

# 滋賀大学データサイエンス学部 佐藤健一教授 の学術的研究および教育指導に関する包括 的調査報告書

## 1. 序論：データサイエンスにおける理論と実践の融合

現代のデータサイエンス教育および研究において、高度な数学的理論に裏打ちされた統計モデルの開発と、それを医学・社会科学といった複雑な実領域に応用する実践力との結合は、極めて重要な課題である。日本初のデータサイエンス学部を設置した滋賀大学において、佐藤健一 (Kenichi Satoh) 教授はこの「理論」と「実践」の交差点に位置する核心的な研究者の一人である。本報告書は、佐藤教授のこれまでの研究軌跡、独自に開発した統計手法、多岐にわたる応用研究、そして教育者としての指導方針について、公開された学術文献、研究データベース、およびソフトウェアリポジトリに基づき、包括的かつ詳細に調査・分析したものである。

佐藤教授の研究は、数理統計学、特に多変量解析を基礎としつつ、非負値行列因子分解 (NMF) の拡張モデルという独自の理論体系を構築している点に最大の特徴がある。同時に、その理論を武器として、認知症患者の予後予測、COVID-19パンデミック下の行動変容、福島原発事故後のリスクコミュニケーション、さらには文学作品の計量分析に至るまで、驚くべき広範さで社会課題の解決に取り組んでいる。

本報告書では、佐藤教授の研究活動を「方法論的革新 (Methodological Innovation)」と「領域横断的応用 (Interdisciplinary Application)」の二軸で捉え、その詳細を紐解く。また、これらの分析を通じて、同教授が主宰する研究室 (ゼミナール) において学生がどのような研究に従事し、どのようなスキルセットを獲得し得るかについて、具体的かつ客観的な予測を提供する。

## 2. 学術的キャリアの軌跡と専門性の形成

佐藤教授の現在の研究スタイルを理解するためには、そのキャリアがいかにして「純粋数学」から「計量生物学」、そして「データサイエンス」へと展開してきたかを時系列に沿って分析する必要がある。このキャリアの変遷は、データサイエンスという学問領域自体の進化とも軌を一にしている。

### 2.1 数理統計学の基礎形成：広島大学時代

佐藤教授の学術的基盤は、広島大学大学院理学研究科 (数学専攻) における厳密な数理統計学の訓練にある。1996年に修士課程を修了し、1999年に博士 (理学) を取得した<sup>1</sup>。指導教官は、多変量解析における世界的権威である藤越康祝教授であった。

この時期の研究は、統計モデルの選択基準、特に「赤池情報量基準 (AIC)」や「Cp基準」の多変量回帰モデルへの拡張に関するものであった。1997年に英国の統計学誌『Biometrika』に掲載された論文 "Modified AIC and Cp in Multivariate Linear Regression"<sup>2</sup> は、モデルの適合度と複雑さのバランスを評価するための新たな指標を提案したものであり、佐藤教授が研究の初期段階から「モ

デル選択の最適化」という視点を持っていたことを示している。これは、現在の機械学習における「過学習の回避」や「正則化」に通じる本質的なテーマである。

## 2.2 計量生物学と国際的研究機関での研鑽

博士号取得後、佐藤教授は広島大学原爆放射能医学研究所(現・原爆放射線医科学研究所)に所属し、助手、講師、准教授として約20年間にわたり医学・生物統計学の実践に従事した<sup>1</sup>。この期間における経験が、現在の「データサイエンス」への志向を決定づけたと言える。

特筆すべきは、以下の海外研究機関での客員研究員としての活動である。

- 英国 **Medical Research Council (MRC)** 放射線ゲノム安定性部門 (2001年)<sup>1</sup>: MRCは英国の医学研究の中核機関であり、ここではゲノムデータ等の高次元バイオデータの解析に関与したと推測される。生物学的データはノイズが多く、変数間の複雑な相関構造を持つため、従来の線形モデルでは捉えきれない現象が多い。ここでの経験が、後の非線形モデルやカーネル法の導入への動機付けになった可能性が高い。
- 豪州 ニューサウスウェールズ大学 (UNSW) 統計学科 (2006-2007)<sup>1</sup>: UNSWは統計学、特に海洋学や気象学を含む応用数学の分野で世界的な評価を受けている。ここで佐藤教授は、より広範なデータモデリングの視点を獲得したと考えられる。

## 2.3 滋賀大学データサイエンス学部への着任と現在

2019年、佐藤教授は滋賀大学データサイエンス教育研究センター教授、2021年にはデータサイエンス学系教授および副学部長に就任した<sup>1</sup>。これまでの医学・生物統計の枠を超え、経済、教育、人文科学などあらゆるデータを対象とする「データサイエンス」の旗手として、理論構築と社会実装の両輪を推進している。

---

# 3. 中核的研究: 拡張NMFモデルの理論と実装

佐藤教授の研究業績において最も独創的かつ中心的な位置を占めるのが、非負値行列因子分解 (**NMF: Non-negative Matrix Factorization**) に対する理論的拡張と、その実装であるRパッケージ `nmfkc` の開発である。これは、既存の統計手法では解決困難であった「高次元性」「非線形性」「解釈可能性」のトレードオフを解消しようとする野心的な試みである。

## 3.1 非負値行列因子分解 (NMF) の基本原理と限界

NMFは、すべての要素が非負(0以上)であるデータ行列  $Y$  を、二つの非負行列(基底行列  $W$  と係数行列  $H$ )の積として近似する手法である( $Y \approx WH$ )。

- 利点: 主成分分析(PCA)などと異なり、「負の値」を含まないため、抽出された特徴(基底)が「部分の足し合わせ」として表現され、人間にとって解釈しやすい。例えば、画像データであれば「目」や「鼻」といったパーツに、文書データであれば「トピック」に分解される。
- 従来の限界:

1. 共変量の統合困難: 観測データ以外の外部情報(年齢、性別、気温、経済指標などの共変量)をモデルに組み込む統一的な枠組みが乏しかった。
2. 線形性の制約: データ構造が複雑な場合、単純な行列積(線形結合)では表現力が不足する。
3. 時系列構造の欠如: 時間的な順序や依存関係(自己相関)を考慮する機能が標準的なNMFには備わっていなかった。

### 3.2 独自開発手法: nmfkc (NMF with Kernel Covariates) の革新性

佐藤教授が開発・公開しているRパッケージ nmfkc<sup>3</sup> は、これらの限界を突破するために設計された包括的な統計解析フレームワークである。その核心は、カーネル法(Kernel Methods)と共変量(Covariates)の融合にある。

以下の表は、標準的なNMFと、佐藤教授が開発した拡張NMF(nmfkc)の機能比較である。

機能・特性	標準的なNMF	佐藤教授の拡張NMF (nmfkc)
モデル構造	$Y \approx$ (2因子分解)	$Y \approx$ (3因子分解 / Tri-factorization)
外部情報の利用	基本的に不可	共変量行列 $A$ を明示的にモデル化可能
非線形性の表現	困難(線形近似のみ)	カーネル法(ガウスカーネル等)により非線形構造を捕捉
時系列解析	各時点を独立と仮定	NMF-VAR(ベクトル自己回帰)による動的モデリング
因果推論	不可	NMF-SEM(構造方程式モデリング)による因果構造推定
ランク選択	手動設定が多い	クロスバリデーション、エルボー法等の診断ツールを実装

#### 3.2.1 理論的枠組み: Tri-Factorizationとカーネル共変量

佐藤教授のモデルでは、観測行列  $Y$  を  $Y \approx XB$  と分解する際、係数行列  $B$  をさらに  $B \approx CA$  と分解する構造をとる<sup>3</sup>。ここで、 $A$  は既知の共変量行列、 $X$  は未知の基底行列、 $C$  は共変量と基底をつなぐパラメータ行列である。さらに、共変量  $A$  に対してカーネル関数 (Kernel Function) を適用することで、共変量と観測データの間の非線形な関係をモデル化することに成功している。

### 3.2.2 時系列データへの応用: NMF-VAR

特筆すべき成果の一つが、**NMF-VAR (NMF with Vector Autoregression)** である<sup>2</sup>。これは、多変量時系列データの解析において画期的な手法である。通常、変数が多数ある時系列データに対してベクトル自己回帰 (VAR) モデルを適用すると、推定パラメータ数が変数の二乗に比例して増大し、推定が不安定になる (次元の呪い)。NMF-VAR では、まず NMF を用いてデータを少数の「潜在基底」に圧縮し、その基底の時系列変化に対して VAR モデルを適用する。

- 応用事例: カナダの経済指標、国際航空旅客数、日本の COVID-19 感染者数データ等の解析において、この手法が景気循環や季節トレンド、流行の波を効果的に抽出しつつ、将来予測を行えることが実証されている<sup>4</sup>。

### 3.2.3 構造方程式モデリングへの展開: NMF-SEM

さらに、**NMF-SEM** という手法も提案されている<sup>3</sup>。これは、NMF で抽出された潜在変数間の因果関係を、構造方程式モデリング (SEM) の枠組みで推定するものである。

- ブラインド入出力解析: システムへの入力と出力の関係から、システム内部のフィードバック構造や外部からの影響を推定する。
- 平衡解釈 (**Equilibrium Interpretation**): 動的なシステムが定常状態に達したと仮定した際の構造を明らかにすることで、複雑な現象の「骨格」を理解することを可能にする。

### 3.2.4 教師あり分類: NMF-LAB

分類問題に対しては、ラベル情報を共変量として扱う **NMF-LAB** が用意されている<sup>3</sup>。これにより、データの次元圧縮 (特徴抽出) とクラス分類を同時に最適化することが可能となる。

---

## 4. 応用研究領域: データサイエンスによる社会課題の解明

佐藤教授の研究活動のもう一つの柱は、開発した手法や高度な統計知識を、医学、公衆衛生、災害心理学、人文科学といった具体的なフィールドに応用することである。ここでは、主要な応用研究事例を詳細に分析する。

### 4.1 医学・公衆衛生学領域: エビデンスの創出

滋賀大学着任前後を通じて、医学研究者との共同研究は佐藤教授の活動の根幹をなしている。

#### 4.1.1 認知症・MCI患者の予後と意欲の関連解析

2024年に『Geriatrics and Gerontology International』に発表された研究<sup>4</sup>は、高齢化社会における重要な課題に取り組んだものである。

- 研究目的: 軽度認知障害 (MCI) および認知症患者において、「意欲 (Vitality)」の低下が死亡リスクにどのような影響を与えるかを解明すること。
- データと手法: 3,731名の患者データを約3.6年間にわたり追跡調査した大規模レトロスペクティブ研究。Vitality Index (Vix) を指標として用いた。
- 分析結果: 意欲が中等度～重度に低下している患者群では、死亡率が有意に高いことが判明した。さらに重要な発見として、この関連性が BMI (体格指数)、IADL (手段的日常生活動作)、BADL (基本的日常生活動作) によって媒介 (mediate) されていることを統計的に明らかにした。
- 意義: これは、認知症ケアにおいて「意欲の向上」を目指す介入が、身体機能や栄養状態の改善を通じて生存期間の延長につながる可能性を示唆するエビデンスであり、臨床現場へのインパクトが大きい。

#### 4.1.2 COVID-19パンデミック下の行動変容とBPSD

2022年の研究<sup>4</sup>では、COVID-19による社会的隔離が認知症患者の行動・心理症状 (BPSD) に与える影響を調査した。

- 知見: デイケアサービスの利用自粛や中断が、患者のBPSD悪化と強く関連していることを定量的に示した。また、栄養状態の悪化もリスク因子として特定された。これは、感染対策と認知症ケアのバランスという公衆衛生上のジレンマに対し、データに基づいた議論を提供するものである。

#### 4.1.3 婦人科領域およびがん研究

卵巣がんの化学療法奏功性予測や、子宮頸がん登録の有効性評価など、がん疫学分野でも多数の実績がある<sup>2</sup>。

- 政策的提言: 2023年の『The Lancet Regional Health』における論考<sup>4</sup>では、日本における経口中絶薬 (ミフェプリストン・ミソプロストール) の承認プロセスについて分析している。承認条件 (入院の必要性など) について、国際的な基準や女性の健康の観点から統計家の視点で考察を加えており、科学と社会制度の接点における発言を行っている。

### 4.2 災害科学とリスクコミュニケーション: 福島からの知見

広島大学での被爆者データ解析の経験は、福島第一原発事故後の研究にも引き継がれている。ここでは、定量的なデータだけでなく、テキストや描画といった定性データの解析に挑戦している点が特徴的である。

#### 4.2.1 小学生の「食事」描画のテキストマイニング的解析

2023年の『Frontiers in Communication』に掲載された研究<sup>4</sup>は、極めてユニークなアプローチをとっている。

- 対象: 福島県の小学生が描いた「日常の食事」の絵。

- 手法: 絵に含まれる要素(主食、主菜、副菜など)をタグ付けし、共起ネットワーク分析(Co-occurrence Network Analysis)等のテキストマイニング手法を適用した。
- 発見: 「副菜(side dishes)」を描く児童は、主食や主菜もバランスよく描く傾向があり、さらに地域社会への満足度が高いことが判明した。
- 意義: 「絵」という非言語的な情報をデータ化し、子供たちの健康意識や社会心理的状态を可視化することに成功した。これは、アンケート調査が難しい対象に対する新たな調査手法としての可能性を拓くものである。

#### 4.2.2 放射線不安の語彙解析

原発作業員や住民の放射線に対する不安を、インタビューや記述回答のテキストデータから解析する研究も行っている<sup>2</sup>。ここでも、単語の頻度分析やクラスター分析を用いることで、漠然とした「不安」の構造(例えば、健康への直接的な不安なのか、社会的差別への不安なのか)を客観的に抽出している。

### 4.3 人文社会科学への展開: 計量文献学

佐藤教授の研究対象の多様性を象徴するのが、『カンタベリー物語(The Canterbury Tales)』の写本系統分析である<sup>2</sup>。

- アプローチ: 中世の写本には、筆写の過程で生じた誤記や異同が無数に存在する。これを生物のDNA変異に見立て、統計的な系統樹作成手法やクラスタリング技術を適用することで、写本間の系譜関係(どの写本がオリジナルに近く、どの写本が派生したものか)を推定する試みである。
- データサイエンスの汎用性: この研究は、医学や経済学のために開発された統計モデルが、文学や歴史学といった全く異なる分野のデータに対しても有効であることを実証している。

## 5. 教育活動とゼミナール(研究室)の展望

佐藤教授のゼミナールへの配属を希望する学生にとって、教授の研究スタイルや教育方針を知るとは極めて重要である。公開されている担当講義情報や研究実績から、その教育環境を分析する。

### 5.1 担当講義から読み解く教育カリキュラム

佐藤教授が担当する主要な講義<sup>5</sup>は以下の通りである。

講義名	内容と教育目的の推察
医療統計論・医療統計演習	臨床試験デザイン、生存時間解析、疫学研究の手法など、医学統計の専門的知識を講義と実践(R言語等)を通じて習得させる。EBM(根拠に基づく医療)を支える統計家の育成を目

	指す。
データサイエンス実践価値創造演習	滋賀大学の特徴であるPBL(Project Based Learning)。企業や自治体から提供される実データを扱い、分析から課題解決の提案までを一気通貫で行う。佐藤教授の実社会での豊富な共同研究経験が、学生の指導に直接反映される場である。
基礎統計活用演習	統計解析ソフトウェア(主にR)の操作と、基礎的な統計理論の適用方法を学ぶ。プログラミングと数学的理解の橋渡しを行う。

## 5.2 ゼミナール(研究室)で期待できる研究と成長

佐藤研究室は、理論的な数理統計学と、泥臭い実データ解析の両方を重視する環境であると推察される。学生には、以下の3つの方向性での研究が期待できる。

### 5.2.1 方向性1: 新規統計手法・アルゴリズムの開発(Theory & Engineering)

- 内容: nmfkc パッケージの機能拡張や、新たな統計モデルの提案。
- 具体的なテーマ例:
  - 「NMFにおける新たなカーネル関数の導入と性能評価」
  - 「大規模スパースデータに対するNMFアルゴリズムの高速化」
  - 「時系列クラスタリング手法の改良と経済データへの適用」
- 得られるスキル: R/C++によるパッケージ開発能力、行列代数、最適化アルゴリズム、シミュレーション実験の設計能力。

### 5.2.2 方向性2: 医学・健康データの高度解析(Medical Data Science)

- 内容: 病院や自治体から得られる臨床データ、あるいはNDB(レセプト情報・特定健診等情報データベース)オープンデータ等の解析。
- 具体的なテーマ例:
  - 「ウェアラブルデバイスデータを用いた高齢者のフレイル検知」
  - 「気象データと救急搬送データの連動解析による疾患リスク予測」
  - 「特定健診データに基づく生活習慣病の因果構造探索(NMF-SEMの適用)」
- 得られるスキル: 医療情報の取り扱い倫理、疫学研究のデザイン、生存時間解析、ロジスティック回帰、因果推論の基礎。

### 5.2.3 方向性3: 非構造化データの探索的解析(Exploratory Data Analysis)

- 内容: テキスト、画像、音声などの非構造化データを統計的アプローチで解析する。
- 具体的なテーマ例:

- 「SNS等のテキストデータを用いた社会動向の時系列モニタリング」
- 「歴史的資料のデジタルアーカイブ化と計量分析」
- 「教育現場における学習ログデータの解析と教育効果の測定」
- 得られるスキル: 自然言語処理(NLP)の基礎、データ前処理(クリーニング)の高度な技術、多変量解析によるパターンの可視化。

### 5.3 指導方針と求められる学生像

佐藤教授の研究背景から、ゼミ生には以下のような姿勢が求められると考えられる。

1. プログラミングへの意欲: 独自の解析手法を実装するため、R言語を中心としたプログラミングスキルは必須となる(あるいは、習得する強い意欲が必要)。
2. 数理的な探究心: 既存のツールをブラックボックスとして使うのではなく、「なぜその手法が適切なのか」「数理的にどのような意味があるのか」を理解しようとする姿勢。
3. 社会課題への関心: 医学、教育、災害など、データ分析を通じて解決したい具体的な問いを持つこと。

## 6. 特筆すべき業績と客観的評価

佐藤教授の研究の質は、学术界から高い評価を受けている。その証左として、以下の受賞歴が挙げられる。

### 6.1 主要な受賞歴

特に日本応用統計学会における評価が際立っている。

受賞年	賞の名称	受賞対象・内容	意義
2023年	応用統計学会 学会賞(優秀論文賞)	NMFを用いた縦断的データのソフトクラスタリングに関する研究 <sup>2</sup>	最新のNMF研究が、統計学の専門家コミュニティで高く評価されたことを示す。
2015年	応用統計学会 学会賞(優秀論文賞)	混合効果モデルを用いた半パラメトリック変量係数の統計的推測 <sup>1</sup>	複雑なデータ構造に対する推測理論の精緻さが評価された。
2010年	応用統計学会 学会賞(優秀論文賞)	離散分布の縦断的データにおける線形変量係数の推測 <sup>1</sup>	医学データ等に多い離散値(カウントデータ等)の解析手法へ

			の貢献。
2017年	BCD Best Paper Award	大規模テキストデータの可視化手法 <sup>1</sup>	ビッグデータ解析や情報工学分野の国際会議(BCD)における評価。
2023年	Outstanding Presentation Award (IAGG)	老年学分野での共同研究 <sup>1</sup>	統計手法の医学応用が、医学・老年学の国際会議で評価された事例。

これらの受賞歴は、佐藤教授が10年以上にわたり、理論統計学と応用研究の両面でトップレベルの成果を出し続けていることを客観的に裏付けている。特に、同じ学会賞を3度受賞することは稀であり、その研究の継続性と信頼性の高さを物語っている。

## 7. 結論

本調査の結果、滋賀大学データサイエンス学部の佐藤健一教授は、「高度な数理統計モデルの開発者」と同時に、「データを用いて社会の複雑な問題を解き明かす実務家」とであると結論付けられる。

佐藤教授の研究内容の要点は以下の通りである。

1. 独自性: 既存のNMF(非負値行列因子分解)にカーネル法と共変量を導入した nmfkc パッケージを開発し、高次元時系列データや非線形データの解析に新たな道を拓いた。
2. 多様性: その応用範囲は、認知症やがん研究といった医学分野から、パンデミック下の行動分析、さらには文学作品の解析まで極めて広範に及ぶ。
3. 教育的価値: 学生は、単なるツールの使い方ではなく、データの背後にある数理構造を理解し、自らモデルを構築・実装する力を養うことができる。

ゼミへの参加を検討する学生への提言:

佐藤研究室は、「数学やプログラミングが好きで、新しい分析手法を作りたい」という学生にも、「医療や社会の課題をデータで解決したい」という学生にも、その両方のニーズに応える環境を提供している。特に、R言語によるパッケージ開発や、医学・異分野との共同研究という実践的なプロセスを経験できることは、将来データサイエンティストとして活躍する上で極めて大きな資産となるであろう。

以上を以て、佐藤健一教授の研究内容およびゼミナールに関する調査報告とする。

(報告書終了)

## 引用文献

1. 佐藤健一@滋賀大DS - 略歴, 1月 26, 2026にアクセス、  
<https://sites.google.com/view/ksatoh/c-v>
2. 佐藤健一@滋賀大DS - English - Google Sites, 1月 26, 2026にアクセス、  
<https://sites.google.com/view/ksatoh/english>
3. ksatohds/nmfkc: R package for Non-Negative Matrix ... - GitHub, 1月 26, 2026にアクセス、  
<https://github.com/ksatohds/nmfkc>
4. Kenichi Satoh's research works | Shiga University and other places, 1月 26, 2026にアクセス、  
<https://www.researchgate.net/scientific-contributions/Kenichi-Satoh-2187911814>
5. Details of a Researcher - SATOH Kenichi, 1月 26, 2026にアクセス、  
[https://researchers.shiga-u.ac.jp/html/100002895\\_en.html](https://researchers.shiga-u.ac.jp/html/100002895_en.html)