

Title	行動の数値モデルとシミュレーション
Sub Title	Mathematical model of behavior and simulation
Author	神谷, 直樹(Kamiya, Naoki) 坂上, 貴之(Sakagami, Takayuki)
Publisher	慶應義塾大学大学院社会学研究科
Publication year	1998
Jtitle	慶應義塾大学大学院社会学研究科紀要：社会学心理学教育学 (Studies in sociology, psychology and education). No.48 (1998. ), p.1- 22
JaLC DOI	
Abstract	
Notes	論文
Genre	Departmental Bulletin Paper
URL	<a href="https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN0006957X-00000048-0001">https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN0006957X-00000048-0001</a>

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

# 行動の数値モデルとシミュレーション<sup>注1)</sup>

## Mathematical model of behavior and simulation

神谷直樹\*・坂上貴之\*\*

*Naoki Kamiya and Takayuki Sakagami*

Traditional behavior analytic approach has alarmed to use mathematical and quantitative models for a long time, because they introduce unnecessary intervening variables and oversimplification. In spite of these shortcomings, it seems to be many merits to use mathematical models, such as an integration of a variety of experimental results, metaphorical extensions to other areas out of psychology, and tools for exploratory data analysis. According to what fundamental mechanisms underlie in behavioral phenomena, we classified these mathematical models into five types of models, these are descriptive, optimization, (information) process, dynamical, and probabilistic models. To examine the effectiveness of building models and analyzing by models, we conducted a simulation of choice behavior under several uncertain conditions using one of probabilistic models. Results showed that the simulation made clear the issues about what occurs when a response is reinforced and is not reinforced, and what components are enough to realize "preference reversal" and probability matching.

Key words: mathematical model, quantitative model, operant behavior, preference reversal, probability matching, simulation, uncertainty, choice

### 1. 行動の数値モデルの功罪と分類

行動分析は、数値モデル (mathematical model) を使用して行動を理解することに長い間懸念を抱いてきた。例えば、Skinner (1950, 1984) は、その論文で対象としている「理論」という用語を、「異なる観察のレベルのどこか別のところで起こっている出来事に訴えてなされた観察事実についての何らかの説明」であり、「このレベルで観察された出来事は、異なる用語によって記述され、(測定されているとすれば) 別の次元で測定されている」と定義した。その上で、行動分析がとるべきではないとする理論として、神経系による説明、心的出来事による説明、解釈のための直接観察できない概念による説明の3つを挙げた。この最後に挙げられた「説明」は、観察不可能な仲介変数を積極的に取り入れたために、最終的に

は検証不可能な理論を構成してしまった Hull, C. L. の仮説演繹法のアプローチを指していると考えられる (例えば、Schultz, 1981)。なぜならば、行動生起に対する予測力と制御力が増すことはないにも関わらず、このアプローチは直接観察可能な現象の生起原因を観察不可能な他の次元に求めているためである。その後、このような「説明」に対する警戒感から、実験的事実の積み重ねに研究の比重が置かれるようになっていった (Staddon, 1993; Staddon & Zanutto, 1998)。

このような懸念がどのような経緯を経て現在に至っているのか、そしてそれらをどのように現在の行動分析が捉えているのかは、それ自体大変興味深い問題であるが、この点に関しては既にいくつかのところで論じられてきている (例えば、Plaud & O'Donohue, 1991; Staddon, 1983)。したがって、本論ではむしろ、現在の行動分析による数値モデルの取り扱い方に焦点を当てて論じてみたい。またここで取り扱いたいと考えている行動の

\* 慶應義塾大学大学院社会学研究科 (実験的行動分析)

\*\* 慶應義塾大学文学部教授 (実験的行動分析)

数理モデルを考えるにあたっては、行動の科学的分析の中でモデルそのものが果たす役割について議論する必要があるだろう。しかしこの問題も、現実になされているモデルをめぐっての行動分析の科学的行動を観察することが先決と考えられる。本論ではこうした点から、数理モデルの功罪や分類について述べるだけでなく、現在までに作り出されてきた数理モデルをできるだけ多く挙げた。その上で、実際に 1 つのモデルに基づいたシミュレーションを例に取りながら、モデルの運用について考えてみる。

数理モデルは、工学分野では数学モデルと翻訳されることもある。近藤(1976)によれば数学モデルは、次のような特徴を持った、数式で表された仮定であると定義される。第 1 の特徴は、現象の特徴または本質を数学的に表現した、モデル(模型)の一種類であるという点である。しかし、数式で表される事柄には限りがあるから、多くの場合には現象に完全に一致した数学的表現はできない。そこで第 2 の特徴として、現象の「簡単」化や抽象化が、数式化する以前に行われている点が挙げられる。

このような工学分野での数学モデルの考え方は、行動分析における数理モデルについても当てはまる。Herrnstein(1961, 1970)のマッチング法則は並立 VI VI スケジュール下で見出されたものであるが、その法則が導出される以前に、いくつものこのスケジュールによる選択実験(Cantania, 1963a, 1963b; Catania & Cutts, 1963; Findley, 1958; Herrnstein, 1961)の結果が検討され、各選択肢における反応数と獲得強化子数という量に焦点が当てられて現象が単純化された(前述の第 2 の特徴)。続いてこれらの量を 2 つの選択肢についての比の形式で表し、最後に反応比と獲得強化子比が一致するという現象の特徴(前述の第 1 の特徴)が記述されたとみることができる。

前述の 2 つの特徴を持つ数理モデルは、しばしば非数量的モデル(non-quantitative model)から区別する目的で、特に数量的モデル(quantitative model)と呼ばれることがある。以下では、Coombs, Dawes, and Tversky(1970)にしたがって、この 2 つについて整理する。彼らによれば、非数量的モデルは経験的な関係系(relational system)と非数量的な関係系との対応を表すものと定義される。経験的な関係系とは、個体や環境の特徴を表す用語を指し、非数量的な関係系とはグラフ理論、集合論や計算論といった考え方を指す。この 2 つの系の対応の例として、Heider による対人認知の P-O-X モデ

ルを挙げることができる。Cartwright and Harary(1956)は、Heider のモデルにおける均衡概念がグラフ理論という均衡という性質を完全に表現していると指摘する。Heider のモデルでは、相手に対し好意を持っている、あるいは好意をもっていない個人を P と O で、ある社会現象を X で表す。この X に対して P, O は肯定的な態度あるいは否定的な態度を取ると仮定する。「社会的均衡」を表現するために、このときの 3 者の関係(経験的な関係系)をグラフの形で対応させるが、そのグラフの性質はグラフ理論における均衡概念(非数量的な関係系)となっているのである。

これに対して数量的モデルという用語は、経験的な関係系と数量的な関係系との対応を表すものと定義される。つまり、経験的な関係系の用語に数値が割り当てられ、用語間の関係が数の間の関係として表されている。そして、数の関係を表しているという点から、数量的モデルは予測が量的であるモデルを総称するものとして使われていると考えられる(Church, 1997)。しかし現実には、行動の研究において、これらの違いに特に目を向けなければならない場面は少なく、むしろ数理的予測を可能にする道具としての数理モデル、非数量的モデル、数量的モデルを捉えることの方が有意義であると思われる。なぜならば、非数量的モデルの中でも特に、計算論に依拠した計算モデル(model of computation)は、アルゴリズムによって現象を定式化したもの(小野, 1991)として定義されるので、記号で表記された変数に数値を割り当て、シミュレーションという方法で数量的予測も可能であると考えられるからである。したがって本論では、数理モデルを経験的な用語を数学的な考え方と対応させたものであると定義し、重要な機能の 1 つである数理モデルの数量的予測について考えてみたい。

## 1.1 数理モデルの功罪

### 1.1.1. 初期の行動の数理モデルと行動分析における警戒

初期の行動の数理モデルは、数理心理学における数学的学習理論(mathematical learning theory)に代表される。数学的学習理論は、学習曲線を数学的方程式によって予測しようとする Thurstone の確率的アプローチ(1930)を発展させ、そこから新しい 2 つのモデルを作り出した。Coombs *et al.* (1970)によれば、数学的学習理論の特徴は 1950 年頃に現れ、1970 年代後半にはほぼ成立したと考えられている。この 2 つのモデルとは、両者のその後の展開からみて相互に関連していたと考えられている、作用素(operator)モデルと状態(state)

モデルをいう。ここで作用素モデルとは、学習段階の数が無限にあることが許容される Bush and Mosteller の確率論的学習理論 (1951) をさす。そして状態モデルとは学習段階の数が有限の Estes の刺激サンプリング理論 (1950) をさす。両者の重要な違いは、学習過程に対する考え方に表れている。後者は、刺激、反応、連合、強化といった用語によって、学習過程を説明する Guthrie の流れを汲んだ刺激・反応連合理論を詳細に展開したものである。それに対して前者では、これらの用語は反応確率の増減を関数（作用素と呼ばれる）表現するために便宜上用いらただけで、経験的事実や理論との対応は重要ではなかった。

さて、Skinner によって導入されたフリーオペラント事態における行動の分析は、数学的学習理論における分析となじむことはなかった。なぜならば、数学的学習理論では、離散試行に基づいた離散的系列における学習過程を想定して、そこでの選択行動の分析がなされていたからである。したがって、新たに導入されたオペラント実験箱でのスケジュールによって制御されたオペラント行動は、このような古典的な道具的条件づけの枠組みからはみ出す結果となった。その上、フリーオペラント事態で測定された反応率などの量を数学的学習理論に合致するように組み替えていく努力も、その後明示的にはなされなかったのである (Schultz, 1981)。

当初 Skinner は、食餌行動に関する数理モデル (Skinner, 1932) を提案していた。このモデルでは、摂食量が摂食開始時点からの時間の指数関数になっていることを仮定している。しかし、既に述べたように 1950 年以降、Skinner は次第に数理モデルを警戒するに至った。彼は、行動の次元を考慮しないで構成された数理モデルは不要であるが、実験データの数学的処理は必要であると考えた。そして「非現実的世界 (dream world) に逃げ込むことなしに、非常に整然としたデータを適切な方法で引き出すことに成功するとき、数学は行動分析の中でその独自の位置を占めることになる」と考えたのである (Skinner, 1984)。

さらに Catania (1981) は、数学的処理そのものが変数操作の影響を見過すことにつながると主張している。彼は絶対反応率の相対反応率への変換という数学的処理によって、実際の観察データの情報が失われる危険性があると考えた。例えば、多元 (multiple) VI VI スケジュールから多元 VI EXT スケジュールに移行すると、一般に第 2 成分スケジュールでは、消去手続きに移行したことにより反応率はほぼゼロになる一方、第 1 成分ス

ケジュールにおける反応率は移行する前より増加する。このような現象を行動対比というが、このとき第 1 成分での相対反応率 (第 1 成分の反応率 / (第 1 成分の反応率 + 第 2 成分の反応率)) を指標とすると、相対反応率は 1 に限りなく近づくが、実は第 2 成分の反応率がゼロであれば、第 1 成分の反応率がゼロでない限り、相対反応率は必ず 1 になってしまう。したがってこの場合、相対反応率では実際の変数操作による影響を適切に表現できないとした。彼はこの他に、2 つの選択肢に割り当てられた異なる強化子による影響、観察データそのものが持つ変動性、選択行動場面の種類 (強制選択か自由選択か) による影響を挙げて、これらを数学的処理によって失われてしまう観察データの具体的情報の例として指摘した。

#### 1.1.2. 行動分析への数理モデルの導入

行動分析の数理モデルへの主要な警戒は、実際の行動の次元から離れて行動の予測と制御から無縁の虚構的な説明を作り出すことと、数学的な処理過程の中で観察データの現象を適切に表現できない可能性があることであったといえる。しかし、Nevin (1984) は、数式間の記号が行動的・環境的事象を記述している限り、記号間の関係が容易に理解できるため、さらに新しい経験的分析を導くことが可能になると考えた。そして、その意味で数式的分析は行動分析の統合的側面を担っていると述べている。例えば、並立 VI VI スケジュール下で見出されたマッチング法則は、選択肢間の反応比と獲得強化子比とが一致するというその単純な形式への要約によって、その後多くの実験を生み出すと同時に、異なる実験場面の結果を同じ形式の下にまとめ上げた。それについては様々な種 (ヒト、ハト、サル、ラット)、様々な反応様式 (レバー押し、キーつつき、眼球運動、言語反応)、様々な強化子 (食物、脳内刺激、貨幣、コカイン、言語的賞賛) でその一般性が確かめられている (Williams, 1988)。

また、Nevin (1984) は、数式間の記号が誤解を生みにくい科学的用語の明瞭性をもっていていると考えている。Shull (1992) も、行動分析が独立変数と従属変数の機能的な関係を記述することを重視していることを挙げ、日常言語をいくら洗練させてもその関係の記述が十分な一般性を持つことは不可能だが、数学的言語ではそれが可能になると考えている。例えば、前述のマッチング法則では、数式間の記号が科学的用語として明瞭であったがゆえに、最終的に上記のような一般性を見出すことができたと考えることができる。また Staddon (1983) は、

数理モデルの機能が説明の明瞭性を保証しているだけでなく、言語的説明や質的なモデルでは十分導き出すことが出来ない予測からの逸脱を、数理モデルが一貫した論理で指摘可能である点を高く評価している。

さらに Shull (1992) は、Nevin のいう統合性を一歩進めて、数学的言語の比喩的拡張性 (metaphorical extension) を認めている。彼は数理モデルを構築する共通の方法として、行動的事象を他の現象と類似させ概念化することを挙げている。そして適切な調整と言ひ換えによって、数理モデルが更に新たな現象へ応用され得ることを主張している。そしてこのような比喩的拡張の一例として、Herrnstein (1970) の反応率関数を挙げている。この反応率関数は、並立 VI VI スケジュール下の選択行動についてのマッチング法則を単一 VI スケジュール下の行動に当てはめた結果得られる関数である。単一 VI スケジュール下の行動として、反応率が測定されているオペラント行動とそれ以外の行動の 2 つを考え、それぞれによって獲得される 2 つの強化子を考える。こうすることで単一 VI スケジュール下の反応率が、オペラント行動に対する強化率と測定されない反応に対する強化率との和でオペラント行動に対する強化率を割った値に比例するという関係式を導くことができる。彼はこの他に比喩的拡張の例として、Estes の刺激サンプリング理論、工学からの拡張である線形システム理論 (McDowell, 1980, 1987; McDowell & Wixted, 1988; McDowell, Bass, & Kessel, 1992), 経済理論や調整理論を応用したモデル (例えば, Baum, 1973; Collier, Johnson, Hill, & Kaufman, 1986; Hursh, Raslear, Shurtleff, Bauman, & Simmons, 1988; Lea, 1978; Rachlin, Battalio, Kagel, & Green, 1981; Staddon & Motheral, 1978; Timberlake, 1984), 時間をベースとした確率モデル (Gibbon, 1977; Killeen & Fetterman, 1988; Luce, 1986) を挙げている。この比喩的拡張により、現象の基底に働く同一のメカニズムを仮定することで、個別に分離している諸現象に対してある法則の下に一貫した説明を与えることが可能となる。

いかにして観察データから有用な結果を引き出すかという点から、Iversen (1988) は、近年の数学と統計学の発展 (Tuft, 1983; Tukey, 1977) を踏まえ、データを視覚化することの効用を主張している。彼は分析がデータや実際の実験から離れることに対して Skinner も同様に警戒している点を指摘し、科学は理論に縛られない研究者間の情報伝達ルール (数学的処理によるデータの視覚化) を備えた経験的分析を通して発展すると考えるに

至った。したがって、数学や統計学の発展を背景に、数学的処理を行った上でデータを視覚化することを推奨するという点で前述の Catania (1981) とは異なる。

数理モデルは、数式で表現できるため、コンピュータによる数値計算が可能である。このような実験を現在ではシミュレーションと呼ぶことが多いが、広い意味ではフリーオペラント事態における動物を用いた実験もそれに含めることが出来る (例えば, Epstein, 1981)。なぜなら、シミュレーションは現象の部分または全体を表現している模擬実験をさし、どちらもこれに当てはまるからである。しかし 1980 年代以降、次第にその試みが増えてきたのは、コンピュータによるシミュレーション研究であった (Hinson & Staddon, 1983; Rachlin, 1982; Shimp, 1984a, 1984b; Ziriax & Silberberg, 1984)。Shimp (1984c) は、行動分析の理論的發展を促す手法として、認知的概念を取り入れた分析、行動の数理的分析、行動の特徴を完全に再現したコンピュータ・シミュレーションによる分析の 3 つを挙げているが、Skinner の立場とは異なる方向にすすむ現在の一部の傾向を予測している点は興味深い。

## 1.2. 行動の数理モデルの分類

行動分析で提案されてきたモデルの分類を試み、それによってこれらを整理することは、この領域における数理モデルの全体的な把握のために有効であると考えられる。数理モデルの分類・整理の試みは、行動分析以外の領域でも何人かの心理学者によってなされてきた (例えば, Estes, 1993; Luce, 1995)。行動分析における数理モデルを整理するために、ここでは行動の数量的モデルを考える代表的研究者である Church (1997) の分類を検討する。

### 1.2.1. Church の分類

Church (1997) は、心理学における数理モデルについて次の 4 つの分類を提案した。(1) 規範的モデル、(2) 心理学的モデル、(3) 生物学的モデル、(4) 形式的モデルである。

規範的モデルとは、環境の条件と個体の目標を述べたものを指す。このモデルは、一定の環境条件下で、目標を最大限に実現するような行動を記述している。規範的モデルを現実の場面に適用する場合、個体がこの目標に漸近的に達するように行動していると仮定する。このモデルの例としては、最適採餌理論 (Krebs, 1978; Stephens & Krebs, 1986), 信号検出理論 (Green & Swets, 1966) が挙げられている。

心理学的モデルとは、2 つまたはそれ以上の心的媒介

過程と、入力（刺激）と出力（行動）との間での変換規則からなるものを指す。代表的な心的媒介過程としては、表象、記憶、意思決定などがある。ここで行動は、何らかの刺激を環境から受け取り、内的な表象としてそれを表した後、既に記憶として蓄積された過去の経験や知識と照らし合わせられ、ある意思決定の基準に従って出力されるもの、と考えられている。この分類に当てはまる例としては、スカラー計時理論 (Gibbon, 1977)、行動計時理論 (Killeen & Fetterman, 1988) が挙げられている。

生物学的モデルとは、心的媒介過程の代わりに生物学的媒介過程（電気的あるいは化学的活動）が用いられたものを指す。行動の生物学的基礎に関する多くの研究は、神経系に障害を作ったり刺激を与えたりしたときに生物の行動に及ぼす効果や、脳の構造や機能と行動との相関について行われている。そして、現在もっとも分析が盛んに行われているモデルとして、ニューラル・ネットワーク (Levine, 1991; Moore & Desmond, 1992) が挙げられている。

形式的モデルとは、明確な心理学的・生物学的対応のない数量的媒介過程が使われているものを指す。条件づけ過程に関する媒介変数を有するほとんどの数量的モデル（例えば、Rescorla-Wagner モデル）は、その意味で形式的モデルであると言える。例えば、反応強度という測度が用いられていた場合、それを心理学的あるいは生物学的媒介変数として考えることも可能であるが、モデルを理解する上でその媒介変数に対する特別な仮定は必要としない。このモデルの例としては、時間と空間の次元を基礎として連合学習における随伴性の働きを導き出す、行動の計算論的分析 (Gallistel, 1990) などが挙げられている。

### 1.2.2. 新たな分類軸の提起

現象をどのように捉えるかによって、同じ現象でもいくつかの数理モデルが存在し得る (近藤, 1976)。それは、現象の持つ全ての様相を簡単な式で表現することが多くの場合困難だからである。例えば、現象の因果関係に注目すると確定的な (deterministic) モデルとなり、現象のばらつきに注目すると同じ原因からでも期待される結果が変動することがあるため確率的な (probabilistic) なモデルになるという。

1.2.1. でみたような Church (1997) の分類は、現象の捉え方によったものではなく、理論の性質や形式による分類である。つまりそれは、実際の行動とは異なる観察レベルのどこか別のところで起こっている出来事に訴え

てなされた観察事実についての説明 (Skinner, 1950) の分類となっている。その上、この分類を使って現時点の行動分析における数理モデルが、対象としている行動の性質や様相を知ることは難しい。そこで、実際の現象の捉え方によって分類すると、今度は以下に見るように Church の分類ではモデルが複数のカテゴリーに重複して現れてしまうという問題が起こる。したがって、Church の分類を考慮しつつ、現象の捉え方による分類を提案し、1.1. でみたような数理モデルの予測に関して理論面と実践面において指摘されてきた短所を数理モデルがどのように克服してきているのかも合わせて整理することが望ましい。

そのような現象の捉え方による分類としては次の5つが考えられる。(1) 現象記述モデル、(2) 過程モデル、(3) 最適化モデル、(4) 力学系モデル、(5) 確率的モデルである。

現象記述モデルとは、以下で述べるような均衡状態、処理過程、因果関係そして不確定性のいずれについても注目せず、現象の数学的表記をめざしたものをいう。このモデルでは、現象を何らかの仮定の下に捉えない。例として、観察された事象の特徴を数式によって単に簡潔に記述している前述のマッチング法則 (Herrnstein, 1961, 1970) や、測定値によって表すことが可能な媒介変数を使用した指数型加重移動平均モデル (Killeen, 1981, 1994) などを挙げることが出来る。これらのモデルは Church の分類では、形式的モデルの一部に相当する。

過程モデルでは、いくつかの段階に分かれた処理過程を経て行動的事象が結果されたと捉える。想定している各処理過程における変数が密接に結合しており、変数が一定の基準を越えると個体は行動を出力する。例えば、スカラー計時理論 (Gibbon, 1977) では、個体が強化までの時間を記憶し、累積された時間のパルス数を過去と現在で比較し、この比較に基づいて行動が生起すると仮定している。また、実際の強化の提示により、累積されたパルスの時間分布が更新される。このモデルは、Church の心理学的モデルや生物学的モデルに相当する。

最適化モデルとは、複数の行動的事象間の均衡状態に注目したものを指す。環境側の出力、例えば提示強化子数を環境の様々な制約とその結果である行動的事象の目的関数として表し、この目的関数を最大または最少にするような行動的事象の値を決めることを目指す。この行動的事象の値は均衡点と呼ばれることがある。例えば、

複数選択肢への各選択行動を時間的に配分する場面において、巨視的強化最大化理論 (Rachlin, 1982) では、長期的に見て強化量を最大にするような行動配分点を均衡点として捉える。このように最適化モデルにおける目的関数は、ある制約条件を持つ環境にさらされた個体の目的を数量化し、関数式で表現したものである。これは Church の規範的モデルに相当するが、個体の目的を行動的事象の関数式で表したものに限定している点でそれとは異なる。

力学系モデルとは、環境的事象と行動的事象との因果関係に注目したものを指す。このモデルでは、個体を 1 つのシステムとして捉え、一定の環境的事象からは決まってある行動的事象が出現することに注目している。システムを表現する方程式の変数には連続量が割り当てられる。また、実数平面上の任意の直線の像が直線、または特別な場合には点になる関数式を線形方程式といい、直線の像が直線や点にならない関数式を非線形方程式という (Gulick, 1992) が、後者の方程式ではその振る舞いが線形よりも豊かになるため、線形方程式では記述できないと指摘されている行動的事象の特徴、例えば FI スキャロップ (Hoyert, 1992) を表現することが可能となる。Church の分類では、形式的モデルに相当する。

確率的モデルでは、同じ原因の環境的事象から期待される結果である行動的事象が変動して観察されることに注目している。このモデルでは、この変動性を誤差と捉えるのでもなければ、必然的なふるまい (例えば、カオス) と捉えるのでもなく、現象の本質として確率過程で表すことのできる不確定性を持っていると考えている。確率論をその背景にもつことから、この偶然性を表現するためにその方程式には確率変数が使用される。Church の分類にはこのようなモデルは明確に表れていない。一方 Luce (1995) は、モデルを分類するために作られた 5 つの対立軸の最後のノイズ対構造 (noise vs structure) という対立項に統計モデリングと確率モデリングを立てている。これに基づけば、ここで述べている確率的モデルは後者と深く関わっている。

### 1.3 行動分析における既存の数理モデル

前述の分類基準にしたがい、(1), (2), (3), (4), (5) の順に行動分析における数理モデルを分類したのが、表 1 である。表 1 では、あわせて各モデルを次のような 2 点から整理している。第 1 は数理モデルの特徴に関するもので、媒介過程の採用・不採用、モデルが再現する行動の状態 (安定・変動)、モデルが再現しようとしている行動局面 (獲得相・変容相・維持相・消失相) である。第 2

は、その数理モデルを検証するために使用した実験手続きとその実験結果、またはそのような実験が行われていない場合にはモデルの妥当性の根拠となる事実という 2 つについて整理している。

ここで、その定式化が表 1 に挙げたモデルに何点か変更を加えたにすぎないモデルであるという理由から、以下に挙げるモデルは表 1 に載せなかった。現象記述モデルでは、マッチング法則に基づいてセッション間の変動の予測をしたモデル (Davis, Staddon, Machado, & Palmer, 1993)、一般化マッチング法則におけるフリーパラメーターである感受性 (sensitivity) に関する分析 (Schofield & Davison, 1997)、ウェーバーの法則を援用して強化の遅延の比に関する項を一般化マッチング法則に加えたモデル (Grace, 1996) である。力学系モデルでは、一般化マッチング法則に基づいて短い集成水準で選択肢間の時間配分の変動を予測するモデル (Mark & Gallistel, 1994) である。また、ベイズ統計モデルの 1 つとして Killeen, Palombo, Gottlib, & Beam (1996) もあるが、ここでの論評の範囲を越えているので、これもとりあげなかった。

表 1 にみるように、行動分析における数理モデルは Herrnstein のマッチング法則 (Herrnstein, 1961, 1970) を出発点とみることが出来る。70 年代後半は、それまでの現象記述モデルに加え、最適化モデル、過程モデルという 3 つの形態に、現象の捉え方が分かれてくる。80 年代前半には新たに力学系モデルが提案され、後半には数学的学習理論を背景に持つ確率的モデルが提案され始める。

これら数理モデルは、ほとんどのモデルの再現しようとする行動局面が維持相であるという特徴が目立っている。全局面を表現しようと試みているのは、スカラー計時理論 (Gibbon, 1977) と行動計時理論 (Killeen & Fetterman, 1988) のみである。

多くの研究でモデルの妥当性を検証する実験が行われているが、いくつかのモデルについてはそのような実験が行われていない。その代わりに、過去のデータとモデルの予測との適合度、あるいは広く認められている行動的事実との一致を検討して、その妥当性を主張している。例えば、巨視的強化最大化モデル (Rachlin, 1982)、行動計時理論 (Killeen & Fetterman, 1988)、連結制御モデル (Hanson & Timberlake, 1983)、非線形システムアプローチ (Hoyert, 1992)、一般化マッチング法則 (Baum, 1974, 1979)、Killeen (1981, 1994) の指数型加重移動平均モデルは、過去のデータとの適合をみることで、その

表 1 行動分析における数理モデルの分類

行動分析における 数値的モデル	明示的な 媒介過程	モデルが 再現する もの	行動の 局面	モデルの妥当性を確かめ るために使用された 実験手続き	モデルの予測 または実験結果	文献
現象記述モデル						
1) $B_1/(B_1+B_2)=R_1/(R_1+R_2)$	不採用	安定	維持相	COD 付き並立 VIVI スケジュール	反応数比と強化数比が一致 する。	Herrnstein (1961, 1970)
2) $B_1/(B_1+B_2)=(T-t_1)/\{(T-t_1)+(T-t_2)\}$ , ( $t_1 < T$ ; $t_2 < T$ のとき) $=1$ , ( $t_1 < T$ ; $t_2 > T$ のとき) $=0$ , ( $t_1 > T$ ; $t_2 < T$ のとき)	採用	安定	維持相	並立連鎖スケジュール	刺激の条件性強化子として の強度は、その刺激の開始 から相関した強化までの時 間の遞減関数になる。	Fantino (1969, 1977)
3) $B=kR/(R+R_e)$	不採用	安定	維持相	VI スケジュール	標的行動と非標的行動の相 対反応率が獲得強化率と一 致する。	Herrnstein (1970)
4) $B_1/B_2=k(R_1/R_2)^a$	不採用	安定	維持相	— (過去のデータ)	マッチング関係からの逸脱 を含めた行動量と強化量と のマッチング関係。	Baum (1974, 1979)
5) $R_1/B_1=R_2/B_2$	採用	変動	獲得相	並立 VIVI スケジュール 4 分単位で局所強化率を評価	各選択肢のコスト(強化数/ 反応数)を等しくするよう に行動を配分する。	Herrnstein & Vaughan (1980), Vaughan (1981)
6) $B=\zeta/\delta-N/a$ ( $N < a/\delta$ ; $\delta, a > 0$ のとき) $=0$ ( $N \geq a/\delta$ ; $\delta, a \leq 0$ のとき) ただし, $\zeta=\rho M_n$	採用	変動	維持相	— (過去のデータ; 比率(ratio) スケジュール)	記憶された強化を受けた反 応間隔は定常であり、そ の記憶に基づいて反応が生 成される。	Killeen (1981, 1994)
7) $B=kR/(R+1/a)-R/\lambda$ ( $\lambda, a, \delta > 0$ ; $R < \lambda k - 1/a$ のとき) $=0$ ( $\lambda, a, \delta \leq 0$ ; $R \geq \lambda k - 1/a$ のとき) ただし, $k=p/\delta$	採用	変動	維持相	— (過去のデータ; 時隔 (interval) スケジュール)	記憶された強化を受けた反 応間隔は定常であり、そ の記憶に基づいて反応が生 成される。	Killeen (1981, 1994)

<sup>①</sup> 一般に  $B, B_1, B_2$  は、オペラント行動量を表し、 $R, R_1, R_2$  は強化量を表す。添え字は、選択肢 1 あるいは 2 に対する反応と強化を表している。以下、それ以外の記号について概説する。

<sup>②</sup>  $T$  は初環と終環の長さをあわせた平均強化間隔、 $t_1, t_2$  は終環の平均時間間隔を表す。

<sup>③</sup>  $\zeta$  は標的行動と強化に関する記憶の程度、 $\delta$  は最小反応間隔、 $N$  はある比率スケジュールが要求する反応数、 $a$  は反応していた秒数（反応は誘因によって誘発されると仮定されている）、 $M_n$  は  $n$  番目の反応後に指数的に重みづけされた反応間隔の記憶の総和を表す。

<sup>④</sup>  $k$  はある時隔スケジュール下における反応数の予想される漸近値、 $\lambda$  は短期記憶の減衰率、 $\rho$  は標的行動を標的行動以外の反応数の関数として表したときの関数値を表す。



表1 行動分析における数理モデルの分類(続き)

行動分析における 数量的モデル	明示的な 媒介過程	モデルが 再現する もの	行動の 局面	モデルの妥当性を確かめ るために使用された 実験手続き	モデルの予測 または実験結果	文献
過程モデル						
8) [ペースメーカー—累積記録器]→ 反応決定ルール→反応生起	採用	変動	獲得相, 維持相, 変容相, 消失相	時間般化法, 残存時間法, ピーク法	強化までの時間を記憶し, 累積されたペースメーカー が発信したパルス数との比 較に基づいて行動を生起す る。	Church & Gibbon (1982), Gibbon (1977), Gibbon & Church (1981), Roberts (1981)
9) 時間知覚→ [ペースメーカー—行動の状態]→ 反応決定ルール→反応生起	採用	変動	獲得相, 維持相, 変容相, 消失相	— (過去のデータ)	知覚した強化までの時間 と, ペースメーカーが発信 したパルスと結びつけて記 憶した行動の状態との比較 に基づいて行動を生起す る。	Killeen & Fetterman (1988)
最適化モデル						
10) $T_1/T_2 = (m_1/m_2)(x_1/x_2)(C_1/C_2)^y$	採用	安定	維持相	— (過去のデータ)	全体強化率(効用)を最大 にするように, 行動を時間 的に配分する。	Rachlin (1982)
11) $S_i = p_i[\exp(-qT_i) + 1/T_i]/(I_i + T_i)$ $B_1/(B_1 + B_2) = S_1/(S_1 + S_2)$	採用	安定	維持相	— (過去のデータ)	ある選択肢に対応した強化 率はそのアローザルの程度 を決定し, オペラント反応 に随伴した強化はそのア ローザルをオペラントに 対する反応に向けさせる。	Killeen (1982)

<sup>注)</sup> 一般に  $B, B_1, B_2$  は, オペラント行動量を表し,  $R, R_1, R_2$  は強化量を表す。添え字は, 選択肢 1 あるいは 2 に対する反応と強化を表している。以下, それ以外の記号について概説する。

<sup>10)</sup>  $T_1, T_2$  は選択肢 1 あるいは 2 に対するオペラント行動の単位時間の配分量を表す ( $T_1 + T_2 = 1$ )。また,  $C$  を単位時間  $T$  だけ行動したときに受ける結果事象であるとする。このとき, 単位時間  $T_1$  との関係は  $C_1 = a_1 T_1^{m_1}$  (ただし,  $0 \leq a_1, 0 \leq C_1, T_1, m_1 \leq 1$ ) となり, 同様に  $C_2$  についても定式化し  $T_1, T_2$  の関係式に代入すると,  $(C_1/a_1)^{1/m_1} + (C_2/a_2)^{1/m_2} = 1$  となる。次に無差別曲線を求めるため, 効用関数を  $U = x_1 C_1^y + x_2 C_2^y$  で表す。  $U$  は選択肢 1 と選択肢 2 で得られる結果事象の効用,  $x_1, x_2$  はフリーパラメーター,  $y$  は  $C_1$  と  $C_2$  の代替率を表している。これに先程の  $C_1, C_2$  を代入する。ただし,  $T_2 = 1 - T_1$  とすると,  $U = x_1 a_1 y T_1^{m_1 y} + x_2 a_2 y (1 - T_1)^{m_2 y}$  となる。  $U$  の最大値を求めるために  $T_1$  について微分し, それを 0 とおいて整理すると表中の形式になる。

<sup>11)</sup>  $S$  は 1 次性強化子の強度,  $p$  は遅延のある 1 次性強化子の強度,  $I$  は初環,  $T$  は終環の長さ,  $q$  はフリーパラメーターを表す。

表 1 行動分析における数値モデルの分類 (続き)

行動分析における 数値的モデル	明示的な 媒介過程	モデルが 再現する もの	行動の 局面	モデルの妥当性を確かめ るために使用された 実験手続き	モデルの予測 または実験結果	文献
最適化モデル						
12) $V = A/(1 + kD)$	採用	安定	維持相	調整スケジュール	強化子の価値は、強化の遅延時間の双曲線関数になる。	Mazur (1987)
力学系モデル						
13) $B(t) = \int_0^t G(t-t')R(t')dt'$	採用	変動	維持相	— (過去のデータ)	ある時点より前のすべての行動がその時点の行動に、それぞれ異なる強度で線形的に反映される。	McDowell (1980), McDowell, Bass, & Kessel (1992)
14) $dt_1/dt = b(s_0^1 - t_2)t_1 - t_1/s_0^1 a_1$ , $dt_2/dt = -\{(s_0^1 - t_1)t_2\}/b - t_2/s_0^2 a_2$	不採用	安定	維持相	— (過去のデータ)	オペラント反応と強化子として随伴する反応について、互いに一方の行動量を他方の行動量の変化のパラメーターとした常微分方程式の解を行動量の予測値とする。	Hanson & Timberlake (1983), Timberlake (1984)
15) $d\lambda/dt = kR_2(c - \lambda) - kR_1\lambda$	不採用	変動	維持相	並立 VRVR スケジュール 並立 EXTEXT スケジュール	強化量は選択肢間の切り替え反応率を減衰させる。	Myerson & Meizin (1980), Myerson & Hale (1988), Nevin (1992)
16) $m_1/m_2$ $= (r_{c1}r_{s2}/r_{s1}r_{c2})^a (r_{k1}r_{s2}/r_{s1}r_{k2})^b$	不採用	安定	維持相	多元並立スケジュール	2つの成分スケジュールの行動の変化抵抗比と強化率比は一致する。	

注) 一般に  $B, B_1, B_2$  は、オペラント行動量を表し、 $R, R_1, R_2$  は強化量を表す。添え字は、選択肢 1 あるいは 2 に対する反応と強化を表している。以下、それ以外の記号について概説する。

<sup>12)</sup>  $V$  は強化価値、 $A$  は強化量、 $D$  は遅延時間、 $k$  はフリーパラメーターを表す。

<sup>13)</sup>  $t'$  は実際に強化を受けた時間、 $t$  は強化までの時間間隔を、 $G$  は重み付け関数を表す。

<sup>14)</sup>  $s_0^1, s_0^2$  はそれぞれオペラント反応と強化子として随伴する反応それぞれのセットポイント (至高点) を表す。セットポイントとは、ある環境の制約下で最も効率の良い反応時間の配分点を意味する。そして、 $t_1, t_2$  はそれらの反応の時間配分量、 $a_1, a_2$  はそれらの反応の変化抵抗、 $b$  はフリーパラメーターを表す。

<sup>15)</sup>  $\lambda$  は選択肢 1 で強化を受けるときに選択肢 2 からの切り替え反応率、 $k$  は別の選択肢への切り替え反応の減少率、 $c$  はフリーパラメーターを表す。

<sup>16)</sup> このモデルは古典力学を背景に持つ。 $m_1, m_2$  は行動質量、 $r_{c1}, r_{c2}$  は 2 つの成分スケジュールの強化比率に対する変化抵抗、 $r_{k1}, r_{k2}$  は強化比率、 $r_{s1}, r_{s2}$  はセッション全体の強化率を表す。 $a, b$  は刺激制御の程度を表すパラメーターである。

表1 行動分析における数理モデルの分類(続き)

行動分析における 数値的モデル	明示的な 媒介過程	モデルが 再現する もの	行動の 局面	モデルの妥当性を確かめ るために使用された 実験手続き	モデルの予測 または実験結果	文献
力学系モデル						
17) $B = (S_r - S_a)p$ $\delta S_r = a S_r (S_{rmax} - S_r)$ $\delta S_a = d S_a (S_{amax} - S_a)$ ただし, $S_r = 1 - S_a$ $= j[\sin(k\pi(MI + SI + T(R)))] + 1$ $T(R) = f(R + m + n)$ また, $T_{n,m}(R) = 1$ $(O + o(m+n) < R < m + o(m+n);$ $o = 0, 1, 2, \dots \text{のとき})$ $= 0$ $(m + o(m+n) < R < (o+1)(m+n);$ $o = 0, 1, 2, \dots \text{のとき})$	採用	変動	維持相	FI スケジュール	スキヤロップの周期性	Hoyert (1992)
確率的モデル						
18) $\delta s = s(p + q - 2w) + w - q$	採用	安定	維持相	並立 VRVR スケジュール (2つの VR スケジュールサ イズは同じ)	2つの選択肢の強化確率を 等しくしたとき, その確率 が低いと選好が偏り, 高い と無作為な反応の配分にな くなる。	Horner & Staddon (1987), Staddon (1988), Staddon & Horner (1989)

<sup>17)</sup> 一般に  $B, B_1, B_2$  は, オペラント行動量を表し,  $R, R_1, R_2$  は強化量を表す。添え字は, 選択肢 1 あるいは 2 に対する反応と強化を表している。以下, それ以外の記号について概説する。

<sup>17)</sup>  $B$  は測定可能な反応率,  $S_r$  はオペラント反応の反応強度,  $S_{rmax}$  は当該強化間隔におけるオペラント反応の最大反応強度,  $S_a$  はそれ以外の反応の反応強度,  $S_{amax}$  は当該強化間隔におけるそれ以外の反応の最大反応強度,  $p$  は反応強度を測定可能な数値に変換するパラメーターを表す。また,  $a, d$  はフリーパラメーターである。 $\delta S_r, \delta S_a$  をそれぞれ時間  $t$  について積分することで,  $S_r, S_a$  を求める。ここで,  $S_r = 1 - S_a$  という関係を仮定する。このとき  $j, k, l, m, n$  をフリーパラメーターとし, 単一 FI スケジュールにおいて  $MI$  を 3 時隔前までの反応強度の平均,  $SI$  を直前の時隔までの反応強度の合計,  $T(R)$  を 1 セッション中に獲得した総強化子数に関する矩形波関数,  $o$  を現時点を含めた 4 時隔で提示された強化子の周期性を示すパラメーターとすると,  $S_r$  は表中のように表される。

<sup>18)</sup>  $s$  は選択肢 1 に対する相対反応配分量を表し,  $p, q$  はそれぞれ選択肢 1, 選択肢 2 における強化確率を表す。 $w$  は各選択肢において強化による反応の増分の比 ( $a(s)/b(s)$ ) が一定であると仮定したときに見出されるパラメーターを表す。

表 1 行動分析における数理モデルの分類 (続き)

行動分析における 数値的モデル	明示的な 媒介過程	モデルが 再現する もの	行動の 局面	モデルの妥当性を確かめ るために使用された 実験手続き	モデルの予測 または実験結果	文献
確率的モデル						
19) $B_w/B_x$ $= c \{ (d_s d_r R_w + R_z) / (d_r R_z + d_s R_w) \},$ $B_y/B_z$ $= c \{ (d_r R_w + d_s R_z) / (d_s d_r R_z + R_w) \}$	不採用	安定	維持相	見本合わせ手続き	信号検出理論における弁別 容易度という指標を刺激— 反応弁別容易度と反応—強 化子弁別容易度という2つ に分け、これらがそれぞれ 一般化マッチング法則にお けるフリーパラメーター (反応のバイアス、強化子 に対する感受性) に対応す る。	Davison (1991), Davison & Jenkins (1985), Godfrey & Davison (1998)
20) $\delta p / \delta n = \alpha p (\beta - p)$	不採用	変動 (変動 そのものを 1 単位のオ ペラントと 捉え、この オペラント 内の反応系 列における 変動)	維持相	反応頻度依存型 スケジュール	2 選択肢に対して、反応の ランダムネスを 1 単位のオ ペラントとして形成する と、その反応系列はマルコ フ性を持つ。	Machado (1997)
21) $p_{i,t+1} = 1 - [(1 - \alpha)^k (1 - p_{i,t})],$ $p_{i,t+1} = p_{i,t} (1 - \alpha)^k$	不採用	変動	維持相	— (行動的事実)	選好逆転現象	March (1996)

注) 一般に  $B_1, B_2$  は、オペラント行動量を表し、 $R, R_1, R_2$  は強化量を表す。添え字は、選択肢 1 あるいは 2 に対する反応と強化を表している。以下、それ以外の記号について概説する。

<sup>19)</sup>  $B_w, B_x$  は選択肢 1 における行動量で、前者が強化を受けた反応、後者が非強化の反応について表している。 $B_y, B_z$  は選択肢 2 における行動量で前者が非強化の反応、後者が強化を受けた反応について表している。これらの値の比がそれぞれ、信号検出理論で仮定されている 2 つの正規分布上の確率を決定する基準になる。また、強化あるいは非強化量は  $R$  で表し、添え字は前述の反応の分類に対応している。信号検出理論における弁別容易度が 2 つに分けられ、 $d_s$  は刺激-反応弁別容易度、 $d_r$  は反応-強化弁別容易度を表す。

<sup>20)</sup> 1 単位のオペラントとして見なされる反応系列において、選択肢 1 から選択肢 2 へ反応を切り替える確率を  $p$ 、その切り替えの回数を  $n$  で表す。 $\alpha, \beta$  はフリーパラメーターである。

<sup>21)</sup>  $p_{i,t}$  は選択肢  $i$  における  $t$  時点での反応生起確率、 $k$  は一度の強化で提示される強化量、あるいは一度に損失する損失量、 $\alpha$  はフリーパラメーターを表す。

表 1 行動分析における数理モデルの分類 (続き)

確率的モデル	行動分析における 数量的モデル	明示的な 媒介過程	モデルが 再現する もの	行動の 局面	モデルの妥当性を確かめ るために使用された 実験手続き	モデルの予測 または実験結果	文献
22)	$A_{i,t} = S_{i,t}/m_{i,t}$ $p_{i,t} = A_{i,t}/\sum A_{i,t}$ $1 - p_{i,t} = A_{i,t}/\sum A_{i,t}$	不採用	変動	維持相	— (行動的事実)	選好逆転現象が観察された が、損失場面ではリスク選 択の割合がやや少ない。	March (1996)
23)	$A_{i,t+1} = \beta O_{i,t} + (1 - \beta) A_{i,t}$ $p_{i,t} = A_{i,t}/\sum A_{i,t}$ $1 - p_{i,t} = A_{i,t}/\sum A_{i,t}$	不採用	変動	維持相	— (行動的事実)	選好逆転現象が観察された が、損失場面ではリスク選 択の割合がやや少ない。	March (1996)

注) 一般に  $B, B_1, B_2$  は、オペラント行動量を表し、 $R, R_1, R_2$  は強化量を表す。添え字は、選択肢 1 あるいは 2 に対する反応と強化を表している。以下、それ以外の記号について概説する。

<sup>22)</sup>  $p_{i,t}$  は選択肢  $i$  における  $t$  時点での反応生起確率、 $k$  は一度の強化で提示される強化量、あるいは一度に損失する損失量を表す。 $S_{i,t}$  は選択肢  $i$  における  $t-1$  時点までの総強化量、あるいは総損失量、あるいは一度に損失する損失量、あるいは一度に損失する損失量を表す。

<sup>23)</sup>  $p_{i,t}$  は選択肢  $i$  における  $t$  時点での反応生起確率、 $k$  は一度の強化で提示される強化量、あるいは一度に損失する損失量、あるいは一度に損失する損失量を表す。 $O_{i,t}$  は選択肢  $i$  における  $t$  時点での強化量、あるいは一度に損失する損失量、あるいは一度に損失する損失量を表す。

妥当性を主張している。また March (1996) の 3 つのモデルは、その妥当性を不確実状況下において利得場面では高確率・低利得のリスクレス選択肢を、損失場面では低確率・高利得のリスク選択肢を偏好するという選好逆転現象との一致を検討することに拠っている。

## 2. 数理モデルの運用

本節以降では、表 1 で紹介したモデルのうち、March (1996) のモデルに焦点を当て、実際にこれを運用して行うことで行動分析における数理モデルの予測と制御について検討していきたいと考える。このモデルを検討の材料とした理由は、次の 3 点である。

- 1) 媒介過程を採用しているモデルは、それを採用している理論やモデルがその過程をその後の論理展開の出发点としているために、直接観察可能な現象の生起原因を観察不可能な他の次元に求めてしまっている可能性がある。その点、March のモデルは観察可能な変数しか扱っておらず、Epstein (1981) が「パフォーマンス・ルール」と呼んで批判するモデル固有の説明概念を特別に想定することなく、行動分析をはじめとする行動研究が積み上げてきた考え方を基本としてシミュレーションができる。
- 2) このモデルの 1 つは、March も述べているような数学的学習理論における初期のモデルを改編したモデルである。また、このモデルは実際に観察可能な要素によって構成され、その構造が単純であるため、数理モデルの運用に関して検討を加えやすい。
- 3) 「変動」する行動の性質を再現しようとしているモデルは、強化随伴性の時系列的な影響について検討できる可能性がある。今まで不確実性についての行動研究では、時間の流れに沿った選択行動の変容過程にはほとんど注意が向けられてこなかった(坂上, 1994) ため、このモデルのように直接、変容過程を検討できることは、今後の研究に資するところが大きいと考えられる。

### 2.1. 数学的学習理論における定式化の方法

March (1996) では、不確実状況における選択行動において次節でみるような 3 つの確率的モデルを作成し、利得・損失場面で起こる選好逆転の予測に関してこれらのモデルを比較・検討しているが、ここで確率的モデル、特に深く関連していると考えられる数学的学習理論における 2 つのモデルの定式化の方法について整理しておきたい。この方法は、後述する March の部分的調整モデルを考える上で重要である。

Coombs *et al.* (1970) によれば、数学的学習理論が生みだした2つの展開、刺激サンプリング理論における状態モデルと確率論的学習理論における作用素モデルは、いずれも学習過程を確率的に捉え同一の方法で定式化しているとされる。どちらのモデルでも学習過程は離散試行の系列と見なされていて、各試行において刺激事態が与えられ、その事態に対して被験体は、各選択肢に付与されている確率で反応選択肢の集合から1つの選択肢を選び、反応すると仮定されている。反応には結果が伴い、それによって次の試行が始まる前に反応確率の更新が行われる。

## 2.2. March のモデルのシミュレーション

### 2.2.1. March のモデルの定式化

さて、March (1996) の3つのモデルは次のように定式化される。まず、部分的調整モデル (fractional adjustment model) は  $p_{i,t}$  を選択肢  $i$  における  $t$  時点での反応生起確率比とし、 $k$  を一度の強化で提示される強化量、 $\alpha$  ( $0 \leq \alpha \leq 1$ ) をフリーパラメーターとすると、強化時と非強化時における予測式は順に、

$$(1) \quad p_{i,t+1} = 1 - [(1-\alpha)^k (1-p_{i,t})]$$

$$(2) \quad p_{i,t+1} = p_{i,t} (1-\alpha)^k$$

となる。部分的調整モデルは、強化量による効果を扱うことができるように Bush and Mosteller の作用素モデルや Estes の状態モデルを改編したものであって、強化量の違いを反映した項  $(1-\alpha)^k$  を除いて基本的には過去の数学的学習理論の伝統をそのままこの場面に適用したものと考えることができる。

平均返戻モデル (average return model) は、それぞれ  $S_{i,t}$  を選択肢  $i$  における  $t-1$  時点までの総強化量（あるいは総損失量）とし、 $m_{i,t}$  を選択肢  $i$  に対して  $t-1$  時点までに選択した総数とすると、 $t-1$  時点における平均強化量  $A_{i,t}$  は、

$$(3) \quad A_{i,t} = S_{i,t} / m_{i,t}$$

と表わされ、強化時と非強化時における予測式はそれぞれ、

$$(4) \quad p_{i,t} = A_{i,t} / \Sigma A_{i,t}$$

$$(5) \quad p_{i,t} = 1 - A_{i,t} / \Sigma A_{i,t}$$

となる。このモデルはマッチング法則の考え方を取り入れたもので、 $t-1$  時点の選択肢  $i$  における強化量を  $t-1$  時点までの総強化量で割った値を、 $t$  時点の選択肢  $i$  における反応生起確率として表している。

加重平均返戻モデル (weighted return model) では、 $t$  時点の選択肢  $i$  における強化量（あるいは損失量）を  $O_{i,t}$  として、(3) 式の代わりに

$$(6) \quad A_{i,t+1} = \beta O_{i,t} + (1-\beta) A_{i,t}$$

を用いて選択肢  $i$  の選択回数あたりの強化量（あるいは損失量）を求める。ここで、 $\beta$  ( $0 \leq \beta \leq 1$ ) はフリーパラメーターとする。以降の予測式は平均返戻モデルと同じになる。別の言葉で言えば、 $t+1$  時点の強化量は、 $t$  時点の強化量の「記憶」と選択肢  $i$  における  $t-1$  時点までの総強化量の「記憶」にそれぞれ重み付けした値の和として表されている。

### 2.2.2. シミュレーションの手順

March (1996) では、ほとんどシミュレーションの手順、具体的設定については触れられていない。そこで論文に示された結果を参考にしながら、2.2.1. で述べた定式的具体化を試みた。

図1は、前述の数学的学習理論のモデルの定式化に沿って、シミュレーションの手順を示している。シミュレーションでは、各試行毎に必ず標的行動が生じ強化される離散試行での行動が再現され、次のようなモジュールから出来ていると想定された（図1）。

- (1) 強化子供給装置：強化子用乱数発生装置と強化子供給基準からできている。強化子用乱数発生装置は0から1までの乱数を発生し、強化子供給基準よりも高い値の時に強化子を発生する。強化子供給基準は初期値として外部から与えられ、その後変化しない。
- (2) 反応出現装置：反应用乱数発生装置と反応出現基準からできている。反应用乱数発生装置は0から1までの乱数を発生し、反応出現基準よりも高いか、低いかによって2種類の反応（選択肢1に対する反応、選択肢2に対する反応）のどちらかを出現させる。3つのモデルでは、各試行毎に必ずどちらかの反応が出現する。反応出現基準はシミュレーションの開始時に初期値（両選択肢共に0.5）として外部から与えられるが、その後、反応出現基準変更計算によって試行毎に更新される。
- (3) 反応増減ルール：以下に述べる反応増減ルールの組み合わせは、論理的には18通りあるが、例えば強化を受けた反応の生起確率が減少し、強化を受けない反応の生起確率が増加することは強化の一般的概念から外れるので除外した。除外されなかった妥当な反応増減ルールは、以下の通りである。出現した反応  $X$  に強化子の提示が伴った場合には、(1) 反応  $X$  は増加させて反応  $Y$  は減少させる、(2) 反応  $X$  は増加させて反応  $Y$  は変化させない、(3) 反応  $X$  は変化させず反応  $Y$  は減少させる、(4) 反応  $X$  も反応

Y も変化させない、の 4 通り考えられる。また、出現した反応が強化子の提示を伴わない場合について、(1') 反応 X は減少させて反応 Y は増加させる、(2') 反応 X は変化させず反応 Y は増加させる、(3') 反応 X は減少させて反応 Y は変化させない、(4') 反応 X も反応 Y も変化させない、の 4 通り考えられる。こうして (1) から (4) の内の 1 つのルールと、(1') から (4') の内の 1 つのルールとの組み合わせによって、反応増減ルールが出来上がる (ただしこの中には (4) と (4') のような意味のないルールは含まないものとする)。シミュレーションの結果から、反応増減ルールはモデルによって適切なものを設定しないと、その結果が不安定になることが判明した。

(4) 反応出現基準変更計算: モデルの中核部をなし、反応増減ルールにおける増加や減少の値を決定する具体的な数値計算を行う。すでに述べた部分的調整モデル、平均返戻モデル、加重平均返戻モデルの 3 つの計算式がこれにあたる。

続いてモデルにおける計算の流れについて述べる<sup>注2)</sup>。

- (1) 「単位時間」は試行であり、試行の終わりで反応出現基準の時系列データを我々は観察している。
- (2) 初期値として、強化子供給基準、反応出現基準にある決まった値が与えられ、反応出現基準変更計算に用いられた計算式のフリーパラメーターの値 (部分的調整モデルでは  $\alpha$ 、加重平均返戻モデルでは  $\beta$ ) が与えられる。
- (3) 反应用乱数発生装置を駆動し、2 種類の内のどちらかの反応を出現させる。
- (4) 強化子用乱数発生装置を駆動し、強化子の提示、非提示を決定する。
- (5) 反応増減ルールに従い、強化子の提示か非提示かで、2 種類の反応の増減もしくは無変化を決める。
- (6) 増加、減少の値を各モデルでの反応出現基準変更計算に基づいて行い、その結果を次の試行の反応出現基準とする。
- (7) シミュレーションは 300 試行行い、これを 100 個体分実行したものの平均値をグラフに表す。

数理モデルは、特定の 2, 3 の事実だけではなく、多くの事実を説明できることも求められる (近藤, 1976)。March のモデルに関して言えば、原論文で検討している選好逆転現象の他に、確率マッチングという現象について検討することでその妥当性を確かめようと考えた。この確率マッチングという現象は、確率学習の実験で見出された、選択肢における反応生起確率の割合がその反

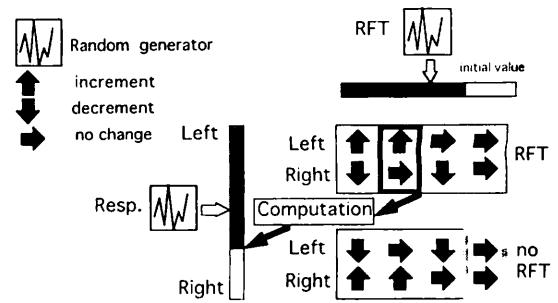


図1 行動選択場面におけるシミュレーション  
シミュレーションにおける計算方法の概略。図中の上向きの矢印は、反応生起確率が計算によって増加することを、下向きの矢印は減少することを表している。この図の例では、強化時には濃い太枠で囲んだ計算方法が取られ、消去時には薄い太枠で囲んだ計算方法が取られるモデルのシミュレーションを示している。1つのモデルをシミュレートするときには強化時 (損失時)、消去時の計算方法は一貫して同一のルールが採用されていた。

応の強化確率と一致する現象をいう。その実験では、点灯確率がそれぞれ  $p, q$  (ただし  $p > q$ ) の 2 つのライトを提示し、どちらのライトが点灯するかを被験者に言わせ、ライトの点灯によってその当否を知らせるまでを 1 試行とし、これを離散試行で行う。このとき、必ずどちらかのライトが点灯することになっている。すると、点灯確率  $p$  のライトを確率 1 で選べば全体としては正答の確率が高いにもかかわらず、多くの試行を重ねどちらのライトがどの程度の確率で点灯するか見当がついたとしても、点灯確率  $p$  のライトを確率 1 で選択しないで確率  $p$  で選択する確率マッチングがヒトを被験者としたいくつもの実験で見出された。しかし、確率学習を強化スケジュールに置き換えた並立 VRn VRm 強化スケジュール (ただし  $n < m$ ) の場合、多くの動物実験からより値の小さい VRn 側に確率 1 で選択がなされることが観察されている。そしてこのような VRn 側を排他的に選択する選択は、逐次改良理論によっても強化最大化理論によっても予測される合理的な選択であるといえる。

ところが、確率学習に分類されてきた手続きの中でも、どちらかが排他的に選択される現象が観察されないことがある (Robbins, 1980)。一般に確率学習の実験は離散試行で行われ、報酬は各試行で必ずどちらかの選択肢で提示されることが保証されていた。また、反応が強化されないエラーと呼ばれる試行も 1 試行として数え

られ、反応確率が計算される。この手続きは、矯正法 (correction method) と呼ばれる。これに対して、同じく離散試行で報酬が各試行で必ずどちらかの選択肢に提示されることを保証せずに、個体が報酬を得られなかった試行を無効とする手続き (非矯正法, non-correction method) では、強化確率の高い選択肢が排他的に選択されることが観察される。後者の手続きとフリーオペラント手続きの共通点は、どちらかの選択肢に対して反応した時点で、報酬 (強化子) が選択肢のどちらかで提示されるとは限らない点と、強化事象の確率が被験体の反応に依存している点であるといえる。しかし、被験体の反応を1試行毎に区切る離散試行と1セッション毎に区切るフリーオペラント手続きの違いを越えて、確率マッチングの成立・非成立を論じた試みは今のところ見あたらない。

Coombs *et al.* (1970) は、確率マッチングという現象が、次の5条件を満たすときに生じるとしている。(1) 強化事象の確率が被験体の反応に依存しない。(2) 一方が強化されるとき他方は強化されない2つの選択肢しか含んではならない。(3) インストラクションによって、全ての試行で正しい反応をするよう被験者を導き、系列がランダムであると暗示してはならない。(4) 分析は最初の数百試行に限る。(5) 個人の行動よりも集団平均の方が良く、また、顕著にこの現象が生じる種が存在する。したがって、少なくとも我々は、上記の5条件のうち(1)~(4)の条件を満たすことで確率マッチングについて March のモデルを検証することにより、モデルの妥当性の検討に加えて、そのモデルが強化事象の確率に反応が依存、あるいは非依存の事態のどちらを想定しているのかも知ることができよう。

### 2.2.3. 部分的調整モデルのシミュレーションの設定

このモデルのシミュレーションでは、(1)式、(2)式におけるフリーパラメーター $\alpha$ の値、強化確率と一度の強化量あるいは損失確率と一度の損失量の値を次のように設定した。利得・損失場面では、 $\alpha$ の値を0.2、選択肢1においては強化 (損失) 確率1、一度の強化量を1または損失量を-1とし、選択肢2においては強化 (損失) 確率0.1、一度の強化量を10または損失量を-10とした。また、確率マッチング現象の検証のために $\alpha$ の値を0.2、選択肢1では強化確率0.8, 0.5, 0.2で一度の強化量を1、選択肢2では強化確率0.2, 0.5, 0.8で一度の強化量を1とした。

このモデルでは、原論文の結果を導くために反応生起確率の変化の仕方を次のように設定した。利得 (損失)

場面では、一方の選択肢が強化を受けた場合、その選択肢に対する反応生起確率を増加させ、他方の選択肢に対する反応生起確率を減少させた。損失場面では、一方の選択肢で損失した場合、その選択肢に対する反応生起確率を減少させ、他方の選択肢に対する反応生起確率を増加させた。また、どちらの反応も強化を受けなかったとき、あるいは損失しなかったときは、前回の反応生起確率の値を維持させた。したがって、このモデルでは利得場面で強化を受けず、あるいは損失場面で損失しないときは、前回の反応生起確率を維持することになる。

### 2.2.4. 平均返戻モデル、加重平均返戻モデルのシミュレーションの設定

これらのモデルのシミュレーションにおける強化確率と一度の強化量あるいは損失確率と一度の損失量の値は、部分的調整モデルと同じ値を設定した。また、加重平均返戻モデルのフリーパラメーター $\beta$ の値は、0.2とした。

これらのモデルでも、原論文の結果を導くために反応生起確率の変化の仕方を次のように設定した。利得 (強化) 場面で一方の選択肢が強化を受けた場合、その選択肢に対する反応生起確率を増加させ、他方の選択肢に対する反応生起確率は前回強化を受けた時の値をそのまま維持した。損失場面では、一方の選択肢で損失した場合、その選択肢に対する反応生起確率は減少し、他方の選択肢に対する反応生起確率は前回の値をそのまま維持した。また、どちらの反応も強化を受けなかったとき、あるいは損失しなかったときは、前回の反応生起確率の値を維持させた。これらのモデルでは、利得場面で強化を受けず、あるいは損失場面で損失しないときの反応生起確率だけでなく、一方の反応に強化あるいは損失事態が随伴したときの他方の反応の反応生起確率までも前回の反応生起確率の値を維持することになる。

## 2.3. シミュレーションによって見出された結果

### 2.3.1. 利得・損失場面における選好逆転現象の予測

図2は、Marchのモデルのシミュレート結果を示している。図2では、上段から順に部分的調整モデル、平均返戻モデル、加重平均返戻モデルのシミュレート結果を示している。また、図2の左側は選好逆転現象に関して、右側は確率マッチング現象に関する結果を表している。

2.2.3., 2.2.4. で述べたように、Marchのモデルを設定した上で、実際にシミュレートしてみると、3つのモデルとも選好逆転現象を予測した (図2左側)。すなわち、不確実状況下において利得場面ではリスクレス選択、損



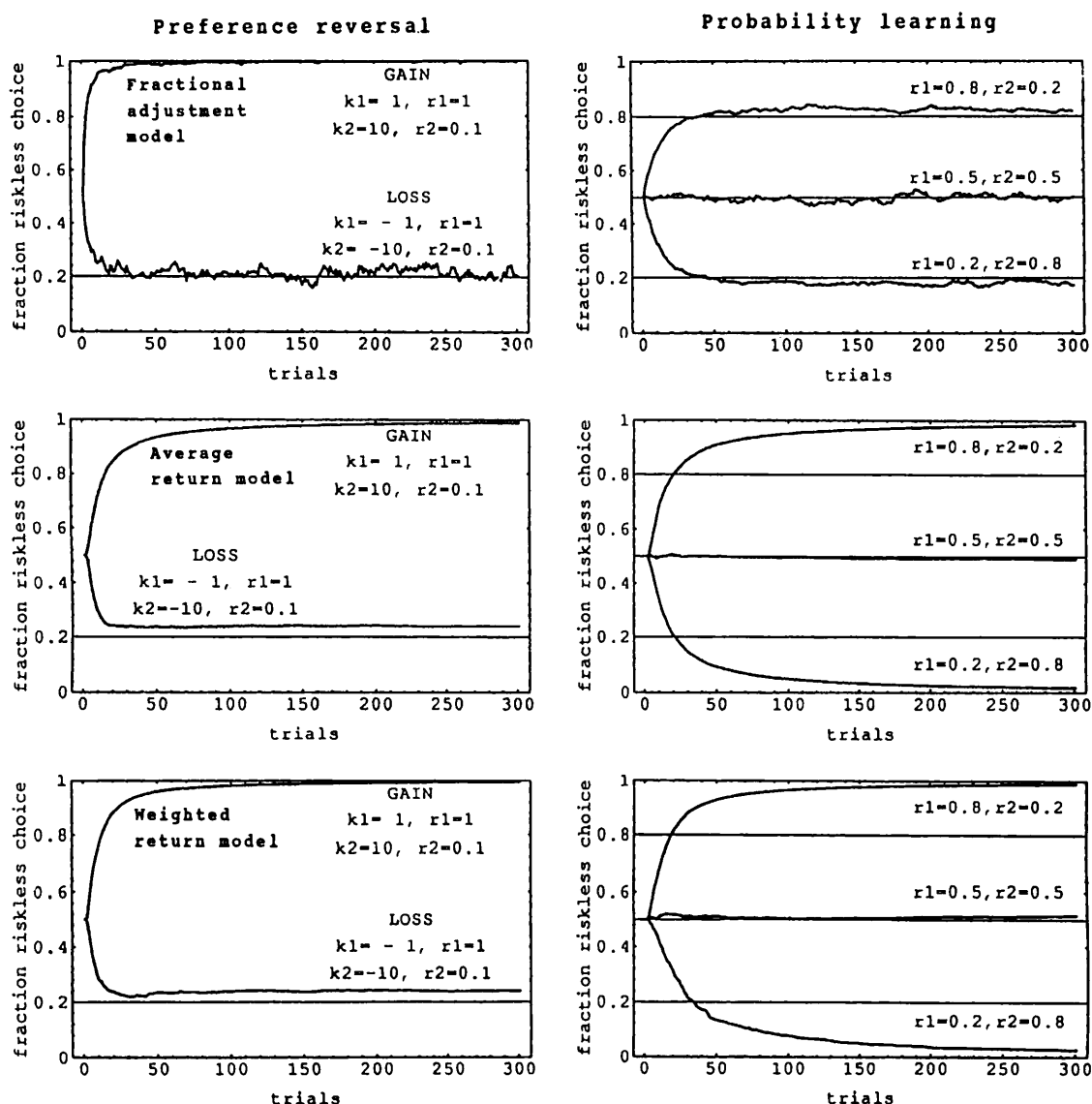


図2 改編した March (1996) の3モデルのシミュレーション結果

上から順に、部分的調整モデル、平均返戻モデル、加重平均返戻モデルのシミュレーション結果を示している。また、左列は利得・損失場面における選好逆転現象に関して、右列は確率マッチング現象に関するシミュレーション結果を示している。横軸は試行数を表し、縦軸は選択肢1（リスクレス選択肢、選択肢2はリスク選択肢）を表している。 $r_1, r_2$ は、それぞれ選択肢1, 2における強化確率を表している。 $k_1, k_2$ は、それぞれ選択肢1, 2における1度の強化量または損失量を表している。利得の場合は正の値で示し、損失の場合は負の値で示した。

失場面ではリスク選択が起こることを予測した。今回みられた選好逆転現象は、原論文の指摘と同様、部分的調整モデルの予測に最も顕著に表れた。また、1個体毎に300試行日の反応生起確率を計算したところ、全ての個体の反応生起確率が図2左側で示されている漸近値付

近に収束した。

しかし、原論文と同じように強化確率と一度の強化量あるいは損失確率と一度の損失量の値を選択肢1においては強化（損失）確率1、一度の強化量を1または損失量を-1とし、選択肢2においては強化（損失）確率

0.01, 一度の強化量を 100 または損失量を -100 とすると, 損失場面に関する部分的調整モデルの予測は, 明確な選好逆転現象を示さなかった。すなわち, そのように設定すると部分的調整モデルは, 損失場面でリスク選択とリスクレス選択がそれぞれ確率 0.5 で起きる試行もあれば, 図 2 上段左側のシミュレート結果よりはリスク選択の確率が低いがおよそ確率 0.7 でリスク選択が起きる試行があることを予測した。また, 個体数を原論文と同様に 1000 個体に増やしたり, フリーパラメーター  $\alpha$  の値を変更して計算しても, 同様な結果を示した。さらに, 反応増減ルールの設定を変更しても結果は同じであった。

一方, 平均返戻モデル, 加重平均返戻モデルでは, 強化(損失)確率と一度の強化(損失)量を原論文と同じ値にすると, 図 2 中・下段左側のシミュレート結果よりも明確な選好逆転現象を予測した。すなわち, 図 2 よりも早い試行で, 利得場面におけるリスクレス選択の生起確率が 1 に漸近し, 損失場面におけるリスク選択の生起確率がどちらのモデルでも約 0.9 となった。しかし, 反応生起確率の変化の仕方の設定を, 例えば部分的調整モデルのような設定に変更すると, 損失場面の生起確率がマイナスの値を示した。また, どちらのモデルについても 1 個体毎に 300 試行目の反応生起確率を計算したところ, 全ての個体の反応生起確率が図 2 左側で示されている漸近値付近に収束した。

### 2.3.2. 確率マッチング現象の予測

前述のように, March のモデルを設定した上で, 実際にシミュレートしてみると, 部分的調整モデルのみ確率マッチング現象を予測した(図 2 上段右側)。すなわち部分的調整モデルでは, 選択肢における反応生起確率の割合がその反応の強化確率と一致することを予測した。一方, 平均返戻モデルや加重平均返戻モデルは, 排他的な選択を予測した。

平均返戻モデル, 加重平均返戻モデルで一度の強化量の設定値の値を変えたり, 反応増減ルールを変更しても確率マッチング現象を予測しなかった。さらに, 加重平均返戻モデルでフリーパラメーター  $\beta$  の値を変更しても, 結果は同じであった。したがって, 部分的調整モデルでは強化事象の確率が個体の反応に依存せず, 平均返戻モデル, 加重平均返戻モデルでは強化事象の確率が個体の反応に依存することを暗に仮定していると考えられる。また, 2.3.1. と同様に 1 個体毎に 300 試行目の反応生起確率を計算したところ, 全ての個体の反応生起確率が図 2 右側で示されている漸近値付近に収束した。

### 2.4. 行動の数量モデルの改善点と今後の展開

March の 3 つのモデルのシミュレーションの運用から, 行動の数理モデルに以下のような改善点と今後の展開が考えられる。

#### 1) 時系列的な反応の生起に関する予測

強化随伴性は, 反応と強化子との時間的關係性と将来の行動の生起確率の変容との機能的関連についての記述でしかないので, 反応がいつ起きるのかを予測することはできない。また, March の 3 つのモデルでは, 1 試行毎に必ず反応が起きることになっていたため, これらのモデルでは反応がいつ起きるかということを予測・制御しているということにはならない。

これを克服するために, 次に述べるような変更や考え方の導入が考えられる。

#### 1a) フリーオペラント実験への変更

March のモデルをフリーオペラント実験に適応させるためには, (1)「単位時間」を試行から微小時間に変えること, (2) 2 種類の反応を反応 X とそれ以外の反応の 2 つに置き換えること, (3) その場合, 反応 X とそれ以外の反応との重み付けを検討すること(なぜかといえば, ここで想定しているような対称的な 2 つの反応ではなくなるからである), などの変更を加えればよいと考えられる。

#### 1b) 強化スケジュールの導入

March のモデルは, いずれも確率的強化であり反応の出現と共に乱数発生装置が働いていることから, フリーオペラント実験でいう乱比率(random ratio)強化スケジュールに読み変えることができる。環境側のモジュール(Church, 1997)である強化子供給装置を改編することで, 時間経過を強化子提示の要件に関係させることはそれほど難しいことではない。ところが, このようなモジュールを用意しても, すでに一般的な事実として確立している, 強化スケジュールに応じた特有の行動パターンを得ることができるかどうかの確証はない。個体側のモジュールである反応出現装置・反応増減ルール, 環境と個体の間にあるモジュールである反応出現基準変更計算の簡単な手直しで, 様々な強化スケジュールに特有な行動パターンを作り出せるのか, それともこのような 3 つの枠組みでは作り出せないのかについて, Kubota (1997) を除き徹底的になされた研究は現在のまでのところない。

#### 1c) 時間経過の導入

強化スケジュールに応じた特有の行動パターンを得るために提案されるもう 1 つの手段は, 時間経過を何らか

の形でモデルに導入することである。現在、明示的に時間経過を導入したモデルとしては、線形あるいは非線形モデル (Hanson & Timberlake, 1983; Hoyert, 1992; McDowell, 1980; McDowell, Bass, & Kessel, 1992; Myerson & Hale, 1988; Myerson & Meizin, 1980; Timberlake, 1984) が提出されている。これに対し、暗黙的に時間経過を導入したモデルとしては、スカラー計時理論 (Gibbon, 1977) でのパルス累積量や行動計時理論 (Killeen & Fetterman, 1988) での出現行動の累積量によって表す方法が考えられてきた。1b) と共に、時間経過をどのモジュールにどのように導入するのかという点に関して、検討が必要である。

### 2) 反応出現基準変更計算の様々なバリエーション

表 1 でみたように現象の捉え方によって、様々なモデル (反応出現基準変更計算) が提出されてきた。March が取り扱った 3 つの反応出現基準変更計算の他に、我々は逐次改良 (melioration) モデル (Herrnstein & Vaughan, 1980; Vaughan, 1981) で用いられている等コスト配分モデルや、Horner and Staddon (1987; Staddon, 1988; Staddon & Horner, 1989) が提案している強化変化率比一定 (ratio invariance) モデルを検討中である。等コスト配分モデルでの計算は、各選択肢の局所強化率が等しく、あるいは各選択肢のコスト (強化子数/反応数) を等しくするように行動を配分する。一方、強化変化率比一定モデルでは、各選択肢における前回と今回の強化確率の差分の比を一定にするように行動を配分する。

### 3) 乱数発生からの脱却

March のモデルでは、反応の発生メカニズムとして、乱数生成と反応出現基準が用いられている。一方、現象の不確定性を本質としたメカニズムを採用しない反応の発生に基づいた考え方もあり得る。個体の行動を複雑系とみなすと、その行動出現の予測の難しさは確率的なノイズではなくモデル構造が特定できていないことにも帰因できてしまう。しかし現在まで、例えばオペラント行動の非整数次元について言及した研究はないし、その発生メカニズムに非線形力学系を応用したモデル (Hanson & Timberlake, 1983; Hoyert, 1992; Timberlake, 1984) もごく少ない。

### 4) 強化学習を初めとする適応的ネットワークモデルとの整合性

強化子の提示、非提示に基づく 2 種類の反応についての反応増減ルールと反応出現基準変更計算は、別の見方をすれば、強化子の有無による 1 入力 2 出力を持った単純なネットワークでの重み付け変更のための学習規則と

考えることができる。通常のニューラルネットワークモデルでは、学習目標として与えられる教師信号とネットワークの実際の出力との差が、何らかの学習規則に従ってネットワーク内のノード間の重み付けを変更するのに用いられる。しかし一般に、強化学習と呼ばれる適応的ネットワークモデルでは、学習目標としての教師信号は与えられず、その代わり出力が強化されるかされないかのどちらかが起こって、それに基づいて重み付けが変更される (Sutton & Barto, 1998)。このように、後続事象によってシステムの振る舞いを変更したり、従来のネットワークシステムと比較すればずっと単純なシステムで複雑な行動を示すことができることが、最近様々なところで関心を持たれている (Bozinovski, 1995; Donahoe & Palmer, 1994)。このような研究の文脈からも、March のモデルのもととなった考え方を見直すことには意義があろう。

## 3. おわりに

本論は、数理モデルに対する行動分析による警戒から始まった。最後に、1) 科学的行動としての数理モデルの構成やシミュレーションの運用にはどのような意義があるのか、2) それらによって行動の予測と制御は高めることができるのか、について再度検討してみることにしよう。

1) 数理モデルは行動的現象の表現法としての特徴を見る限り、日常言語による表現に比べ、より厳密で縮約された表現を提供するものでしかない。しかし、この表現をコンピュータなどの論理機械で実行することで数理モデルを構成したことの利点が現れる。つまり、第 1 に数理モデルは数量的関係を具体的に記述しているので、日常言語による表現に比べてこのようなシミュレーションを実現しやすい。第 2 に、行動的現象の生成アルゴリズムを定義するシミュレーションの「仕様」の決定過程で、数理モデルの中では暗黙的に仮定されていた諸前提を明示的に述べることが要請される。こうして数理モデルは、シミュレーションを通してその機能や構造が調べられるのであれば、より整合性のある記述体系に高められていくことがわかる。そしてこのような精度の高い記述体系は、日常言語によって表現されたモデルに比べ、より高いレベルで科学的共同体の吟味に耐えうるという点で、科学的行動としての意義があるといえるだろう。

2) 数理モデルの構成やシミュレーションの運用が科学的行動として意義があるとしても、こうした行動が行動分析が求めてきた予測と制御を高めるという観点から

の科学的行動であるかどうかとは別の問題である。Skinner (1984) や Sidman (1960) らの、実験とそこから得られるデータに依拠した科学的行動を尊重する背景には、科学的行動そのものを研究対象である現実の環境や行動の場に直接さらすことで随伴性形成行動として成立・維持させるという観点があると思われる。その観点から見れば、数理モデルの構成やシミュレーションの運用は、実験データを基にしているが、実験場面という科学的環境下の強化随伴性に直接さらされてはいない。しかし、そのような随伴性とは別に、それと機能的に関連した随伴性に支えられた諸行動を我々は見出すことができる。例えば、データを集約し、選別し、再構成すること、それに基づいて予定された結果以外の新しい結果を読み出すこと、他の現象とのアナロジーから新たに研究法や分析法を導入することなどの行動は、いずれも実験データを獲得する際に働いている随伴性とは別の随伴性に支えられているが、現在の行動分析の諸研究によって生み出された予測や制御を高める可能性をもつ科学的行動である。そして、ある随伴性に支えられた行動が別の随伴性の支える行動と深く関連したり、その行動を生成する機会を与える行動となっているような科学的行動を育む環境をそこに観察することができる。数理モデルとシミュレーションという組となった科学的行動は他の科学的行動と同様に、そのような環境下の行動とみなすことができる。以上のことから、行動分析学が挙げてきた行動の予測と制御をめぐる数理モデルの危険性—実験データと行動の次元から乖離したモデルを構成すること—を十分に考慮する限り、行動の研究における数理モデルの構成とその吟味は、上述の意味から決して負の遺産とならないと考えられる。

注 1) 本研究の一部は 1998 年 10 月に開催された日本心理学会第 62 回大会シンポジウム「心理学方法論の基礎にある数理と計量」で報告された。本研究は文部省科学研究費補助金 (基盤研究 (c)(2) 一般 10610083) の援助を受けた。

注 2) 実際のプログラムは Appendix を参照された。Appendix は部分的調整モデルのプログラムである。プログラミングには、インタープリタ言語である Wolfram Research Inc. の Mathematica を使用した。

## References

- Baum, J. W. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, 20, 137-153.
- Baum, J. W. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 22, 231-242.
- Baum, J. W. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, 12, 861-874.
- Bozinovski, S. (1995). *Consequence driven systems*. Bitola, Macedonia: GOCMAR.
- Bush, R. R. & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, 58, 313-323.
- Cartwright, D. & Harary, F. (1956). Structural balance: A generalization of Heider's theory. *Psychological Review*, 63, 277-293.
- Catania, A. C. (1963a). Concurrent performances: Reinforcement interaction and response independence. *Journal of the Experimental Analysis of Behavior*, 6, 253-263.
- Catania, A. C. (1963b). Concurrent performances: A baseline for the study of reinforcement magnitude. *Journal of the Experimental Analysis of Behavior*, 6, 299-300.
- Catania, A. C. (1981). Discussion: The flight from experimental analysis. In C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), *Recent developments in the quantification of steady-state operant behavior* (pp. 49-64). Amsterdam: Elsevier.
- Catania, A. C. & Cutts, D. (1963). Experimental control of superstitious responding in humans. *Journal of the Experimental Analysis of Behavior*, 6, 203-208.
- Church, R. M. (1997). Quantitative models of animal learning and cognition. *Journal of Experimental Psychology: Animal Behavior Processes*, 23, 379-389.
- Church, R. M., & Gibbon, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes*, 8, 165-186.
- Collier, G. H., Johnson, D. F., Hill, W. L., & Kaufman, L. W. (1986). The economics of the law of effect. *Journal of the Experimental Analysis of Behavior*, 46, 113-136.
- Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical psychology: An elementary introduction*. Englewood Cliffs, NJ: Prentice-Hall. (クームス C. H., ドーズ R. M., トヴェスキー A. 小野 茂 (監訳) (1979). 数理心理学序説. 新曜社.)
- Davis, D. G. S., Staddon, J. E. R., Machado, A., & Palmer, R. G. (1993). The process of recurrent choice. *Psychological Review*, 100, 320-341.
- Davison, M. (1991). Stimulus discriminability, contingency discriminability, and complex stimulus control. In M. L. Commons, J. A. Nevin, & M. C. Davison (Eds.), *Signal detection: Mechanisms, models, and applications* (pp. 57-78). Hillsdale, NJ: Erlbaum.
- Davison, M. & Jenkins, P. E. (1985). Stimulus discriminability, contingency discriminability, and schedule performance. *Animal Learning & Behavior*, 13, 77-84.
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Boston, MA: Allyn and Bacon.
- Epstein, R. (1981). On pigeons and people: A preliminary look at the Columbian simulation project. *The Behavior Analyst*, 4, 43-55.
- Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review*, 57, 94-107.
- Estes, W. K. (1993). Mathematical models in psychology. In

- G. Keren & C. Lewis (Eds.), *A Handbook for data analysis in the behavioral sciences: Methodological issues* (pp. 3-19). Englewood Cliffs, NJ: Lawrence Erlbaum Associates.
- Fantino, E. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, 12, 723-730.
- Fantino, E. (1977). Conditioned reinforcement: Choice and information. In W. K. Honig & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 313-339). Englewood Cliffs, NJ: Prentice-Hall.
- Findley, J. D. (1958). Preference and switching under concurrent scheduling. *Journal of the Experimental Analysis of Behavior*, 1, 123-144.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84, 321-340.
- Gibbon, J., & Church, R. M. (1981). Time left: Linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 87-108.
- Godfrey, R. & Davison, M. (1998). Effects of varying sample- and choice-stimulus disparity on symbolic matching-to-sample performance. *Journal of the Experimental Analysis of Behavior*, 69, 311-326.
- Grace, R. C. (1996). Choice between fixed and variable delays to reinforcement in the adjusting-delay procedure and concurrent chains. *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 362-383.
- Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York, NY: Wiley.
- Gulick, D. (1992). *Encounters with chaos*. New York, NY: McGraw-Hill. (ギュリック D. 前田恵一・原山卓久(訳) (1995). カオスとの遭遇—力学系への数学的アプローチ. 産業図書.)
- Hanson, S. J. & Timberlake, W. (1983). Regulation during challenge: A general model of learned performance under schedule constraint. *Psychological Review*, 90, 261-282.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, 4, 267-272.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243-266.
- Herrnstein, R. J. & Vaughan, W., Jr. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143-176). New York, NY: Academic Press.
- Hinson, J. M. and Staddon, J. E. R. (1983). Hill-climbing by pigeons. *Journal of the Experimental Analysis of Behavior*, 39, 25-47.
- Horner, J. M. & Staddon, J. E. R. (1987). Probabilistic choice: A simple invariance. *Behavioural Processes*, 15, 59-92.
- Hoyert, M. S. (1992). Order and chaos in fixed-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, 57, 339-363.
- Hursh, S. R., Raslear, T. G., Shurtleff, D., Bauman, R., & Simmons, L. (1988). A cost-benefit analysis of demand for food. *Journal of the Experimental Analysis of Behavior*, 50, 419-440.
- Iversen, I. H. (1988). Tactics graphic design: A review of Tufte's *The visual display of quantitative information*. *Journal of the Experimental Analysis of Behavior*, 49, 171-189.
- Killeen, P. R. (1981). Averaging theory. In C. M. Bradshaw, E. Szabadi, & C. F. Lowe (Eds.), *Recent developments in the quantification of steady-state operant behavior* (pp. 21-34). Amsterdam: Elsevier.
- Killeen, P. R. (1982). Incentive theory II: Models for choice. *Journal of the Experimental Analysis of Behavior*, 38, 217-232.
- Killeen, P. R. (1994). Mathematical principles of reinforcement. *Behavioral and Brain Sciences*, 17, 105-172.
- Killeen, P. R. & Fetterman, G. (1988). A behavioral theory of timing. *Psychological Review*, 95, 274-295.
- Killeen, P. R., Palombo, G., Gottlob, L. R., & Beam, J. (1996). Bayesian analysis of foraging by pigeons (Columba livia). *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 480-496.
- 近藤次郎 (1976). 数学モデル—現象の数式化—。丸善株式会社。
- Krebs, J. R. (1978). Optimal foraging: Decision rules for predators. In J. R. Krebs & N. D. Davies (Eds.), *Behavioral ecology: An evolutionary approach* (pp. 23-63). Oxford, London: Blackwell.
- Kubota, A. (1997). Recursive Bayesian formulation of operant behavior -A frame work. 慶應義塾大学大学院社会学研究科紀要, 45号, 41-59.
- Lea, S. E. G. (1978). The psychology and economics of demand. *Psychological Bulletin*, 85, 441-446.
- Levine, D. S. (1991). *Introduction to neural and cognitive modeling*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. New York, NY: Oxford University Press.
- Luce, R. D. (1995). Four tensions concerning mathematical modeling in psychology. *Annual Review of Psychology*, 46, 1-26.
- Machado, A. (1997). Increasing the variability of response sequences in pigeons by adjusting the frequency of switching between two keys. *Journal of the Experimental Analysis of Behavior*, 68, 1-25.
- March, J. G. (1996). Learning to be risk averse. *Psychological Review*, 103, 309-319.
- Mark, T. A. & Gallistel, C. R. (1994). Kinetics of matching. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 79-95.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior. Vol. 5: The effect of delay and of intervening events on reinforcement value* (pp. 55-73). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McDowell, J. J. (1980). An analytic comparison of Herrnstein's equations and a multivariate rate equation. *Journal of the Experimental Analysis of Behavior*, 33, 397-408.
- McDowell, J. J. (1987). A mathematical theory of reinforcer value and its application to reinforcement delay in simple schedules. In M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of*

- behavior*, Vol. 5: *The effect of delay and of intervening events on reinforcement value* (pp. 77-105). Hillsdale, NJ: Lawrence Erlbaum Associates.
- McDowell, J. J., Bass, R., & Kessel, R. (1992). Applying linear systems analysis to dynamic behavior. *Journal of the Experimental Analysis of Behavior*, 57, 377-391.
- McDowell, J. J. & Wixted, J. T. (1988). The linear system theory's account of behavior maintained by variable-ratio schedules. *Journal of the Experimental Analysis of Behavior*, 49, 143-169.
- Moore, J. W. & Desmond, J. E. (1992). A cerebellar neural network implementation of a temporally adaptive conditioned response. In I. Gormezano & E. A. Wasserman (Eds.), *Learning and memory: The behavioral and biological substrates* (pp. 347-368). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Myerson, J. & Hale, S. (1988). Choice in transition: A comparison of melioration and kinetic model. *Journal of the Experimental Analysis of Behavior*, 49, 291-302.
- Myerson, J. & Meizel, F. M. (1980). The kinetics of choice: An operant systems analysis. *Psychological Review*, 87, 160-174.
- Nevin, J. A. (1984). Quantitative analysis. *Journal of the Experimental Analysis of Behavior*, 42, 421-434.
- Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, 57, 301-316.
- 小野寛晰 (1991). 計算論. 広中平祐 (編集委員代表) 現代数理解科学事典 (pp. 333-338) 丸善株式会社.
- Plaud, J. J. & O'Donohue, W. (1991). Mathematical statements in the experimental analysis of behavior. *The Psychological Record*, 41, 385-400.
- Rachlin, H. (1982). Economics and matching law. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior. Vol. 2: Matching and maximizing accounts* (pp. 347-374). Cambridge, MA: Ballinger.
- Rachlin, H., Battalio, R. C., Kagel, J. H., & Green, L. (1981). Maximization theory in behavioral psychology. *Behavioral and Brain Sciences*, 4, 371-388.
- Robbins, D. (1980). Mathematical learning theory: W. K. Estes and stimulus sampling. In G. M. Gazda & R. J. Corsini (Eds.), *Theories of learning: A comparative approach* (pp. 295-331). Itasca, IL: F. E. Peacock Publishers.
- Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 242-268.
- 坂上貴之 (1994). 不確実性をめぐる動物行動研究. 心理学評論, 37, 294-316.
- Schofield, G. & Davison, M. (1997). Nonstable concurrent choice in pigeons. *Journal of the Experimental Analysis of Behavior*, 68, 219-232.
- Schultz, D. (1981). *A history of modern psychology*. New York, NY: Academic Press. (シュルツ D. 村田孝次 (訳) (1986). 現代心理学の歴史. 培風館.)
- Shimp, C. P. (1984a). Relations between memory and operant behaviour, according to an associative learner, A.L. *Canadian Journal of Psychology*, 38, 269-284.
- Shimp, C. P. (1984b). Timing, learning and forgetting. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (pp. 346-360). New York, NY: New York Academy of Sciences.
- Shimp, C. P. (1984c). Cognition, behavior, and the experimental analysis of behavior. *Journal of the Experimental Analysis of Behavior*, 42, 407-420.
- Shull, R. L. (1992). Mathematical description of operant behavior: An introduction. In I. H. Iversen & K. A. Lattal (Eds.), *Experimental analysis of behavior, Part 2* (pp. 243-282). Amsterdam: Elsevier.
- Skinner, B. F. (1932). Drive and reflex strength: II. *Journal of General Psychology*, 6, 38-48.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, 57, 193-216.
- Skinner, B. F. (1984). Methods theories in the experimental analysis of behavior. *Behavioral and Brain Sciences*, 7, 511-546.
- Sidman, M. (1960). *Tactics of scientific research*. New York, NY: Basic Books.
- Staddon, J. E. R. (1983). *Adaptive behavior and learning*. New York, NY: Cambridge University Press.
- Staddon, J. E. R. (1988). Quasi-dynamic choice models: Melioration and ratio invariance. *Journal of the Experimental Analysis of Behavior*, 49, 303-320.
- Staddon, J. E. R. (1993). *Behaviorism: Mind, mechanism and society*. London: Duckworth.
- Staddon, J. E. R. & Horner, J. M. (1989). Stochastic choice models: A comparison between Bush-Mosteller and a source-independent reward-following model. *Journal of the Experimental Analysis of Behavior*, 52, 57-64.
- Staddon, J. E. R. & Motheral, S. (1978). On matching and maximizing in operant choice experiments. *Psychological Review*, 85, 436-444.
- Staddon, J. E. R. & Zanutto, B. S. (1998). In praise of parsimony. In C. D. L. Wynne & J. E. R. Staddon (Eds.) *Models of action: Mechanisms for adaptive behavior* (pp. 239-267). Mahwah, NJ: Lawrence Erlbaum Associates.
- Stephens, D. W. & Krebs, J. R. (1986). *Foraging theory*. Princeton, NJ: Princeton University Press.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: A Bradford Book, MIT Press.
- Thurstone, L. L. (1930). The learning function. *Journal of Genetic Psychology*, 3, 469-493.
- Timberlake, W. (1984). Behavior regulation and learned performance: Some misapprehensions and disagreements. *Journal of the Experimental Analysis of Behavior*, 41, 355-375.
- Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- Vaughan, W., Jr. (1981). Melioration, matching, and maximization. *Journal of the Experimental Analysis of Behavior*, 36, 141-149.
- Williams, B. A. (1988). Reinforcement, choice, and response strength. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Sterens' handbook of experimental psychology. Vol. 2. Learning and cognition* (2nd ed., pp. 167-244). New York: John Wiley & Sons.
- Zirix, J. M., & Silberberg, A. (1984). Concurrent variable-interval variable-ratio schedules can provide only weak evidence for matching. *Journal of the Experimental Analysis of Behavior*, 41, 83-100.

## Appendix

(\* fam is a fractional adjustment model proposed by March (1996). \*)

```
fam[kv1_,rv1_,kv2_,rv2_,av_,t_,n]:=Module[
{k1,r1,k2,r2,r,problast,templist,a,num,
prob01,prob02,prob11,prob12,prob21,prob22,time},
k1=kv1;
```

```
r1=rv1;
```

```
k2=kv2;
```

```
r2=rv2;
```

```
a=av;
```

```
time=t;
```

```
num=n;
```

```
problast=Table[0,{time+1}];
```

```
Do[
```

```
prob01=.5;
```

```
prob02=.5;
```

```
prob11=prob01;
```

```
prob21=prob02;
```

```
templist={prob01};
```

```
Do[
```

```
r=Random[Real];
```

```
If[
```

```
Random[Real]<prob11,
```

```
{Which[
```

```
r<r1,{Which[
```

```
k1>0,prob12=1-(1-prob11) (1-a)^k1,
```

```
k1<0,prob12=prob11 (1-a)^(-k1)
```

```
]],
```

```
r>r1,{Which[
```

```
k1>0,prob12=prob11 (1-a)^k1,
```

```
k1<0,prob12=1-(1-prob11) (1-a)^(-k1)
```

```
]]
```

```
];
```

```
prob22=prob21;
```

```
prob11=(prob12)/(prob12+prob22);
```

```
prob21=(prob22)/(prob12+prob22)
```

```
},
```

```
{Which[
```

```
r<r2,{Which[
```

```
k2>0,prob22=1-(1-prob21) (1-a)^k2,
```

```
k2<0,prob22=prob21 (1-a)^(-k2)
```

```
]],
```

```
r>r2,{Which[
```

```
k2>0,prob22=prob21 (1-a)^k2,
```

```
k2<0,prob22=1-(1-prob21) (1-a)^(-k2)
```

```
]]
```

```
];
```

```
prob12=prob11;
```

```
prob11=(prob12)/(prob12+prob22);
```

```
prob21=(prob22)/(prob12+prob22)
```

```
]
```

```
];
```

```
templist=AppendTo[templist,prob11],
```

```
{time}
```

```
];
```

```
problast=problast+templist,
```

```
{num}
```

```
];
```

```
problast/num <
```

```
];
```

fam 関数は、大括弧内の変数から成る。kv1、rv1 は選択肢 1 における 1 度の強化（損失）量とその強化（損失）確率を表す。kv2、rv2 は選択肢 2 における 1 度の強化（損失）量とその強化（損失）確率を表す。av は部分的調整モデルのフリーパラメーター、t は試行数、n は個体数を表す。

最終出力に利用するリストの定義をする。

反応出現基準の初期値を与え、選択肢 1 において計算された反応生起確率を一時保存するリストを定義する。

強化子用乱数発生装置

反应用乱数発生装置と反応出現基準

反応増減ルールと反応出現基準変更計算  
反応出現基準によって選択肢 1 または 2 のどちらに対する反応が生起するか決定される。決定後、具体的な数値計算を行う。

反応出現基準変更計算によって得た、選択肢 1 に対する反応生起確率を一時保存リストに格納する。これを各個体毎にも試行ずつ行う。

n 個体分の反応生起確率を最終出力用リストに格納する。

n 個体分の反応生起確率の平均を算出し、その値を返す。