

統計科学の文法と論理的推論

椿 広計†

(受付 2019 年 1 月 5 日; 改訂 3 月 24 日; 採択 3 月 26 日)

要 旨

本論文は, Goertz and Mahoney (2012)に従って, 法則を秩序とする科学の主要な研究プロセスが統計的・定量的探求プロセスであるのに対して, プログラムを秩序とする科学の主要な研究プロセスが, 論理的・質的探求プロセスであるとの仮説を提示する. その裏付けのために, 統計科学的立場から, 定量的プロセスと論理的プロセスの差異が顕著となる幾つかの状況を示す. 更に, プログラム科学と法則科学との融合可能性についても議論する.

キーワード: 科学の文法, プログラム, 法則, 定量的研究, 定性的研究.

1. 問題意識の変遷—法則の認識科学と法則科学から設計科学とプログラム科学へ

本論は, 吉田民人が主導した「新しい学術の体系」で示された「プログラム科学」(日本学術会議運営審議会附置新しい学術体系委員会, 2003, 吉田, 2013)の主要な研究プロセスが, Goertz and Mahoney (2012)が示した社会科学における論理的・質的探求プロセスであるとの仮説を提示する. この仮説を補強するために, 統計科学によって実証される法則科学的探究と論理的なプログラム科学の探究プロセスとの方法論的差異についても幾つかの事例を示す. 特に, 集合・数え上げ・分類といった質的・離散数理技法の多くが, 万象を記述する法則の科学的認識ではなく, 人間が生存の便宜のために開発した情報縮約的便法と考えられることも指摘する.

筆者は科学哲学の専門家ではなく統計家であり, 統計的实践の成功や失敗の経験を通じて, 科学的法則の探究プロセスやその支援技法について暗黙知を形成してきた. 従って, 本論では, 筆者の統計家の経験を基に, 法則科学ないしは統計科学の視点から, プログラム科学ないしは論理的推論の特性を指摘しただけである. 決して統計的推論が論理的推論に優越するという主張を意図してはいない.

筆者は, 横断型基幹科学技術連合活動の中で, 「新しい学術の体系」の解釈作業(日本学術会議自動制御研究連絡委員会工学共通基盤研究連絡委員会自動制御学専門委員会, 2005)に参加し, 吉田の科学論を知った. その中で, 価値に依存しない「あるものの追求」を目的とする「認識科学」形成の標準プロセスは, 統計科学創成を興した Pearson (1892)の「科学の文法」が原点であると再認識した. 吉田は, 価値に依存した「あるべきものの追求」を目的とする「設計科学」の標準プロセスを「情報循環」とした. しかし, 当時の筆者には, 設計科学のための情報循環プロセスが, 明確に標準化されているとは考えられなかった.

そこで, Tsubaki et al. (2008)は,

1) 価値(関係者の声)の選択

† 統計数理研究所: 〒190-8562 東京都立川市緑町 10-3

- 2) 関係者の声を技術の声に翻訳(価値実現システムの選択)
- 3) 技術システムの最適化
- 4) システムが実現する価値の社会への注入(実装)

からなる、「技術開発の文法」を提案し、そのプロセスを支援する数理技術の適用例を示した。

次に、筆者は、新技術・新製品開発の標準プロセスモデルと、その各フェイズにおける支援方法群を示す指針を国際標準とする活動を ISO/TC 69 “Applications of statistical methods” で開始した。2009 年に ISO/TC69/SC8 “Application of statistical and related methodology for new technology and product development” が設置され、2015 年には ISO16355 シリーズの第 1 部を発行した(ISO, 2015)。標準化活動と並行して、設計科学の文法を産学共同研究する「価値創生プロセス実践知開発ネットワーク(VCP-Net 研究会)」も組織した。この研究会では、管理技法、数理技法、統計技法の目的と機能から統計的に分類すると共に、技術開発プロセスのあるフェイズで活用される技法が、前のフェイズから引き渡される入力情報、後のフェイズに引き渡す出力情報の要求品質明示化などを研究した(VCP-Net 研究会編, 2014)。このような標準化や研究が、吉田が提起した情報循環を高質化・効率化すると考えたからである。

吉田(2013)の「プログラム科学」を再考するきっかけとなったのは、2016 年から開始した「顧客の期待を超えたエクセレント・サービス」を実現するプロセスの標準化活動である。物質やエネルギーを支配する秩序と、サービスを制御する秩序とでは、それに対処する数理科学的方法が本質的に異なることをどうプロセスに反映させるかが課題なのである(椿, 2017)。実際、1980 年代からソフトウェアやサービスの品質経営分野では、事実に基づく問題解決に資する統計的方法よりも、あるべき姿を求めて、言語情報を論理的に分析する技術が有効と考えられてきた。

椿(2018)では、プログラムを記述する記号列の事象への当てはめでは、連続的回帰モデルより、離散層別モデルの方がプログラム科学のためには有効ではないかと考えた。顧みれば、土橋 他(1985)では、サービス分野での言語情報収集とその定量的分析の必要性、統計的方法に限定しても、層別原理の徹底、特に今日では第 2 世代人工知能と呼ばれる Breiman et al.(1984)の CART による層別の自動化への注目を呼び掛けている。

こうした中、Goertz and Mahoney(2012)が、数学が定性的(質的)研究、統計学が定量的研究のエンジンであり、因果推論も両者では本質的に異なるという指摘に触れた。それを通じて数理的学術形成プロセスの 2 つのパラダイムの乖離が、人工的・離散的なプログラム科学と、天然的・連続的な法則科学との数理的差異に起因するという仮説を形成するに至った。本論は、そのような問題意識を試論としてまとめたものである。

第 2 章でこれまでの筆者の「科学の文法」に関わる認識を総括する。3 章では、吉田が提唱したプログラム科学における認識が、法則科学における認識とは異なる離散的選択原理によって支えられていることを示す。4 章では、生物の生存にとって必要な、認識の離散化を支援する様々な数理的方法が人間によって生み出されていることを述べる。5 章では、これらの考察を受けてプログラム科学の文法生成の課題と法則科学文法との融合可能性について議論する。

2. 科学の文法

2.1 法則認識科学の文法

Pearson(1892)は、「人間が法則(Law)を自然(Nature)に付与する」、事実に基づく科学形成のプロセスモデルを「科学の文法」として提示した。観察に基づく記述に基づき現象への法則を付与する手続き自体を科学的行為とした。そのプロセスは、単純化すれば、観察に基づく事実の秩序に関する分析、その総合、批判的検証からなる一連の追究プロセスである。Pearson は、

科学的法則記述の文法としての標準プロセスを支援する数理的方法群の開発も開始した。ヒストグラム、標準偏差、相関分析、重回帰分析、主成分分析、モーメント法(ピアソン系分布族への当てはめや混合正規分布)、適合度などである。研究者が、法則記述のプロセスと記述に用いる数理技法体系を同時提供したのである。

20世紀前半には、計量生物学、計量心理学、計量経済学といった計量諸科学が創生され、統計的記述は自然現象のみならず社会現象までその適用範囲を広げると共に、技法も進化を続けた。特に、相関概念や適合度概念は、人間が付与する法則のパフォーマンス尺度であり、より便宜性の高い法則を求める実証科学の営みを支える基幹概念となった。

有限のデータがもたらす有限の情報量に基づいて、統計的方法が人間の法則に関する認識の地平を広げるためには、3つのステップがある。

第一ステップでは、関心のある現象に対する既存知識を表現した統計モデルによる予測値とは乖離する実測値、いわゆる外れ値を検出し、それを問題視するか否かを検討する。この異常検知操作が、統計的方法による問題発見の基本形である。

第二ステップでは、データが外れ値となったことに影響を与えたと考えられる原因候補(要因)を考察し、要因に関わる変数を新たな入力データとして統計モデルに追加する。この際、必要に応じて新たなデータを統計モデル同定にとって最も情報効率が高くなるように採取する。

第三ステップでは、統計モデルをデータに当てはめ、その結果を解釈し、要因から原因を絞り込み(モデル選択)、現象に対する知識を更新する。

なお、既存知識の地平を表現し、現象に当てはめられる統計モデルには、物理学のような演繹的科学を基に、現象の静態や動態を記述する数理モデルを原型としたモデルも含まれる。計量科学は、理論なき実証と批判されることも多いが、「科学の文法」は、法則進化を目指すプロセスモデルであり、統計的方法を数理技術として制御や予測に応用する技術操作とは目的が異なる。

第二、第三のステップが、統計分析ないしは要因分析と呼ばれ、統計教育・研究の中心となった。この統計分析のあるべきプロセスを Cox and Donnelly (2011)は、次のようにまとめ“Ideal Sequence”と呼んでいる。

1. 研究すべき問題、仮説の定式化
2. 関連するデータの探索と適切なデータを採取する研究の計画と実施
3. データ解析
4. 適切な意思決定に繋がる結果の解釈

確かに、研究すべき問題や仮説の発想は、研究者の先見性や多様性受容能力も肝要である。しかし、上記第一ステップで示したように既存統計モデルからの外れ値という事実に着目すれば、新たな「突き止められる原因(Assignable Cause)」が明らかになる可能性が強い。この常識を基本原理として明示化し、社会実装したのは統計的品質管理学の創成者、Shewhart (1939)であり、実証科学の自律的改善のサイクル形成に寄与した。

2.2 設計科学の文法—統計数理的マネジメントの社会実装

吉田に主導された日本学術会議運営審議会附置新しい学術体系委員会(2003)は、価値に依拠した科学を「設計科学」と呼び、従来の価値に依存せず現象の背景原理に関する解釈を追究する科学を「認識科学」と呼んだ。吉田は設計科学の基本的方法を「情報循環」とした。

一方、Shewhart (1939)は、表1に示すように統計科学的マネジメントを大量生産のマネジメント更には一般社会のマネジメントに進化させた。Shewhartの影響を受け Deming, 石川馨らは、PDCA サイクルと「問題解決型 QC ストーリー」という統計的改善活動の標準プロセスを

表 1. Shewhart による統計科学プロセスの社会実装.

フェイズ	適用範囲		
	統計科学	狭義品質管理	マネジメント全般
Plan (計画)	仮説の提示	仕様の提示	達成すべき 目的の提示
Do (実施)	観察・実験	生産	目的を達成しよう とする行為
Check (検証)	仮説の検定	仕様適合性の検査	目的が達成された か否かの検証

社会実装した。これらは、まさに価値を追求するための汎用的プロセスモデルであり、「設計科学の文法」と呼ぶべきものであった(椿, 2019)。

設計科学が認識科学と異なるのは、価値関数の選択, 社会的制約条件の選択, 価値関数の制約付き最適化といった, 個々の人間の価値観や倫理観に依存するプロセスが明確に存在することである。このため, 数理計画や統計的決定理論が重要な位置づけになる。

認識科学の進化のためにも最適化や制約が用いられる。しかし, 認識科学ないしは法則科学では, 物質や情報の法則の背後にある変分原理や尤度原理に基づく天然の最適化や, 既存法則成立を制約とした現象認識といった, 価値には依存しない最適化や制約を用いるのである。

ここで, 価値最適化を実現する条件設定のための統計的方法開発の歴史を振り返る。多くは, 法則を進化させるための統計モデル当てはめという意識よりは, 必要な比較評価技法の開発である。Gosset は, 平均値の比較のための t 統計量を開発し, ギネスビールの品質改善を行った (Student, 1908; Box, 1987)。Fisher (1935) は, 実験計画法を提唱し, 農産物の収量や品質の改善に寄与した。田口 (2000) による機能性評価実験 (動特性 SN 比) は, 利用したい理想の法則 (機能) に現実がどのくらい近いのか, それをどう改善するかの体系的評価技法を示したものであり, 設計科学の発展に横断的に寄与した (椿・河村, 2008)。このようにデータに基づく方法が, 現象の認識だけでなく, 様々な要因に駆動されたシステムを統計的に比較して, 人間が利用するシステムの最適化にも利用されてきた。

実社会における最適化の対象は, 支配秩序としての物質やエネルギーなどの法則を利用する現象には限らない。整数計画法に代表される離散最適化を必要とする秩序体系が厳然と存在する。これを以下では, 吉田に従ってプログラムと呼ぶことにする。

3. プログラムの認識科学

3.1 秩序原理としてのプログラム

吉田 (2013) は, 設計科学と認識科学, 物質科学, 生物科学, 社会科学など様々な科学の分類軸を提案した。ここでは, 「法則科学」と「プログラム科学」という分類に注目する。物質科学は, 「法則」という違反もなければ例外もない秩序に支配されている。法則は, 物理学モデルやその影響をうける統計的定量モデルで近似可能である。

一方、生物科学や社会科学は「プログラム」と呼ばれる、違反事象や例外事象を許容する秩序原理にも支配される。生物科学は遺伝子のようなシグナル列(シグナル性プログラム)に支配され、社会科学はシンボル列で記述されるプログラム(シンボル性プログラム)に支配されるとした。

Pearson (1892)は、20世紀には物理学は数学との融合が進み、生物科学は物理学との融合が進むと述べた。一方、Pearson (1892)はlawという用語が「法」という社会秩序にも用いられていることを意識しており、それが「法則」とは異なることを強調していた。また、Pearsonは、現実に認識される法則には、誤差があることも指摘した。これは様々な未知原因が観察現象を駆動するからであり、それら未知要因に起因する変動を記述する数理ツールとして確率モデル利用を提案したのである。なお、竹内(2010)も人間の「無知」の部分としての偶然概念を取り上げており、この種の認識は確率モデルの利用を推進した近代統計学の立場だと考えられる。

Pearsonにとっての科学の進化は、利用可能な法則が、人間科学・社会科学も含む多くの分野に広がり、法則の現象近似パフォーマンスが上がることだった。これに対して、吉田のプログラムは、未知の原因が交絡していなくても、例外や違反が存在することを前提とする。逆説的だが、プログラム秩序を利用するものからすれば、プログラムの有用性は、例外や違反が生じる割合が小さいこととなる。しかし、プログラムの違反や例外は、無知のために確率変動と見なした原因で起きるのではない。異なる論理を共有する少数派と呼ばれる人間の判断には確信的違反や例外も存在する。

3.2 プログラム認識科学成立に必要な統計的方法と論理学的方法

プログラム科学を実効化させるには、符号列を原因情報として、結果系パフォーマンスを認識評価する方法が必要となる。プログラム科学が事実に基づく方法を必要とするならば、統計的には2.2節で述べた離散化されたグループ間の事実に基づく評価・比較という方法体系が必要となる。

プログラムが何らかの法則を駆動し結果として量的パフォーマンスが得られる場合には、Fisher (1935)の実験計画法と多元配置分散分析を異なるシンボルやシグナルの作用といった質的制御因子に限定して適用すればよい。プログラム間の定量的結果比較を行えば、パフォーマンスの効率的認識と共に最適化まで行える。時系列的シンボル列の順序構造がパフォーマンスに与える影響評価は、医学統計ではFieler (1940)、農業統計ではCochran et al. (1941)のクロス・オーバー計画と呼ばれる実験計画法技術に埋め込まれた。

一方、離散的入力 of 質的出力パフォーマンスを評価する離散データ解析も統計学の枠組みで整備されてはいた。符号系列の質的パフォーマンスに関する影響評価についてもクロス・オーバー計画は利用可能であり、符号列投入の時期・順序・直前符号の残存影響など様々な検討が可能である。広津・椿(1984)は、3種の薬剤のうち2種類を様々な順序で一定期間ずつ投与し、前期・後期の有効性の優劣比較を行うクロス・オーバー試験から有効な治療法を定める統計モデルを議論した。表2のように薬剤名をシンボル化すれば、これは最も単純なシンボル性プログラム評価のためのデータと考えられる。空間的二値シンボル列(0-1列)と離散的パフォーマンスとの関係性を認識することは、文字認識など、背景法則が存在しない人工概念に対する統計的パターン認識は、統計的機械学習分野で日常的に用いられている。

しかし、離散現象の背後に連続的な法則を想定して、離散観測量から法則の性能評価することと、その種の法則科学的前提無しにプログラムのパフォーマンスを例外や違反の割合で評価することには、本質的な差異がある。

ある記号列のパフォーマンスに対する影響を評価する最も単純な状況は、ある原因系記号の有無と結果の成否をデータ化した表3のような2×2分割表で表現される。この分割表のある記

表 2. 3 薬剤 2 期間クロス・オーバー臨床試験治療時期優劣比較結果(広津・椿, 1984).

	AB	AC	CB	BC	CA	AC
記号 1 勝	3	7	10	4	8	4
記号 2 勝	9	9	12	13	13	10
引き分け	8	2	4	4	2	5

表 3. シンボル評価の最も単純な例.

		結果の成否		
		結果+	結果-	行合計
原因の有無	シンボル有	N_{++}	N_{+-}	N_{+}
	シンボル無	N_{-+}	N_{--}	N_{-}
	列合計	N_{+}	N_{-}	$N_{..}$

表 4. 例外のないシンボル性プログラム.

		結果の成否		
		結果+	結果-	行合計
原因の有無	シンボル有	N_{+}	0	N_{+}
	シンボル無	0	N_{-}	N_{-}
	列合計	N_{+}	N_{-}	$N_{..}$

号の有無, 結果の成否に任意の数量化を行い決定係数の二乗(相関係数)を求めたのが, 2×2 分割表の関連性尺度である, Pearson の ϕ 係数の二乗となる. 更にその $N_{..}$ 倍が, 分割表の独立性に関する Pearson カイ二乗適合度検定統計量である. Pearson は, 離散変数の世界でも連続量の世界同様, 相関分析で背後の法則を性能評価できると考えていた. 実際, メンデルの法則のような客観科学的離散法則を検討するには, 統計モデルによる法則評価は有効であった.

一方, 表 3 で入力記号のプログラムとしての有効性を評価するならば, プログラムが例外なく意図通り動作した割合を評価することになる. 例外の無い理想状態とは表 4 に示したような状態である. 確かに表 4 は ϕ 係数が 1 になる場合である. しかし, 独立性を統計的に評価するのではなく, 理想状態からのずれを測定するのがプログラム評価である. 論理的には表 4 の第 1 行が成り立つことは記号の存在が, 結果成立の十分条件, 第 2 行が成立するのが必要条件である. その種の論理がどの程度成り立っているかを定性的・論理的研究では重視してきたことは Goertz and Mahoney (2012) が, 強調した通りである. 表 3 で, 十分性に関するプログラムの性能は N_{++}/N_{+} , 必要性に関する性能は N_{--}/N_{-} で評価すればよい. 偽陽性率, 偽陰性率

といった検査性能評価が、医学では行われてきた。パターン認識の閾値評価にも ROC 曲線を用いることも多いが、これがパフォーマンスの背後にある法則を評価したものではなく、プログラムのパフォーマンスの論理性能評価と考えられる。

実際、Goertz and Mahoney (2012)は、科学的接近を統計的な量的方法と数学的・論理的な質的方法に峻別した。彼らが指摘したように、社会科学でも統計的接近が徐々に重視されてきている。しかし、筆者は、離散的シンボルやシグナルが本質的な秩序であるプログラム科学では質的方法が、法則科学では量的方法が適しているという仮説を持っている。4章では、筆者の限られた統計的体験の中で、統計的方法周辺に存在する量と質との問題の差異が、鮮明に現れた状況を紹介する。

4. 人間の生存に必要なだった離散化とそれに対峙する統計的方法

4.1 離散的認識における量的情報の捨象—離散的認識の人為性

人間の認識する事実は、時間的・空間的に何らかのバンドパスフィルタリングが施されている。実際、人間は極めて微小な領域の事実は認識できない。椿他 (1982)は、視覚も帯域上のバンドパス特性を持ち、それが錯視現象の原因と考えた。生物全般が生存に適した進化の中で、適切な情報の強調や不要な情報を捨象するための帯域特性を保有するようになったと推察する。このことが所定の主要特性を共有する集合の同一視と、その要素の数え上げという人間あるいは生物特有の操作を導いている。プログラム科学の背景にある離散性は、人工物としての数学に基づく認識と強く関連する。

数学が最初に扱うのは、自然数を用いた数え上げである。統計教育でもヒストグラムなどの度数分布の表示は第一歩である。更に理論度数分布としてのポアソン分布は、小数の法則として事故件数の従う分布として多用されている。

しかし、数え上げが可能なのは、集合としての要素が数学的に定義されているからである。リンゴが3個、ミカンが2個あったとき、果物が何個あるかという問いと、ミカンが何個あるかという問いとでは答えは異なる。約100gのミカンは何個あるかと問われれば、人間は $100 \pm \varepsilon$ gのミカンを同一視して数え上げを行う。事故であるか否かの判断も人間が事故の範囲を定めていることによって可能となる。数え上げ操作は、包含関係を人間が定義し、有限集合論的認識を支援するための重要な人為的操作である。

しかし、法則科学的観点からすれば、離散確率分布も、連続確率分布の極限分布として表現可能なことに留意すべきである。統計学で人間が量的現象を捨象し、数え上げの対象にしていることを端的に示すのが、ポアソン分布には「散らばり母数 (Dispersion Parameter)」が存在しないとの誤解である。Morris (1982)が、自然指数型分布族で分散関数が期待値 μ の1次式となるのは、離散分布であるポアソン分布に限ることを証明したことも誤解を助長させている。もちろん、これは非負整数値分布を前提にした証明である。

実際に、ポアソン変量の σ 倍の確率変数は、 $0, \sigma, 2\sigma, \dots$ といった離散的値をとり、分散関数 $V(\mu) = \sigma^2 \mu$ となる指数型分布族である。そのような状況を量的観測ではなく、量を捨象して数え上げとするのが人間の判断としては経済的なことは否定しない。

しかし法則科学にとって重要なのは、この散らばり母数を有するポアソン分布は、散らばり母数を持つ指数型連続分布族の極限として理想化された分布だという認識を持つことである。すなわち、分散関数 $V(\mu) = \sigma^2 \mu^p$ で特徴づけられるガンマ複合ポアソン分布($1 \leq p \leq 2$)で非負連続値分布である Tweedie 分布の極限である (Jørgensen, 1987)。

Tweedie 分布は、同一分布に従うガンマ変量を散らばり母数の無いポアソン変数個を加える、量の加法操作から得られる。この量的操作結果の観測が、より法則科学に近い認識行為であ

る。ポアソン変量とは、分散 0 のガンマ変量を確率的個数で加え合わせているから離散変量となっているだけである。

非負整数値をとる 2 変量ポアソン分布に相関係数概念を導入することは、通常、擬似尤度あるいは一般化推定方程式でしかできない。同一のポアソン変数個、別々のガンマ変量を加えた 2 変量 Tweedie 分布は、実在する指数型分布族であり、母相関係数がパラメータとして存在し、量的関連性を記述することができる確率モデルとなっている (Iwasaki and Tsubaki, 2005)。

本来、統計モデルによる法則科学的認識は、主として人間の周辺に内在する量的性質に関わる方法として開発されてきたことを意識すべきである。これは、人間の認識の特性・捨象・価値観・論理に依存しない法則を追究する営みであった。

4.2 論理的測定モデルと統計的測定モデル—実在する概念とそうでない概念

Bollen (1989) は、統計的測定モデルの第 1 段階としての概念の理論的定義を挙げた。その好例として、CIA (1982) の「テロリズム」の定義が示されている。当時の CIA の定義が、「暴力と関連する行為」かつ「政治的意図を持つ行為」かつ「直接の被害者を越えた範囲に影響を与えることを意図した行為」という 3 特性を有することが、テロリズム事象を定義する必要十分条件であることが明確になっているからである。この種の考察は用語学 (Terminology) から導かれる離散的かつ集合論的なものである。ISO 1087-1:2000 は、国際標準化活動における用語の定義に関わる理論と方法を標準化したものである (ISO, 2000)。実際、そこでは、概念は特性群の特定の結合から構成される知識、内包 (Intension) は、概念を形成する特性の集合として定義されている。従って、CIA (1982) の理論的定義から Bollen (1989) が特性を列挙した方法で、概念の成立は複数特性を共有する必要十分条件として定義され、定義された概念に対象が属するか否かは論理的に定まる。概念に対応する英語なり日本語等の用語が存在しても、概念の定義は数理的・論理的に確定する。しかし、唯名論の立場をとらずとも、テロリズムといった概念は、物質科学の世界には実在せず、人間が認識の便宜のために対象を分類した数学的概念にすぎない。

それにも関わらず、Bollen (1989) は、用語学的概念定義の次の段階として、統計的測定モデル、例えば因子モデルに関連する事実には当てはめる手続きを与えている。このとき特性と概念との関係は量的法則モデルで記述される。例えば、テロリズムという論理的概念の背景に、テロらしさといった潜在標準正規確率変数 X を想定し、暴力行為らしさ Y_1 、政治的行為らしさ Y_2 、範囲の広範性らしさ Y_3 を目的とした行為らしさという特性をアナログスケールなどで近似的に連続変数と考えられる尺度で観測することで、連立方程式モデルである因子モデル

$$(4.1) \quad Y_j = \alpha_j + \beta_j X + \varepsilon_j, j = 1, \dots, 3, \varepsilon_j \sim N(0, \sigma_j^2)$$

を当てはめ、潜在因子 X の測定方法が導かれるのである。

確かに、この種の考え方は、背後にある潜在因子が実在する概念の場合には有効である。Spearman (1904) は、一般知能因子が存在し、それが所定の試験項目に正の影響を与えると考えた。実在する潜在変数を客観測定するために、適切な観測項目が設計されることは、計量心理学分野だけでなく物理測定分野ですら行われてきた法則科学的営みである (例えば、牧他, 2004)。

Bollen (1989) の理論的用語定義は、論理学・用語学的概念測定にとって重要だが、この唯名論的概念の測定に、法則科学的統計モデル (4.1) を適用するには、次の問題がある。統計的測定モデル (4.1) の下では、全くテロらしさが無い行為、すなわち連続変数として X がマイナス無限大になったとき、暴力行為・政治的行為・行為が影響する範囲もマイナス無限大になる。すなわち、テロらしさが全くなければ、暴力行為らしさも全くなくなることを意味する。しか

し、全くテロらしくない行為であっても政治家の演説のように政治的行為らしさをもつ事象は存在する。

ここでも、Goertz and Mahoney (2012)が指摘したように、連続的な法則に基づく統計的推論と、必要条件や十分条件といった論理的推論には乖離が生じている。

以下では、具体的に論理的推論モデルと統計的離散測定モデルとの乖離を例示したい。そのために、ある母集団に属する人が、ある事象を q 個の特性から定義される概念に適合すると判断する確率を P_0 、その事象が特性 j ($j = 1, \dots, q$) を有すると判断する確率を P_j とし、どのような確率間の構造を考えるのが妥当かを調べる。このとき、次の3条件を要求するのが論理的条件と整合的である。

条件1: $P_0 = f(P_1, P_2, \dots, P_q)$ は、 P_j について単調増加関数である。

条件2: ある特性 j について $P_j = 0$ ならば、 $P_0 = 0$ となる。

条件3: $P_0 = 1$ ならば、全ての j について $P_j = 1$ となる。

条件2は、ある特性の完全な欠落は概念の完全に不成立となることを意味する。条件3は、完全な概念の成立は、全特性が完全に成立することを意味する。注意が必要なのはある特性 j について $P_j = 1$ となっても、必ずしも P_0 は1とはならないことである。

この種の条件を満たす関数形 f は、一意に定まるわけではない。例えば条件3は、

$$(4.2) \quad \log(1 - P_0) = \alpha_j + \beta_j \log(1 - P_j), \quad \beta_j > 0$$

が満たされれば成立する。(4.2)は、 $\log(1 - P_j) = \alpha_j^* + \beta_j^* \log(1 - P_0)$ と書き換えれば、あまり自然ではないが、 P_0 の統計的測定モデルと見なせる。一方、条件2は、例えば

$$(4.3) \quad \log P_0 = \gamma + \delta \sum_{j=1}^q \log P_j, \quad \delta > 0$$

が満たされれば成立する。そこで、離散統計モデルとしては全く異形だが

$$(4.4) \quad \log(1 - P_0) = \alpha_j + \beta_j \log(1 - P_j) + \delta \sum_{k=1}^q \log P_k \quad \beta_j > 0, \quad \delta > 0 \quad j = 1, \dots, q$$

というモデルを考えると、条件1から3は満たされる。これを概念が原因で特性が結果とする測定モデル(4.1)に類似させて書き直せば、

$$(4.5) \quad \log(1 - P_j) = \alpha_j^* + \beta_j^* \log(1 - P_0) + \delta^* \sum_{k=1}^q \log P_k$$

となる。(4.5)は右辺第3項の存在のため、統計的測定モデルとはならない。(4.5)式左辺並びに右辺2項をロジットに置き換えることが統計的には自然だが、これを入れると $P_0 = 0$ ならば、全ての特性 j について $P_j = 0$ を要求することになり、論理的な条件を不必要に追加したことになる。

このように論理的測定を記述する確率モデルの定式化は一意ではないが、いずれにしても統計的測定モデルとは本質的に異なる。従って、論理的概念測定と統計的概念測定は異なる行為と考えなければならない。

4.3 論理的分類と統計的分配

ある特性の有無に関する認識を統計的に扱うために、4.2節ではその特性が有るか否かの回答に曖昧さが無く、各個人については True か False が定まることを前提にした。このとき、有

限混合分布といった頻度論的確率分布が用いられる。しかし、各個人が True なり False と事象を判断することは、人為的論理操作と考えられる場合が多い。

公的統計分野では、収集したデータを集計するために、家計統計では家計簿に記述された情報から商品分類、事業所統計では調査票情報から事業所の産業分類を行わなければならない。しかし、家計統計では、被調査者が 200 円支払った商品が、発泡酒なのかビールなのかは、商品の定義が明確かつ網羅性があるのならば、統計的教師付き分類問題として明確に定式化できる。一方、調査対象事業所の売り上げが 1500 万円として、その事業所は製造業としての売り上げもあれば、サービス業としての売り上げもあるという状況がある。このとき事業所の産業分類は、売り上げを各業種に按分して按分率の高い業種に形式的に分類を行うこととなる。この按分率は、論理的分類を前提とする有限混合分布の混合確率ではない。背後にあるのは連続量の分配である。

このように人間が介在する分類問題の背後には様々な曖昧さがあるため、Goertz and Mahoney (2012) が、示唆するようにテロらしさ、暴力行為らしさといった曖昧性をファジイ集合論で表現しようという立場も生じる。

しかし、事象自体が頻度論的かつ論理的に確定するのではなく、法則科学の実態を持つ量が加法的に観測されていることを統計モデル化することは可能である。椿 (1999), Tomosada and Tsubaki (2011) は、画像処理の Mixed Pixel 問題が、分類問題ではなく按分問題(粒子混合)であるという立場の統計モデルを提案した。以下にその概要を示す。

p 次元観測変量 \mathbf{Y} は、 K 個の細分された変量の和であるとする。按分対象となるカテゴリは Q 個あり、細分された変量は、あるカテゴリ q に属し、 K_q 個存在すると仮定する。このとき、 \mathbf{Y} は、次のように量の加法操作で表現される。

$$\mathbf{Y} = \sum_{q=1}^Q \sum_{k=1}^{K_q} \mathbf{X}_{qk}$$

但し、 \mathbf{X}_{qk} は、カテゴリ q に属する細分変量を意味する。さらに、 \mathbf{X}_{qk} が互いに独立に、モーメント母関数 $\phi_q(\mathbf{t}) = E[\exp(\mathbf{t}^T \mathbf{X}_{qk})]$ を持つ同一分布に従うと仮定する。このとき、カテゴリ q のデータだけからなる観測変量 \mathbf{Y}_q のモーメント母関数は $\psi_q(\mathbf{t}) = \phi_q(\mathbf{t})^{1/K}$ となる。

従って、 $p_q = K_q/K$ と置けば、観測変量 \mathbf{Y} の従う分布のモーメント母関数は(4.6)のようになる。

$$(4.6) \quad \psi_{\mathbf{Y}}(\mathbf{t}) = \prod_{q=1}^Q \phi_q(\mathbf{t})^{p_q}$$

従って、そのキュムラント母関数は、純粹観測ベクトルの従う分布のキュムラント母関数が、比率 p_q で荷重和されたものになる。通常の頻度論的混合では、密度関数あるいはモーメント母関数が混合確率で荷重和されるのと、期待値構造は一致するが、2 次のモーメント構造は異なる。

このように按分的混合は、量の加法操作を背景機序として前提にする、論理的分類ではなく量の分配率の推定問題である。ファジイ理論が扱うような人間の分類行為には、論理学と法則科学との中間的状況も存在すると考えられる。そのような場合への一つの対処法として、確率変数 X_i のモーメント母関数 $\varphi_i(\theta) = E[\exp(\theta X_i)]$, $i = 1, \dots, I$ に対して、一般化キュムラント母関数を

$$(4.7) \quad \eta_i(\theta, t) = \{\varphi_i(\theta)^t - 1\}/t$$

と定義することが考えられる。このとき、

$$\begin{aligned}\eta_i(\theta, t) &= \left[\left\{ 1 + \theta E[X_i] + \frac{\theta^2}{2} E[X_i^2] + \dots \right\}^t - 1 \right] / t \\ &= \theta E[X_i] + \frac{\theta^2}{2} E[X_i^2] + \dots + \theta^2 (t-1) / 2 \left\{ E[X_i] + \frac{\theta}{2} E[X_i^2] + \dots \right\}^2 + \dots \\ &= \theta E[X_i] + \frac{\theta^2}{2} \{ E[X_i^2] + (t-1) E[X_i]^2 \} + o(\theta^2)\end{aligned}$$

と展開されるので、一次一般化キュムラントは、 $E[X_i]$ 、二次一般化キュムラントは、 $E[X_i^2] + (t-1)E[X_i]^2$ となる。

この一般化キュムラント母関数の混合を $\eta(\theta, t) = \sum_{i=1}^I p_i \eta_i(\theta, t)$ を確率モデルとして採用すれば、頻度論的(論理的)分類と連続的分配とを一つのパラメータで繋ぐことができる。なお、混合後の一般化キュムラントは、一次一般化キュムラントが $c_1 = \sum_{i=1}^I p_i E[X_i]$ 、二次一般化キュムラントが $c_2 = \sum_{i=1}^I p_i \{ t E[X_i]^2 + \text{Var}[X_i] \}$ となる。

量の加法的操作は法則科学を記述する方法論であり、4.1節で論じたように、ポアソン分布をガンマ分布も含まれる Tweedie 連続分布の極限として捉えれば、量の按分率推定問題がベータ分布やディリクレ分布と同等の意味を持つ、二項分布や多項分布を極限として導いていることが分かる。

統計的分類問題、パターン認識問題は、論理的推論を背景に持つ、プログラム科学的方法論である。しかし、逆に法則科学的の立場から、分類問題と分配問題との差異に留意する必要がある。

5. おわりに：プログラム科学の文法を目指して

5.1 法則科学的検討を進めるべきプログラム科学的問題

一見、プログラム秩序、離散的シンボルで表記されている現象であっても、背後に連続量が存在し、閾値を設定した人為的切断でシンボルが形成されている場合が多い。このような量的情報を論理的に扱う必然性はない。第一段階で法則科学的洞察を行い、第2段階でそこに閾値を入れたときのパフォーマンス評価を行えば良い。

一方で、この種の人為的切断がない、純粋な記号と量的パフォーマンスの関係は法則科学的には、どのように考えれば良いのだろうか。これは質と量との混合問題である。これが、端的に現れるのが遺伝子情報といった純粋な記号列が、物質科学的法則に対してシグナルとして作用する生物学である。

先ず量的科学として、関心のある量に影響を与える量を網羅し、記号が関心のある量に直接影響を与える可能性だけでなく、関心のある量に直接影響する量に影響を与えている可能性をも探求することは最低限必要である。ただし、これは法則科学的研究である。

一方、 M 個の純粋記号からなるシグナル記号列が、量にどのような影響を与えるかは、論理的・演繹的考察も重要である。実際、シグナル記号列の量への影響は論理のパズル、鍵をあけるような組み合わせ最適化を必要とする。しかし、統計的方法が開錠を効率化する可能性は大きい。それが、3章で紹介した M 水準の質的要因からなる実験計画法で問題を定式化することである。ただし、連続要因の場合と異なり、高次の交互作用効果が、鍵を解くのに重要となることは容易に予想される。従って、高次交互作用項(論理型入力変数の積の項)も考慮したモデルをデータに当てはめることが必要となる。

この種のモデル同定を限られた標本から効率的に行う方法は、少なくとも2つ現存すると考えている。一つはカーネル法の適用により高次交互作用をモデルに投入することである。もう

一つは尺度化と非線形性に関する射影追跡回帰 (Friedman et al., 1981)により, 尺度間交互作用の消去を試みることである. 前者は, サポートベクターマシン, 後者はニューラルネットワークなどの統計的機械学習技術として, 今や日常的に用いられるようになった. この種のセミパラメトリックモデルを効率的に同定するデータ収集技法としての逐次実験計画法も強化学習と呼ばれるようになった. 実験計画法の配置理論にとって, 重要な研究分野であるが, その最適計画については筆者には専門性はない.

5.2 純粹プログラム科学とその文法

人間社会を動かす, 法や政策は吉田のシンボル性プログラムであり, その成果も経済的効果のような量に変換されない, 社会構成員の感性的反応となることも多い. この種の純粹プログラム科学に近い問題については, ロジックツリーと MECE (Mutually Exclusive and Collectively Exhaustive) と呼ばれる本来純粹数学的接近を適用することが学術のみならず, 社会全般に広がっている. 日本の EBPM (Evidence Based Policy Making) 推進においても, 統計的方法と共に大きな位置づけを占めるようになってきている (行政改革推進本部, 2017). この種の方法論が統計科学の文法とは一線を画するものであることが, 本論の主張だった. しかし, 品質経営分野で統計的問題解決の標準シナリオ以外に, 論理的方法が主眼となる「課題達成型 QC ストーリー」と呼ばれる標準シナリオが 1980 年代から普及していることを再度強調しておきたい (狩野, 1993).

- 1) テーマの選定
- 2) 課題の明確化と目標設定
- 3) 方策の立案・選定
- 4) 成功シナリオの追究と実施
- 5) 効果の確認
- 6) 歯止め

これを支える論理的技法群 (ロジックツリーをも含む新 QC 七つ道具など) も整備されていた. 仮に統計的方法がその有効性を発揮するとすれば, プログラムの質的パフォーマンスを評価するシンボル列に関する If-Then 論理の発見であり, これは Breiman et al. (1984) の CART, すなわち自動層別技術の適用である. 今日的には, 層別原理をアンサンブル学習で高性能化する Breiman (2001) の Random Forest も有力な記号要因解析技法となろう.

一方で, プログラム科学が科学として成立する論理条件は, 科学的プロセスとしての Pearson の科学の文法, あるいは Shewhart, Deming, 石川の PDCA サイクルとの整合性である. なぜならば, この種の実証的プロセスが, 論理的方法論ですら自律的進化と再現性の確保のために重要だからである. 例えば,

- 1) あるべきプログラム・パフォーマンスの明確な定義
- 2) プログラム・パフォーマンスにおける看過できない違反や例外の発見
- 3) パフォーマンスに影響をあたえる可能性のある記号列の論理的網羅
- 4) 記号列の必要十分性の観点からの比較・評価
- 5) 最適な記号列プログラムの実装
- 6) プログラムの実施とそのパフォーマンスのモニタリング

といった課題達成型 QC ストーリーと PDCA サイクルを結合したプロセスモデルが, プログラム科学の文法として有効な可能性はある. しかし, それも論理科学的方法の実践の中で実証されるべき仮説にすぎない. 統計数理科学が将来 2 つの科学的方法のスコープマネジメントを

適切に実現できる日がくることを期待して、本論を閉じたい。

査読者から示された論点と謝辞

本ノートに対しては、査読者より本質的な3つの論点を提示いただいている。第一は、プログラム科学と法則科学とが、科学の特徴づけというよりは、外界との関係性の違いであるとの指摘である。自身の論理ないしはモデルを変更しつづける、すなわち、外界からのフィードバックが不可欠か否かということではないかという見解である。第2は、プログラム科学においても、現実ないしは観察から独立ではなく、出発点は帰納的観察に基づいているのではないかという指摘である。この2つの指摘は、行動モデルの観点からは明らかに正しく、筆者も同意する。生物の進化ですら、外界とのフィードバックがトリガーで起きていることに間違いはない。また、プログラム科学が支配的とされる生物科学、社会科学においても、法則はその基盤的秩序の一部を担っていることは、吉田の学術の体系でも前提になっている。筆者自体の疑問は、法則を認識し、それに基づいて何らかの設計に反映する統計科学的方法論と、プログラムを設計する方法論とに、科学的探究プロセスが少し異なるのではないかということである。もちろん、後者に対しても帰納的接近が有用であることについては、何ら疑いを持ってはいない。

第3の論点は、一般化キュムラント関数のように、頻度論的分類と連続的分配とを接続すること自体が、プログラム科学と法則科学との区別の必要性がないことを意味しないかという疑問である。この指摘は、筆者の視点が統計科学に近すぎる、すなわち法則科学的立場で頻度論的分類を眺めていることに気づかされた。科学的法則と論理的秩序を混合させること自体に実学的意義が有るかについては、査読者から提起された課題として強く意識したい。以上、本ノートに貴重な論点を示唆すると共に、様々な誤りを指摘していただいた査読者に深甚の謝意と敬意を捧げたい。

参 考 文 献

- Bollen, K. A. (1989). *Structural Equations with Latent Variables*, Wiley, Hoboken.
- Box, J. F. (1987). Guinness, Gosset, Fisher, and small samples, *Statistical Science*, **2**(1), 45-52.
- Breiman, L. (2001). Random forests, *Machine Learning*, **45**(1), 5-32.
- Breiman, L., Friedman, J., Stone, C. J. and Olshen, T. A. (1984). *Classification and Regression Trees*, Wadsworth, Belmont.
- Central Intelligence Agency (1982). Patterns of international terrorism: 1981, <https://www.cia.gov/library/readingroom/document/cia-rdp85m00363r000901960025-4>.
- Cochran, W. G., Autrey, K. M. and Cannon, C. Y. (1941). A double change-over design for daily cattle feeding experiments, *Journal of Dairy Science*, **24**(11), 937-951.
- Cox, D. R. and Donnelly, C. A. (2011). *Principles of Applied Statistics*, Cambridge University Press, Cambridge.
- Fieler, E. C. (1940). The biological standardization of insulin, Supplement to *the Journal of the Royal Statistical Society*, **7**(1), 1-64.
- Fisher, R. A. (1935). *The Design of Experiments*, Oliver and Boyd, Edinburgh.
- Friedman, J. H. and Stuetzle, W. (1981). Projection pursuit regression, *Journal of American Statistical Association*, **76**, 817-823.
- Goertz, G. and Mahoney, J. (2012). *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*, Princeton University Press, Princeton. (西川賢, 今井真士 訳 (2015). 『社会学のパラダイム論争：2つの文化の物語』, 勁草書房, 東京.)

- 行政改革推進本部 (2017). 「EBPM の試行的検証」EBPM 推進に向けた取組み・ロジックモデルなどについて, https://www.cas.go.jp/jp/seisaku/gyoukaku/H27_review/H29_fall_open_review/z7.pdf.
- 広津千尋, 椿広計 (1984). クロス・オーバー計画でとられた臨床データの統計モデル, 「実験データの解析の理論的背景」, 京都大学数理解析研究所講究録, **526**, 46-70.
- ISO (2000). *Terminology Work — Vocabulary — Part 1: Theory and Application*, ISO 1087-1:2000, International Organization for Standardization, Geneva.
- ISO (2015). *Application of Statistical and Related Methods to New Technology and Product Development Process — Part 1: General Principles and Perspectives of Quality Function Deployment (QFD)*, ISO 16355-1: 2015, International Organization for Standardization, Geneva.
- Iwasaki, M. and Tsubaki, H. (2005). A new bivariate distribution in natural exponential family, *Metrika*, **62**, 323-336.
- Jørgensen, B. (1987). Exponential dispersion models, *the Journal of the Royal Statistical Society, Series B*, **49**(2), 127-162.
- 狩野紀昭 監修 (1993). 『QC サークルのための課題達成型 QC ストーリー』, 日科技連, 東京.
- 牧雅康, 沖一雄, 椿広計, 安岡善文 (2004). 潜在要因の影響を考慮した水質計測の因果モデリング, 日本リモートセンシング学会誌, **24**(2), 153-162.
- Morris, C. (1982). Natural exponential families with quadratic variance functions, *Annals of Statistics*, **10**(1), 65-80.
- 日本学術会議自動制御研究連絡委員会工学共通基盤研究連絡委員会自動制御学専門委員会 (2005). 横断型基幹科学技術としての制御学の役割—「知の統合」を目指す研究・教育の促進に向けて, 自動制御研究連絡委員会工学共通基盤研究連絡委員会自動制御学専門委員会報告, <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-19-t1031-14.pdf>.
- 日本学術会議運営審議会附置新しい学術体系委員会 (2003). 新しい学術の体系—社会のための学術と文理の融合—, 日本学術会議, <http://www.scj.go.jp/ja/info/kohyo/18pdf/1829.pdf>.
- Pearson, K. (1892). *The Grammar of Science*, Walter Scott, London.
- Shewhart, W.A. (1939). *Statistical Method from the Viewpoint of Quality Control* (eds. W. E. Deming), The Graduate School, the Department of Agriculture, Washington D.C. (坂元平八 訳 (1960). 『品質管理の基礎概念—品質管理の観点からみた統計的方法』, 岩波書店, 東京.)
- Spearman, C. (1904). 'General intelligence,' objectively determined and measured, *The American Journal of Psychology*, **15**(2), 201-292.
- Student (1908). The probable error of a mean, *Biometrika*, **6**(1), 1-25.
- 田口玄一 (2000). 『ロバスト設計のための機能性評価—効率的開発の方法』, 日本規格協会, 東京.
- 竹内啓 (2010). 『偶然とは何か—その積極的意味』, 岩波新書 1269, 岩波書店, 東京.
- Tomosada, M. and Tsubaki, H. (2011). Novel mixture model for mixed pixel classification of multi-spectral image data, *International Journal of Mathematical Models and Methods in Applied Science*, **5**(1), 175-183.
- 椿広計 (1999). リモートセンシングにおける分類 (2) ミクセルの識別に関する統計的考察, 日本リモートセンシング学会誌, **19**(1), 75-80.
- 椿広計 (2017). サービスと品質管理, 品質管理学会等の活動を振り返る, 標準化と品質管理, **70**(10), 2-8.
- 椿広計 (2018). データ駆動型社会の人と品と質とのマネジメント, 応用統計学, **47**(2&3), 89-98.
- 椿広計 (2019). 第 1 部日本社会における統計科学の展開, 第 2 章日本の品質管理活動と統計科学 (国友直人, 山本拓 編), 『統計と日本社会 データサイエンス時代の展開』, 東京大学出版会, 東京.
- 椿広計, 河村敏彦 (2008). 『設計科学におけるタグチメソッド—パラメータ設計の体系化と新たな SN 比解析』, 日科技連, 東京.
- 椿広計, 鈴木守, 花泉弘, 藤村貞夫 (1982). 視覚の帯域通過特性に基づく錯視現象の解析, 計測自動制御学会, 日本機械学会, 日本自動制御協会第 25 回自走制御講演会要旨集, 509-510.

- Tsubaki, H., Nishina, K. and Yamada, S. (eds.) (2008). *The Grammar of Technology Development*, Springer, Tokyo.
- 土橋俊人, 高須久, 椿広計 (1985). どう解析するかこのデータ, どう収集するかこの言語情報, 品質, **15**(3), 56-66.
- VCP-Net 研究会 編著 (2014). 『知の巡りをよくする手法の連携活用—サービス・製品の価値を高める価値創生プロセスのデザイン』(大藤正, 黒河英俊 監修), 日本規格協会, 東京.
- 吉田民人 (2013). 『近代科学の情報論的転回 プログラム科学論』(吉田民人論集編集委員会 編), 勁草書房, 東京.

Grammar of Statistical Science and Logical Reasoning

Hiroe Tsubaki

The Institute of Statistical Mathematics

Although the standardized research process of science should be a statistical and quantitative exploration process in order to give appropriate laws to the Nature or societies, the main research process of science with program as an artificial and weak order may be logical and qualitative, as suggested by Goertz and Mahoney (2012). From a statistical scientific point of view, I illustrate some situations in which the difference between quantitative and logical processes is noticeable. I also discuss the possibility of unifying the two scientific approaches along the Grammar of Science by Pearson (1892).