

データ科学的手法によるヒト認知機能の統計モデリング



技術解説

朝倉 暢彦*

Statistical Modeling of Cognitive Functions Based on
Principles and Methods of Data Science

Key Words : Data science, Cognitive function, Statistical modeling

はじめに

近年の急激な情報通信および計測技術の進展により、多種多様な大規模・大量データ（ビッグデータ）が集積されるようになり、これを活用した新たな科学的発見や価値創造、そして革新的な新サービスの創出が期待されている。そして、多様化・高度化したビッグデータを適切に扱うための方法論として、AIと並んでその重要性が認識されるようになってきたのがデータ科学（データサイエンス）である。データ科学は統計学を基礎として、情報学および計算機科学が融合した学際分野であり、ゲノム科学や天文学、高エネルギー物理学などの最先端の実験科学におけるビッグデータ解析において今や欠かせないものとなっている。

このようにデータ科学が統計学を内包しながらも、新たな学問分野として成立してきた一つの要因として、ビッグデータ解析では従来の統計学が扱ってこなかったタイプのデータを扱う場面が多いということが挙げられる。その典型が、自明でないデータ取得環境から得られた非構造化データである。従来の自然科学的研究では、ある目的のために厳密に構築された実験系と統制された実験条件のもとで得られたデータが統計的解析の対象となる。これは多くの場合、古典的統計学におけるフィッシャーの3原則：1) 局所管理、2) 無作為化、3) 反復に沿った実験計

画に基づいて計測されたデータであり、系統誤差がランダム誤差に転化されるとともに、その誤差の定量的評価が可能である。一方、インターネットやウェアラブルセンサーから得られる個人の活動データに代表されるように、近年の急速な計測技術の進展によって蓄積されてきたビッグデータは、そのデータ取得環境が自明でないばかりか、何らの統制もされていないことが多い。これはランダム誤差のみが存在することを前提とした統計的解析には不向きなデータであり、データ取得環境とそれに起因するあらゆる不確定性をモデリング（統計モデリング）することがデータ解析に先立って必要である。データ科学では、このデータ生成過程の統計モデリング、すなわちデータ生成モデル（generative model）の構築と不確定性の定量的評価をベイズ推定の枠組みで行うのが主流のアプローチとなっている。

このような観点からデータ科学の課題を考えると、それは不確定性をもったデータからの有効な情報の抽出、その確からしさの定量的な評価に基づく解析対象の理解、およびデータを取り巻く環境において適切な意思決定を導くことと言える。そしてこの課題は、世間で注目されているビッグデータ解析のみならず、ヒトの認知システムにおいても解決すべき同型の課題であり、データ生成モデルを用いたベイズ推定というデータ科学の枠組みがヒト認知機能のモデリングにも極めて有効であるということが本稿における主張である。

実際、ヒト認知システムのベイズモデリングは近年の認知科学研究において標準的なアプローチである。ただ、その端緒が1980年代半ばの第2次AI（ニューロ）ブームにあったにもかかわらず、現在の第3次AIブームがAIやデータ科学技術の社会実装に注力するあまり、その技術の脳科学や認知科学との接点が希薄となっているように思われる。従って、



* Nobuhiko ASAKURA

1969年10月生まれ
京都大学大学院 文学研究科 心理学専攻
攻博士後期過程（1998年）
現在、大阪大学数理・データ科学教育研究センター 特任准教授（常勤）
博士（文学） 専門／認知科学、計算論的神経科学、データ科学
TEL : 06-6850-6468
FAX : 06-6850-6092
E-mail : asakura@sigmath.es.osaka-u.ac.jp

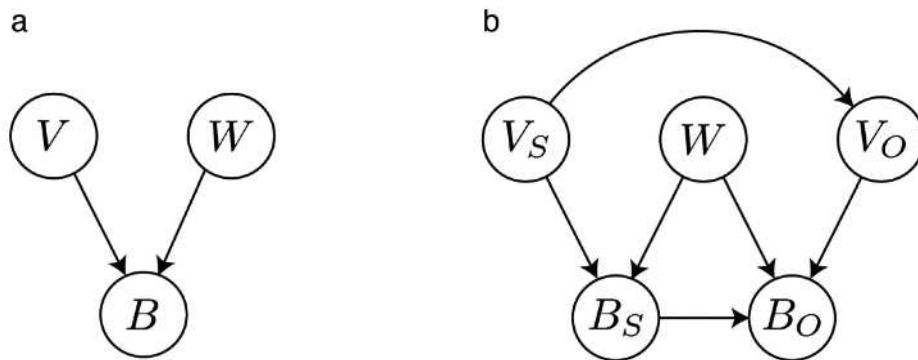


図1 心的状態のベイジアンネットワーク。
 a 外界の知識 (W) と視覚経験 (V) および信念 (B) の因果関係のグラフ。
 b 誤信念課題に対するベイズモデル。添字の S は自己, O は他者。

本稿の内容は決して最新の技術について述べるものではないが、現在のデータ科学に対して新しい視点を提供できるのではないかと考え、関連する筆者の研究を紹介する、

心の理論のベイズモデリング

心の理論 (theory of mind) とは、直接観察できない他者の心的状態を推測し、それに基づいて他者の行動予測を可能とする認知的枠組みのことを指す。これは知識と欲求および信念に基づいて行動が選択されるという心の働きに関する理解とも言える。心の理論を適切に運用することは、円滑な対人コミュニケーションを実現する上で重要である。また、自閉症スペクトラムにおけるコミュニケーション障害の背景要因として、心の理論の機能不全があると考えられている。

心の理論は生得的なものではなく、乳幼児からの発達に伴って成立していく認知機能である。それゆえ発達心理学においては重要なトピックであり、誤信念課題 (false belief task) と呼ばれる課題を用いて幼児を対象とした数多くの研究が行われてきた。誤信念とは、事実とは異なる誤った信念のことを指す。例えば、特徴的なパッケージデザインの菓子箱 (例、ポッキーの外箱) を見れば、通常その中には対応するお菓子 (ポッキー) が入っているという信念をもつが、もし中に鉛筆が入れられていたとすると、これは現実とは異なった誤信念をもったことになる。この事実を受け入れた上で、それを知らない他者の信念状態、すなわち他者が箱の中に何が入っていると考えているのかを問うのが誤信念課

題である。ここで「お菓子」と正答すれば、他者の誤信念を心の理論を用いて正しく推測していることが示される。一方、「鉛筆」と答えるのは、心の理論がうまく機能しておらず、自己と他者の信念が混同され誤った推測を行っていることになる。これまでの研究から、誤信念課題に正答できるようになるのは4歳以降であり、その正答率も年齢に従って上昇していくことが明らかにされている。

我々は、心の理論の機能が心的状態変化の内部モデルを用いたベイズ推定として実現されていると考え、そのモデリングを行なった¹⁾。この内部モデルが心的状態の生成モデルであり、外界の事前知識 (W) と視覚経験 (V) から外界に関する信念 (B) が形成される過程の心的シミュレータの役割を果たす。そしてこの過程を条件付き確率 $P(B|V,W)$ に対応するベイジアンネットワーク (BN) として表現した (図 1a)。このモデルでの仮定は、心の理論が成立していれば、自己と他者の心的状態がそれぞれ独立した BN として脳内で実装され、他者の内部モデルを適切に用いた誤信念推定が可能となるというものである。一方、自己と他者の内部モデルが未分化な状態が心の理論の発達前の誤信念推定を反映していると考えた。この状態は、自己と他者の BN に対して、自己 BN の変数 V_S, B_S を親ノードとして他者 BN の変数 V_O, B_O へとそれぞれリンクが付与される形で統合された構造に対応する。これは、自己の視覚経験と信念が他者の心的状態推定に影響を与える構造となる (図 1b)。

以上の定式化を踏まえると、心の理論の発達過程は、自己 BN から他者 BN へのリンクの強さが減弱

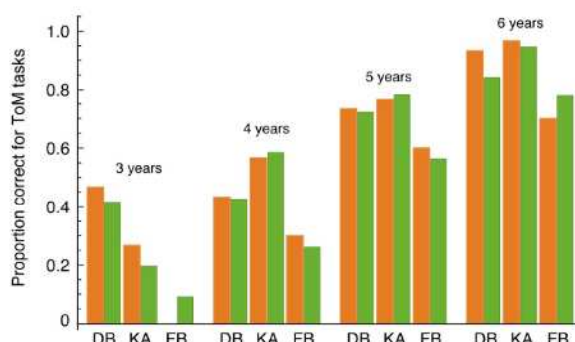


図2 モデルと東山による実験結果³⁾との比較。
DB: 信念課題, KA: 知識アクセス課題, FB: 誤信念課題。縦軸は課題の通過率 (全参加者中, 正答した幼児の割合)。オレンジの棒が実験データ, 緑の棒がモデルからの予測。

していく過程として捉えることができる。すなわち、自己の視覚経験が他者のものと異なること、および自己の信念が他者のものと異なることをそれぞれ理解できるようになって他者の誤信念理解が実現されることになる。実際、最初の2つの心的状態の理解は、それぞれ信念課題 (diverse belief task)、知識アクセス課題 (knowledge access task) として誤信念課題とともに検討されており、誤信念の理解に先立って発達することが明らかにされている²⁾。モデルにおいては、自己BNから他者BNへのリンクの強さのパラメータとして信念課題と知識アクセス課題の正答率を取り入れ、他者信念B₀の事後確率に基づいて誤信念課題での意思決定が行われるとした。図2はモデルを東山³⁾による3~6歳児を対象とした実験データに当てはめた結果であり、年齢に伴う発達過程をうまく説明できている。また、同じモデルが様々な文化的背景の幼児データおよび自閉症児や聴覚障害児のデータにもうまく当てはまることも示されている。

アイオワギャンブル課題における健常者の選択行動のクラスターリング

意思決定の神経心理学的検査として用いられるアイオワギャンブル課題 (IGT) では、報酬と損失の金額およびその頻度が異なる4つのカードの山を逐次選択する (図3)。4つのうち2つの山 (AとB) は1枚あたりの報酬・損失が高額ながら、選択し続けると損となる期待値が設定された「悪い山」であり、残りの2つの山 (CとD) は1枚あたりの報酬・損失は低額であるが、長期的には得となる

	A	B	C	D
Gain per card	\$100	\$100	\$50	\$50
Loss per card	\$150-\$350	\$1250	\$50	\$250
Loss (10 trials)	\$1250	\$1250	\$250	\$250
Loss frequency	50%	10%	50%	10%
Net (10 trials)	-\$250	-\$250	\$250	\$250

図3 アイオワギャンブル課題

期待値が設定された「良い山」である。そして良い山と悪い山の双方において、損失が高頻度で生じる山 (AとC) と低頻度で生じる山 (BとD) が設定されている。このようなカードの山の構成は実験参加者には知らされず、試行を繰り返して課題の原理を学習しつつ、最終的な獲得利益が最大となるような選択を行うことが要求される。IGTを考案したBacharaらによると、健常者は長期的利益に基づいた合理的な選択が可能であり、試行が進むにつれて次第に良い山からカードを引くようになる。一方、脳の腹内側前頭前野を損傷し意思決定に障害のある患者群では短期的利益をのみを考慮して「悪い山」のカードを引き続ける傾向が顕著であることが報告されている⁴⁾。しかし、近年の研究では最終的に悪い山を選好する健常者も多数存在することが報告されており、意思決定の障害に対する検査法としての妥当性に疑義が呈されている。

そこで我々は、IGTにおける健常者の選択行動の生成モデルに基づいてIGTの公開データベース⁵⁾のクラスター分析を行い、健常者の特徴的な選択パターンを同定した⁶⁾。用いたデータベースでは、10事例614名の健常者が各試行において選択したカードの山とその際に得られた報酬・損失の時系列が全て記録されている。その中で総試行数が100以上で標準的な報酬・損失額の設定を用いている6事例405名のデータを取り上げ、各カードの山の期待報酬をほぼ学習したと考えられる終盤25試行における選択パターンを分析対象とした。

まず個人の選択行動の生成モデルを、K種類の可能な選択パターン (各カードの山の選択確率の組み) から採用確率 π で1つの選択パターンを採用し、これに基づいて1つのカードの山を選択するものとした。IGTは4枝強制選択課題であるので、個人のデータの尤度は4つのカードの山それぞれの選

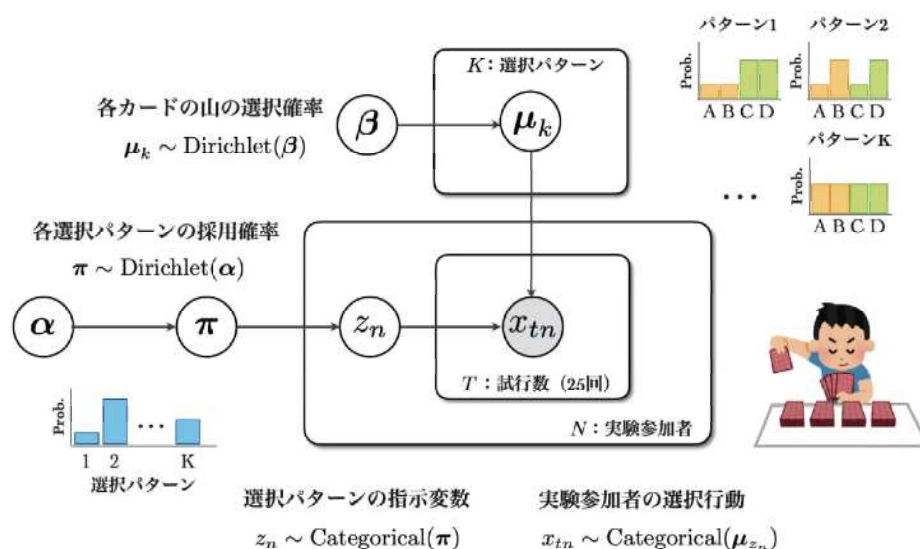
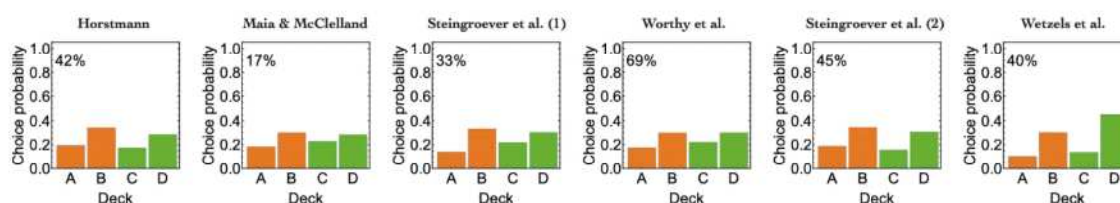


図4 全実験参加者のカード選択行動の生成モデル。

図5 クラスタリングから得られた典型的な選択パターン。
 各グラフ中の数値は各研究事例においてこの選択パターンを採用した
 実験参加者の割合

択確率をパラメータとした多項分布となる。これより、各研究事例の全実験参加者のデータの尤度が、 K 種類の多項分布とそれぞれの生起確率 π から構成される混合多項分布として表現される。さらに、各カードの山の選択確率と各選択パターンの採用確率の事前分布としてディリクレ分布を導入することにより、全参加者のカード選択行動の生成モデルが構成される (図4)。これはテキスト処理における機械学習アルゴリズムの1つである混合ユニグラムモデルと等価なものとなっている。この生成モデルを用いて、特定の選択パターンに対応する個々の多項分布のパラメータとそのパターンの生起確率をベイズ推定することでクラスタリングを行った。また、最適なクラスタ数は周辺完全尤度を用いた情報量基準 (ICL) により決定した。

その結果、6 事例のデータに対して ICL より 4 ～9 個のクラスタが決定された。そして各事例における上位クラスターを確認したところ、全事例に共

通して多数の参加者 (平均 41%) が示すのは、長期的な利益の基準ではなく、損失の頻度の低い山 (B と D) を選好するパターンであった (図5)。これは報酬が得られればその山のカードを引き続け、損失を被るとカードを引く山を変えるという Win-Stay, Lose-Shift の行動戦略に対応するものである。一方、これまで主張されてきた長期的な利益をもたらす山 (C と D) を選好するパターンはほとんど見られなかった。以上の結果は、IGT に対する健常者の選択が必ずしも合理的意思決定の方略に従わないことを示すとともに、意思決定の神経心理学的検査としての妥当性を再考すべきであることを示唆している。

おわりに

本稿では、データ科学と認知システムに課せられた課題が同型であることを指摘し、データ生成モデルを用いたベイズ推定の枠組みからヒトの認知機能をモデリングする試みについて紹介した。実際、不

良設定問題として捉えられる感覚・運動系の情報処理(乾⁷⁾、川人⁸⁾)から、他者の表情や動作の認識などの対人コミュニケーションに必須な高次認知処理に至るまで、不確定性をもつ観測データから有効な情報を抽出してその確からしさを適切に評価するという点で、認知システムが対処すべき課題はデータ科学が扱う課題と同型である。

現在、様々な業界を席卷しているAIや機械学習そしてデータ科学であるが、いわゆるブームとして捉えられる時代は過ぎ去りつつあると思われる。今後の発展の1つの形として、本稿で紹介したような脳科学や認知科学との結びつきを重視したアプローチが有望なのではないかと考えている。

参考文献

- 1) Asakura, N & Inui, T: A Bayesian framework for false belief reasoning in children: a rational integration of theory-theory and simulation theory. *Frontiers in Psychology*, 7:2019 (2016).
- 2) Wellman, HM & Liu, D: Scaling of theory-of-mind tasks. *Child Development*, Vol.75, pp.523-541 (2004).
- 3) 東山薫: “心の理論”の多面性の発達—Wellman & Liu 尺度と誤答の分析, *教育心理学研究*, Vol.55, pp.359-369 (2007).
- 4) Bechara A, Damasio AR, Damasio H & Anderson SW: Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, Vol.50, pp.7-15 (1994).
- 5) Steingrover, H, Davis, H, Fridberg, DJ, Horstmann, A, Kjome, KL, Kumari, V, Lane, D, Maia, TV, McClelland, JL, Pachur, T, Premkumar, P, Stout, JC, Wetzels, R, Wood, S, Worthy, DA & Wagenmakers, EJ: Data from 617 Healthy Participants Performing the Iowa Gambling Task: A “Many Labs” Collaboration. *Journal of Open Psychology Data*, 3e5 (2015).
- 6) 朝倉暢彦, 乾敏郎: アイオワギャンブル課題における健常者の選択行動のクラスタ分析, *日本認知心理学会第15回大会発表論文集*, (2017).
- 7) 乾敏郎: 「Q & Aでわかる脳と視覚」, サイエンス社 (1993).
- 8) 川人光男: 「脳の計算理論」, 産業図書 (2006).

