### FEBRUARY. 2024

# THE INSTITUTE OF STATISTICAL MATHEMATICS

# NEWS

### CONTENTS

02 響き合う人とデータ - 統数研プロジェクト紹介 第36回「時空間データを用いた方向統計モデルの構築」

### 06 シンポジウム報告

「Applied Statistical Analysis using R Software for Forestry Studies 2023」 共同開催報告 ワークショップ 「科学モデル論は何を目指せばよいのか」 開催報告 共同研究集会 「極値理論の工学への応用」 開催報告

### 07 研究教育活動

2023年10月-12月の公開講座実施状況

統計数理セミナー実施報告(2023年11月~2024年1月)

統計数理研究所夏期大学院「データ同化夏の学校」の開催

「探索的ビッグデータ解析と再現可能研究」実施報告

統計思考院人材育成ワークショップ「生物多様性と群集動態:定量化の数理と統計的推定法」開催報告

### 09 | 統数研トピックス

数学・数理科学5研究拠点合同市民講演会「数学・数理科学の未来!?」を開催

江村剛志教授が台湾の国立中興大学を訪問

村上大輔准教授が2023年度坂下賞を受賞

統計数理研究所子ども見学デー 2023 を開催

日本学術会議見解「大学における数理・データサイエンス・AI 教育の中での統計科学の教育について」の 公表について

### 11 | 共同利用

2023年度共同利用公募追加課題

11 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

### 12 人事

13 会議開催報告

令和5年度第3回運営会議の開催

14 刊行物

Research Memorandum (2023.11  $\sim$  2024.1)

統計数理/ Annals of the Institute of Statistical Mathematics

15 お知らせ

2024年度公開講座(一般講座)

16 コラム



# 統計数理研究所ニューマ

大学共同利用機関法人

・システム研究機構

# 事故や災害にかかわる「方向

# の解析で損保会社と協働



▲加藤昇吾准教授

### 事故や災害の補償から社会課題の解決へ データサイエンスで保険が進化する

現代の保険システムの基になった「相 互救済」の考え方は、古代バビロンの時 代に生まれたと言われる。日本でも奈良・ 平安時代にすでに、飢饉に備えて穀類を 備蓄しておく「義倉」と呼ぶ制度があった。 今では災害や事故などさまざまなリスクを 補償する損害保険が存在し、企業や個 人の夢と挑戦を支えている。

日本を代表する損害保険会社の一つ、 三井住友海上はこれからの自社の役割に ついて、「事故や災害による被害を補償 するだけでなく、これまで蓄積した知見を 生かして補償前後のソリューションを提供 し、お客さまや社会課題を解決すること」 を掲げている。事故や災害のリスクを評価



▲石川雅之氏(三井住友海上火災保険)

する技術は、補償の査定の精度向上はも ちろん、事故の予防や防災にも役立てら れる可能性が大きいからだ。

今回紹介するプロジェクト「時空間データを用いた方向統計モデルの構築」は、そうした目標のもと、三井住友海上が統計数理研究所の産学連携プログラムに、方向統計学(directional statistics)の専門家である加藤昇吾准教授との共同研究を申し入れて実現したものだ。

同社のデータサイエンティストであり、ビジネスデザイン部 データサイエンスチーム 主任スペシャリストの石川雅之さんは、「当社のデータ分析業務では、気象災害や自動車の事故をはじめとした物理的な事象に対するリスク評価や、事故の予防策を講じることを目的とした分析の機会が多く

あります。特に気象や移動情報などを分析対象として扱う場合には方向データを扱う機会も少なくないのですが、定量的な活用や有効な統計的解釈ができていない点を課題として感じていました」と経緯を語る。

方向データに対して、既存では抽出できていなかった特徴の抽出や定量的評価を可能とすることで、リスク評価や予測の改善、高度化を図りたいというのが、三井住友海上の目的だ。

# 実数値データの定義がそのままでは通用しない「方向データ」の難しさ

方向統計学は、「方向」として表される 観測値を含むデータ、いわゆる「方向データ」を扱う統計学の研究分野だ。例えば、 2次元における方向としては、石川さんが 扱う気象災害に影響を与える風向や、自 動車事故における自動車の移動方向、 生物学における地上動物の移動方向な どが挙げられる(図1)。

「アナログ時計の短針からイメージを広げて、24時間で一周する時計の針のようなものを想像してみてください。医学における『心臓発作が起こる時刻』なども、2次元平面における方向として表現可能であることが分かるでしょう」と加藤は説明する。

私たちの身の回りは、さまざまな「方向」

# 第 36 回

# 「時空間データを用いた方向統計モデルの

# データ」

### 構築」

三井住友海上火災保険株式会社(以下、三井住友海上)は、同社が保有している多くの事故や災害に関連するデータを活用した分析により蓄積した知見を、社会課題の解決に生かす方針を打ち出している。その一環として、気象や移動情報の分析で扱う「方向データ」の定量的な活用や有効な統計的解釈の実現を目指し、統計数理研究所の産学連携プログラムを利用。現在、方向統計学を専門とする教職員と進めている共同研究の様子を紹介する。

で溢れている。3次元ユークリッド空間では化学に現れる分子のねじれ角、地震学で扱う断層の動く方向、天文学で観測される星の方角など枚挙にいとまがない。さらに、多次元ユークリッド空間における方向の観測も、言語処理や人口統計学などで議論されている。

こうした方向データを扱う方向統計学は、近年、機械学習などにも広がりを見せている活発な研究分野だ。だが、統計解析をするうえで、一筋縄ではいかないのが方向データの特性でもある。「じつは、方向データには、統計学が主に対象としている実数値データのための解析手法をそのまま使うことができないという問題があります」と加藤は言う。

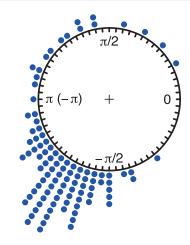


図1:方向データの例。風向や動物の移動方向などの方向データは角度で表される。

例えば、統計学における最も基本的な統計量に平均や分散があるが、実数値データの平均や分散の定義をそのまま方向データに用いると、不自然な定義となってしまう(図2)。確率分布、回帰モデル、時系列モデルなど実数値データのためのほとんどの統計モデルも、方向データにそのまま応用することは一般には適切ではない。方向データを解析するための統計手法の発展は、統計学における重要な研究課題となっている。

特に、方向データについては、正規分 布に相当するような実用性と統計的な扱 いやすさを併せ持つ確率分布がなく、方 向データのための確率分布の研究は、 方向統計学において最も重要な研究

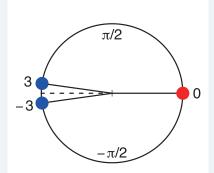


図2:方向データ解析における問題点。方向データの統計解析では、実数値データのための統計的手法をそのまま用いることができない場合がある。例えば図のように、2つの角度 (-3と3) の実数値データにおける標本平均は3<sup>13</sup>=0と計算できるが、0を "角度の平均" とするのは自然な定義とは言えない。

テーマの一つと見なされてきた。

加藤は、今まで着目されていなかった「コーシー型」の確率分布に着目し、既存の確率分布では成り立たない多くの扱いやすい性質を持つ確率分布を提案した。また、他にも方向データのための回帰モデルや時系列モデルに関する研究結果がある。

大学院でマクロ経済の実証研究や、 民間企業でデータ分析に従事してきた石 川さんは、統数研の公開講座に参加す ることも多く、その際に統数研が共同研 究を募集していることを目にし、方向統計 学の分野において先端的な研究に取り 組んでいる加藤に共同研究のパートナー として白羽の矢を立てたという。

### 方向統計・時空間データ解析・ベイズ統計 3つの分野のクロスポイント

今回のプロジェクトではまず、方向データの一つである「風向データ」を用いて時空間の関係を定量的に評価できる手法の開発を進めている。ここでは、2023年5月に日本気象学会の春季大会で発表した成果を紹介しよう。観測機器が設置された地点における気象情報から、未観測地点における気象情報を補完する問題を考えるものだ。

この研究では、国内の風向データに対して射影ガウス過程のパラメータの事後

### 響き合う人とデータ ▶ 統数研プロジェクト紹介

分布を推定し、推定対象地点から除外した検証地点に「クリギング」と呼ばれる補完手法を適用することで、射影ガウス過程による日本の風向の傾向と補完値の精度について考察した(図3、4)。

クリギングは、観測されたデータ間の空間的自己相関関係から予測や補完を行うもので、以前から風速などの実数値データには適用されていたが、方向データへの適用理論は2012年に初めて海外で発表された。統計ソフトウェアRのライブラリとして一般での利用が可能となったのは、2019年と近年のことだ。今回は、この最新の理論を国内の風向データへ適用した点で画期的と言える。

「射影ガウス過程を用いて仮想的に作

成した未観測地点に対して風向を補完 し、精度評価を行う取り組みは、われわれ の調べた限り、国内で初めてです」と加 藤は話す。

研究の結果、射影ガウス過程によるクリギングを実施することで、未観測地点に対する風向の補完値の算出が可能であることと、その補完精度の水準について確認ができた。だが、二人はこれだけではまだ納得していない。現行の手法では、風向の補完に用いる情報は各地点の風向、位置および時間のみであり、その他の情報を活用することで、補完の精度を改善できる余地があるからだ。

石川さんは「推定されたパラメータの事 後分布や補完精度の傾向から判断する と、風向における空間的自己相関は距離に対して弱まりやすい傾向にあると考えられます。このような風向に対する地理的情報や、風速などの他の情報との関係を捕捉し、補完精度をより改善できるような統計手法の開発を行いたい」と先を見据える。

「風向」とその観測された「位置」「時間」から、その関係を「ベイズモデル」として表現する。それは言うなれば、「方向統計」と「時空間データ解析」、「ベイズ統計」という3分野の交差点にあたることから、広い知識が必要となる。

「私自身も勉強になるし、石川さんにも 知見を深めてもらっています」と加藤が言 えば、石川さんも 「統計理論の数理的研

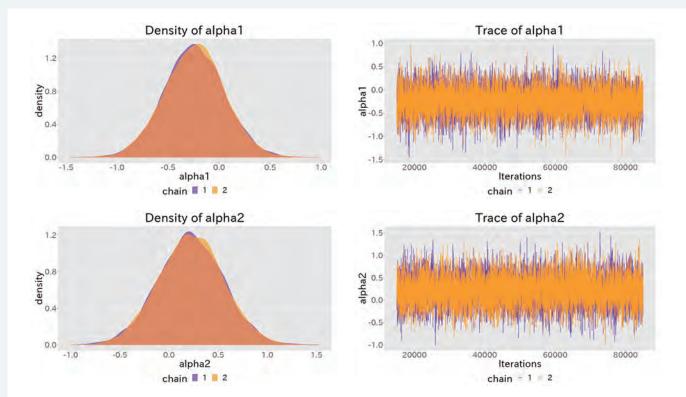


図3:パラメータの事後分布とトレースプロット。気象庁が公開しているアメダスの日次単位での地点別最多風向データを対象に、MCMC (マルコフ連鎖モンテカルロ法)を用いて射影ガウス過程と時空間相関係数のパラメータを推定した。上段は射影ガウス過程における横方向、下段は縦方向の平均パラメータの事後分布(左)とそのサンプルを示すトレースプロット(右)。横方向では、事後中央値は-0.243となっており、観測期間において全体的に風は西から吹く傾向であることを捕捉している。また、サンプルは-0.243付近で集中しており、収束した事後分布を得ることができていると考えられる。縦方向での事後中央値は0.221となり、観測期間において風は北から吹く傾向であることを捕捉している。サンプルは0.221付近で集中しており、縦方向においても収束した事後分布を得ることができていると考えられる。

# 第 36 回「時空間データを用いた方向統計モデルの構築」

究を行われている加藤先生から、厳密な統計的解釈や統計理論の最先端の状況を教えていただきながら理論研究と分析を進められる点に、協働の意義と感じています」と応える。

### データドリブンな民間の研究動機が 理論研究者にも好刺激に

統計数理研究所は研究活動のみならず、産業界との連携を積極的に推進している。共同研究以外にも、教職員による面談を中心とした学術指導や、研究所内に社員の座席を確保する「データサイエンス・リサーチプラザ」など、相手企業のニーズに応じてさまざまなプログラムが選べることが特徴だ(https://www.ism.ac.jp/sangaku/sangakurenkei.html)。

こうした活動は、社会貢献を目的として 実施されるものである一方、教職員自身 にもよい影響を与えている。

加藤は、このプロジェクトで初めて民間 企業との連携を経験した。「若手の頃は 大学や研究所との共同研究に忙しく、民間との共同研究まで手を広げる余裕はありませんでした」と明かす。だが、統数研の同僚には産学連携に参加している研究者も多く、次第に経験してみたいと考えるようになったという。そのような中、石川さんから提案された研究が興味深かったこともあり、今回の共同研究に取り組むことを決めた。

「大学や研究所の理論研究者は、過去の理論研究の結果をベースにして次の研究テーマを設定することも多い。これに対し、石川さんは初めに実データから、データの特徴を抽出する統計手法を開発したい、というデータドリブンな研究動機がありました。そのようなアプローチから理論を構築するのは民間研究者らしい視点

で、私自身刺激になっています」と加藤 は感想を述べる。

プロジェクトの今後について、石川さんは「今後、情報技術が発展するにつれ、より多様で多くの情報が蓄積され、活用されていくことが考えられます。多様なデータに対しても有効に活用できるように、より次元を拡張した3次元単位球面などの方向データに対する統計手法の開発や応用にも取り組みたいと考えています。ゆくゆくはこのプロジェクトで得た知見を自社事業のサービス向上や社会貢献に役立てられれば」と抱負を語る。

加藤もまた、「今回のプロジェクトで扱う

データは、方向だけでなく、さまざまな変数が同時に観測されている『多変量データ』であると同時に、実数値ではない『非ユークリッドデータ』でもある。今後は、未発展な分野であるこれらのデータのための統計理論の発展を目指していきたい」と語り、「論文になりそうなテーマが見つかったことも、当初は予期していなかった収穫です」と微笑む。

統数研と民間企業の協働によって生まれたシナジーが、社会課題の解決に大きな貢献を果たすことに期待したい。

(広報室)

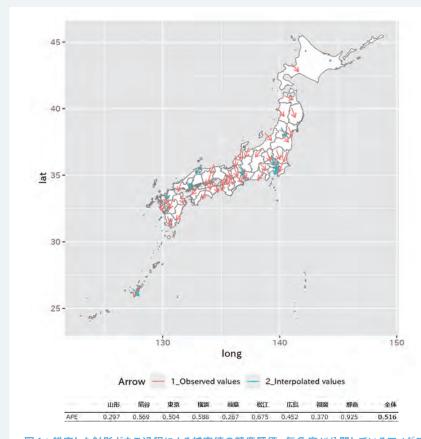


図4:推定した射影ガウス過程による補完値の精度評価。気象庁が公開しているアメダスの 都道府県別の38地点の最多風向データ (2022/4/1 ~ 4/7) から推定した射影ガウス過程を 用いて、2022/4/1 の検証地点9地点の風向を補完した。地図上の赤い矢印は観測された最 頻風向、青い矢印は射影ガウス過程のクリギングにて補完した最頻風向を意味する。下表は 検証地点9地点での補完精度を評価する評価指標APE (Average Prediction Error) の値。 値が小さいほど実際に観測された値に対して補完の誤差が小さいことを意味する。他の観測地 点から距離の遠い那覇では、他の地点と比較して補完精度は低い水準であることが分かる。

### シンポジウム報告



### 「Applied Statistical Analysis using R Software for Forestry Studies 2023」 共同開催報告

2023年11月15~17日の3日間に渡り、カンボジア森林研究所と統計数理研究所リスク解析戦略研究センターとの共催により、プノンペンの森林研究所にてワークショップ "Applied Statistical Analysis using R Software for Forestry Studies 2023" をコロナ後、初めて対面形式で開催しました。今回、統計ソフトRの初学者が参加者の半数以上を占めたため、1日目は、現地スタッフによりデータハンドリング等の統計ソフトRの基礎的な操作方法を講義しました。2日目以降、統計数理研究所側スタッフにより、グラフィックス

に関するコマンドの取り扱い方などを講義し、演習を通じて様々なグラフ作成能力を高めました。最終日である3日目には、回帰分析の基礎的な内容について講義したのち、Rによる実習を行いました。参加者には、過去のワークショップ経験者や既に独学でRの自習を行った者もおり、参加者同士で協力して問題解決に取り組む姿も見られました。なお、今回のワークショップの運営は、カンボジア森林研究所所長Dr. Heng Sokhを中心に行い、参加者は30名でした。

(吉本 敦)





### ワークショップ 「科学モデル論は何を目指せばよいのか | 開催報告

重点型共同利用研究「データサイエンスからみた統計数理科学と統計数理科学からみたデータサイエンス」の「科学哲学の視点からの統計教育:入門レベルの確立と発展レベルの探索」について、日本科学哲学会(12月3日、つくば)でワークショップ「科学モデル論は何を目指せばよいのか」を行いました。本研究所と科学哲学は古くからつながりが深く、元所長の故林知己夫氏や故赤池弘次氏も哲学系の学術誌に論文業績を有します。本代表の島谷は、公開講座「科学

哲学の視点からの統計学再入門」で本分担者でもある森元 良太講師から教わった「集団的思考と誤差論」など、学部 レベルの統計教育に生かせる科学哲学ネタを紹介しました。 ワークショップのオーガナイザー・松王政浩氏と尾崎有紀氏、 髙橋和孝氏(分担者、北海道大)は、それぞれ先端の機 械学習も視野に入れたモデル論について話しました。いず れも近い将来、大学院レベルの統計教育に生かせることを 期待させる内容でした。 (島谷 健一郎)



### 共同研究集会「極値理論の工学への応用」開催報告

2023年10月6日・7日に共同研究集会「極値理論の工学への応用」を開催しました。

今年度は、コロナ禍下での3回のオンライン開催を経た後のハイブリッド方式(対面+ZOOM配信)による開催で、会場での研究者同士のコミュニケーションの良さと来場できない方々の参加とを両立できたように思います。

講演は新たな推測やモデル化や分布の性質などの統計 的問題・確率場など数学的問題・土木系や生物系やOR など新たな方法や方面への応用があり、この分野がいまだ 大きく発展していることを感じさせるものでした。

今回は運営側の予想外のトラブルにより開催が危ぶまれま したが、統計数理研究所の加藤昇吾先生・栗木哲先生・ 椿広計先生をはじめとする先生方・スタッフの皆様・都立 大の吉羽要直先生らの献身的なご助力を得て、成功裏に 終えることができました。

ハイブリッド方式は予期せぬ障害が起こり困難も少なくありませんが、研究所の皆様の開催経験をもとに大過なく運営できました。ここに記して感謝申し上げます。

来年以降も開催を企画しておりますので、ぜひ皆様のご 参加をお願いいたします。 (山梨大学 西郷 達彦)



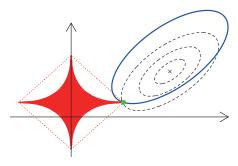
# 研究教育活動

# Repor

### 2023年10月-12月の公開講座実施状況

11月30日(木)に、当研究所の二宮 嘉行講師による「情報量規準 ~スパース推定や因果推論での活用も視野に~」の講義がオンラインにて行われました。統数研ならではのトピックとして、AICタイプの情報量規準を扱い、僅かに一般化した設定でAICを導出し、解析例や利用の注意点を挙げました。また、基本的なスパース推定や因果推論(傾向スコア解析)において、AIC元来の考え方で情報量規準を導くとどうなるかを紹介し、形式的なAICとは大幅に異なることを見ることで、情報量規準に関する研究の重要性を伝えました。

12月12日(火)~13日(水)に、国立環境研究所の深谷 肇一講師と早稲田大学の野村俊一講師によるリーディング DAT講座L-B「統計モデリング入門」の講義がオンラインに て行われました。現代のデータ解析で常用される手段として、一般化線形モデル・階層ベイズモデルを用いた統計モデリング、時系列データを柔軟に扱う状態空間モデリングの2つを解説しました。 (情報資源室)



公開講座「情報量規準 ~スパース推定や因果推論での活用も視野に~」資料より



### 統計数理セミナー実施報告(2023年11月~2024年1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11月~2024年1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11月 1日	瀧澤 由美	時空間位置推定法の基礎研究と極地での環境計測への適用研究
11月 8日	上野 玄太	極値時系列の状態空間モデル
11月15日	逸見 昌之	統計学のための無限次元情報幾何
11月15日	島谷 健一郎	動物のルート選択:GPS 軌跡というpresence-only data
11月22日	加藤 昇吾	Weighted estimation with copula-based divergence
11月22日	川崎 能典	歪対称誤差分布による財務データのモデリング
11月29日	小山 慎介	スケートボードにおける重心運動の最適化
11月29日	池田 思朗	自己矯正法とALMA望遠鏡の画像化アルゴリズム
12月 6日	江村 剛志	コピュラの概要とベイズリッジ回帰
12月13日	三分一 史和	直線プラズマの周波数モード発現メカニズムの解明:多変量自己回帰モデルを利用した因果関係の定量化
12月20日	João Pedro Pedroso	Improving kidney exchange programs
2024年 1月10日	野間 久史	WedderburnのQuasi-likelihood理論に基づく多変量解析と不完全データの解析
1月24日	伊庭 幸人	Wカーネル:事後分布の情報を利用した頻度論の展開

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)



### 統計数理研究所夏期大学院「データ同化夏の学校」の開催

統計数理研究所夏期大学院として、2023年8月7日から10日までの日程で2023年度データ同化夏の学校を開催しました。データ同化夏の学校は、データ同化に関わる若手研究者の育成と実践的な技術の普及を目的として毎夏開催されており、今年で27回目となります。今年は青森県むつ市のプラザホテルむつで開催し、59名が現地に集まった他、36名がオンラインで参加しました。

今年は、初日の夕方に開校式と自己紹介を行い、2日目、3日目に基礎講義、研究発表、データ同化に関するプログラミング演習を実施しました。最終日には、演習の成果発表会を行い、各グループの代表者が演習の成果を発表し合いました。演習では各自が自分の設定した課題に熱心に取り組んでおり、また研究発表や成果発表会では活発に意見が交わされ、盛況のうちに終わりました。オンライン参加者への対

応も含めて多大なご助力、ご支援をいただいた日本海洋科学振興財団、会場の確保にご支援いただいたデータサイエンス共同利用基盤施設、および演習のチューターや講師としてご協力いただいた諸先生方に感謝申し上げます。

(中野 慎也)



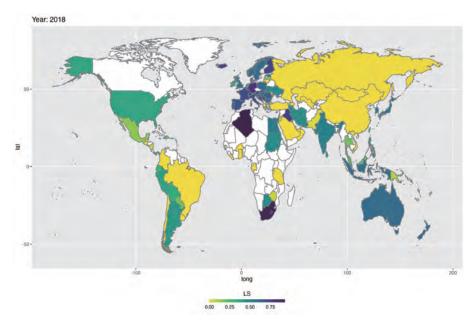


### 「探索的ビッグデータ解析と再現可能研究」実施報告

統計数理研究所統計思考院2023年度公募型人材育成事業(ワークショップ)に採択された「探索的ビッグデータ解析と再現可能研究」が、2023年8月27日(日)にオンラインで開催され、176名の参加者を得ました。

3年目となった今年度は、地道正行先生・阪智香先生 (関西学院大学)、宮本大輔先生・小池祐太先生(東京大学)のご講演による、財務・金融関連のビッグデータ処理関連のトピックに加え、生物医学データ・医療用ビッグデータの前処理に関連して植木優夫先生(長崎大学)、荒牧英治先生・西山智弘先生(奈良先端科学技術大学院大学)に講 演いただき、応用領域のカバレッジは更に広がりました。また、SQLを用いた大量データ分析の高速化に関して海外浩平先生(HeteroDB)に講演をお願いし、計算技術的側面からのトピックも充実したものとなりました。また、「次の一歩を踏み出すためのtidyverse入門」と題するチュートリアル講演を瓜生真也先生(徳島大学)にお願いしました。

3年続けたことで、公募型人材育成事業ワークショップとしては一区切りと考えており、次年度は研究寄りにシフトした研究集会を企画中です。 (川崎 能典)



地道正行先生·阪智香先生の講演スライドから。世界120カ国の非上場企業の労働分配率を世界地図で可視化。

# Report

### 統計思考院人材育成ワークショップ「生物多様性と群集動態:定量化の数理と統計的推定法」開催報告

統計思考院人材育成ワークショップ「生物多様性と群集動態:定量化の数理と統計的推定法」を、筑波大学菅平センターの協力のもと、長野県菅平高原にあるプチホテル・ゾンタックにおいて11月8-10日に開催しました。計45名の参加があり、そのうち26名が大学院生でした。野外データと統計モデルのほか、観察されなかった種数の推定法、生物多様性プロジェクト、方程式系から導出される生物多様性

維持機構、野外調査・統計・数理モデルをセットにする授業実践、草原生態系でのエクスカーション等々、提供される話題は多岐に渡り、質疑は延々と続いて食事時間を遅らせました。24時間稼働のポスター会場は文字通り24時間営業でした。写真を撮る余裕もないほど参加者それぞれ議論に夢中で、解散後の上田駅でようやく電車利用の人と集合写真を写せた次第です。 (島谷 健一郎)



## 統数研トピックス



### 数学・数理科学5研究拠点合同市民講演会「数学・数理科学の未来!?」を開催

2023年11月18日(土)に数学・数理科学5研究拠点合同市民講演会「数学・数理科学の未来!?」をハイブリッドで開催しました。南和宏副所長の司会のもと、椿広計所長が開会挨拶を行い、続いて文部科学省研究振興局大学研究基盤整備課の柳澤好治課長からのビデオメッセージを公開しました。

最初の講演は「拡散現象の不思議」と題し、明治大学 先端数理科学インスティテュートの俣野博研究特別教授より アラン・チューリングのアイデアやその歴史的意義について お話をいただきました。次に「秘密計算と数学」と題し、九 州大学マス・フォア・インダストリ研究所の縫田光司教授より秘密計算についての紹介や、その技術への数学の活用 についてご講演いただきました。休憩を挟み、京都大学数 理解析研究所の牧野和久教授より「楽して計算するには 一計算を科学する一」と題し高速アルゴリズム設計の意義や 重要性などについてご講演いただき、大阪公立大学数学研 究所古澤昌秋特任教授からは「先のことはわからない?!」 と題しご自身の研究生活や数学の未来についてお話いただ きました。最後に本研究所の鎌谷研吾教授より「ベイズ計算・対称性と非対称性」と題しベイズ統計学についての説明とその対称性・非対称性の最新動向についてご講演いただきました。各講演の最後には会場とオンラインから質問が寄せられ、活発な質疑応答が交わされました。

本講演会にはオンライン424名、会場30名と合わせて 454名の多くの方にご参加いただき、数学・数理科学の最 先端の研究の一端に触れていただきました。 (広報室)





### 江村剛志教授が台湾の国立中興大学を訪問

2023年11月19日から23日まで、江村剛志教授(データ科学研究系)が、台湾の国立中興大学の応用数学系と統計学研究所に招待されました。同大の応用数学系は1961年に設置され、台湾で最も歴史が長い応用数学系です。その中の統計学研究所(大学院)は、統計学を専門とする教授5名、准教授2名、博士学生4名、修士学生39名を有します。

招待者はTsung-Jen Shen (沈宗荏) 教授で、同大学の 教員と大学院生と学術交流を行いました。Tsung-Jen Shen 教授は、生物多様性・生体統計で著名な統計学者で、その他、有限混合モデル・多変量解析で著名なTsung-I Lin教授、同大の副学長 Yin-Tzer Shih教授、応用数学系主任 Jung-Miao Kuo教授らと講演会・夕食会で交流しました。 江村教授は統計学研究所の大学院生を対象とした講演を、20日 (演題「Overview of copulas and Bayesian ridge estimator」) と22日 (演題「Factorial survival analysis for treatment effects under dependent censoring」) に行いました。 (江村 剛志)





### 村上大輔准教授が2023年度坂下賞を受賞

村上大輔准教授 (データ科学研究系) が2023年12月 9-10日に開催された第37回応用地域学会研究発表大会 にて、2023年度坂下賞を受賞しました。本賞は、地域科学 研究の発展に顕著な貢献をした満40歳以下の若い研究者 を顕彰することを目的としたもので、時空間データの統計解析に関する一連の研究活動が評価され受賞に至りました。

(村上 大輔)



### 統計数理研究所子ども見学デー2023を開催

2023年10月29日(日)統計数理研究所子ども見学デー2023が開催されました。立川市と一般社団法人立川観光コンベンション協会主催の「立川体験デジタルスタンプラリー2023」との同時開催で、4年ぶりの対面イベントとなる本イベントには73名の来場者を迎えました。

ハニカムペーパーを使ってグラフを立体にする「統計グラフの立体工作キット体験」や、数字や漢字で書かれた数の単位でかるた取りをする「カードゲーム"大きな数"」、10万個のBB弾でサンプリング実験をする「白玉・黒玉数当て実験」、キャラクターの「ぬり絵」のコーナーが設けられ、多くの

家族連れに楽しみながら統計数理に親しんでもらう機会となりました。 (広報室)





# 日本学術会議見解「大学における数理・データサイエンス・AI教育の中での統計科学の教育について」の公表について

第25期日本学術会議数理科学委員会数理統計学分科会は、2023年9月26日に見解「大学における数理・データサイエンス・AI 教育の中での統計科学の教育について」を公表いたしました[1]。

同分科会は本研究所の教員、名誉教授6名を含む22名から構成されています。2014年に提言「ビッグデータ時代における統計科学教育・研究の推進について」を公表していますが、この10年の間のデータサイエンス、AIを巡る情勢の激変にともない、改めて統計科学の教育についての提言をまとめたものです。構想から、執筆・編集、査読対応、受理・公表までに約1年の時間を要しました。

本見解の主張は以下の通りです。(1) 数理・データサイエンス・AI分野の理論的基礎としての統計科学の重要性、

(2) 数理・データサイエンス・AI分野の再教育(リスキリング)の推進、(3) 学士課程及び大学院教育が必要とする統計教員の育成の必要性、(4) 初等・中等教育における教材、ソフトウェア、デジタル環境の整備と統計教育充実の必要性。

また付録では、(i) ビッグデータの統計解析の実例、(ii) 外国におけるデータサイエンス教育の例、(iii) 諸外国の教科書と日本の指導要領、(iv) 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モデルカリキュラム、(v) 統計関連の学会と科学研究費の状況、などについてリサーチした結果を参考文献、URLとともに付しています。

[1] https://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-25-k230926-24.pdf (栗木 哲)

# 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)
i2	Domain-Independent Dynamic Programming (DIDP) solver	品野	勇治(統計数理研究所•客員教授)
	の大規模並列化		

(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	確率的推論によるデータ取得とデータ 解析の包括的最適化	R5.10.1~ R6.3.31	2,600,000	坂田 綾香 准教授
	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	機械学習による液相分離場の予測と設計	R5.10.1~ R6.3.31	5,200,000	吉田 亮 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
郭 一村	中国科学院大学・助教	深さと震源メカニズムの両方を組み込んだ高次 元時空間 ETAS モデル	2023.10.23~ 2023.12.22	庄 建倉 教授
Gianpaolo Zammarchi	University of Cagliari, Research Fellow in Statistics	気候変動に関する公開データセットの統計的モデリング	2023.11.15~ 2023.12.15	川﨑 能典 教授
Wei Zhang	東北師範大学・博士課程学生	RKHSに基づくノンパラメトリック回帰モデルにおける基底関数の選択	2023.10.11~ 2023.10.21	福水 健次 センター長
市原 沙也	国立遺伝学研究所・特任研究員	発生の進行に伴うクロマチン高次構造の構築メ カニズムの理解	2023.10.5~ 2024.3.31	栗木 哲教授
Stefania Gentili	National Institute of Oceanography and Experimental Geophysics, Researcher	実用的な地震特性を把握するための統計・物 理ベースのリアルタイムツールの開発	2023.11.6~ 2023.11.18	庄 建倉 教授
園田 桂子	内閣府・政策統括官 (経済財政分 析担当) 付参事官 (総括担当) 付	公的データと民間データのリンケージによるミクロ データ構造化とリンケージによる効果の分析	2023.11.2~ 2024.3.31	山下 智志 教授
詹 成祥	中国地質大学・大学院生	データ科学に基づく地震予測の高度な精度向 上を目指す統合的アプローチ	2023.12.1~ 2024.11.30	ウ・ステファン 准教授
João Pedro Pedroso	University of Porto, Associate Professor	腎臓生体肝移植における再照会のもとでの最 大期待値マッチング	2023.12.18~ 2023.12.22	池田 思朗 教授
田沼 巌	株式会社ドワンゴ	多様なユーザアイテムフィードバックデータモデリ ングの研究	2023.11.2~ 2024.3.31	松井 知子 教授
安藤 誠	国立天文台ハワイ観測所・特任研 究員	大規模観測データに基づく銀河進化の統計的 研究	2023.11.2~ 2024.3.31	栗木 哲 教授
李 永波	中国地震局地球物理研究所・大学 院生	ベイズ型重力インバージョンにおけるモンテカル 口法の応用	2023.12.20~ 2024.2.13	庄 建倉 教授
Jessica Elizabeth Allen	University of Otago, PhD Course Student	ETASモデルによる群発地震の識別	2024.1.3~ 2024.1.18	庄 建倉 教授
胡 虞锴	香港中文大学・博士課程学生	誘発地震研究における統計的手法の応用	2024.1.11~ 2024.3.10	庄 建倉 教授
王 婷	University of Otago, Associate Professor	ニュージーランド、ヒクランギ沈み込み帯における 将来の巨大地震予測	2024.1.2~ 2024.1.19	庄 建倉 教授
Mark Stephen Bebbington	Massey University, School of Mathematical and Computational Sciences, Professor	ニュージーランド、ヒクランギ沈み込み帯における 将来の巨大地震予測	2024.1.9~ 2024.1.18	庄 建倉 教授

(研究推進課·共同利用係)



### 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
ワン カルロス ヒメネス・ソブリノ Juan Carlos Jiménez-Sobrino	Senior Professor	Dept. Interdisciplinary Mathematics Institute of Cybernetics, Mathematics and Physics, Cuba	客員教授	プラズマ物理における 確率微分方程式の同 定と非線形モデリング	2023.9.25~ 2023.11.24	三分一 史和 准教授
ミルボルトル アンドレ Myrvoll Tor Andre	准教授	ノルウェー科学技術大学	客員教授	統計的機械学習によるマルチモーダルデー タ処理	2023.10.6~ 2023.10.22	松井 知子 教授
シャベンコ パベル Shevchenko Pavel	教授	Macquarie University	客員教授	気候変動のための統 計手法の研究	2023.11.3~ 2023.12.13	村上 大輔 准教授
ダ・シルベイラ・ゴウベイア ジョアン エドワルド Da Silveira Gouveia João Eduardo	准教授	University of Coimbra	客員准教授	凸錐の幾何学につい て	2023.11.8~ 2023.12.7	Figueira Lourenco Bruno 准教授

(総務課·統数研人事·給与係)

# Repa

### 外国人客員紹介



### Juan Carlos Jiménez-Sobrino 客員教授

My main research focuses on the area of approximation and identification of stochastic dynamic systems and their applications. Specifically, I have contributed to the development of a new type of numerical integrators for various classes of differential equations, as well as inference methods for the estimation of these equations from actual data. I have been involved in applying these mathematical and statistical tools to solve complex problems in neuroscience, immunology and finance. As during my previous visit to the ISM, I will work with my host, Prof. Miwakeichi, on the identification of plasma models.



### ●Tor Andre Myrvoll 客員教授

今年度もまた2023年10月にTor Andre Myrvoll 先生(ノルウェー科学技術大学)が2週間ほど統数研に滞在された。Myrvoll 先生は2005年度以降、ほぼ毎年来所くださっており、時系列データの統計手法、および大規模データの計算・実装手法を駆使して、様々な統計的機械学習の方法・応用研究を幅広く共同推進している。今回は昨年度に引き続き、科研の挑戦的研究(開拓)「統計・機械学習による異分野相関を俯瞰する方法論の確立」に関連して、気候変動と経済の地域的な統合評価モデルの構築を試みている。推定パラメータ数が多いモデルなので、統数研のスーパーコンピュータを利用することを考えている。



### João Eduardo da Silveira Gouveia 客員准教授

It was with great pleasure that I visited the Institute of Statistical Mathematics for one month in November 2023, to work with Professor Bruno Figueira Lourenço. The institute provided excellent conditions conducive to a productive research effort. Our research focused on fundamental questions regarding the geometry of convex cones, leading to significant advances on multiple fronts. Not only was it a very fruitful visit with the prospect of several new papers, but it also marked my first visit to Japan. It was wonderful to have the opportunity to enjoy the culture and sights of Tokyo and the surrounding areas. I am sincerely grateful for this opportunity and eagerly look forward to revisiting ISM in the future.

# 会議開催報告



### 令和5年度第3回運営会議の開催

2023年12月22日(金)に、ハイブリッド形式で令和5年度第3回の運営会議が開催されました。

はじめに、研究系の改組や教育研究職員の選考につい て協議され、原案どおり承認されました。

引き続き、研究所関連として、統計エキスパート人材育成 プロジェクトの中間成果や、子ども見学デーをはじめとする 研究所の主な事業について報告がありました。

その後、改組後の統数研に期待する役割と将来を見据 えた人事企画のあり方、研究所ひいては国内における若手 研究者・外国人研究者を増大するための取組みや現在の 問題点等について、活発な意見交換が行われました。

(統数研事務課)

# 刊行物



### Research Memorandum (2023.11~2024.1)

No.1226: Iwata, T., Mixture of linear and non-linear Hawkes processes and its application to real earthquake sequences

(メディア開発室)



### 統計数理 第71巻 第2号

特集「諸科学における統計数理モデリングの拡がり川」

「特集 諸科学における統計数理モデリングの拡がりⅡ」について
横山 雅之99
行動意思決定研究者の読んだ 『確率の出現』 とその示唆 [研究ノート]
広田 すみれ101
生態学におけるモデルと法則に関する科学哲学論考の意義 [研究ノート]
島谷 健一郎119
果樹豊凶の決定論的非線形予測 ―アンサンブル再構成とLorenz類推法による1年先収量予測― [研究ノート]
酒井 憲司129
力学系を組み込んだ変分自己符号化器による顕微鏡動画モデリング:ゼニゴケ精子の運動を例として [原著論文]
近藤 洋平、南野 尚紀、上田 貴志149
位相的データ解析による銀河分布の定量化とバリオン音響振動抽出 [原著論文]
竹内 努、河野 海、クレ スチェータ、 西澤 淳、 村上 広耶、 馬 海霞、 本武 陽一
(メディア開発室)



# Annals of the Institute of Statistical Mathematics Volume 75, Number 6 (December 2023)

Conor Kresin and Frederic Schoenberg

Parametric estimation of spatial-temporal point processes using the Stoyan-Grabarnik statistic
Hengfang Wang and Jae Kwang Kim
Statistical inference using regularized M-estimation in the reproducing kernel Hilbert space for handling missing
data911
Kuangnan Fang, Jingmao Li, Yaqing Xu, Shuangge Ma and Qingzhao Zhang
Gene-environment interaction analysis under the Cox model931
Xiuli Wang, Jingchang Shao, Jingjing Wu and Qiang Zhao
Robust variable selection with exponential squared loss for partially linear spatial autoregressive models949
Chihiro Watanabe and Taiji Suzuki
A goodness-of-fit test on the number of biclusters in a relational data matrix979
Bruno Ebner, Adrian Fischer, Norbert Henze and Celeste Mayer
Goodness-of-fit tests for the Weibull distribution based on the Laplace transform and Stein's method1011
Yuqing Ma, Peijie Wang and Jianguo Sun
Estimation of complier causal treatment effects with informatively interval-censored failure time data1039

(メディア開発室)

### ●2024年度公開講座(一般講座)

一般社会人・学生を対象に、下記の公開講座 (一般講座)をオンラインにて開催いたします。健康は医療健康データ科学に関わる人材育成事業の推奨講座です。

### A 推移行列モデル:個体群生態学とベイズ推定・ MCMC

日時:6月4日(火)10時~16時(5時間)

講師:島谷健一郎(統計数理研究所)

申込受付:4月15日(月)10時~4月22日(月)10時

受講料:5,500円(税込) 会場:統計数理研究所

定員:80名(応募者多数の場合は抽選)

講義レベル: 初級

### (内容)

- 1. 野外生物の個体数の変動を行列とベクトルで表す
- 2. 個体群行列を作ることでわかってくること1: 個体群成 長率、安定分布、繁殖価
- 3. 個体群行列を作ることでわかってくること2: 感度分析
- 4. 行列要素の推定1:最尤法
- 5. 行列要素の推定2:ベイズ統計とMCMC (Markov chain Monte Carlo)
- 6. MCMCと数理と行列モデル

### (受講者に期待する予備知識やレベル)

線形代数の中の固有値と固有ベクトル

教科書:

資料を配布します。

### B 多変量解析法健

日時:9月5日(木)~6日(金)、9月9日(月)~10日(火)10時~ 16時45分(23時間)

講師:馬場康維·清水信夫(統計数理研究所)今泉忠(多摩大学)

申込受付:7月22日(月)10時~7月29日(月)10時

受講料:22,000円(税込)

会場:オンライン

定員:50名(応募者多数の場合は抽選)

講義レベル:初級

### (内容)

多くの現象は一つの変数で観測されるものではなく多数の 変数の観測、即ち多次元のデータによって把握されます。得 られた多次元のデータを用いて、数量の推測をする、判別を する、あるいは尺度を作る等の手法の総称が多変量解析法 です。

この講座では、重回帰分析、判別分析、主成分分析、因子分析、数量化など、多変量解析の古典的・標準的な手法の解説をします。また、クラスター分析、共分散構造分析などの解説をします。

### (受講者に期待する予備知識やレベル)

平均、分散、標準偏差等、統計学の基礎的な概念を 知っていることを前提とします。手法の数学的な説明より は、用い方に重点をおいた解説を行いますが、大学初級 程度の線形代数と微分積分の知識があると理解しやすく なります。

教科書:

資料を配布します。

●一般講座の他、系統的な講座編成により現代的な統計 科学の姿を示す「リーディングDAT講座」や、医学・ 健康科学分野のデータサイエンスを基礎から最先端まで 詳説する医療健康データ科学研究センターの公開講座 も開講予定です。

詳細は、以下のwebサイトをご覧ください。

https://www.ism.ac.jp/lectures/kouza.html

(情報資源室)

# column 163

# 機内の過ごしかた

上野 玄太

モデリング研究系

海外出張では当然、飛行機に乗る。飛行機は安全な乗り物であるが、旅特有の心細さを感じられるところがよい。心細さとは、もし落ちたらどうしよう、自力で助かることは絶対に不可能だ、といった想像のことである。

中学生の頃、機内の同乗者と、部屋を片付けずに出発してしまった、もし落ちたら、まずいものが明るみに出てしまう、でも今さらどうしようもない、という話をしたことがある。現在でも自分はオフィスや自宅を片付けずに海外出張に出ているあたり、まったく成長が見られていないが、無事に帰国することの強い意思表示である、と前向きに解釈することにしている。

部屋の片付けまでは至らないが、出国時にはある程度の整理はついている。「ある程度」とは、出張前の会議やセミナーの日程を調整し、依頼されている書類を提出し、研究発表の材料を揃え、ポスター印刷を完了し、持参する論文や本を厳選する、ということを指す。厳選する必要が出てくるのは、出張期間中は自由時間がたっぷりあるような気がするため、あれもこれも読みたいと積読状態であった文献をピックアップしていくと大量になるからである。また、本気で読むためには電子ファイルではなく紙でないといけない。

セレクションをくぐり抜けた文献は、収納先が2通りに分かれる。手荷物に入れて機内に持ち込むか、スーツケースに入れて預けるかである。優先順位の高いものから手荷物に入れる。書き込みや手計算をしていかないと頭に入らないし再読も困難になるので、赤ペンも手荷物に入れる。経験上、スーツケースに入った文献は、出張先では手に取らないことが多く、復路の機内持ち込み荷物に割り当てられない限り読まれずに帰国する。

さて、搭乗し、ドアが閉まってシートベルトを締める。ノートパソコンやテーブルは使用できない時間帯なので、ここで、優先順位1位の文献を手にし、赤ペンを片手に読み始める。優先順位1位は、論文ドラフトであったり論文であったり教科書であったりする。論文の査読は後回しである。調べ物ができないので、分かりにくいところがあってもその文献にどんどん目を通す。机はなく、手持ちであるために赤ペンの文字はふにゃふにゃになるが構わず書き込む。そのうちにシートベルトの着用サインが解除されるので、テーブルを引き出して続けて読む。これで手計算が可能になる。この時点で、シートモニターで映画を見ないことが肝要である。それは、ここが出張期間を通じて最も集中できる時間であるためである。復路のフライトでは疲労と気の緩みのためか集中力に欠ける。往路のここが勝負の時間と言ってよい。

つづいて、機内食の準備時間である。飲み物のサービス が始まり、食器やテーブルクロスが配られる。そのため、手元 の文献は座席前のポケットにしまい、テーブルを空ける。テー ブルクロスが敷かれたら、テーブルの上での作業は一時中断 である。ところが、そこから実際に機内食が出てくるまでには 意外に時間がある。この時間が重要である。ただ座って待っ ているだけの、ある意味で不自由な状態なのだが、何も見ず に考えられることに思いを馳せていることが多い。これまでに よくわからなかったことや、考えなくてはいけないことへの閃き が得られることがある。そんなときは、テーブルの向こうにしまっ た紙とペンを取り出してメモを取る。確かに書き留めた充実 感に浸ったところで機内食が出てくれば完璧である。



今回の行先はデンバーである。スーパーコンピュータの会議 と展示会が同時に開かれる。



統数研のデータ同化スーパーコンピュータシステム (HPE Superdome Flex)。展示会で紹介した。

統計数理研究所ニュース No.163 (2024年2月25日発行)

発行: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 広報委員会

〒190-8562 東京都立川市緑町10-3

TEL 050-5533-8500(代) https://www.ism.ac.jp/

制作:統計科学技術センター

●送付先変更は、Fax(042-526-4334)にてお知らせ下さい。

