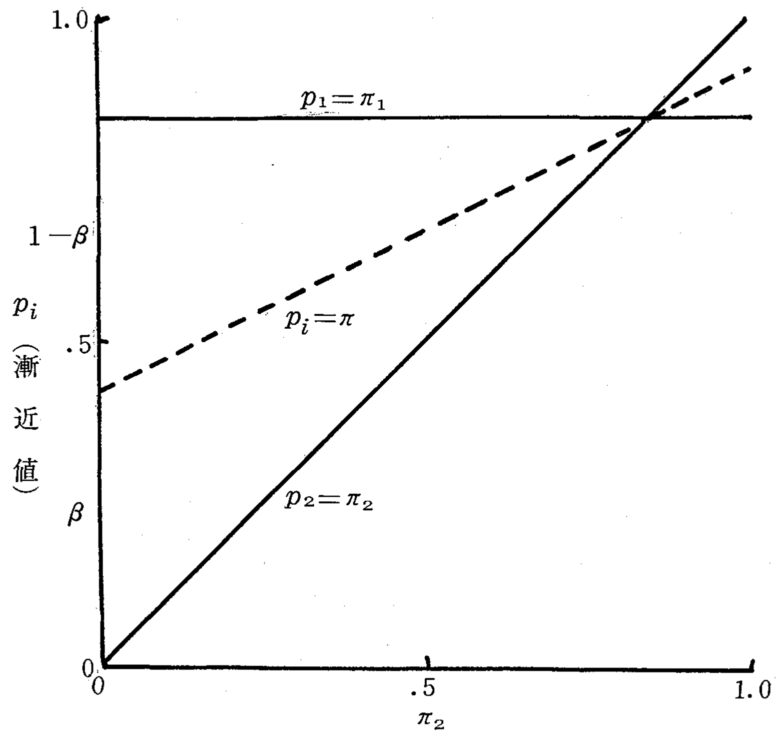

図1　弁別的確率学習において確率一致現象が生じたときの，π_2 の関数としての $p_i(i=1, 2)$。

まず考えつくことは，p_i ($i=1$, 2) において確率一致現象が生じるのではないか，ということである。すなわち

$$p_1 \longrightarrow \pi_1$$

$$p_2 \longrightarrow \pi_2$$

となるのではないかという予測が成立つ。p_1 は，π_1 が固定されている条件のもとでは，π_2 の値のいかんによらずある一定の値をとり，p_2 のみが π_2 の関数として変化するのであろう。このことが，図1に示されている。(図1において，ここでは，2つの実線 $p_1=\pi_1$，$p_2=\pi_2$ のみに注目してほしい。点線で示した $p_i=\pi$ については後で触れる。)

確率一致現象が生じるとすれば，手がかり C_i ($i=1$, 2) が提示されたとき A_1 と予想する確率 p_i は，上述のように，試行を重ねるにつれて π_i に漸近するであろう。では，実際の実験結果はどうであろうか。Popper, J. & Atkinson, R. C. (1958) は，次のような実験を行なっている。手がかり C_1，C_2 の提示確率は等確率，すなわち $p(C_1)=p(C_2)=.50$ ($\beta=.50$)

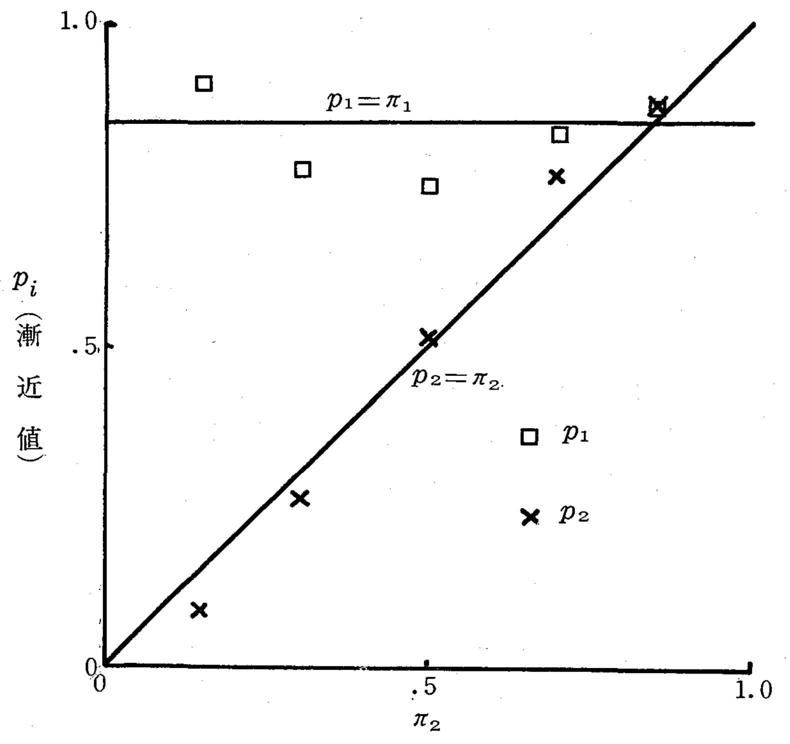

図2　π_2 の関数としての p_i($i=1$, 2)。
〔Popper & Atkinson〕

とする。π_1 については，$\pi_1=.85$ と固定する。この実験の変数は π_2 であって，$\pi_2=.15$, .30, .50, .70, .85と5条件を設ける。各条件とも320回の試行を行った。p_i（漸近値）は，320試行中，最終120試行の結果に基づいて算出した。π_2 を変化したとき，p_1, p_2 がどのようになるかを示すと，図2のようになる。

この実験結果は，p_1, p_2 ともに，ごく大ざっぱに見れば，確率一致現象を示していて，$p_1 \fallingdotseq \pi_1$, $p_2 \fallingdotseq \pi_2$ であると云ってもよかろう。しかし，仔細に見れば，p_1 は下に凸のU字型曲線であり p_2 はやや上に凸の曲線である。

このような結果になった理由としては，確率の混同が考えられる。実験条件としては，手がかり C_1 が提示されたとき確率 π_1 で事象 E_1 が生起し，手がかり C_2 が提示されたときには，確率 π_2 で事象 E_1 が生起する。そこで，これら2つの確率 π_1, π_2 が識別されたときに，手がかり C_1 に対して確率 π_1 で A_1 と予想し，C_2 に対しては確率 π_2 で A_1 と予想することになる。

π_1 と π_2 の差が大であれば，両者は，はっきり識別されるが，その差が小さいときは，両者の混同が生じ，C_1 に対し確率 π_2 で A_1 と予想したり，C_2 に対して確率 π_1 で A_1 と予想したりする。そこで，π_1 と π_2 の差が大のときは p_1 と p_2 は分離するが，差が小さくなるにつれて，p_1 は p_2 に近づき，p_2 も p_1 に近づいて，その結果，p_1 と p_2 は彎曲した曲線になるものと思われる。

手がかりの事象に対する関連性 π_i に関する考察は，これくらいにとどめよう。予想確率 p_i に及ぼす，手がかりに関する条件として，次に取上げなければならないのは，2つの手がかり C_1, C_2 の間の類似度である。

以上に述べた確率の混同に関する考察は，2つの手がかり C_1 と C_2 がはっきり弁別できる場合の話である。2つの手がかりが類似していて弁別困難である場合にはどのようになるであろうか。

その最も極端な場合は，2つの手がかりが全く等しいときである。この場合は手がかりを提示しないのと同じである。そこで，もし確率一致現象が生じるとすれば，p_1, p_2 はともに $\pi(=p(E_1))$ に近づくであろう。すなわち

$$p_1,\ p_2 \longrightarrow \pi$$

になると期待される。

ここで π は，π_1, π_2 と次のような関係がある。前述の C と E の関連表（表1）に示した関係から

$$p(E_1)=p(C_1,\ E_1)+p(C_2,\ E_1)$$

である。したがって

$$\pi=\beta\pi_1+(1-\beta)\pi_2$$

すなわち，π は，β，$1-\beta$ を重みとする π_1 と π_2 の加重平均になる。

標準的な実験事態（π_1 を固定し，π_2 を動かすような事態）においては，p_i（$i=1, 2$）が確率一致現象を起して π に等しくなるとすれば，π は，前掲の図1で点線により示した直線になる。

とくに，2つの手がかり C_1，C_2 が等確率で提示されるときには，$\beta=1/2$ であるから，π は π_1 と π_2 の算術平均となり，$\pi=(\pi_1+\pi_2)/2$ である。

このように，2つの手がかり C_1 と C_2 が類似していて弁別できないときには，手がかりは与えられなかったのと同様であるから，p_1 と p_2 はともに同じ π に等しくなるであろう。手がかり C_1 と C_2 の類似度が減少し，2つの手がかりの間の相異が増すと，p_1 と p_2 は分離し，p_1 は π_1 の値へ，p_2 は π_2 の値に近づくであろう。もっとも，正確に云えば，p_i は π_i と完全に一致するのではなくて，確率の混同により，図2に示されているように，p_1 は下に凸，p_2 はやや上に凸の曲線になるであろう。

手がかりの類似度を変えた実験としては，Massaro, D. W., Halpern, J. & Moore, J. W.（1968）の研究がある。この研究は Exp. I，IIの2つのパートから成っているが，ここでは Exp. IIを取上げる。

ライトが左右2つあって，手がかりが提示されたあと，どちらのライトがつくか（事象 E_1，E_2）予想する。それぞれのライトの下にはボタン・スイッチがあって，被験者は，つくと予想したライトの下のボタン・スイッチを押す（予想 A_1，A_2）。このような試行を300回くりかえす。

手がかり C としては，800Hzの純音を用いる。ここでの主変数は，2つの手がかり C_1 と C_2 の間の類似度であって，音の強度を変えて類似度を変える。類似度については3条件を設けて，それは

73—74.5dB　($\Delta I=1.5$)
73—76 dB　($\Delta I=3$)
73—79 dB　($\Delta I=6$)

であった。

手がかりの提示確率については等確率，すなわち $p(C_1)=p(C_2)=.50$（$\beta=.50$）とする。手がかりの関連性 π_1 と π_2 については，標準的な実験手つづきにしたがい，$\pi_1=.80$ と固定し，π_2 を .20，.50，.80 と変える。

300試行中，最終100試行の p_i（$i=1, 2$）を，π_2 の関数として図示すると，図3のように

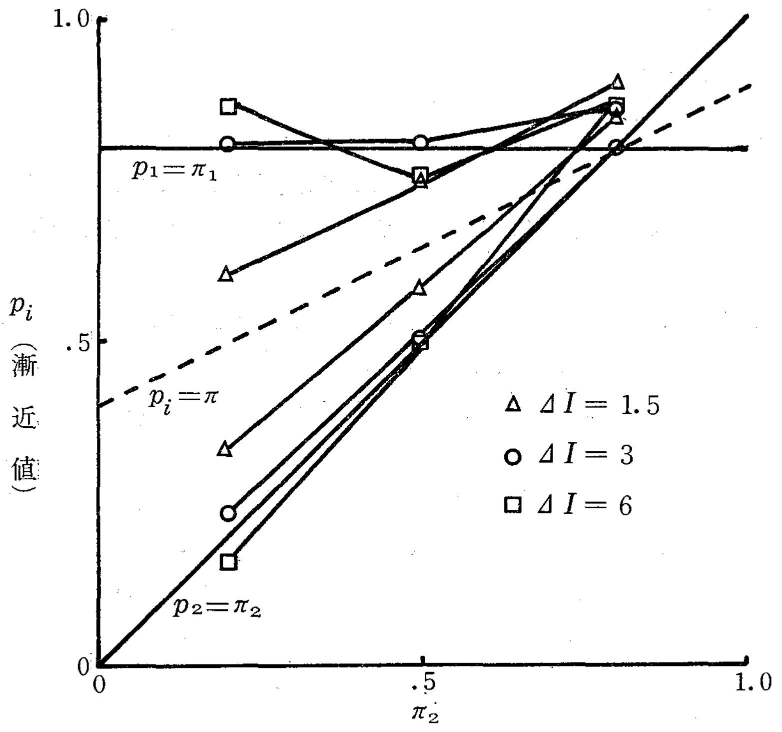

図 3　2つの手がかりの類似度が異なるときの，π_2 の関数としての $p_i(i=1,\ 2)$。
〔Massaro, Halpern & Moore〕

なる。2つの手がかりの間の差 ($\varDelta I$) が大きくなるにつれて，p_1 と p_2 は，$p_i=\pi$ をはさんで分離してゆく様子が示されている。

以上，手がかりに関して，予想確率 p_i に影響する条件として，手がかりの事象に対する関連性 π_i と，手がかり C_1 と C_2 の間の類似度を取上げて考察した。最後に，第3の条件として，手がかりの提示確率 β について述べることにしよう。

手がかり C_1 が提示されたとき事象 E_1 が生じる確率は π_1 である ($\pi_1=p(E_1|C_1)$)。同様に，手がかり C_2 が提示されたとき事象 E_1 が生じる確率は π_2 である ($\pi_2=p(E_1|C_2)$)。そこで，多数回の試行のあとには，手がかり C_1 が提示されると，確率 π_1 でもって A_1 (事象 E_1 が生じるという予想) が生じる傾向が形成されるものとしよう。そして，このような予想傾向を R_1 という記号で表わそう。ここでは，このことを次のように示しておこう。

$$R_1 : C_1 \xrightarrow{\pi_1} A_1$$

同様に，手がかり C_2 が提示されたときに確率 π_2 でもって A_1 が生じる傾向を R_2 という記

号で表わす。すなわち

$$R_2 : C_2 \xrightarrow{\pi_2} A_1$$

とする。

弁別的確率学習においては，このような予想傾向 R_1，R_2 が形成され，手がかりの提示確率 β は，この R_1 と R_2 の形成に影響すると考えてみる。

もっと具体的に云えば，$p(C_i)$ $(i=1,\ 2)$ が大であるとき R_i も増大する，と仮定するのである。手がかり C_1 について考えてみよう。$p(C_1)$ が大ということは，“C_1 にひきつづいて確率 π_1 で E_1 が生じる” ということが多数回生じる，ということである。したがって，このような場合には，予想傾向 R_1 も大になるであろう。手がかり C_2 についても同様である。$p(C_2)$ が大ということは，“C_2 にひきつづいて確率 π_2 で E_1 が生じる” ということが頻繁に生じることである。このときは予想傾向 R_2 が大になるであろう。このようなことから，$p(C_i)$ が大ならば R_i も大であると仮定する。

以上の仮定のほかに，ここでもう一つ考慮に入れたいことがある。それは，前述の確率の混同である。手がかり C_i $(i=1,\ 2)$ が提示されたときには，確率 π_i で事象 E_1 が生じる。そこで，確率の一致が生じることを前提にして考えるならば，手がかり C_i が提示されると，確率 π_i でもって A_1 という予想がなされることになろう。しかし，ここで2つの確率 π_1 と π_2 の間で混同が生じると，C_1 が提示されたとき，確率 π_2 で A_1 と予想したり，また，C_2 が提示されたときに確率 π_1 で A_1 と予想したりする。このような確率の混同が生じると，p_1 は p_2 に，また p_2 は p_1 に接近するであろう。この機制は前に述べた通りである。

p_1 と p_2 の間での接近の仕方であるが，$\beta(=p(C_1))$ が大であれば，上述の仮定により R_1 が強い為に，p_1 はそれほどの影響を受けず（p_2 にそれほど接近しない），反対に p_2 の方が大きな影響を受けて p_1 に大きく接近するであろう。また，反対に β が小（$p(C_2)$ が大）であれば，R_2 が強い為に，p_2 の p_1 への接近の程度は小さく，むしろ p_1 の方が大きく p_2 に接近するであろう。

この仮定に基づいて，提示確率 β が異なるとき，π_2 の関数としての p_i $(i=1,\ 2)$ がどのようになるか，理論的に期待される傾向を図示すると図4のようになる。もっとも，この期待曲線は，確率が混同されないときには確率一致現象が生じることを前提にして描かれている。確率の一致ではなくて，極大化の傾向があるときには，p_1 については，$\pi_1(>1/2)$ よりも大きくなり，p_2 については，$\pi_2<1/2$ のとき π_2 よりも小さく，$\pi_2>1/2$ のとき π_2 よりも大になる傾向が現われるであろう。

提示確率 β を変えたとき，p_1, p_2 は理論的には図4のようになることが期待されるのであるが，実験データの上では，どのような結果が得られているだろうか。ここでは，Meyers,

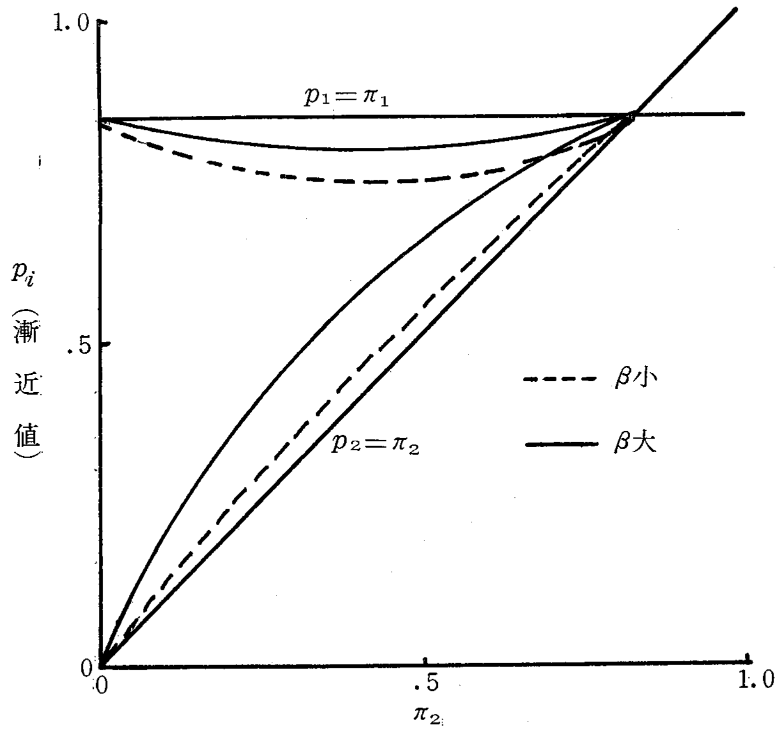

図 4　手がかりの提示確率 β が異なるとき，理論的に期待される，π_2 の関数としての $p_i(i=1,\ 2)$。

J. L. & Cruse, D. (1968) が行っている実験を取上げよう。手がかり C は，上下に並んだ2つのライトであって，毎試行，上または下のいずれかのライトがつく (C_1 または C_2)。その後，左右に並んでいる2つのミドリ色のライトのいずれか一方が点灯する (E_1 または E_2)。いずれの事象が生じるかを予想し，ミドリ色の事象ライトの下にあるスイッチを動かす (A_1 または A_2)。このような試行を500回くりかえす。

この実験での主変数は，手がかりの提示確率 β であって，$\beta=.25,\ .50,\ .75$ と3通りの条件を設ける。実験手つづきは標準的方法にしたがい，$\pi_1=.85$ と一定に保ち，π_2 の方を変えて，$\pi_2=.15,\ .50,\ .85$ と変化する。

500試行中，最終100試行の $p_i\ (i=1,\ 2)$ を π_2 の関数として示すと，図5のようになる。この結果はそれほどきれいではないが，ここでの仮定から期待されるような傾向が生じているように思われる。すなわち，手がかり C_1 の提示確率 β が小さければ (C_2 の提示確率が大きければ)，p_2 は，基準になる p_2 (確率の一致が生じるときには π_2) から離れること少なく，p_1 の方が，基準となる p_2 に近づく。反対に，β が大きければ，p_1 は，基準となる p_1 (確率の一致が生じるときには π_1) から離れること少なく，p_2 の方が，基準となる p_1 に

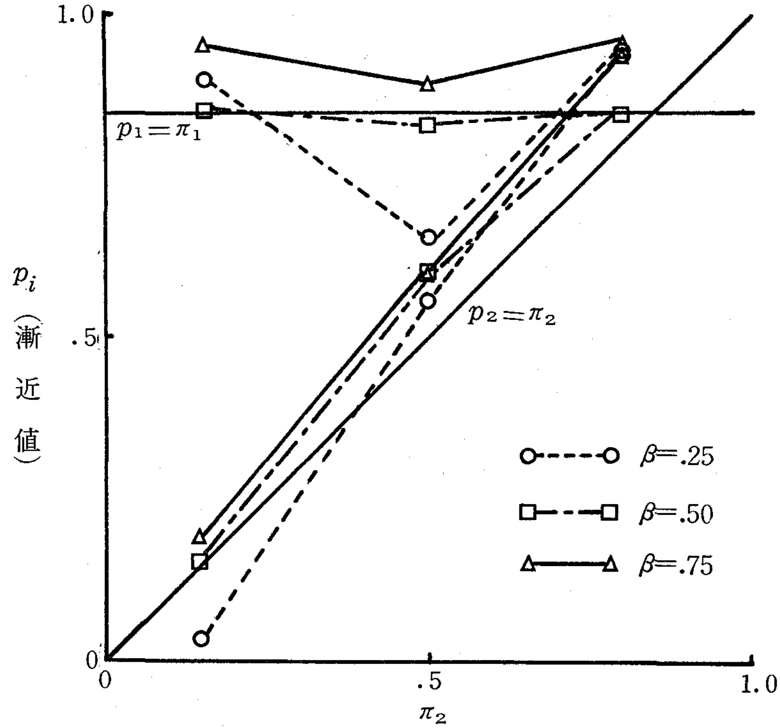

図5　手がかりの提示確率 β が異なるときの，π_2 の関数としての $p_i(i=1,\ 2)$。〔Meyers & Cruse〕

近づく。ただし，このデータを見るかぎり，この実験では，確率の一致というよりは，むしろ極大化傾向が生じている。そこで，基準となる p_1 は，π_1 ではなくて，π_1 よりも1に近い値になり（$\pi_1>1/2$ であるから），また，基準になる p_2 は，$\pi_2<1/2$ のとき，π_2 よりも0に近い値になり，$\pi_2>1/2$ のとき，π_2 よりも1に近い値になる。

3）弁別的確率学習の機制

以上のような，手がかりに関しての，予想確率に影響を及ぼす諸条件の分析に基づき，弁別的確率学習における，手がかりに対する予想出現の機制を，次のような3段階に分けて考えることにしたい。

1．手がかり C_1 と C_2 の弁別。この弁別に対しては，C_1 と C_2 の間の類似度が影響する。もちろん，類似度が大であるほど弁別は困難である。

2．確率 π_1 $(=p(E_1|C_1))$ と π_2 $(=p(E_1|C_2))$ の識別。まず2つの手がかり C_1 と C_2 を弁別し，そのあと C_1 または C_2 が提示されたときに，どの程度の確率で事象 E_1 が生じるかを識別する。この識別は，π_1 と π_2 の差に依存し，差が大であるほど識別は容易であろう。

確率 π_1 と π_2 はしばしば混同される。手がかり C_1 が提示されているときならば，混同により確率 π_2 で A_1 と予想し，手がかり C_2 が提示されているときならば確率 π_1 で A_1 と予想する。このような混同は π_1 と π_2 の差が小さいときに著しいものと思われる。

3．予想傾向 R_1 と R_2 の成立。手がかり C_1 が提示されたとき確率 π_1 でもって A_1 と予想する（これは R_1 である）。また，手がかり C_2 が提示されたときには確率 π_2 でもって A_1 と予想する（これは R_2 である）。手がかり C_1 と C_2 が弁別され，また，確率 π_1 と π_2 が識別されたあと，予想傾向 R_1，R_2 が成立する。この成立に対しては，手がかりの提示確率 β $(=p(C_1))$ が影響し，β が大ならば R_1 が強くなり，β が小ならば（$p(C_2)$ が大ならば），R_2 が強くなると考えるのである。

弁別的確率学習において，このような3つの段階を想定するモデルを構成すると，これまでの分析から分かるように，少くとも手がかりに関する変数の効果は，よく説明できる。

なお，本稿では詳論しないが，弁別的確率学習においても，系列依存関係は存在する。系列依存関係については，単純確率学習の項で行ったのと同様の分析を，弁別的確率学習についても行なうことができ，そのような分析から予測されるような実験結果が得られている（Massaro, D. W. 1969)。手がかり C は，予想に対して確かに機能している。しかし，それと同時に，前試行の諸事象（手がかり，予想，結果の事象）も，現試行の予想に影響する。

6．結　語

本論文の要約によって結語にかえよう。本稿では，不確実な手がかりのもとでの選択を，多重手がかりの確率学習（MCPL）の面から考察することを目的にしている。

MCPL は，一方では概念学習を拡張したものである。すなわち，通常の概念学習では手がかりは完全に妥当であるが，概念学習において，手がかりの妥当性を不完全にすれば MCPL になる。（概念学習と MCPL の関係についての詳細は，本稿では触れることができなかった。この問題については，稿を改めて考察してみたいと思っている。）

他方，MCPL は弁別的確率学習を拡張したものとも考えられる。すなわち，通常の弁別的確率学習では手がかりは1次元である。しかし，これを n 次元に拡張すると MCPL になるのである。

この弁別的確率学習は単純確率学習を拡張したもの，すなわち，単純確率学習に手がかりを導入したものが弁別的確率学習である。

本稿では，不確実な手がかりのもとでの一般的な選択状況（すなわち MCPL）の分析に先だち，MCPL の特殊形態である確率学習（単純確率学習と弁別的確率学習）を取上げた。

とくに弁別的確率学習については，その機制に関して，3つの段階を区別するモデルを提

案した。

文　　献

Brehmer, B. Hypothesis about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance*, 1974, 11, 1-27.

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

Castellan, N. J. Jr. Decision making with multiple probilistic cues. In Castellan, N. J., Pissoni, D. B. & Potts, G. R. (ed.) *Cognitive theory*. vol. 2. 1977, Chap. 5 (p.117-147).

Jones, M. R. From probabilistic learning to sequential processing; A critical review. *Psychological Bulletin.* 1971, 76, 108-185.

Massaro, D. W. A three state Markov model for discrimination learning. *Journal of Mathematical Psychology.* 1969, 6, 62-80.

Massaro, D. W., Halpern, J. & Moore, J. W. Generalization effects in human discrimination learning with overt cue identification. *Journal of Experimental Psychology.* 1968, 77, 474-482.

Meyers, J. L. & Cruse, D. Two-choice discrimination learning as a function of stimulus and event probabilities. *Journal of Experimental Psychology* 1968, 77, 453-459.

小野　茂　学習実験．情報科学講座E17・3, 1966, 共立出版。

Popper, J. & Atkinson, R. C. Discrimination learning in a verbal conditioning situation. *Journal of Experimental Psychology* 1958, 56, 21-25.

Suppes, P. & Atkinson, R. C. *Markov learning models for multiperson interactions.* 1960, Stanford Univ. Press.

CHOICE UNDER UNCERTAIN CUES : PRELIMINARY CONSIDERATION —ON PROBABILITY LEARNING—

Shigeru ONO

Multiple cue probability learning (MCPL) may be considered as choice situation under uncertain cues. In this paper, probablity learning is analized as preliminary consideration, though the final purpose is to clarify properties of MCPL. The MCPL is a generalized form of discriminative probability learning. Probability learning including simple and discriminative one is taken up and discussed, as the result, a model of discriminative probability learning is proposed. In this model, it is assumed that learning proceeds through the following three stages: (1) discrimination between two cues C_1 and C_2, (2) identification of two probabilities π_1 and π_2 (event probabilities conditional to cue C_i, $i=1, 2$), (3) formation of response tendencies R_1 and R_2 (event anticipation under cue C_i, $i=1, 2$). This three-stage model can explain the main experimental findings relating with cue variables.