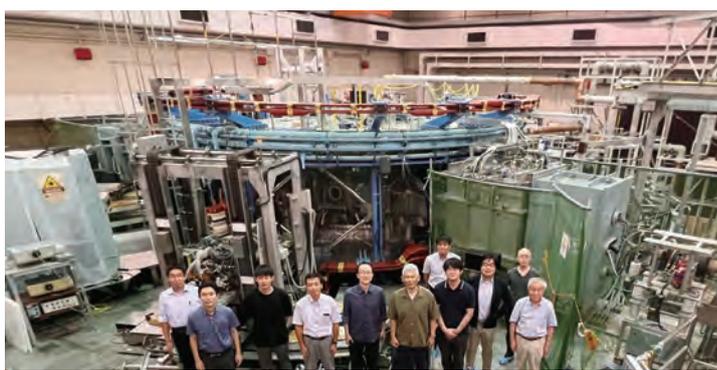


## ▼ CONTENTS

- 02 響き合う人とデーター 統数研プロジェクト紹介  
第35回「従属性のあるデータへの不確実性の定量化」
- 06 シンポジウム報告  
「リスク解析戦略研究センターシンポジウム」の開催報告  
特別セミナー「最近の実験計画法の発展～Special Seminar on Recent Advances in DoE～」開催報告  
「Applied Statistical Analysis using “R” Software in Makassar」共同開催
- 07 研究教育活動  
新入教員紹介  
2023年7～9月の公開講座実施状況  
統計数理セミナー実施報告(2023年9月～10月)  
「統計サマーセミナー 2023」実施報告  
連続最適化および関連分野に関する夏季学校2023実施報告
- 09 統数研トピックス  
有竹俊光元特任助教、日野英逸教授、赤穂昭太郎特任教授が2023年度日本神経回路学会論文賞を受賞  
情報・システム研究機構「第4期戦略的研究プロジェクト」における、  
京都大学エネルギー理工学研究所 核融合実験装置(ヘリオトロンJ)見学、意見交換会の報告  
第3回 統計エキスパート育成に向けたワークショップ  
「データサイエンス・統計分野のエキスパートをオールジャパンで育成」を開催  
気象大学校学生訪問プログラムを実施  
立川高校向け統数研データサイエンス・ハイスクールを実施  
兵庫高校オンライン講義を実施  
愛媛県立松山南高等学校の訪問
- 12 総合研究大学院大学統計科学コース関係  
大学院説明会について
- 13 共同利用  
2023年度共同利用公募追加課題
- 13 外部資金・研究員等の受入れ  
受託研究・受託事業等の受入れ  
外来研究員の受入れ  
寄附金の受入れ
- 14 人事
- 15 刊行物  
Research Memorandum (2023.8～2023.10)  
研究教育活動報告  
Annals of the Institute of Statistical Mathematics
- 16 コラム

大学共同利用機関法人  
情報・システム研究機構  
統計数理研究所  
二ユース



# ベイズ統計の適用範囲拡大 目指し新たな計算手法の



▲鎌谷研吾教授

## ベイズ統計による時系列データ解析 「二重の難題」への挑戦

今回紹介するプロジェクトのテーマは「従属性のあるデータへの不確実性の定量化」だ。「ここでいう『不確実性の定量化』とはすなわち、ベイズ統計学のことで」。統計数理研究所モデリング研究系複雑構造モデリンググループの鎌谷研吾教授は、そう説明する。

統計学とはまさに不確実な事象を定量化することであり、その定量化の意味を解釈することだ。さまざまな統計手法がある中で、ベイズ統計を使うメリットは何か。「ベイズ統計の特徴は、どのような指標で定量化しても解釈が一通りしかないこと。定量化の意味が一つに決まるという意味では理想的な手法です」（鎌谷）。

ただ、ベイズ統計は、計算量が膨大になるというデメリットも内包している。この問題に対処するため、一般的となっているベイズ計算法（ベイズ統計学での不確実性の計算法）が「マルコフ連鎖モンテカルロ法（MCMC）」だ。MCMCはランダムなデータを発生させるアルゴリズムで、ベイズ統計で計算が困難な場合に、乱数を無作為に抽出して概算する手法だ。

「近年のMCMCは機械学習と結びついて、大雑把な近似をして計算を簡単にするアプローチの研究が盛んですが、近似をすれば、統一的な解釈というベイズ統計の良さは薄まってしまふ。近似をしなくてもいい計算手法を開発するのが私たちの目標です」と鎌谷は話す。

一方、「従属性のあるデータ」とは、主に時系列データなどを指す。例えば日経平均株価のような金融時系列データが典型だ。「こうした従属性のあるデータは、たとえシンプルなモデルでも特微量、いわゆる『尤度』を簡単には計算できません。独立したデータなら前のデータを無視して計算できますが、時系列ということはデータがつながっているわけですから、それを切り離すのがとても難しい。理論的にも計算機的にも難易度が高いのです」（鎌谷）。

機械学習を取り入れながらベイズ計算法を進展させ、従属性のあるデータに適用して統一的な定量化の解釈を可能に



▲増田弘毅教授（東京大学）

することが、このプロジェクトの狙いなのである。もともと計算が困難な性質を有するベイズ統計を用いて、扱いの難しい時系列データを解析しようという、二重の難題への挑戦だ。

## 従属性データを扱う諸分野に インフラを提供する

このプロジェクトは、科学技術振興機構（JST）の推進するチーム型研究プログラム「CREST」の2021年度採択課題「大規模時空間従属性データ科学へ向けた先端的確率統計学の新展開」（研究代表者：吉田朋広東京大学大学院数理科学研究科教授）の取り組みだ。

時空間従属データに関わる多様な科学分野や社会事象に応用できる汎用的

# 大を 開発へ

時系列データのような従属性のあるデータの解析では、近似計算が行われることが多い。しかし、結果にはさまざまな解釈が生まれ、扱いが難しくなるデメリットがある。そこで、どのような指標で定量化しても解釈が一通りに統一されるベイズ統計の特徴を生かし、従属性データに適用することで、信頼性を向上させようという研究が進んでいる。機械学習を取り入れつつ、ベイズ計算の新手法の確立に挑むプロジェクトを紹介する。

方法を創出し、インフラとして提供することで、課題解決へ貢献することが、5年半に及ぶCREST研究における達成目標となっている。

ベイズ統計における計算法を専門とする鎌谷と現在、共同研究を進めているのは、東京大学大学院数理科学研究科の増田弘毅教授と、東京大学大学院総合

文化研究科附属先進科学研究機構の仲北祥悟特任助教やイギリス、ドイツの研究者らだ。

増田教授は確率過程モデル、特にレヴィ過程におけるジャンプの大家。「増田さんとは以前から共同研究をしてきました。標準的なブラウン運動ではなく、ジャンプを表現するモデルをぜひ使いたいと考え、

今回もチームに入っていました」(鎌谷)。仲北助教は、鎌谷が大阪大学に所属していた頃の上司の教え子。「機械学習と時系列データの両方に詳しい仲北さんのような人は珍しい。ぜひにと声をかけました」(鎌谷)。

鎌谷が計算手法を考案し、増田教授がその理論的正当性を確立する。仲北

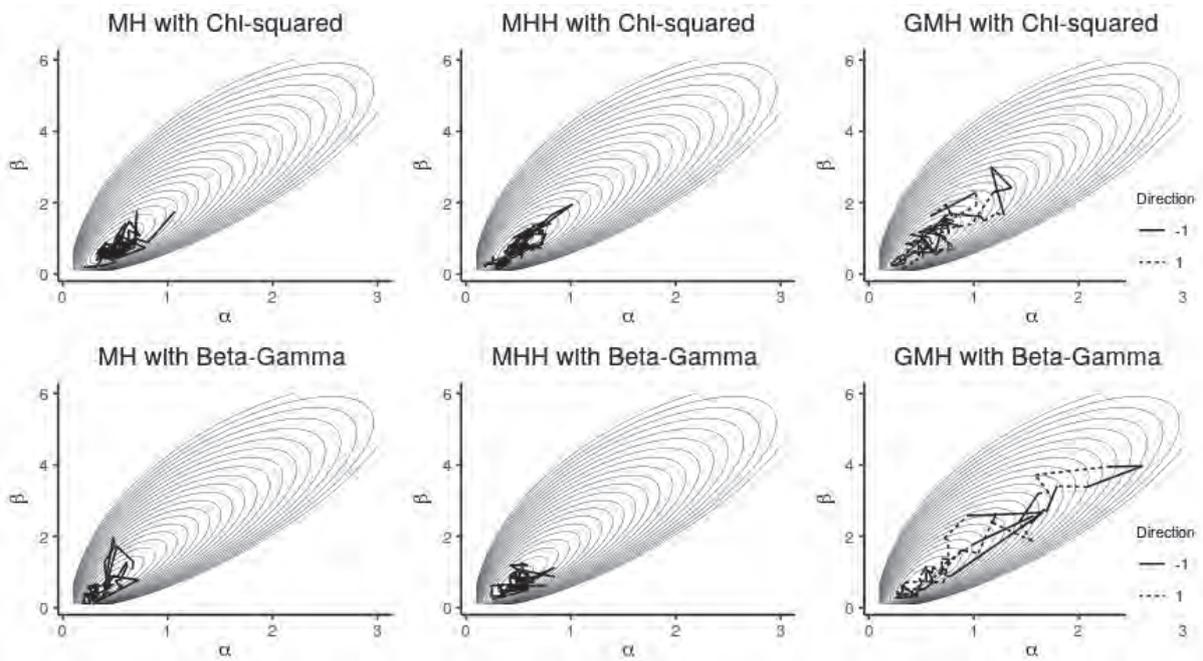


図1：対称性、非対称性をもつマルコフ連鎖モンテカルロ法の比較。左から2つの列は対称、一番右の列は非対称なマルコフ連鎖モンテカルロ法でベイズ計算を行った例。移動距離が大きいほどよい。対称なアルゴリズムは行きつ戻りつ進むので散漫であるが、非対称な方法は制約がないので移動距離が大きい。出典 Kengo Kamatani, and Xiaolin Song. Non-reversible guided Metropolis kernel. *Journal of Applied Probability*, 60 (3), 2023, pp. 955 - 981。



▲ 仲北祥悟 特任助教 (東京大学)

助教は機械学習的手法の展開や理論のアイデアを出す。そんな体制で研究を進めているという。

バイアスのない、すなわち適当な条件下で、計算で出力される値の期待値がちょうど計算したい対象と一致するベイズ計算の手法における今のトレンドは、「非対称性」の手法だ。「法則の対称性とは、ある経時的遷移の仕方と、その逆の遷移の仕方が同じ割合で起こることを言います。従来の手法の多くは、対称性の制約の中で発展してきたものです。しかし最近では、非対称性を含む手法でも実装できることが分かってきました(図1, 2)。うまく非対称を作ることができれば、効率よく安定的な状態にたどり着ける。われわれも非対称性に着目した研究を進めています」(鎌谷)。

増田教授が鎌谷と共同研究を始めたのは、5年ほど前のことだ。「正規分布ではなく非正規ノイズ過程で駆動されるランダムなダイナミクスの統計に興味を持っています。現世に実在するデータでは、往々にして正規性は棄却されるわけですが(図3)、非正規分布として扱えば推定の精度が上がるし、正規分布に基づく統計

モデルでは抽出できない情報を搾取できるからです」と増田教授は振り返る。

そんなとき、MCMCの漸近挙動(データ数が増えていくときのMCMCの性質)に関する鎌谷の理論に刺激を受けたことが、共同研究のきっかけだった。シンガポール国立大学のJasra准教授(現サウジアラビア・アブドラ王立科学技術大学教授)を加えた3人は、その成果を2019年に共同論文として発表した。「今回のCRESTプロジェクトで、私のグループでは複数の時系列データの相互作用を考慮

した統計手法の開発を手がけていますが、そこにも非ガウスノイズとベイズ統計で切り込んでいければ」(増田教授)。

一方で、機械学習のコミュニティでは比較的良好に研究されている、近似を行うアプローチでベイズ計算を研究しているという仲北助教は、「近似がうまく働くためには、ユーザーの問題設定に制約を強いることになります。データ間に従属性があるなど、複雑な構造の場合にはそうした仮定を満たしづらく、近似がうまく働かない状況が生じます」と指摘する。

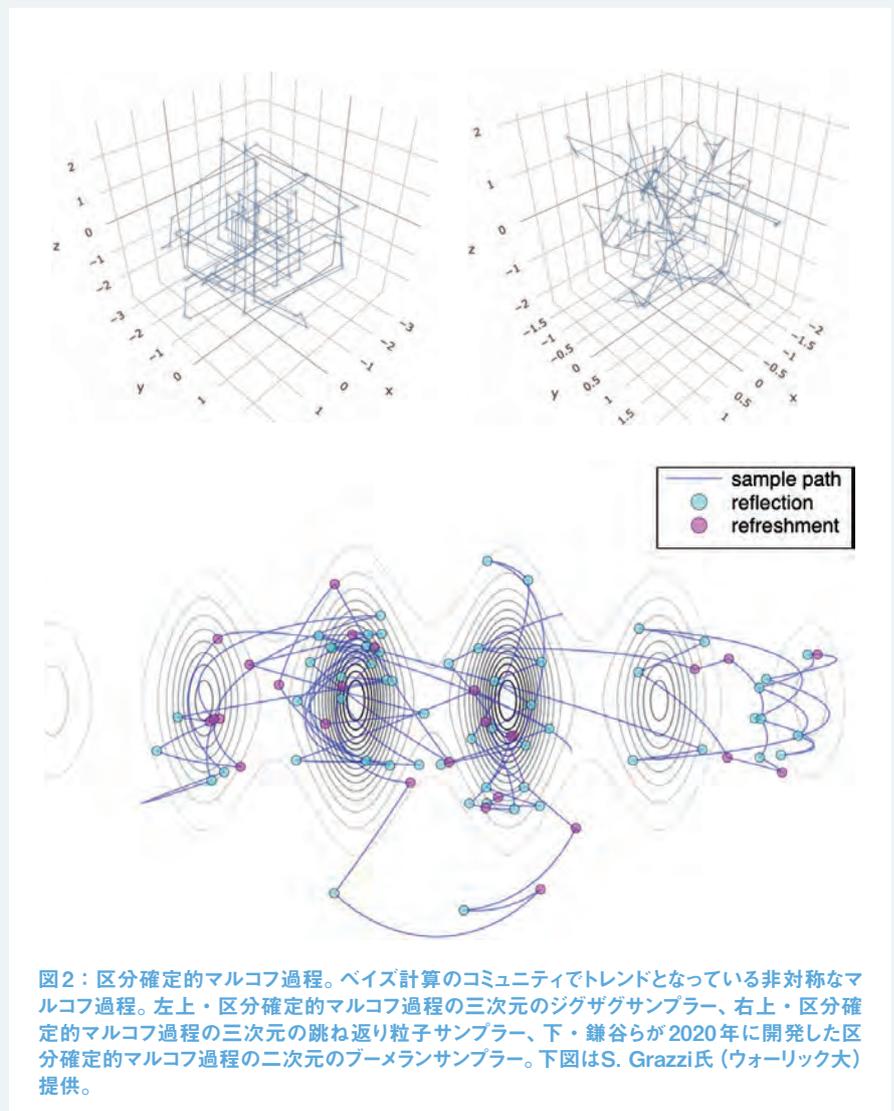


図2：区分確定的マルコフ過程。ベイズ計算のコミュニティでトレンドとなっている非対称なマルコフ過程。左上・区分確定的マルコフ過程の三次元のジグザグサンプラー、右上・区分確定的マルコフ過程の三次元の跳ね返り粒子サンプラー、下・鎌谷らが2020年に開発した区分確定的マルコフ過程の二次元のブーメランサンプラー。下図はS. Grazioli氏(ウォーリック大)提供。

「鎌谷さんのアイデアを組み入れることで、ユーザーがより広範な問題設定をしても、現実的に厳密な計算ができるような手法を開発できるのではないかと考えました。それが私にとって、このチームに参加する重要なモチベーションになっています。設定の妥当さと計算の簡単さのバランスを追究していきたいと思っています」(仲北助教)。

### 統計の根本はベイズにあり。計算手法の進化にかかる期待

CRESTの研究は、理論統計学、ベイズ統計学、機械学習、漸近分布理論、従属系における情報量規準およびスパース推定論など、これまでの研究を進展させ、時空間従属性解析の新しい地平を開くものだ。アウトカムの一つとして、確率過程に対する統計推測およびシミュレーションのためのRパッケージ「YUIMA」の開発も手がけており、鎌谷のチームではここにベイズ関数群を加える。

ベイズ統計のアプローチを採用することで、機械学習のブラックボックス問題や、いわゆる「P値の過信」といった問題の解決にもつながる可能性がある。例えば、2016年アメリカ統計協会は統計的仮説検定やP値（設定した仮説が正しいかを判定する基準となる値）について声明を出し、誤用への注意を喚起した。どんなデータでも、どんな解析をしても、P値が0.05を下回りさえすればよしとする風潮に警鐘を鳴らした形だ。

「難しい問題に対しては、どうしても計算が遅くなる。だから近似に頼りすぎてブラックボックス化が進むのです。ベイズ統計は正直な解釈のまま計算していくので、ブラックボックスがない。出てくる数字は

頭の中で直感的に浮かぶ確率とほぼ一致しています。機械学習を使っても、きちんと計算できる裾野を広げていきたい」と鎌谷は話す。

増田教授も「われわれが立脚している確率過程の統計学は、すでに半世紀以上にわたり練り上げられた歴史があります。微分方程式で記述される物理現象のモデルを機械学習で解析するといった最近の手法とは、統計的な取り扱いが根本的に違う。確率過程の統計学を土台にしつつ、機械学習的な要素を取り入れ、ベイズ計算の裾野を広げるというのは、かなりユニークなプロジェクトだと言えるでしょう」と補足する。

日本ではベイズ統計の経済・金融分野での研究が盛んだが、鎌谷のようにベイズ計算を専門とする研究者は少ないと

いう。手法を進展させるには、研究人材を確保する必要がある。そう考えた鎌谷は、2023年4月から統数研でベイズ計算法をテーマとするセミナーを開始した。

「不確実性を定量化するのに、ベイズ計算では必ず『確率』を使います。一般の人から見たら、統計で確率を使うのは当たり前と思うでしょうが、じつは他の統計手法ではそうでもなく、さまざまな解釈が生まれてしまう。定量化にはすべて統一に確率を使うベイズ計算は、非常にシンプルなので、最終的にすべての統計はベイズに行き着くしかないと思っています」と鎌谷の口調に熱がこもる。定量化の根本はベイズにあり。古典的で本来的なベイズ統計は、計算手法の進化次第で大きな可能性を秘めている。

(広報室)

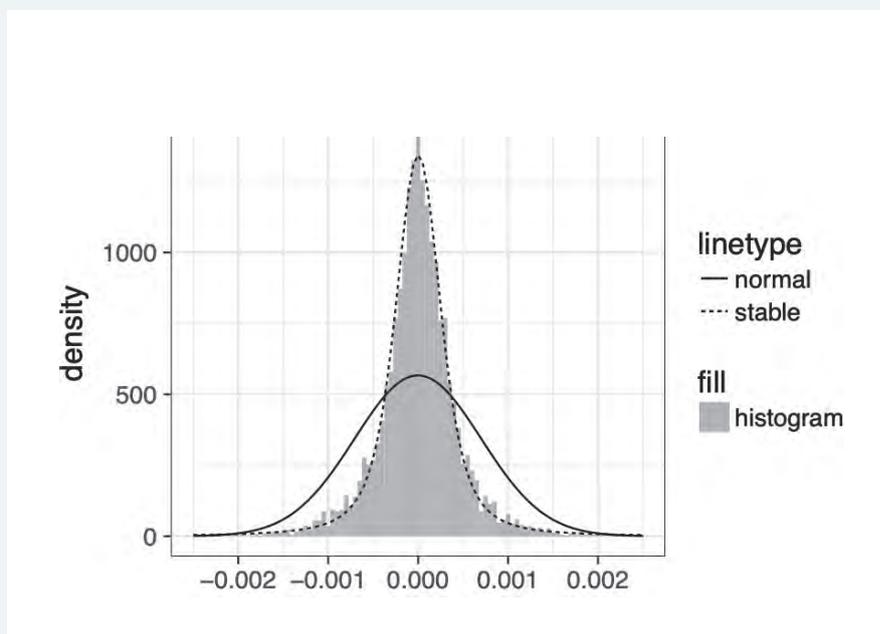


図3：非正規なモデル。鎌谷らの現在のキーワードの一つは「非正規性」。この図は、実データはしばしば非正規であることを示す例（2014年4月1日から2014年4月30日までのIBM株の1分ログリターン）のヒストグラム。データは強い非正規性を有している。出典 A. Jasra, K. Kamatani, and H. Masuda. Bayesian inference for stable Levy driven stochastic differential equations with high-frequency data, *Scandinavian journal of Statistics*, 46 (2), 2019, pp. 545 - 574.

## 「リスク解析戦略研究センターシンポジウム」の開催報告

2023年7月25日に、「リスク解析戦略研究センターシンポジウム」がオンラインで開催されました。同シンポジウムは、リスク解析戦略研究センターが毎年開催している恒例の学術イベントとなります。今年は「データサイエンスが切り拓く資源管理のフロンティア」というテーマで、同センターの資源管理リスク分析プロジェクトに関連した講演が行われました。加藤昇吾センター長による開催挨拶とセンターの活動紹介の後、チュートリアルセッションとして、加茂憲一准教授（札幌医科大学）、富田哲治教授（県立広島大学）、木島真志教授（琉球大学）による資源管理リスク分析で用いられる統計的手法の基礎について発表がありました。その後の特別講演では資源管理リスク分析の最先端の研究発表があり、光田靖教授（宮崎大学）による「土砂災害リスクを考慮した林業適地抽出技術の開発」と吉本敦教授（統計数理研究所）

による「森林資源ランドスケープ管理の最適化とリスク制御」の講演がありました。最後に、椿広計所長により閉会挨拶が行われました。参加者数は約97名で、盛況なシンポジウムとなりました。（加藤 昇吾）



## 特別セミナー『最近の実験計画法の発展～Special Seminar on Recent Advances in DoE～』開催報告

台湾中央研究院統計科学研究所（The Institute of Statistical Science, Academia Sinica, Taiwan : ISSAS）



講演を行うPhoa教授

のFrederick Kin Hing Phoa教授を講演者に迎えた特別セミナー『最近の実験計画法の発展～Special Seminar on Recent Advances in DoE～』が2023年8月4日（金）の午後、統計数理研究所セミナー室1で開催されました。

このセミナーは日本品質管理学会の後援を受け、ハイブリッド形式で行われ、統数研の二宮嘉行教授が司会進行役を務めました。産官学から多数のオンライン参加の申込があり、当日は、会場参加、オンライン参加併せて80名以上が、Phoa教授の実験計画法の発展、最新動向についての講演を熱心に聴講しました。講演後は、会場、オンラインともに予定時間を超過して活発な質疑応答が展開されました。

（国際連携推進室）

## 「Applied Statistical Analysis using “R” Software in Makassar」共同開催

2023年8月23日と24日の2日間に渡り、インドネシア・マカッサル市のMuhammadiyah University of Makassarにてワークショップ“Applied Statistical Analysis using R Software”を開催しました。このワークショップはインドネシア

農業経済学会（マカッサル委員会）とリスク解析戦略研究センターとの共催により、対面形式で開催しました。参加者は、自身のパソコンを用いて、統計ソフトRの基礎を学びました。初日、AICを含め統計数理研究所の概要を報告した後、

データハンドリングのベーシックなコマンドを学びました。2日目には、グラフィックスのコマンドを学び、様々なグラフを作成したのち、単回帰、重回帰、そして非線形回帰の基礎的な内容について学びました。今回のワークショップの運営は、農業経済学会のBakri氏をはじめ、Muhammadiyah University of MakassarのAsse教授、Khaeriyah氏らが中心に行い、参加者は27名でした。 (吉本 敦)



## 研究教育活動



### 村上 隆夫

データ科学研究系 准教授

新入 1  
教員紹介

本年の7月よりデータ科学研究系に着任いたしました。6月まで産総研（産業技術総合研究所）に在籍し、データ解析・機械学習におけるプライバシー保護、特に差分プライバシーの研究を行っていました。例えば、ソーシャルグラフにおいて、差分プライバシーを用いてユーザのプライバシーを強固に保護しながら、部分グラフの数え上げを行うアルゴリズムを確立し、その有用性の理論解析や評価実験などを行ってきました。また、差分プライバシーでは多くの応用先において、有用性が大幅に低下する問題にも着眼し、安全性と有用性を両立する新しい安全性指標の研究にも取り組んできました。今後も引き続き、プライバシー保護の研究を推進する予定です。どうぞよろしくお願いいたします。

## 2023年7-9月の公開講座実施状況

7月4日（火）に東京農工大学農学研究院の酒井憲司講師による「カオス農学実践編—トラクタのカオス動力学と農作業安全対策—」が開講されました。講義では、カオス理論の農学応用の実践編として、現代の機械化農業の成り立ちを農耕の起源から概観し、トラクタのカオス動力学について説明しました。対面での開催となり、モーション型トラクタドライビングシミュレーターでの仮想テストドライブも行い受講生には大変好評でした。

8月3日（木）から8日（火）まで、「多変量解析法」がオンラインにて配信されました。多摩大学の今泉忠講師と、当研究所の馬場康維講師、清水信夫講師が、多変量解析の基礎的・標準的な手法の解説をしました。また、クラスター分析、共分散構造分析などの解説もしました。質問も多数寄せられ、大変活発な講座となりました。

9月11日（月）から9月12日（火）まで、当研究所の伊庭幸人講師、二宮嘉行講師、日野英逸講師、矢野恵佑講師、慶應義塾大学医学部医療システムイノベーション寄附講座

所属で当研究所客員の立森久照講師、岡山大学の大久保祐作講師による、リーディングDAT 講座「L-A.現代統計学の基礎」がオンラインにて配信されました。現代的な統計学の基礎として、最尤法から、モデル選択、一般化線形モデル（GLM）までの内容を2日間にコンパクトにまとめて講義する講座です。 (情報資源室)

### 実機再現によるシミュレータの検証・高精度化

農研機構・農業機械研究部門で実施した実機実験をCarSimにより再現（農工大）



公開講座「カオス農学実践編—トラクタのカオス動力学と農作業安全対策—」資料より

## 統計数理セミナー実施報告(2023年9月~10月)

毎週水曜16時から所内研究教育職員および外部の方が1人40分ずつ、1日に2人の講演を行っています。2023年9月~10月のセミナーは下記の通り行われました。統計数理セミナーは現在オンラインで開催しています。

日程	氏名	タイトル
2023年 9月20日	南 和宏	匿名化データに対する差分プライバシー適用の検討
9月27日	熊澤 貴雄	能登半島群発地震活動の時空間的特徴
9月27日	服部 公平	Greedy Optimistic Clustering法を用いた銀河系における力学時計の構築
10月 4日	松井 知子	Closeness analysis of sequential data: A case study on COVID-19
10月 4日	間野 修平	A sum formula of GKZ hypergeometric polynomials at a specific value associated with decomposable graphical models
10月11日	持橋 大地	ホログラフ埋め込みに基づくCCG構文解析
10月18日	中野 慎也	磁気圏電離圏系の再解析データ作成に向けた取り組み
10月18日	豊田 祥史	分布外リスク最小化問題とドメイン不変性によるアプローチ
10月25日	村上 隆夫	差分プライバシーを満たす部分グラフ数え上げ

セミナーの開催予定はホームページにてご案内しています。 <https://www.ism.ac.jp/>

(メディア開発室)

## 「統計サマーセミナー2023」実施報告

統計思考院公募型人材育成事業の一環である統計サマーセミナー2023が、2023年8月7日から9日にかけて福井県のあわら温泉清風荘で開催されました。このセミナーは、若手研究者たちの交流の場として設けられ、彼らが研究発表や討論を通して相互に刺激を受け取ることを主旨としています。現代の学界で統計科学の研究が多様化していることを考慮し、このセミナーではその多様な領域を早期に知る機会を提供しており、60名の研究者が参加する中、実り多い討論が展開されました。

セミナーでは、2名の招待講演者を迎え、26件の口頭講演と16件のポスター発表が行われました。同世代の研究者が集まることで、オープンで活発な議論が交わされ、ナイトセッションでは討論が深夜まで続くなど、参加者の熱意が伝わる場となりました。招待講演では、天文学に関する最新の

統計的手法や従属性のモデリング技術に関する最先端の研究が紹介され、若手研究者への貴重なアドバイスも交えられました。総じて、このセミナーは参加者にとって学びの多い時間となりました。(矢野 恵佑)



## 連続最適化および関連分野に関する夏季学校2023実施報告

2023年8月9日から11日にかけて、統計思考院公募型人材育成事業として、連続最適化および関連分野に関する夏



季学校が開催されました。本夏季学校は、連続最適化とその関連分野における基本的な事項から最先端の動向までを整理・理解し、学生を含む若手研究者の基礎力の養成および新たな研究テーマの発見を目的として2021年から開催しているものです。今年は講師として東京大学の佐藤峻先生と統計数理研究所のProf. Bruno F. Lourençoをお招きして、それぞれ連続最適化への応用に向けた常微分方程式の数値解析入門、A Conic Smörgåsbordと題した講義と演習をしていただきました。また、現地会場での演習やポスターセッションでは参加者同士の議論が盛んに行なわれました。現地参加は遠隔参加よりも大きな効果があるように思

います。一方で、遠隔参加は気軽に参加できるという長所もあります。現地と遠隔を合わせて86名と多くの方にご参加

いただくことができました。

(田中 未来)

## 統計研トピックス

### 有竹俊光元特任助教、日野英逸教授、赤穂昭太郎特任教授が2023年度日本神経回路学会論文賞を受賞

有竹俊光元特任助教(統計的機械学習研究センター)、日野英逸教授(モデリング研究系/統計的機械学習研究センター)、赤穂昭太郎特任教授(大学教員育成センター)らが2023年度日本神経回路学会論文賞を受賞しました。日本神経回路学会では、一年間の研究会発表論文の中から優れたものを数件選び、日本神経回路学会論文賞を授与

しています。今回の受賞は「Neural Networks, Volume 149 (2022)」に掲載された論文『Detecting cell assemblies by NMF-based clustering from calcium imaging data』に対するものです。授賞式は9月6日(水)に日本神経回路学会第33回全国大会(会場 東京大学弥生講堂)にて行われました。(日野 英逸)

### 情報・システム研究機構「第4期戦略的研究プロジェクト」における、 京都大学エネルギー理工学研究所 核融合実験装置(ヘリオトロンJ)見学、意見交換会の報告

情報・システム研究機構(ROIS)の「第4期戦略的研究プロジェクト」に採択された課題「プラズマ物理と相補的なプラズマデータに対する統計数理モデリング」の一環として、7月24日から25日にかけて、京都大学エネルギー理工学研究所ヘリオトロンJ装置見学、ならびに意見交換会を、JSPS研究拠点形成事業「磁場の多様性が拓く超高温プラズマダイナミクスと構造形成の国際研究拠点形成」との共催で開催しました。

この企画は3月の九州大学応用力学研究所で開催したプラズマ実験装置見学・意見交換会に続くもので、統計学とプラズマ/核融合科学の双方のコミュニティの相互理解と協調的発展を目的としています。

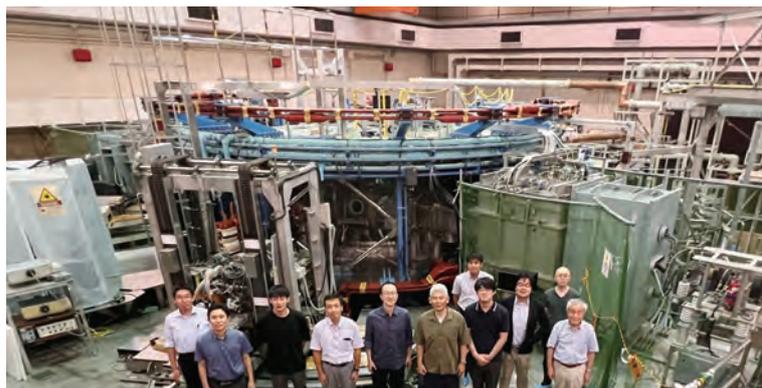
7月24日は核融合科学研究所(統計研客員)横山雅之教授による開会の挨拶の後、京都大学エネルギー理工学研究所長崎百伸教授によるJSPS研究拠点形成事業とヘリオトロンJ装置についてご紹介頂きました。続いて、三分一

史和代表による本プロジェクトの趣旨説明の後、時系列解析、機械学習など統計学関係の発表が行われました。その後、長崎教授のご案内で、ヘリオトロンJの見学をすることができました。これは京都大学が独自に開発した装置で、従来に比べ良好な粒子閉じ込めと電磁流体的安定性の両立を目指す研究が進められています。実際の装置を見学し、プラズマデータの生産現場に触れ、プロジェクトで取り組む研究課題へ具体的なイメージを醸成することができました。

7月25日は午前と午後にプラズマ・核融合の専門家にインタラクティブセッション形式で8件のご講演を頂きました。予定時間を大幅に超え、データ解析や制御における統計学上の問題点とその解決法について非常に活発な議論が続きました。

閉会後も双方のコミュニティの研究者による交流が続き、2件の共同研究が開始され、非常に有意義な企画でした。

(三分一 史和/横山 雅之)



京都大学エネルギー理工学研究所 ヘリオトロンJ装置見学

### 第3回 統計エキスパート育成に向けたワークショップ「データサイエンス・統計分野のエキスパートをオールジャパンで育成」を開催

「統計エキスパート人材育成コンソーシアム」では、2023年9月1日（金）、「第3回 統計エキスパート育成に向けたワークショップ」をオンライン形式で開催しました（166名が登録）。

千野雅人センター長の進行のもと、椿広計所長による「データサイエンス・統計分野のエキスパートをオールジャパンで育成～温故知新の人材育成～」の発表の後、「数理・データサイエンス・AI教育強化拠点コンソーシアムの取組」（東京大学 駒木文保教授）、「データ関連人材育成プログラムの取組と統計エキスパート人材育成」（大阪大学 狩野裕教授）、「大学統計教員育成研修の取組―第1期研修を振り返って―」（統計数理研究所 神保雅一特任教授）の発表があり、リテラシー・応用基礎レベルからエキスパートレベルまで、現在、進められている統計・データサイエンス分野の人材育成の取組を総覧し、人材育成に向けたあり方を考察しました。

発表後のディスカッションでは、「プロジェクト間の交流・情報交換が有用であること」、「統計エキスパート人材育成プロジェクトの発展的強化が求められること」、「エキスパートレベルの標準化・認証は困難であること」など、熱心な議論が行われ、参加の方々のご支援とご協力をお願いして閉会となりました。（大学統計教員育成センター）



### 気象大学校学生訪問プログラムを実施

2023年7月11日（火）気象大学校の第4学年の学生11名が、引率の同大学校の大塚道子准教授、齊藤直彬講師と共に当研究所を訪れました。

初めに、南和宏広報室長による挨拶と、統計数理研究所の



概要説明が行われました。次に、中野慎也教授により「電離圏電場分布のベイズ推定に関する事例紹介」と題し、HFレーダーのデータを用いた極域電離圏電場分布の推定について、経験モデルに基づくベイズ推定、磁気流体力学モデルシミュレーションへのデータ同化、機械学習の方法を用いて磁気流体力学モデルを模倣するエミュレータへのデータ同化などの講義が行われ、学生たちは熱心に耳を傾けていました。

その後、中野慎也教授により統計数理研究所のスパコンの説明が行われました。大型計算機の構造から、スパコンの性能、計算処理の仕組みなどの説明の後、計算機室、計算機展示室の見学が行われ、実際のスーパーコンピュータや1950年代からの貴重な計算機を興味深く見学し、写真を撮る様子が見受けられました。（広報室）

### 立川高校向け統数研データサイエンス・ハイスクールを実施

2023年7月6日（木）、東京都立立川高等学校の1年生12名を対象に統数研データサイエンス・ハイスクールをオンラインで実施しました。生徒はミニ探究として調査・研究、ポスター発表を体験しており、これから分野別の探究活動を開始するにあたり「データの取り方」「調査方法の選び方」

「分析の仕方」などを相談したいという目的で本プログラムへ申し込みが行われました。

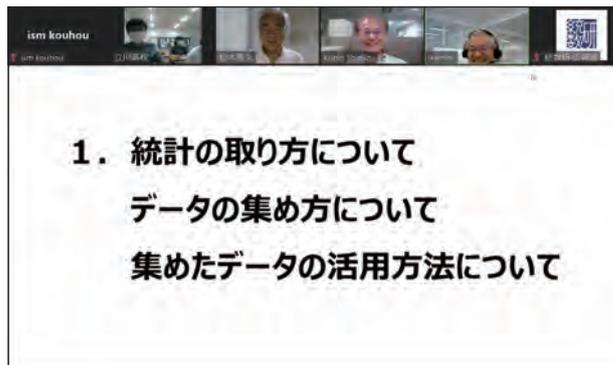
清水邦夫特任教授、柏木宣久特命教授、池森俊文特命教授が事前に提出された11項目の質問・相談事項について、生徒から各質問に含まれる問題意識の詳細を聞きなが

ら一つずつ回答する形で進められました。

先生方からは質問内容への回答のほか、前提にしていることや意識すべきこと、おすすめの公開されている資料の紹介なども行われ、これから探究活動を行う生徒にとって有意義な機会となりました。

最後に追加で「統計学を学んでいくうえでどのような力が必要になるのか」という質問があり、「数学に対する素養、コンピュータが使えることが大前提になる」と回答があり、データ科学の学部がある大学を受験するの一つの手であることや総研大統計科学コースの紹介なども行われました。予定されていた1時間を超過しての熱心な質問・相談が行われ

た本プログラムは生徒にとって貴重な体験となったとコメントがありました。(広報室)

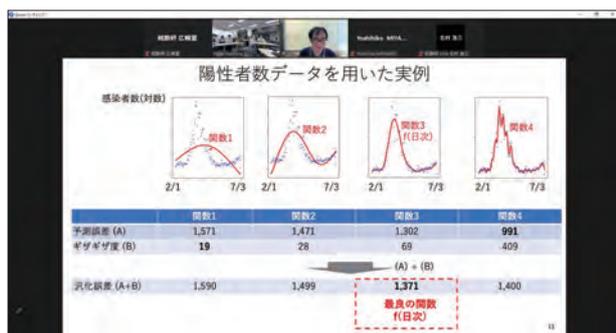


Report

## 兵庫高校オンライン講義を実施

2023年8月3日(木)、兵庫県立兵庫高等学校の2年生24名を対象に統計数理研究所オンライン講義を実施しました。

最初に本研究所の宮里義彦広報室副室長より、研究所の沿革や統計数理科学についての説明、どのような研究が行われているか等の研究所概要説明が行われました。次に



兵庫高等学校の代表生徒が事前学習の資料をまとめたスライドを使用して、「統計に対して抱いているイメージ」や「統計数理の活動に興味を持ったこと」について発表しました。続いて、データ科学研究系 村上大輔准教授が「統計モデルで探るCOVID-19の要因分析」と題した講義を行いました。陽性者数データを用いた実例をあげながら加法モデルについて説明され、イギリスでの大気汚染の解析や都市デザインの最適化など応用例も紹介されました。講義後は「使用するデータは正しいか、データを選別する時に気にしていること」や「統計と複数の分野が混ざった研究はあるか」等の多くの質問が寄せられ、活発な質疑応答が交わされました。

本講義は令和5年度東京未来フロンティアツアーの代替研修としてオンラインで開催されました。(広報室)

Report

## 愛媛県立松山南高等学校の訪問

2023年8月2日(水)、愛媛県立松山南高等学校の2年生、3年生の20名と教諭2名が本研究所を訪れました。同校はSSH(スーパーサイエンスハイスクール)の指定校であり、アドバンスデータサイエンス研修の中で研究施設訪問を通して、科学技術を担う人材として意識を高めることを目的としています。

初めに宮里義彦広報室副室長による挨拶と統計数理研究所の概要説明が行われました。次に計算機室、計算機展示室へ移動し、スパコンや統計数理が過去に使用してきた古い計算機類を熱心に見学しました。

研究紹介では、持橋大地准教授による「短歌の評価をモデル化する」と題した講義が行われました。見えないデータを推定して数学的にモデル化することで、多くのデータを適切に分析することが可能だと説明が行われました。生徒達からは

「なぜ7段階評価なのか」「言語を数式にする面白さは?」等の質問があり、興味の深さが伺えました。

この訪問が今後の研究課題に取り組む姿勢や、将来の進路を考えるきっかけとなることを願っています。(広報室)



## 入学者選抜試験結果

### 【5年一貫制博士課程】

試験年月日	合格者数	
2023年8月8日(火)～8月9日(水)	2023年10月入学(第2回)	－
	2024年4月入学(第1回)	0名

### 【博士後期課程】

試験年月日	合格者数	
2023年8月9日(水)	2023年10月入学(第2回)	3名
	2024年4月入学(第1回)	1名

## コース修了式

令和5年9月20日(水)に、オンライン(Zoom)にて、コース修了式が行われ、4名が本コースを修了しました。

## 秋季学位記授与式

令和5年9月28日(木)に、総合研究大学院大学葉山キャンパスにおいて秋季学位記授与式が行われ、本コースから4名が学位記を授与されました。



秋季学位記授与式

## SOKENDAI賞・研究科長賞授賞式

令和5年9月28日(木)に、総合研究大学院大学葉山キャンパスにおいてSOKENDAI賞・研究科長賞授賞式が行われ、本コースから尾崎凌斗さんが第11回SOKENDAI賞を受賞されました。



左から藤澤コース長、SOKENDAI賞を受賞した尾崎さん、二宮教授(主任指導教員)

## 学位取得者

令和5年9月学位取得者は次のとおりです。

### 【課程博士】

氏名	論文題目
尾崎 凌斗	Cox比例ハザードモデルにおける変化点検出のための情報量規準
園田 桂子	金融政策の波及効果及び企業マイクロデータ構造化のための統計的手法の改善に関する研究
田沼 巖	Discrete Distribution-based Collaborative Filtering Utilizing Various Data Sources for Recommendation Systems
Zhang Qi	Machine learning for de novo design of functional molecules and their synthetic routes

## 秋季入学式

令和5年10月10日(火)に、総合研究大学院大学葉山キャンパスにおいて秋季入学式が行われ、本コースへは3名が入学しました。

## 大学院説明会について

令和5年10月13日(金)に、2023年度第2回大学院説明会をオンラインにて開催しました。

入試ガイダンス、カリキュラムの説明、質疑応答、教員

マッチングの説明を行い、その後修了生2名より講演いただきました。参加者は34名でした。

(統数研事務課・統数研係)

## 共同利用

### 2023年度共同利用公募追加課題

【一般研究1】1件

分野	研究課題名	研究代表者(所属)
d7	人工概念から問い直す福島県の未来に関する意識調査	椿 広計(統計数理研究所・名誉教授)

(研究推進課・共同利用係)

## 外部資金・研究員等の受入れ

### 受託研究・受託事業等の受入れ

委託者の名称	研究題目	研究期間	研究経費(円)	受入担当研究教育職員
国立大学法人京都大学 学長 湊 長博 代理人 医学・病院構内共通 事務部長 河野矢 英成	京都大学大学院における臨床統計家育成推進のための大学院・卒後一貫したプログラム構築に関する研究開発	R5.4.1～ R6.3.31	5,783,700	椿 広計 名誉教授
国立大学法人京都大学 学長 湊 長博 代理人 医学・病院構内共通 事務部長 河野矢 英成	こころの健康の保持増進のための超個別化AIプロジェクト:完全要因ランダム化試験からlivingRCTプラットフォームに至る開発研究	R5.4.1～ R6.3.31	650,000	データ科学研究系 野間 久史 教授
国立大学法人東京大学 総長 藤井 輝夫 代理人 工学系・情報理工学系 等事務部長 櫻井 明	インフラ考慮都市発展モデルへの気候変動リスクの組み込み	R5.4.1～ R6.3.31	0	データ科学研究系 村上 大輔 准教授
国立研究開発法人国立成育医療 研究センター 理事長 五十嵐 隆	小児および若年成人における再発難治CD19陽性B細胞性急性リンパ性白血病に対する同種造血細胞移植後維持療法の確立に関する研究	R5.4.1～ R6.3.31	130,000	データ科学研究系 野間 久史 教授
厚生労働省 厚生労働事務次官 大島 一博	全国がん登録情報の利用及び提供における情報の特性と安全管理措置に関する研究	R5.4.1～ R6.3.31	1,560,000	データ科学研究系 南 和宏 教授
国立研究開発法人科学技術振興 機構 分任研究契約担当者 契約部長 近藤 章博	交通・建築物のエネルギー需要マネジメント手法の開発	R5.7.28～ R6.3.31	3,450,000	データ科学研究系 村上 大輔 准教授

(研究推進課・研究推進係)

## 外来研究員の受入れ

氏名	職名	研究題目	研究期間	受入担当研究教育職員
Charly Andral	Université Paris Dauphine-PSL, Doctoral Course Student	区分確定的マルコフ過程と非可逆的なマルコフ連鎖	2023.7.2～ 2023.8.27	鎌谷 研吾 教授
Aleksandar Pavlović	TU Vienna, Doctoral Course Student	幾何学的ナレッジグラフの埋め込み	2023.6.29～ 2023.8.27	Bruno Figueira Lourenço 准教授
安井 雄一郎	日本経済新聞社日経イノベーション・ラボ・主任研究員	科学技術文献の引用構造に対するモデル化および指標の構築	2023.6.8～ 2024.3.31	田中 未来 准教授
魏 梦怡	Institute of Geology and Geophysics, Chinese Academy of Sciences, PhD. Candidate	龍門山断層帯の地震定位における二重差分手法の適用性	2023.7.1～ 2023.8.1	庄 建倉 教授
王 志峰	China University of Geosciences, Master's Course Student	地震震源過程のベイズ的反転	2023.7.1～ 2023.9.9	庄 建倉 教授
毛 寧	Institute of Geophysics, China Earthquake Administration, Master's Course Student	幾何学的垂直変換関数のベイズ型3Dインバージョン	2023.7.20～ 2023.9.10	庄 建倉 教授
斎藤 大雅	東北大学・博士前期課程学生	空港基礎地盤の土層構成と地盤物性値の推定	2023.7.10～ 2023.8.11	Stephen Wu 准教授
大竹 雄	東北大学工学部・准教授	空港基礎地盤の土層構成と地盤物性値の推定	2023.8.2～ 2023.8.11	Stephen Wu 准教授
中村 淑子	湊運輸倉庫ドローン事業部・データアナリスト	UAV-空撮・レーザ計測技術による森林資源量解析と次世代森林GISの基盤構築	2023.7.6～ 2024.3.31	田中 未来 准教授
司 政亜	Institute of Geophysics, China Earthquake Administration, Doctoral Course Student	複数の地震観測網から得られたマグニチュードのベイズ統合法	2023.9.1～ 2023.10.30	庄 建倉 教授
朴 健	Peking University, Doctoral Course Student	余震の減衰と地震ハザード評価への影響	2023.9.16～ 2023.11.12	庄 建倉 教授

(研究推進課・共同利用係)

## 寄附金の受入れ

受入決定年月日	寄附者	寄附金額(円)	担当教員	寄附目的
2023.7.6	株式会社ブリヂストン 取締役 代表執行役 Global CEO 石橋 秀一	500,000	松井 知子	機械学習及びモデル化技術の基礎研究 ※条件は無し

(研究推進課・研究推進係)

## 人事

### 外国人研究員(客員)

氏名	現職	所属	職名	研究課題	期間	受入教員
シヨン ジョー Shiyong Zhou	教授	北京大学 地球宇宙科学学院	客員教授	情報理論に基づく地震波形極性の確率的自動判断法	2023.8.1～ 2023.8.31	庄 建倉 教授

(総務課・人事係)

## 外国人客員紹介



● Shiyong Zhou 客員教授

It is an honor to have been invited to visit ISM once again from August 1 to 31, 2023. I have been collaborating with the Statistical Seismology Group of ISM since 2003, and both parties have greatly benefited from our longstanding cooperation. During this visit, my focus will be on the development of an “Automatic Probabilistic Polarity Determination Method based on Information Theory.” This method aims to detect the initial arrival times and polarities of seismic waveforms in a probabilistic manner, while also providing a quantitative evaluation of picking uncertainties. Phase detection and polarity determination are crucial and fundamental aspects of seismic data analysis. Although many previous methods have been proposed and proven effective, the uncertainties of the results are rarely given serious consideration. In our pursuit of a quantitative evaluation of picking uncertainty, we are collaborating to develop a new method that calculates the probability distribution of arrival times based on information theory.

I would like to express my deep gratitude to Prof. Jiancang Zhuang for hosting my visit. I greatly appreciate the research environment at ISM, including the Akaike guest house. As always, ISM has provided me with the opportunity to interact with Japanese researchers and other visiting scholars, facilitating progress on exciting research projects.

## 刊行物

### Research Memorandum (2023.8~2023.10)

No.1225: Iba, Y., W-KERNEL AND ESSENTIAL SUBSPACE FOR FREQUENCIST'S EVALUATION OF BAYESIAN INFERENCE

(メディア開発室)

### 研究教育活動報告

No.55: 統計数理研究所, 総合研究大学院大学 先端学術院先端学術専攻 統計科学コース, 2023年 統計数理研究所オープンハウスポスター発表資料集 及び 統計科学コース学生研究発表会資料集 (2023.7)

(メディア開発室)

### Annals of the Institute of Statistical Mathematics Volume 75, Number 5 (October 2023)

Shibin Zhang

A copula spectral test for pair wise time reversibility .....705

Florian Dussap

Nonparametric multiple regression by projection on non-compactly supported bases .....731

Abhijit Mandal, Beste Hamiye Beyaztas, and Soutir Bandyopadhyay

Robust density power divergence estimates for panel data models .....773

Nannan Ma, Hailin Sang and Guangyu Yang

Least absolute deviation estimation for AR(1) processes with roots close to unity .....799

Shaoting Li and Jiahua Chen

Mixture of shifted binomial distributions for rating data .....833

Yasuhito Tsuruta and Masahiko Sagae

Automatic data-based bin width selection for rose diagram .....855

(メディア開発室)



# 歴史の節目に統計学

三分一 史和

モデリング研究系

2019年から流行の新型コロナの分類は5類感染症へ変更になり、テレビでお馴染みだったコメンテーターを目にすることはめっきり少なくなった。連日、感染者数の増減云々がセンセーショナルに報道されていた当時は、アナウンサーやタレントまでも、PCRと抗原検査の性能評価で感度、特異度、検定、有意性などの統計用語を使うようになり、ある意味、統計学がより身近になったのではないであろうか。関連して、2011年の震災と原発事故で“安全神話”が崩壊したとき以来、「可能性はゼロではない」という言い回しを耳にすることが増えた（責任逃れに使われることが多いが）。これは、確率的思考が徐々に一般にも浸透していることなのかもしれない。

話をコロナに戻す。ある時期に様々な感染者数の予測モデルが出され、“予測合戦”の相を呈するようになった。しかし、その予測に使うデータ収集には、FAXや手入力などアナログ的なプロセスがボトルネックとなり、実態の把握が追い付かなく、医療機関がパンクし多くの患者が放置される一因になった。政府の迷走、責任回避、情報と分析の欠如など先の大戦中との類似性を唱え、「コロナ敗戦」と呼ぶ歴史学者もいる。

では、その戦時中ではどのような様子であったのであろうか。日本では第一次世界大戦後に様々なデータの収集と分析が行われはじめ、統計学の重要性の認識が進んでいた。しかし、先の大戦前から統制経済へと移行し、統計は分析ではなく報告程度の扱いになり、データは省庁ごとに囲い込まれ、秘匿や捏造が頻発した。戦略や物資の配分、兵員配置、生産管理などの多岐にわたる意思決定に支障が出るのは必然である。ようやく大戦末期になり、軍部も事の重大性に気づき、

統計数理研究所の設立へと繋がったのであるが、遅きに失していたのは自明である。戦後、マッカーサーがStuart A. Rice博士に日本の統計再建と勧告を委任し、1947年から1951年にかけてライス・レポートと呼ばれる3部の報告書にまとめられた（図1）。

戦後、GHQ、日本政府、学者、官僚が参画する統計委員会が発足し、統計法も公布され、国際的な信用も向上した。しかし、その後、徐々に各省庁での統計業務は一般の官僚が行う閑職へと追いやられ、データ収集の目的意識も希薄になっていった。その結びは2018年と2021年に厚労省と国交省における統計不正問題として発覚し、歴史は繰り返されることになった。

そのようなことが忘れかけられていた2022年11月にOpenAI社がChatGPTをリリースし、連日、黒船来襲であるかのような報道が続いた。大規模言語モデルや統計学についての専門家のインタビューや解説記事が氾濫状態で、再び、「統計学」という語を耳にする機会が多くなった。

統計学は地味な学問であるが、このように、良きにつけ悪しきにつけ歴史の節目に現れる。「未来が現在の成れの果ての姿」にならないように、現在における統計学の学問側面、そして、社会の中での取り扱われ方が歴史的にどのような位置付けになるのか、未来の研究者にどのように引き継いでいくか、節目ごとに立ち止まって考えてみる必要があると思う。図2はAIに“統計数理研究所”というプロンプトを与え生成された画像である。因みに、図1のRice博士の画像はAIによりアップスケールしたものである。

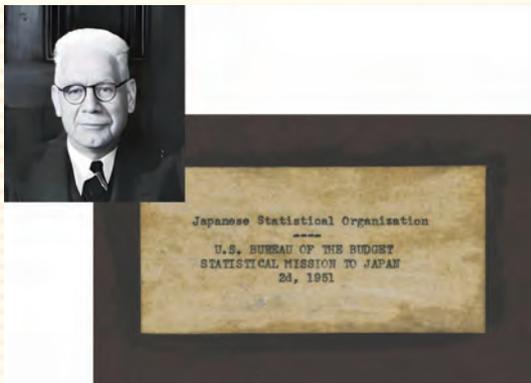


図1 Rice博士とライス・レポートの表紙(部分)

[引用] 顔写真: The International Statistical Institute HPより  
レポート表紙: ウィスコンシン大学図書館より



図2 AIに生成させた“統計数理研究所”の画像