

# 臨床心理学と認知モデリング<sup>1)</sup>

国 里 愛 彦

専修大学

## Clinical psychology and cognitive modeling

Yoshihiko KUNISATO

Senshu University

Studies in clinical psychology, especially abnormal psychology, typically use null hypothesis statistical testing to examine behavioral data that were gathered through cognitive tasks. This paper identifies the problems associated with using conventional research methods and discusses the advantages of using a computational approach to examine mental disorders. The computational approach includes cognitive modeling that estimates latent parameters that cannot be directly observed from behavioral data. Cognitive modeling makes it possible to explain and predict behavioral data in a logically valid way and contributes to the assessment of mental disorders. This paper introduces the best practices in cognitive modeling by using probabilistic reversal learning task as an example. In conclusion, this paper provides examples of research studies that use cognitive modeling in clinical psychology and discusses future directions, including parameter estimation by hierarchical Bayesian estimation and hierarchical Bayesian inference model as a cognitive model.

**Key words:** cognitive modeling, computational clinical psychology, simulation, parameter estimate, hierarchical Bayes model

キーワード：認知モデリング，計算論的臨床心理学，シミュレーション，パラメータ推定，階層ベイズモデル

### はじめに

臨床心理学は、精神疾患や心理的不適応などの問題について、メカニズムを検討する異常心理学、評価を行う心理アセスメント、心理的介入を行う心理療法の3領域からなる研究分野である(丹野ら, 2015)。本論文では、臨床心理学の中でも、異常心理学に焦点をあてる。一般的に、異常心理学研究では、精神疾患や心理的不適応を抱える者とそういった問題のない健常者を対象に、認知課題や行動観察による行動データ、質問紙による主観データ、生理学・神経科学的指標による生理・神経データを測定する。そして、これらのデータが疾患群と健常者群で異なるかどうか、帰無仮説検定を行い、統計学的に有意な平均値差な

どを検討する。例えば、不安症の者は、恐怖対象へ注意が向いてしまい、そこから離れることが難しいと仮説を立てたとする。その場合に、不安症の者と健常者を対象に、注意の切り替えを検討するような認知課題を実施し、そのパフォーマンスの群間差を帰無仮説検定する。その結果、恐怖対象が提示された場合に、恐怖対象とは反対側に出てくる刺激への反応時間が、不安症群において統計学的に有意に長くなったとする。この結果をもって、不安症の者は恐怖対象への注意によって反応が遅くなったと考え、不安症にはネガティブな情動刺激への注意バイアスがあると考え。異常心理学研究の中でも認知課題を用いる実験心理学的指向性をもった研究は、このようなプロセスで行われる。

実験心理学的志向性をもった研究は、異常心理学において重要な役割を担ってきた。しかし、上記のような研究方法では理論に基づいた検討が弱

1) 本論文の作成にあたり、JSPS 科研費 JP16H05957 の助成を受けた。

くなることや理論やモデルで想定する潜在変数を直接検討することができないという問題点がある。実験心理学的志向性をもった異常心理学研究は、思弁的な考察や個人の体験に基づいて心理的問題のメカニズムを検討するのではなく、データに基づいた検討ができるため、根拠のある臨床実践において重要な知見を提供する。しかし、その一方で、実験心理学の研究における理論心理学的側面の弱さについては、これまでも指摘がなされてきている（須賀，1989；須賀，2011）。データに基づいた実験的検証だけでは、臨床心理学において扱う臨床現場での素朴な「なぜ？」や「どのようにして？」といった疑問に解答を与えることが難しい（須賀，2011）。なぜなら、「なぜ？」や「どのようにして？」といった疑問に答えるには、理論的説明が不可欠であり、これらの疑問を検証可能な仮説におとしこむために、理論心理学が不可欠になる。心理学が理論をしっかりと組み立てた上で実証的に検証する科学に発展する上では、データ駆動型の仮説検定研究だけでなく、研究を行う上で理論を意識するような枠組みや工夫が必要かもしれない。また、通常の認知課題を用いた異常心理学研究においては、理論やモデルで想定している潜在変数を直接的に検討することはない。しかし、例えば認知課題などで刺激を提示してから、反応が出力されるまでに、どのような過程があるのかを直接推定できたほうが疾患のメカニズムを捉えられる可能性がある。

本稿では、まず上記の問題について臨床心理学に対する計算論的アプローチ、特に認知モデリングの説明をする。次に、確率的逆転学習課題などの具体例を用いて、認知モデリングにおける推奨実践について論じる。最後に、臨床心理学における認知モデリングと階層ベイズ推定・モデルを含めた今後の展開について論じる。

### 臨床心理学における計算論的アプローチ

計算論的アプローチでは、私たちの脳は、外から入力される刺激に対して、なんらかの計算を行った上で、行動として出力していると考ええる。計算論的アプローチでは、その計算の理論、計算過程を表現する数理モデル、計算を実行する神経基盤について検討する。この計算論的アプローチ

を神経科学に適用すると計算論的神経科学と呼ばれ、精神医学に適用すると計算論的精神医学と呼ばれる（Adams, Huys, & Roiser, 2015; 国里, 2013）。計算論的精神医学は、計算論的アプローチ、具体的には後述する Marr の3つの水準を意識した理論駆動型の研究方法を精神疾患研究に適用する学問領域である（Adams et al., 2015）。計算論的精神医学は、2011年頃より、計算論的神経科学において著名な研究者によって始められ、2017年には、*Biological Psychiatry* 誌で特集が組まれるなど、注目を集めている。なお、著者は、精神医学と同様に、臨床心理学に計算論的アプローチを適用する計算論的臨床心理学を標榜し研究を行ってきている（国里, 2018）。

### Marr の3つの水準

Marr の3つの水準（Marr, 1982 乾・安藤（訳），1987）とは、ヒトの心などの複雑な情報処理システムを理解するために、計算理論、表現とアルゴリズム、インプリメンテーションの3つの水準から検討する枠組みである（図1）。

まず、計算理論の水準では、計算の目標は何か、どうしてその計算が適切かを明らかにする。私たちの心を理解する上で、出力された行動だけを見ても理解は深まらないかもしれない。そういった行動が出力されるにあたって行われる演算過程の目標や解くべき問題が明らかになることで、行動の理解は深まる。また、計算理論の水準において、計算が満たすべき制約条件も明らかにしておく、次の表現とアルゴリズムの水準に制約を課すこともできる。

表現とアルゴリズムの水準は、計算理論の水準で明らかになった制約条件の下で、解くべき問題と処理過程をどのように表現し、どのように具体的なアルゴリズムで問題を解くのかを明らかにする。表現とアルゴリズムの水準では、具体的な問題解決に向けて数理モデルを用いる。問題解決のために用いられる数理モデルが1つとは限らない。実際は、複数の数理モデルが考えられ、計算理論の水準からの制約によって絞り込んだり、実際の行動データへのフィットからモデルが選択されたりする。図1にあるように、表現とアルゴリズムの水準は、認知課題を用いた研究の行動デー

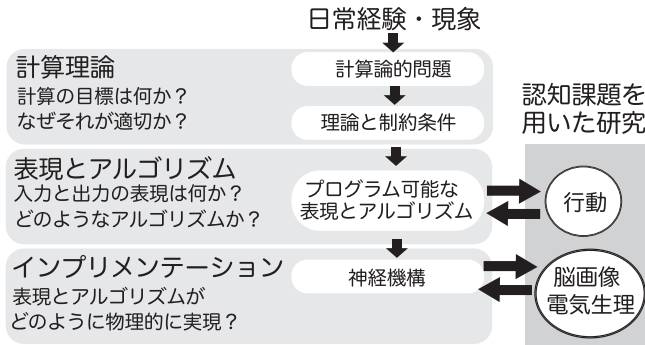


図1 Marrの3つの水準 (Marr, 1982 を元に一部改変)

タと参照関係にある。そのため、実際の行動データから数理モデルが選択されたり、数理モデルから行動データを収集する研究が計画されたりする。

最後のインプリメンテーションの水準は、表現やアルゴリズムがどのように物理的に実現されるのかを明らかにする。私たちが刺激を受け取ってから、なんらかのアルゴリズムを介して、行動を出力する場合、そのアルゴリズムは脳などの生物学的な基盤において実行される。そのため、インプリメンテーションの水準は、認知課題を用いた研究の脳画像や生理データと参照関係にある。インプリメンテーションの水準では、神経系に関するモデルを明らかにし、実際の神経・生理データからモデルの妥当性を検討したり、モデルから神経・生理データを収集する実験が計画されたりする。

Marrの3つの水準は、計算論的アプローチのコアになる考えである。本論文で扱う認知モデリングは、Marrの3つの水準において、表現とアルゴリズムの水準に対応するが、計算理論の水準とインプリメンテーションの水準も考慮する必要がある。

## 認知モデリング

2000年以降、認知モデリングを用いた論文が出版されるようになってきており、認知科学領域における認知モデリングの使用は加速してきている (Palminteri, Wyart, & Koehlin, 2017)。Palminteri et al. (2017) が、認知科学領域においてインパクトの高い雑誌に2009年以降に掲載された論文を

対象に調査したところ、なんらかの認知モデリングを行った研究は半数近くに及んでいる。国内においても、2017年には、豊田 (2017) による認知モデリングに関する書籍や Lee and Wagenmakers (2013 井関 (訳), 2017) の邦訳書が出版されている。このように、日本語でも認知モデリングを学ぶ環境が徐々に整ってきている。

認知モデリングは、認知モデルを用いてモデリングを行う研究手法である<sup>2)</sup>。モデリングとは、モデルの構築や分析を通じて、現実世界を間接的に研究する手法のことである (Weisberg, 2013 松王 (訳), 2017)。そして、認知モデルとは、知覚、学習、記憶、推論などの認知過程について十分な説明をすることとそれらの相互作用を説明することを目的としたモデルである (Busemeyer & Diederich, 2010)。認知過程は潜在的な過程であり、直接観察することはできず、課題のパフォーマンスに与える影響から間接的に観察することができるのみである (Heathcote, Brown, & Wagenmakers, 2014)。例えば、冒頭の注意バイアスの例だと、観察できるのは反応時間であり、条件の差から注意バイアスという認知過程を推測している。認知モデルは数理モデルだが、心理学研究で用いられることの多い回帰モデル、ANOVAのような統計モデルとは異なり、認知に関する原理から作られる (Busemeyer & Diederich, 2010)。

モデルとは、複雑な現象に対して、その現象の一面を人が理解しやすく理想化された形で整理されたものである。現象はデータと置き換えること

2) 認知モデルを含む統計モデリングについては、清水 (2018) に詳しい。

もできる。モデルは抽象的な構成概念であり、データの生成に関わる概念間の構造を捉えるものである。認知モデルは、実際の認知過程よりも単純かつ抽象化される必要があり、実際のシステムの詳細部分は除外した上で、重要な特徴を扱う (Farrell & Lewandowsky, 2015)。そのため、「全てのモデルは間違っているが、いくつかは有用である (Box, 1976)」という言葉に表されるように、1つのモデルで現象の全ての側面を説明することは難しい。そこで、特定の側面から現象を切り取ることで、現象の理解を深めるモデルを作成することが重要である。科学において用いられるモデルには、現象を具体的な事物によって表現する具象モデル、現象を数学的に表現する数理モデル、現象をコンピュータ上の数値計算で表現する数値計算モデルの3つがある (Weisberg, 2013 松王 (訳), 2017)。具象モデルは、あまり心理学ではイメージしにくい、心理学研究で用いられる実験動物 (モデル生物) がこれにあたる。数理モデルの例としては、遅延によって報酬の価値が割り引かれる現象を表現する双曲割引モデルなどの数式を用いたモデルがある。数値計算モデルの例としては、個人の行為と他者との相互作用の数値計算から社会現象を表現するエージェント・ベース・モデルなどがある。

認知過程は、刺激入力からデータ (行動) を生成するまでの内的過程である。そのデータ生成過程を抽象化して数理モデル化したのが認知モデルになる。そして、認知モデルにはパラメータが含まれており、それは個人ごとに異なると考えることが多い。データ、認知モデル、パラメータとの関係は、図2のように表すことができる。パラメータと認知モデルを準備して、認知課題からの刺激入力を与えれば、データが生成される。このように、トップダウン的にパラメータと認知モデルを決めてデータを生成し、シミュレーションをすることができる。一方、データ収集などを行ってデータが手元にある場合は、認知モデルを用いてデータからパラメータを推定するパラメータ推定を行うこともできる。なお、理論とモデルとデータの関係について、須賀 (2003) は、ツァー・モジュール説を提唱し、モデルはデータと関係し、理論はモデルと関係するとしている。つまり、理論はデータとは直接的には関係せず、

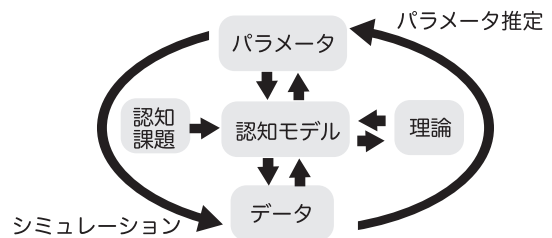


図2 認知モデルとシミュレーション・パラメータ推定

モデルがデータとの類似度によって検証されることを通して間接的に関係すると考える。図1のMarrの3つの水準においても、計算理論の水準は行動や生理データとは直接関係せず、表現とアルゴリズムの水準が行動データと関係し、インプリメンテーションの水準が生理データと関係する。抽象的な計算理論は、表現とアルゴリズムの水準やインプリメンテーションの水準のモデルを通して初めて具体的に理解することができる。このように、理論とモデルを明確に分けた上で、データはモデルの妥当性の検証をしているが、データは理論の検証を直接しているわけでない点を理解することが重要である。そして、研究プロセスにおいては、理論とモデルを行き来する理論的な検証プロセスとデータとモデルの類似度を検討する実証的な検証プロセスの2つがあることを意識することが重要といえる。

認知モデリングを用いる利点としては、(1) モデルの解釈のズレが小さくなる、(2) モデルから論理的に正しい量的・質的な予測ができる、(3) 行動データそのものよりも推定されたパラメータの説明力・予測力が高い可能性があるなどを挙げることができる。

まず、認知モデリングを用いることで、用いるモデルに対する研究者間の解釈のずれが小さくなる。心理学の理論の多くは言語的に記述されたモデルになるが、言語的に記述されたモデルはメタファーを含むその記述の仕方によって読み手の受け取り方が異なることがある。例えば、意味記憶がノード間の相互接続ネットワークによって表現されるとする活性化拡散モデル (Collins & Loftus, 1975) では、その説明を電気が電線を通るように説明されるか、水がパイプを通るように説明されるか、どのようなアナロジーを使うかによって印象が変わってくるかもしれない (Farrell



& Lewandowsky, 2015)。なお、このようなアナロジー自体を排することで厳密なモデルの説明もできるかもしれない。しかし、モデルはデータや現象がどのように生じたのか、どのように変化するのかを記述するものなので、どうしてもアナロジーを含めたモデルの動作の説明が必要になる。それをすべて言語的な記述で行うと研究者間でモデルの理解が異なってしまうこともあるかもしれない。その点、認知モデルで用いるものは基本的には数理モデルなので、研究者間でモデルの解釈が異なることは少なくなり、言語的なモデル記述で生じる問題を減らすことができるかもしれない (Farrell & Lewandowsky, 2015)。また、モデルの動作や予測についても、研究者の主観的な予測ではなく、数理モデルから導かれた論理的に妥当な予測をすることができる。

次に、認知モデルは数理モデルなので、一般化が可能な形で、正確な量的予測が可能である (Busemeyer & Diederich, 2010)。これは、言語的に記述されたモデルでは難しい。なお、認知モデルではない数理モデルでも、正確な量的予測は可能であるが、その数理モデルに認知過程に関する原理が含まれていないと、そのモデルの解釈や一般化が難しくなる。認知モデルでは、認知過程について解釈可能なモデルを扱うので、モデルを作成した状況以外の状況にも一般化することもできる。なお、言語的に記述されたモデルも、そこに認知過程についての原理や制約条件が含まれているならば、数理モデル化し認知モデルとすることもできる。

最後に、行動データ自体よりも認知モデリングによって推定されたパラメータのほうが、より明確に群間差や関連を検討できる可能性がある。例えば、Wiecki, Poland, and Frank (2015) は、認知課題で得られた反応時間と正誤の反応をそのまま使った時と Drift-Diffusion Model (Ratcliff, 1978) という情報蓄積過程から反応時間を説明するモデルを用いて推定したパラメータを使った時とで、患者群と健常者群を想定した異なる 2 群の診断精度を検討するシミュレーションを行った。診断精度の指標として ROC 曲線下面積 (area under the curve) を用いたところ、行動データをそのまま使った場合の 0.68 に比べて、Drift-Diffusion Model を用いた方が 0.91 と診断精度が高くなることが

示された (Wiecki et al., 2015)。このように、認知モデリングで推定されたパラメータを用いることで、臨床上のアセスメントが向上する可能性がある。なお、この Wiecki et al. (2015) のシミュレーションは、設定を変えると Drift-Diffusion Model に有利に働かないこともあり、どういう場合に適切なのかについては今後の課題になる。

## 認知モデリングの推奨実践法

認知モデリングが研究手法として用いられ始めてからまだ日が浅いため、他の研究手法に比べると、認知モデリングを行うための標準的な方法が確立されていない。そのような状況ではあるが、Busemeyer and Diederich (2010) による認知モデリングの手順、Heathcote et al. (2014) による認知モデリングの推奨実践法、Palminteri et al. (2017) による認知モデリングにおけるモデル選択法などが提案されてきている。これら 3 つの提案を総合すると、認知モデリングは、1) 認知課題と認知モデルを準備する、2) 人工データ生成とパラメータ・リカバリーの確認によって認知課題やモデルを修正する、3) データ収集と行動データを確認する、4) パラメータを推定する、5) 相対モデル比較を行う、6) モデル・シミュレーションと選択的影響テスト (Selective Influence Test) を行うという手順に集約できる。以降では、認知モデリングの推奨実践法について、確率的逆転学習課題と Q-learning モデルを例に挙げて解説する。

### 1) 認知課題と認知モデルを準備する

まず、関心のある認知過程が生起するような認知課題を準備する。認知課題の選択にあたり、先行研究でよく用いられている認知課題を用いる場合もあるが、認知モデルの検討を前提として、複数の対立する認知モデルが識別可能な認知課題を設定することもできる。後者のアプローチの方が事前の認知課題の準備は難しくなるが、より明確に認知モデルが想定する認知過程の差異に踏み込むことができる (Palminteri et al., 2017)。また、主な関心が認知過程であったとしても、少なくとも 1 つは条件差や群間差などの行動レベルの効果が生じるような研究計画を作る必要がある (Palminteri et al., 2017)。

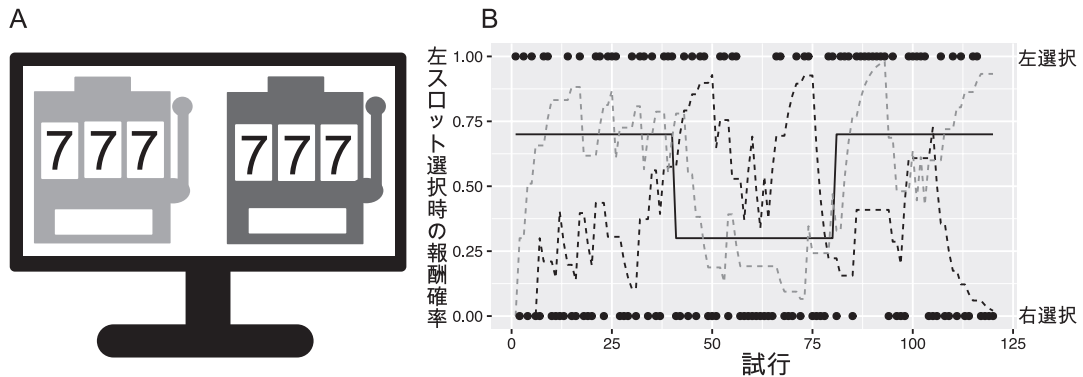


図3 確率的逆転学習課題と Q-learning モデル

例えば、環境が変化する状況において、ある心理の問題を抱える者は、学習の速度や学習したことを行動に反映させる程度に違いがあるのではないかという問題に関心があるとする。このような問題を検討するにあたり、先行研究で用いられることが多い確率的逆転学習課題を用意した。確率的逆転学習課題は、図3Aの2つのスロットマシンのように、複数の選択肢がある認知課題である。参加者が、どちらかの選択肢を選ぶと、その選択に対して、報酬（金銭やポイントの獲得）か罰（金銭やポイントの損失）が与えられる。参加者は、報酬を最大化する選択をするように求められる。参加者の選択に対する報酬や罰のフィードバックは、確率的になっている。図3Bの黒の実線は、左のスロットマシンを選んだ時に得られる報酬の確率を表している。最初の40試行は、左のスロットマシンを選んだ時に70%の確率で報酬が得られる（右を選ぶと30%の確率で報酬が得られる）。そのため、この期間、参加者は左のスロットマシンを選択する可能性が高い。次の41試行から80試行までは、フィードバックが逆転して右のスロットマシンの報酬確率が高くなる。最後の81試行から120試行はまた左のスロットマシンの報酬確率が高くなる。このように、確率的逆転学習課題は、変化する状況における個人の意思決定過程を検討できる。

課題の準備と並行して、関心のある認知過程の認知モデルも準備する。認知課題と同様に、認知モデルの選択では、先行研究において提案されている認知モデルをそのまま用いる場合もあるが、既存の認知モデルを調整したり、言語的・概念的

モデルを認知モデルに落とし込んだり、認知課題やその行動レベルの効果などから新規に認知モデルを作成することもある。モデル比較を行う場合には、複数の認知モデルを準備する。このモデル空間（検討に含める認知モデルの数）の上限を理論的に決めることはできないが、下限としては、少なくとも関心のある行動レベルの効果を生成できるモデルと生成できないモデルの2つを含めるということがある（Palminteri et al., 2017）。なお、人間のようなサブシステム間の関連性が強いシステムの研究では、局所的なモデルを作ってもうまくいかないことが多い（戸田, 1992）。データ駆動でモデルを作ると想定可能なモデルの数が爆発的に増加するため、トップダウン的に理論からモデルを作ることが重要である。このように理論や言語的・概念的モデルからトップダウン的に認知モデルを作成する場合に、制約条件などが少ないために、数理モデルに落とし込めないことや推定が難しくなることがある。このように、事前の理論的な仮定やモデルへの制約条件が少ない場合に、認知モデルを作成しながら、事後的に仮定を追加していく作業が必要になってくる（Busmeyer & Diederich, 2010）。例えば、パラメータとしては無限に値を取りうる場合に、その後のパラメータ推定を考慮して、現実的に取りうるパラメータの範囲を決めて推定を行うことは実践上起こり得る。やむを得ない部分はあるが、事後的な仮定の追加になるので、最小限にする必要があるし（Busmeyer & Diederich, 2010）、論文上で追加した仮定について記載する必要があるだろう。

本稿では、確率的逆転学習課題を認知課題として取り上げる。そして、確率的逆転学習課題の認知モデルとして、行動の価値を表すQ値を更新することで学習を実現するQ-learningモデル(Watkins, 1989)を用いる。確率的逆転学習課題は、選択することで選択肢などの状態が変化するような課題ではないので、各試行が終わると参加者が選択肢についての行動価値を更新する簡易版のQ-learningモデルを用いる。Q-learningモデルでは、参加者が各選択肢(A=左のスロットマシン, B=右のスロットマシン)を選んだ時の行動価値を持っており、選択後のフィードバックに応じて、その価値の更新を

$$Q_{t+1}(A) = Q_t(A) + \alpha(r_t - Q_t(A)) \quad [1]$$

で行う。ここで、 $Q_{t+1}(A)$ は時点 $t+1$ における選択肢Aの行動価値、 $Q_t(A)$ は時点 $t$ における選択肢Aの行動価値、 $\alpha$ は学習率、 $r_t$ は時点 $t$ におけるフィードバックを表している。式[1]は、実際に時点 $t$ で得られたフィードバックから時点 $t$ の選択肢Aの行動価値を引いたものと学習率の積を時点 $t$ の選択肢Aの行動価値に足すことで、次の時点 $t+1$ の選択肢Aの行動価値を更新することを表している。学習率 $\alpha$ は、0から1の範囲の値をとり、大きいほど、学習の更新が速くなる。また、この更新を $Q_t(B)$ に対しても行う。図3Bにおいて、特定の学習率における行動価値をプロットした( $Q_t(A)$ : 灰色破線,  $Q_t(B)$ : 黒色破線)。

次に、式[1]で更新される各選択肢の行動価値から行動選択を出力するモデルとして、

$$P_t(A) = \frac{\exp(\beta Q_t(A))}{\exp(\beta Q_t(A)) + \exp(\beta Q_t(B))} \quad [2]$$

のソフトマックス関数を用いる(今回は2選択肢のため、2選択肢用の式に変更した)。ここで $P_t(A)$ は時点 $t$ における選択肢Aの選択確率、 $\beta$ は選択確率の計算において行動価値の違いを重視するかランダムな選択を重視するかをコントロールするパラメータであり、統計力学領域の呼び方で逆温度と呼ばれる。逆温度 $\beta$ は、0以上の値をとり、0に近くなると行動価値の違いにかかわらずランダムな選択確率になり(2選択肢なので $P_t(A)$ は0.5に近づく)、大きくなるほど行動価値の違いに敏感な選択確率になる( $Q_t(A)$ の行動価値が

大きいなら、 $P_t(A)$ も大きくなる)。なお、 $P_t(B)$ は $1 - P_t(A)$ である。

## 2) 人工データ生成とパラメータ・リカバリーの確認によって認知課題やモデルを修正する

認知課題と認知モデルが準備できたら、人工データ生成とパラメータ・リカバリーの確認をする。認知モデルは、参加者が出力するデータの生成過程のモデルのため、認知課題の情報(提示する刺激やフィードバック)とパラメータがあれば、モデルからデータを生成することができる。その生成されたデータが、人間の示すような振る舞いをするのかどうかを解析前に人工データ生成して確認することが重要である。また、複数の認知モデルで比較をする場合は、それぞれのモデルでパラメータが取りうる範囲の値を用いて人工データ生成を行って、実際にモデルの違いが行動(データ)の違いとして現れるかどうか確認する(Palminteri et al., 2017)。もし、モデルの違いによって行動上の違いが生じないのであれば、実際のデータを用いた時のモデル比較もうまくいかない可能性がある。

人工データ生成ができたら、パラメータ・リカバリーの確認を行う。パラメータ・リカバリーの確認とは、研究で用いる認知モデルとパラメータ推定法によって正確かつ一貫性をもったパラメータ推定ができるのかを検証することである(Heathcote et al., 2014)。特定の認知課題と推定方法においてパラメータ・リカバリーができることが既に論文として発表されていることもあるが、それがない場合は、データ収集の前にパラメータ・リカバリーを調べるシミュレーションを行うことが推奨される。Heathcote et al. (2014)が示した手順としては、まず認知モデルと設定したパラメータの値からデータを生成する(シミュレーション・データの作成)。ここで設定したパラメータの値を、パラメータの真値とする。そして、そのシミュレーション・データに対して、研究で用いるパラメータ推定法でパラメータ推定し、事前に設定したパラメータの真値と一致するかどうかを検討する。パラメータの値、認知課題の設定(試行数など)、参加者の数などを変更しながら検討する。このパラメータ・リカバリーを調べるシミュレーションを通して、真値と離れた

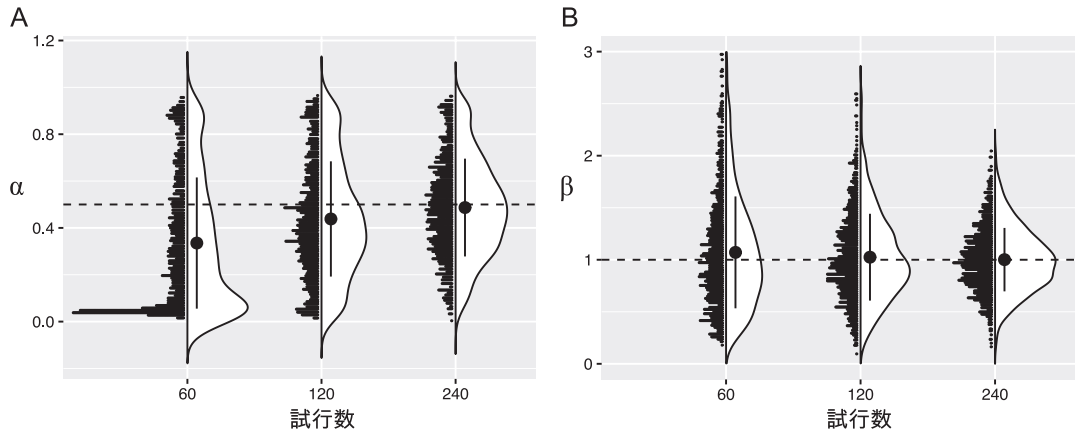


図4 確率的逆転学習課題における Q-learning モデルのパラメータ・リカバリー・シミュレーション

値を推定してしまうなどの問題が認められた場合は、認知課題の修正を行うか、モデルの修正を行う (Palminteri et al., 2017)。

具体的なパラメータ・リカバリーの確認の例として、前述した Q-learning モデルの  $\alpha$  を 0.5、 $\beta$  を 1 に設定してデータを生成し、そのデータからパラメータ推定した (図4)。具体的には、パラメータの真値と Q-learning モデルを用いて、60 試行、120 試行、240 試行の3つの試行数において、それぞれ1000回人工データ生成した上で、ハミルトニアン・モンテカルロ法を使ったベイズ推定でパラメータを推定した。推定されたパラメータの事後分布から、Maximum a posteriori 推定量を求めて、プロットした。人工データ生成とベイズ推定には、R version 3.3.2 (R Core Team, 2016), rstan version 2.16.2 (Stan Development Team, 2017), tidyverse version 1.1.1 (<https://www.tidyverse.org/>) を使用し、プロットの作成にあたり Burant, J.B の Github (<https://gist.github.com/jbburant>) で公開されたコードを参考にした。このシミュレーション条件においては、試行数が240の場合はリカバリーした学習率  $\alpha$  の平均値が真値に近くなっているが、試行数が少なくなるにつれて真値より過小評価されるのが明らかになった (図4A)。また、リカバリーした逆温度  $\beta$  の平均値は、試行数によって真値から遠くなることはあまりないが、その分散は試行数が増えるほど小さくなっていることが明らかになった (図4B)。現実的な制約を考慮しなければ、今回のシミュレーションにおいては、

240 試行でデータ収集を実施するのが良いかもしれない。このように人工データ生成やパラメータ・リカバリーの確認を通して、研究で用いる認知課題やパラメータの推定法がどのくらいの精度でパラメータ推定できるかを確認した上で、実際のデータ収集とパラメータ推定することが望ましい。これらの手続きによって、より精度高くパラメータ推定できる認知課題の選択や課題の修正、その課題に適切な推定法の選択、さらには推定上の工夫からモデルをよりシンプルなものに変更するなどのモデルの修正が可能になる。

### 3) データ収集と行動データを確認する

研究で用いる認知課題と認知モデルが決まったら、データ収集を行う。そして、得られたデータについて、関心のある行動レベルの効果について検討する。

### 4) パラメータを推定する

データにおける行動レベルの効果が検討できたら、パラメータ推定を行う。パラメータ推定法については、最尤法、マルコフ連鎖モンテカルロ法などを用いたベイズ推定、変分ベイズ推定などがあるので、目的に応じて選択する。なお、推定に用いた試行数などに応じて推定されたパラメータの不確実性は異なってくる。そのため、推定されたパラメータの不確実性を評価できるように精度を示す必要がある。具体的には、最尤法などの場合は、ブートストラップを用いて評価し、ベ



イズ推定の場合は事後分布の信用区間 (credible interval) などから評価する (Heathcote et al., 2014)。

## 5) 相対モデル比較を行う

複数の認知モデルがある場合に、相対モデル比較を行って、最も儉約的かつデータにフィットしたモデルを選択する。データ・フィットは、尤度の対数をとった対数尤度で評価できる。しかし、認知モデルに含むパラメータ数が増えるほど、データ・フィットは良くなるが、儉約的ではなくなり、過剰適合を起こす可能性がある。そこで、AIC, DIC, BIC, WAIC, WBIC などの情報量規準やベイズ・ファクターを用いて、相対モデル比較をする<sup>3)</sup>。情報量規準を使うことでモデル比較が簡単にできそうだが、その基準の仮定を考慮した上で、適切な基準でモデル選択するのは実際には難しい (Heathcote et al., 2014)。そのため、後述するモデル・シミュレーションのような情報量規準以外の方法でもモデルの良さを検討する必要がある。他のアプローチとしては、今回パラメータ推定に使ったのとは異なるデータセットでデータフィットを検討することで、一般化可能性を検討する方法もある (Palminteri et al., 2017)。このような交差妥当性の検証もモデルの検討においては有用である。

## 6) モデル・シミュレーションと選択的影響テストを行う

認知モデリングにおいてデータ・フィットなどから相対モデル比較を行うことは多いが、モデル・シミュレーションや選択的影響テストを行う研究は多くない。モデルのデータ・フィットは、モデルの適切さに関して、十分条件を満たすことを意味するが、必要条件を満たすわけではない (Farrell & Lewandowsky, 2015)。さらに、相対モデル比較では、モデル自体の反証は難しいという指摘もある (Palminteri et al., 2017)。相対モデル比較では、モデルが行動データを予測できることを示す予測パフォーマンスは明らかになるが、モデルが関心のある効果を再現できることを示す生成パフォーマンスは明らかにならない (Palminteri

et al., 2017)。生成パフォーマンスとは、与えられたモデルがデータを生成する能力のことである (Palminteri et al., 2017)。生成パフォーマンスを示すには、モデル・シミュレーションを行う必要がある (Palminteri et al., 2017)。モデル・シミュレーションは、パラメータ推定によって最もフィットしたパラメータ値と認知モデルを用いて人工データ生成を行って、実際の行動レベルでの効果が検出できるのかを検討することによって実施する (Palminteri et al., 2017)。

モデル・シミュレーションを行うことのメリットとしては、説明・予測パフォーマンスは良いが生成パフォーマンスが低い認知モデルを除外できる点になる。例えば、確率的逆転学習課題において、1つ前の行動と同じ行動をする反復モデルという認知モデルを作ったとする (Palminteri et al., 2017)。ある程度参加者の行動に一貫性があれば、反復モデルのデータ・フィットは悪くないので、反復モデルでもデータの説明はできているといえる。しかし、反復モデルから人工データ生成をすると、1つ前の行動と同じ行動をするので、初期値から変化せず、実際の行動データとは異なる人工データが生成されてしまう (Palminteri et al., 2017)。反復モデルは少々極端な例ではあるが、モデル・シミュレーションは生成パフォーマンスの低さを検証できる点で重要である。

次に、選択的影響テストは、認知モデルの妥当性を検討するために実験的操作を用いる方法である (Heathcote et al., 2014)。選択的影響テストでは、認知モデルの特定のパラメータに影響すると想定される実験的操作を行って、実際に推定されたパラメータに変化が生じるのかを検討することで、パラメータとモデルの構造の妥当性を検討する。例えば、Voss, Rothermund, and Voss (2004) は、反応時間に関する Drift-Diffusion Model (Ratcliff, 1978) の妥当性を検討するために、色識別課題の条件を操作することで、操作に対応したパラメータの変化を確認している。選択的影響テストを行うことで、パラメータの妥当性の確認ができるとともに、パラメータの解釈も容易になる。

3) ベイズファクターを用いたモデル評価については、岡田 (2018) に詳しい。

## 臨床心理学における認知モデリングと 今後の展開

臨床心理学においても認知モデリングを用いた研究が行われるようになってきている (Busemeyer & Diederich, 2010)。例えば、著者らは、確率選択課題 (Frank, Seeberger, & O'reilly, 2004) と呼ばれる強化学習課題を抑うつつの高い群と低い群に実施し、Q-learning モデルを用いた認知モデリングを行った (Kunisato et al., 2012)。その結果、高抑うつ群では行動レベルにおいて報酬に対する感受性が低く、Q-learning モデルの逆温度パラメータが低いという結果になった。逆温度パラメータが低いということは、それまでの試行で更新してきた行動価値を選択に反映させる程度が低く、選択にランダムさが多く含まれていることを示す。うつ病患者が訴える「なかなか決めることができない」という内的過程を表しているのかもしれない。その後のレビュー論文 (Robinson & Chase, 2017) において、逆温度と抑うつとの関係については、2つの研究では否定されたが、間接的な証拠も含めると著者らの研究以外にも5つの研究で支持されている。まだ、逆温度と抑うつとの関係については知見が少ないが、レビューが行われる程度には増えてきている。今後も認知モデリングを用いた計算論的精神医学研究や計算論的臨床心理学研究は増えると思われるが、まだ研究方法の洗練化も含めて知見の蓄積を進めていく段階といえる。なお、著者が認知モデリングを始めた段階では、認知モデリングの推奨実践についてはまとまっておらず、研究前の人工データ生成やパラメータ・リカバリーの確認を通した検討、競合するモデルを設定した相対モデル比較、モデル・シミュレーションなどは実施していない。今後は、より良い認知モデリング研究のために、推奨実践に沿った研究が求められる。

最後に、認知モデリングの今後の展開について述べる。ベイズ統計学については、パラメータ推定法の1つとして簡単に触れただけであったが、ベイズ統計学を認知モデリングに活用するメリットとしては、推定法におけるメリットと認知モデルの洗練化におけるメリットの2つがある。

まず、認知モデリングにおけるパラメータ推定では、1人の参加者の認知課題データに対して、

最尤法などによるパラメータ推定をする。そして、その参加者ごとに得られた推定値に対して、群間の平均値差や他の変数との相関などを調べる。このようなアプローチは、要約統計量アプローチ (summary statistics approach) と呼ばれる (Daw, 2011)。要約統計量アプローチは、直感的で分かりやすいが、パラメータ・リカバリーの確認で例に挙げたように、個々の参加者の安定したパラメータ推定には多くの試行数が必要になることがある。患者などの臨床群を対象とした時にあまり試行数が多い認知課題の実施は、患者の負担などを考慮すると現実的には難しい。こういった問題に対して、個々の参加者が所属するグループ (研究に参加した人達のグループ、患者のグループなど) を考慮して、そのグループ内で個々の参加者のパラメータが分布するという仮定をおき、グループと個人のパラメータを同時制約的に推定する階層ベイズ推定の利用が提案されている (Ahn et al., 2011)。Ahn et al. (2011) は、アイオワ・ギャンプル課題に対して最尤法を用いた要約統計量アプローチと階層ベイズ推定を比較するパラメータ・リカバリーを調べるシミュレーションを実施した。その結果、最尤法を用いた要約統計量アプローチは参加者とパラメータによっては真値から大きくずれて推定されたが、階層ベイズでは、おおそ真値と一致した推定ができることが示された (Ahn et al., 2011)。なお、階層ベイズ推定を使うと、グループを考慮することによる個人の推定がうまくいくだけでなく、事後分布を示すことができるので、認知モデリングにおける推定したパラメータの精度なども容易に評価できる。認知モデリングにおいて階層ベイズ推定を用いる利点は多いが、実践的には、Stan や JAGS などのプログラミング言語を利用する必要があるため、敷居が高いかもしれない。しかし、最近は階層ベイズや階層モデルに関する優れた解説書が出版されたり (久保, 2012; 松浦, 2016)、複雑なプログラミングをしなくても簡単に特定の認知モデルを用いた階層ベイズ推定をStanを用いて実行できる hBayesDM (Ahn, Haines, & Zhang, 2017) などのRパッケージも作成されており、研究での利用も容易になってきている。今後は、臨床心理学や精神医学における認知モデリングでの階層ベイズ推定が増えてくると考えられる。

ベイズ統計学やベイズの定理は、パラメータ推定法だけでなく、認知モデルの中に組み込むことができる。ベイズの定理は、「事後分布が事前分布と尤度（データに関する分布）の積に比例する」ことを表現している。これを認知モデルに組み込むと、事前にもっていた信念の分布が、環境からの刺激入力から推定される分布によって更新されることで、私たちの知覚や認知などの信念の分布が生じるということになる。認知モデルにベイズの定理を組み込んだモデルは、ベイズ推論モデルと呼ばれる<sup>4)</sup>。ベイズ推論モデルは、私たちの信念を確率分布として表現できるので、ノイズを含むような観察下における私たちの信念の不確実性をモデルに組み込むことができる（O'Reilly & Mars, 2015）。さらに私たちの信念は階層性をもっていることが考えられる。階層性を考慮したモデルは階層ベイズ推論モデルと呼ばれ、代表的なものとして、階層ガウシアンフィルター（Mathys et al., 2014; Mathys et al., 2011）やブレディクティブ・コーディング（Friston, 2005）などがある。階層ベイズ推論モデルなどは認知モデル内に確率分布が含まれており、複雑なモデルにはなるが、階層ガウシアンフィルターなどは統合失調症や自閉スペクトラム症などに対する実証的な研究も行われてきており（Lawson, Mathys, & Rees, 2017; Powers, Mathys, & Corlett, 2017）、今後のさらなる研究が期待される。

## 引用文献

- Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2015). Computational psychiatry: towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery, and Psychiatry*, 87, 53–63.
- Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing neuro-computational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry*, 1, 24–57.
- Ahn, W. Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics*, 4, 95–110.
- Box, G. E. P. (1976). Science and statistics. *Journal of the American Statistical Association*, 71, 791–799.
- Busemeyer, J. R., & Diederich, A. (2010). *Cognitive modeling*. Thousand Oaks, CA: SAGE Publications.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407.
- Daw, N. D. (2011). Trial-by-trial data analysis using computational models. *Decision Making, Affect, and Learning: Attention and Performance XXIII*, 1–26.
- Farrell, S., & Lewandowsky, S. (2015). An introduction to cognitive modeling. In B. U. Forstmann & E. J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 3–24). New York, NY: Springer.
- Frank, M. J., Seeberger, L. C., & O'reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, 306, 1940–1943.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 360, 815–836.
- Heathcote, A., Brown, S. D., & Wagenmakers, E. J. (2014). An introduction to good practices in cognitive modeling. In B. U. Forstmann & E. J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 25–48). New York, NY: Springer.
- 久保拓弥 (2012) データ解析のための統計モデリング入門：一般化線形モデル・階層ベイズモデル・MCMC 岩波書店.
- 国里愛彦 (2013) うつとストレスに対する計算論的アプローチ：計算論的精神医学入門 ストレス科学, 28, 101–107.
- 国里愛彦 (2018) 計算論的臨床心理学からみた認知行動療法 認知療法研究, 11, 2–12.
- Kunisato, Y., Okamoto, Y., Ueda, K., Onoda, K., Okada, G., Yoshimura, S., Suzuki, S., Samejima, K., & Yamawaki, S. (2012). Effects of depression on reward-based decision making and variability of action in probabilistic learning. *Journal of Behavior Therapy and Experimental Psychiatry*, 43, 1088–1094.
- Lawson, R. P., Mathys, C., & Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nature Neuroscience*, 20, 1293–1299.
- Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian cognitive modeling: a practical course*. Cambridge, UK: Cambridge University press. 井関龍太（訳）・岡田謙介（解説）(2017) ベイズ統計で実践モデリング：認知モデルのトレーニング 北大路書房.
- Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman and Company. 乾敏郎・安藤広志（訳）(1987) ビジョン：視覚の計算理論と脳内表現 産業図書.
- Mathys, C., Daunizeau, J., Friston, K. J., & Stephan, K. E. (2011). A bayesian foundation for individual learning

4) ベイズ推論モデルとその動向については、中村（2018）に詳しい。

- under uncertainty. *Frontiers in Human Neuroscience*, 5, 39.
- Mathys, C. D., Lomakina, E. I., Daunizeau, J., Iglesias, S., Brodersen, K. H., Friston, K. J., & Stephan, K. E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Frontiers in Human Neuroscience*, 8, 825.
- 松浦健太郎 (2016) Stan と R でベイズ統計モデリング 共立出版.
- 中村國則 (2018) 高次認知研究におけるベイズ的アプローチ 心理学評論, 61, 67–85.
- 岡田謙介 (2018) ベイズファクターによる心理学的仮説・モデルの評価 心理学評論, 61, 101–115.
- O'Reilly, J. X., & Mars, R. B. (2015). Bayesian models in cognitive neuroscience: a tutorial. In B. U. Forstmann & E. J. Wagenmakers (Eds.), *An introduction to model-based cognitive neuroscience* (pp. 179–197). New York, NY: Springer.
- Palminteri, S., Wyart, V., & Koechlin, E. (2017). The importance of falsification in computational cognitive modeling. *Trends in Cognitive Sciences*, 21, 425–433.
- Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science*, 357, 596–600.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59.
- R Core Team (2016). R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Robinson, O. J., & Chase, H. W. (2017). Learning and choice in mood disorders: searching for the computational parameters of anhedonia. *Computational Psychiatry*, 1, 208–233.
- 清水裕士 (2018) 心理学におけるベイズ統計モデリング 心理学評論, 61, 22–41.
- Stan Development Team (2017). rstan: R interface to Stan. R Package version 2.16.2.
- 須賀哲夫 (1989) 理論心理学アドベンチャー 新曜社.
- 須賀哲夫 (2003) 三つの個性 北大路書房.
- 須賀哲夫 (2011) 実験心理学をリフォームする: 理論心理学からの提言 北大路書房.
- 丹野義彦・石垣琢磨・毛利伊吹・佐々木淳・杉山明子 (2015) 臨床心理学 有斐閣.
- 戸田正直 (1992) 感情: 人を動かしている適応プログラム 東京大学出版会.
- 豊田秀樹 (編) (2017) 実践ベイズモデリング: 解析技法と認知モデル 朝倉書店.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: an empirical validation. *Memory & Cognition*, 32, 1206–1220.
- Watkins, C. J. C. H. (1989). *Learning from Delayed Rewards*. Cambridge, UK: Cambridge University.
- Weisberg, M. (2013). *Simulation and similarity: using models to understand the world*. Oxford: Oxford University Press. 松王政浩 (訳) (2017) 科学とモデル シミュレーションの哲学入門 名古屋大学出版会.
- Wiecki, T. V., Poland, J., & Frank, M. J. (2015). Model-based cognitive neuroscience approaches to computational psychiatry: clustering and classification. *Clinical Psychological Science*, 3, 378–399.

— 2018. 1. 31 受稿, 2018. 2. 27 受理 —