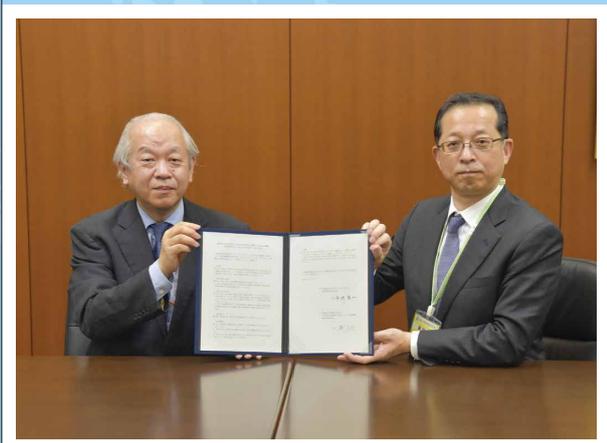


▼ CONTENTS

- 02 響き合う人とデーター 統数研プロジェクト紹介
第24回「新型コロナウイルス対応プロジェクト」
- 06 研究教育活動
新入教員紹介／特任教員紹介
2020年10月ー12月の公開講座実施状況
統計数理セミナー実施報告(2020年11月ー2021年1月)
2020年度若手研究者クロストークの開催報告
- 09 統数研トピックス
統計数理研究所オープンハウスを開催
オープンハウス連携イベント「データサイエンスの現在と未来を解剖する! ~企業向け産学連携シンポジウム~」の開催
総務省統計研究研修所と統計数理研究所との連携に係る協定の締結
統計数理研究所子ども見学デー 2020を開催
- 11 総合研究大学院大学複合科学研究科統計科学専攻関係
大学院説明会について
- 11 お知らせ
一般講座
- 12 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／寄附金の受入れ
- 12 会議開催報告
国際アドバイザーボード：オンライン会議の開催
- 13 刊行物
統計数理／Annals of the Institute of Statistical Mathematics
- 14 コラム

統計数理研究所 ニュース

大学共同利用機関法人 情報・システム研究機構



縦横無尽の統計解析で 感染の実態と要因を炙



▲松井知子教授

新型コロナウイルスの対策に 立ち上がった統計学者たち

「新型コロナウイルス感染症のパンデミックという前例のない、国と世界の危機的事態を前に、統計数理やデータ解析を専門とする研究者として何かできないか、一人ひとりが真摯に考え、迅速に行動に移しました」。統計数理研究所における「新型コロナウイルス対応プロジェクト」のリーダーを務める松井知子教授はそう振り返る。

2020年1月、国内で初めて新型コロナウイルスの感染者が確認された。その2か月後には世界保健機関（WHO）がパンデミック（世界的大流行）を宣言。社会不安が全世界を覆った。

そんなとき、いち早く動き出したのが世界の統計学者たちだ。松井も、その一人だった。椿広計所長に相談し、すぐに統数研として独自にプロジェクトを起し、この未曾有の感染症に立ち向かうことを決定。「プロ

ジェクトの趣旨に賛同する研究者たちが、所内外から自然に集まりました」と椿は言う。現在は疫学モデル、地理、データ計測、検定などさまざまなジャンルを専門とする20人を超えるデータサイエンティストたちがプロジェクトを支える。

とはいえ、誰もが経験したことのない事態だ。どこにどんなデータがあり、どれが信頼できるものなのか。まずはデータを探るところから始まった。解析結果を対策に生かすためには、統計数理だけでなく制度設計の知見なども必要。各自が専門知識を持ち寄り、多角的な視点で進めなければならない。オンライン会議システムを使って当初は毎週行っていた打ち合わせでは、「3時間で議論が尽きることはまれで、5時間に及ぶこともあった」（松井）。今も隔週で情報交換を続けている。

感染症対策の鍵を握るのは、正確な状況把握だ。統数研のプロジェクトでは、東京都公式ホームページ、Jag Japan社などの情報提供ページ、JX通信社の有料ページなどから陽性者数、死亡者数、PCR検査数などのデータを日々収集。また、感染拡大に関係すると考えられる人の移動や密集、土地利用のデータなども集め、ミクロからマクロまでさまざまな階層でデータ分析を進めている。

ホットスポットの正確な範囲や 感染の広がり方を分析

プロジェクトの研究テーマの一つは、「東京都市区町村ごとの感染状況の解析」だ。統数研の鈴木和幸客員教授は、都が市区町村の累積陽性者数のデータを公表し始



▲椿広計所長

めた4月1日以降、1日も欠かさずホームページから手作業で数字を拾い、パソコンに入力してデータを収集、集計してきた。

市区町村ごとに時系列で陽性者数の推移をグラフ化すると、都全体のグラフからは見つけにくい現象を捉えることができる。例えば、4月から8月までの10万人あたりの陽性者数の推移（7日間の移動平均）を折れ線グラフにしたところ、都全体では陽性者が漸増して8月2日にピークを迎えるのに対し、新宿区では6月16日の週を境に急増し、都全体よりも半月早い7月19日（単日では16日）にピークを迎えていたことが明らかになった（図1）。

ものづくりの品質管理や信頼性を専門とする鈴木は「製品の開発・生産工程、検査、市場クレームの関係は、コロナ禍における日常生活、検査態勢、医療体制に置き換えて考えることができます。品質管理の手法が役立つのではないかと思います」と

第24回 「新型コロナウイルス対応プロジェクト」

世界中を不安と混乱に陥れた新型コロナウイルス感染症。2020年3月、統計数理研究所は「新型コロナウイルス対応プロジェクト」を立ち上げた。所内・所外から有志の統計学者約20人が自発的に集まり、さまざまなデータ解析を行っている。

り出す



▲鈴木和幸客員教授

話す。

鈴木が集計したデータは、他にもさまざまな分析に活用されている。岡山大学の石岡文生准教授が手がけるのは、「陽性者の空間集積性に関する分析」だ。空間集積性分析とは、地理・空間データにおいて局所的に高い、または低い観測値を示す場所を統計的根拠に基づき評価するもので、疫学や環境学、犯罪学などさまざまな分野で活用されている。この手法では、高い観測値を示す場所を「ホットスポットクラスター」と呼ぶ。

プロジェクトでは、都内にホットスポットが存在するかどうか、存在するとすればどの範囲までかを明らかにする手法を開発した。

メディアなどでよく目にする累積陽性者の色分け地図は、階級区分の決め方によっていかようにも変わるにもかかわらず、濃い赤色などで示された部分がホットスポットであるかのような印象を与えがち。統計的根拠に

基づき、客観的に判断できる情報はまだまだ世の中に出回っていないのが実情だ。

そこで、プロジェクトでは「空間スキャン検定」の手法を用いて日々のホットスポットの検出を試みている。分析対象地域なら東京都53市区町村の場合と23区に限定した場合、また陽性者数も累積陽性者数と新規陽性者数に対し、その地区の居住人口や労働人口、人流などを加味するなど多面的に分析。空間スキャン検定も数種類を適用している。

この分析から、23区の南西部に有意なホットスポットが検出された。また4~6月と7月以降では傾向に変化が見られ、後者では新宿区や中野区を中心とする少領域がホットスポットとして検出された。

多摩大学の久保田貴文准教授は、統計ソフトRのパッケージ「shiny」を用い、任意の状況下でのホットスポットを検出するためのアプリケーションを開発。これにより、実際の陽性者数の多いエリアだけでなく、その周辺も含むエリアがホットスポットであることが分かった。こうした結果は、政策決定の際のエビデンスともなる。

椿もまた、鈴木の収集した日々陽性者数データを基に、「1人の患者が何人に感染を広げるか」という「実効再生産数の簡易推定」を自ら手がけている。

陽性者数は検査で陽性と判定された人数であり、感染していても検査を受けていない人の数は含まれない。PCR検査数が少なかった当初の陽性者数は、実際の感染

10万人当たりの東京都の主な区の陽性者数(7日間の移動平均)

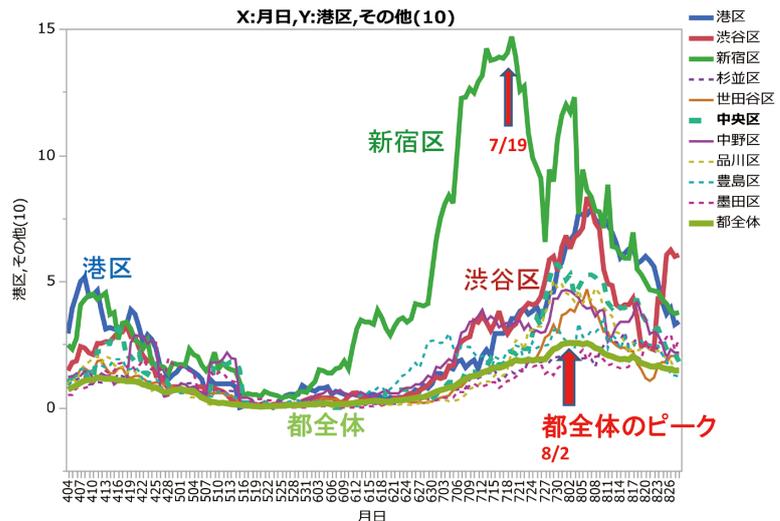


図1：東京都と主な区の陽性者数の推移。新宿区では都全体のピークより約半月前に陽性者の急増が始まっていたことが分かる。



▲下野寿之外来研究員

者数より大幅に少なかったと考えられる。つまり、PCR検査の普及とともに陽性者数が激増したからといって、感染者自体の数が急が増えたとは限らない。「本当の感染者数のうち陽性者数の比率は4月と現在で大きく変化していても、日次変化率は感染者数と陽性者数ではあまり変わっていないとも考えられます」(椿)。

陽性者数の報告は曜日や休日の影響などで大きくばらつくため、まず移動平均値(一定期間の平均値)のような平滑化処理を行い、日次陽性者数の対数変化率を求めたうえで実効再生産数を推定した。

松井は同様に東京都市区町村の陽性者数の日次データを利用し、4月から8月までの各週の陽性者数をマルコフ連鎖でモ

デル化。各週のモデルの近さを検定することによって、週ごとの特徴を分析した(図2)。これにより、「クラスター感染が起きて急に陽性者が増えた場合」「市中感染でじわじわと陽性者が増えた場合」などを区別することが可能になり、感染状況の解析に役立つ。

携帯電話の位置情報から人口密度と感染の関係を解析

このほか、プロジェクトでは「滞留人口データを用いた感染状況の解析」も手がけている。データの可視化を得意とする統数研の下野寿之外来研究員は、NTTドコモから「500mメッシュの空間に1時間に何人がいるか」という携帯電話の位置情報データの提供を受け、新しい集計方法を考案。都道府県ごとに滞在者がどのようなメッシュにいるか、100分位点、20分位点、4分位点に分かる形にグラフ化した(図3)。これによって特定の都道府県において、過密したメッシュに人が出入りしていることが分かり、密集の影響を考慮した新型コロナウイルス感染症対策を考える際の参考になる。

また、統数研の村上大輔助教は、商業地の人口密度と陽性者数のデータから、「人口密度が陽性者数に及ぼした影響」を分析している(図4)。その結果を見ると、3月上旬の段階ではまだほとんどの地域で「密であること」の影響があまり見られないものの、3月末の3連休を経て全国的に密



▲村上大輔助教

な地域が増え、陽性者が急増。その後は緊急事態宣言に伴う自粛により全国的に密であることの影響が少なくなったとはいえ、東京などでは密度が陽性者数の増加に影響を与える傾向が続いていることが分かった。村上は、人口密度ばかりでなく、どんな要因が絡んでいるかの解明に注力している。

刻々と発表される世界の論文を読み解いていく

新型コロナウイルス感染症が流行して以降、膨大な数の学術論文が出版されている。疫学・生物統計を専門とする統数研の船渡川伊久子准教授は、医学総合誌の「NEJM(ニューイングランド・ジャーナル・

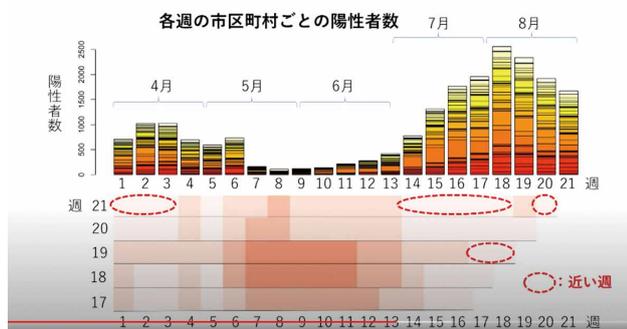


図2: 各週の市区町村ごとの陽性者数。棒グラフのヒートマップは各市区町村を表す。下のグラフがモデルの近さを検定した結果。各週を互いに比べ、差が小さいほど薄い色で示した。例えば21週目は、1~3週目、14~18週目、20週目と特徴が近いことが分かる。

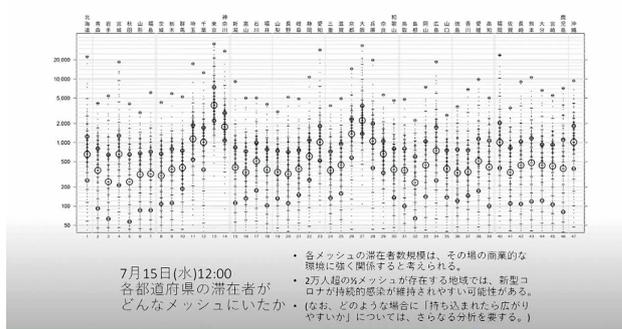


図3: 全国の都道府県ごとの過密状況。縦軸は500mメッシュにいる人数、横軸は都道府県。各都道府県の滞在者がどのようなメッシュにいたかを示す。北海道、東京、神奈川、愛知、大阪、福岡は、2万人以上のメッシュが存在する。



▲船渡川伊久子准教授

オブ・メディシン)」「ランセット」、科学総合誌の「サイエンス」「ネイチャー」などのトップジャーナルに次々と公開される論文を読み解いている。

「各国で感染状況が刻々と変化するなか、査読前論文や有力研究グループの論文が脚光を浴びました。重要な政策決定に感染症数理モデルが大きな役割を果たし、モデリング研究が多数見られます」(船渡川)。

厳格な抑制政策を採った中国、死亡者数が増え続けた米国、厳格でない行動制限でも死亡者数を抑えている日本など、政

策や感染状況には地域差があるため、一般化にも注意が必要だ。「例えばマスクの着用など、科学者間でさえ意見がさまざまで、まだ決定的な結論が出ていない問題が多いのが今回の特徴です」(船渡川)。

ちなみに、マスクの問題については、米国大統領選をめぐる共和党支持者と民主党支持者の意見が対立したことから、それぞれの支持層の多い州の陽性者数を分析する研究を鈴木が進めている。

新たな感染症にも有効なエビデンスとなる中核技術へ

統数研では、ここで紹介した分析研究や文献調査の結果などの成果を日本疫学会や日本感染症学会など関連学会に限定して情報提供していく方針だ。ただ、新型コロナウイルス感染症は現在進行形で状況が変化しており、研究も発展途上であることから、誤った情報を提供することには大きなリスクがある。そこで、統計の専門家である統数研の所員による事前レビューを受けるためのレビューサイトを作成。レビューが完了した成果は、学会報告や公開用サイトを通じて順次発信していく予定だ。

すでに、研究内容の一部は文部科学省科学技術試験研究委託事業「数学アドバンスイノベーションプラットフォーム

(AIMaP)」が2020年10月31日に開催した研究交流会「感染症に立ち向かう数理科学」で発表した。また、今年に統計数理研究所和文誌「統計数理」の特集「公衆衛生(新型コロナウイルス感染症)」や、MDPIの発行するオープンアクセスジャーナル「エントロピー」の特集号にも論文を投稿する予定だ。

さらに、プロジェクトから派生した「COVID-19のマルチスケールな時空間解析と情報提供基盤の構築」と「確率的機構モデルを用いた新型コロナウイルス流行分析」の2件の研究が、情報・システム研究機構(ROIS)のCOVID-19対策研究プロジェクトに採択された。

松井は「このプロジェクトで開発した統計数理の技術が、新型コロナウイルスに限らず、今後発生するであろう新たな感染症対策の中核技術となるように、メンバー一同が協力して活動を進めていきたい」と抱負を語った。(広報室)

新型コロナウイルス感染拡大防止のため、本インタビューはオンラインで行われました。

人口密度の影響 (3月上旬)

都市部で一定の影響みられるが、まだ人口密度はそれほど陽性者数に影響していない

赤：人口密度は陽性者数に影響

青：人口密度は陽性者数に影響しない



人口密度の影響 (3月末)

3連休(3/20-22)などを経て、人口密度の影響が増大。陽性者急増

赤：人口密度は陽性者数に影響

青：人口密度は陽性者数に影響しない



図4：人口密度が陽性者数に与えた影響。3月上旬にはまだ人口密度が陽性者数に与えた影響は小さかったが(左)、3月末の段階では人口密度の影響が増大した(右)。

新入 1 教員紹介



松井 茂之

データ科学研究系 教授

2020年7月にデータ科学研究系教授と医療健康データ科学研究センター長を拝命しました(名古屋大学とのクロス・アポイントメント制度)。ご存知の方も多いかと思いますが、わたくしは、その昔、統数研に勤めていたことがあり(2008年7月から2013年3月まで)、この度、ご縁があって再びお世話になることとなりました。私がいた当時、ちょうど広尾から立川への移転があり、大きな環境の変化があったことが特に印象に残っています。ほう、Mt. Fujiがこんなに大きくみえるのか! その後しばらくして立川を離れましたが、7-8年ぶりに戻ってみると、しっかりこの地に根を下ろした落ち着きを感じました。誰もいない、まっすぐの遠くまで見通せる廊下に立ったとき、思わず時の流れを感じました。

本来なら、懐かしい先生方との再会、そして、新しく来られた先生方、職員の方々との出会いをじっくり楽しむところからはじめたかったですが、リアルの場合は、新型コロナの嵐が過ぎ去った後の楽しみにとっておきたいと思います。皆様方くれぐれもお気をつけください。改めまして、今後ともどうぞよろしくお願いいたします。

新入 2 教員紹介



鎌谷 研吾

モデリング研究系 准教授

大阪府の北の外れにある、北摂地方から来ました。山と川と古墳に囲まれた北摂地方は大変のどかで、通学路にはタヌキやイノシシが出ます。通勤ラッシュのときでさえそれほど混まない北摂と異なり、ウィルス流行中の東京での電車移動は神経をすり減らします。そのなかで、昭和記念公園の緑や、研究室から見る美しい富士山には癒やされます。季節につれて変化する風景が楽しみです。専門はベイズ統計学やそのモンテカルロ法です。最近やや応用確率論にシフトしており、変化を求めてこちらに異動しました。統計数理およびデータサイエンスの日本の中心拠点である統計数理研究所の諸先生方の知識を吸収し、ベイズ統計学への研究による貢献と、人材交流・育成によって今後の発展に寄与したいと思います。

新入 3 教員紹介



奥野 彰文

統計思考院 助教

2020年10月から統計思考院に在籍しております。数理統計を背景に統計的機械学習の研究を行っています。これまでは主に、グラフ埋め込みや相関分析といった特徴量学習の数理的な解析をしていました。

近年では機械学習分野が目覚ましい発展を遂げています。様々なデータ解析手法が提案されている一方で、数理的な観点から見れば未解決の重要な問題が多数残っていて、なぜその手法がうまくいくのか、またより良い推定を行うためにはどうすればよいのか、考えるべき問題は尽きません。

統計思考院では様々な先生とお話させていただく機会があり、日々刺激を受けながら過ごしています。新型コロナウイルスの影響で研究所の方々と直接お会いする機会がほとんど無いのは残念ですが、新時代の生活様式に順応しつつ統計学の発展に寄与できればと考えております。よろしくお願いいたします。

新入 4
教員紹介



服部 公平

統計思考院 助教

2020年9月に着任しました服部公平です。博士号取得後にイギリスで2年、アメリカで4年のポスドクを経て、国立天文台テニュアトラック助教として統計思考院に在籍しています。私の専門は天文学で、特に人類の住む銀河系の構造や歴史を、人工衛星や望遠鏡の観測データを用いて解析しています。具体的には、星の軌道運動をもとに銀河系の暗黒物質の3次元分布をベイズ推定したり、銀河系中心の超巨大ブラックホールが弾き出した星を異常検知アルゴリズムで探索したりといったように、統計学の手法を天文学に応用した研究を行なっています。天文学に新たな統計手法を導入することを目指し、統計数理研究所での研究生生活を通じて様々な統計手法・解析手法に関して見識を深めたいと思っております。コロナ禍の中で気軽に研究の議論が行えない日々が続きますが、研究所の環境を最大限生かして新たな研究の方向性も模索していきたいと思っておりますので、ご指導ご鞭撻のほどよろしくお願いいたします。

新入 5
教員紹介



白崎 正人

統計思考院 助教

2020年10月1日より、国立天文台より出向して統計思考院の一員として勤務しています。私は「観測的宇宙論」と呼ばれる天文学の一領域に興味があります。観測的宇宙論では、天文データの統計解析により、宇宙の起源や構造の進化を明らかにすることが課題です。よってデータ解析の重要性は年々高まっています。私はこれまで宇宙に遍く(全宇宙の物質の8割)存在しつつも光らない「暗黒物質」と呼ばれる謎の物質の正体を解明するために、数百万の銀河の写真を解析し、暗黒物質の地図を描く研究を進めてきました。暗黒物質が観測された銀河の手前にあると、暗黒物質の生み出す重力によって、背景にある銀河はわずかに歪みます。このことを用いて、銀河の歪みから暗黒物質の分布を逆算するのです。この推定問題は、データ科学的な側面が強く、まだ天文業界は知らない(しかし統計数理的には自明な)画期的な手法があるかもしれません。皆さんとの共同研究の可能性にワクワクしています。よろしくお願いいたします。

特任
教員紹介
1



小池 孝明

リスク解析戦略研究センター 特任助教

リスク解析戦略研究センター特任助教に着任いたしました、小池孝明と申します。以前は博士課程の学生としてカナダのUniversity of Waterlooという所におりました。修士課程では慶應義塾大学に所属しつつ、スイスのETH ZurichでVisiting studentをしていました。私が学部生であった頃に統計数理研究所出身の先生方との交流が多くあり、こうして今研究所の一員となれたことを大変嬉しく思っております。専門は統計学および金融リスク管理で、特にリスク尺度、リスク分配、接合関数、従属性尺度の研究を行ってきました。今後の研究に関しましては、分配資本の計算手法や不確実性の下でのリスク分配の理論、和が定数となる分布の解析とシミュレーション、従属性尺度の様々な観点からの比較や公理的特徴付け、裾従属性尺度の一般化などを考えております。皆様とご一緒に研究できる機会を楽しみにしております。何卒よろしくお願いいたします。

青木 祐太

ものづくりデータ科学研究センター 特任助教



物質・材料研究機構でのポストドク、日東電工株式会社での研究員を経て、2020年7月に統計数理研究所に着任しました。元々の専門は第一原理計算を用いた物性物理学・物質科学研究でしたが、学位取得後は材料科学分野にデータ科学的手法を適用するマテリアルズ・インフォマティクス (MI) 分野の研究に取り組んでいます。当初はMIのことを単なる新しい流行の研究分野の一つとしか認識していませんでしたが、現在は科学・技術における研究開発活動のあり方そのものを変革するかもしれない大きな潮流であると感じています。そのような変革の最前線に自身が実際に立ち会い参画することができることを、大変光栄に思っています。第一原理計算という演繹的な手法と、データ科学という帰納的な手法、これらのある意味対極に位置する2つのアプローチを組み合わせ、物性物理学や物質科学の新たな展望を切り開いていきたいと思えます。さらには、これらの分野固有の課題に対応するための、データ科学手法の改善・開発にも取り組んでいきたいと思えます。この分野の研究には、自分が今まで漠然と関心を抱いていた事柄が、全てひとつにつながるような大きな可能性を感じています。

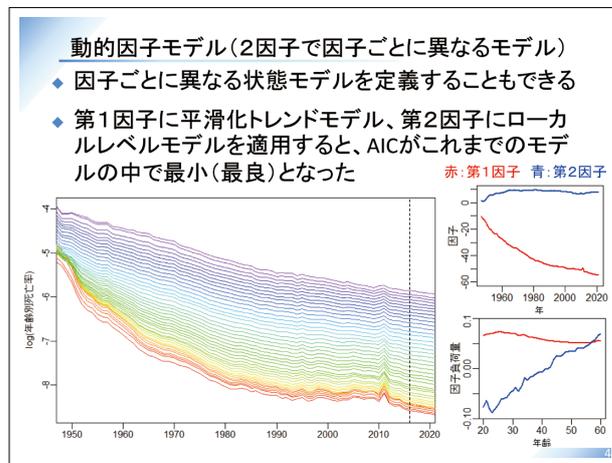
2020年10月-12月の公開講座実施状況

2020年10月29日(木)に、理化学研究所の瀧川一学講師による「決定木とアンサンブル学習の基礎と実践」講義がオンラインにて配信されました。この講座は前年度3月に開催予定でしたが、新型コロナウイルス感染拡大の影響により、中止としたものをオンラインとして開講しました。オンライン講座の開催は公開講座では初めての試みでしたが、大きな問題もなく講義を終えることができました。

11月19日(木)に、当研究所の藤澤洋徳講師による「ロバスト統計～外れ値への対処～」を対面講座にて開催しました。定員を通常の100名から40名に減らし、受講生間の距離も2m間隔となるよう配慮し、新型コロナウイルス感染拡大防止に最大限の注意を払った上での開催でした。受講生からも質問がたくさん寄せられ、少人数ではありましたが大変活発な講座となりました。

12月17日(木)～18日(金)に、国立環境研究所の深谷肇一講師と当研究所の野村俊一講師によるリーディングDAT講座L-B1「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。オンライン講座を2日間連続で行うのは初の試みでした。講義中に疑問に思う事があればすぐに質問ができるような対応としたため、受講生からは大変多くの質問が寄せられ、講師は講義中だけでは回答できないほどの盛況な講座となりました。(情報資源室)



リーディングDAT L-B1講座資料より

統計数理セミナー実施報告(2020年11月～2021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。2020年11月～2021年1月のセミナーは下記の通り行われました。新型コロナ感染拡大防止のため、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
2020年11月 4日	郭 中梁	ベイズ推論に基づく合成経路探索
11月11日	芝井 清久	社会不安と国民性に関する国際比較分析
11月11日	本武 陽一	学習済み深層ニューラルネットワークからの解釈可能な物理法則抽出
11月18日	吉本 敦	空間制約下における野生動物コリドー及び食草域形成と森林資源管理に対する離散最適化モデリング
11月18日	三分一 史和	自発リズム活動を生成するミニマムニューロンネットワークの推定

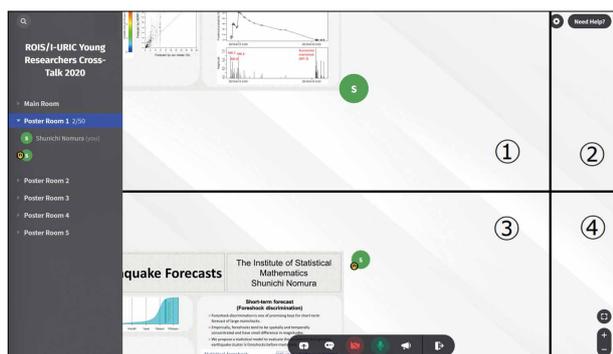
日程	氏名	タイトル
2020年11月25日	南 和宏	機械学習と差分プライバシー
12月 2日	間野 修平	調整尤度によるプラグイン推定量の漸近バイアス削減の幾何
12月 2日	川崎 能典	市場制度変更が金融取引に及ぼす影響の高頻度データに基づく分析
12月16日	小山 慎介	状態空間モデルを用いたCOVID-19実効再生産数の推定
12月16日	加藤 昇吾	球面上のコーシー分布のパラメータ推定
12月23日	持橋 大地	確率的潜在意味スケールリング
12月23日	逸見 昌之	診断研究のメタアナリシスにおける公表バイアスの最悪評価と感度解析
2021年 1月13日	伊庭 幸人	自由エネルギーは暗号化できるか? —計算量制限と熱力学
1月13日	島谷 健一郎	個体群生態学における状態空間モデルの1事例
1月27日	中野 慎也	アンサンブル変分法とその地磁気永年変化予測への応用
1月27日	清水 信夫	集約的シンボリックデータにおける変数の統計量の統一的な表現

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)

2020年度若手研究者クロストークの開催報告

2020年12月16日、オンライン交流スペース“Spatial chat”において、大学共同利用機関法人4機構連携企画の2020年度若手研究者クロストークが開催されました。昨年度まではオンサイトでの開催でしたが、今年度は新型コロナウイルスの感染拡大防止の観点からオンラインでの開催となりました。



今年度は“Spatial chat”を利用したポスターセッションを企画し、大学共同利用機関法人4機構と総合研究大学院大学の大学院生から計20名の方にオンラインでのポスター発表を行っていただきました。参加者たちは各ポスター画像の前で発表者の話を聞き、ディスカッションも交えて盛況な場となりました。さらに、今年度から初めて、参加者からの投票によってポスターセッションの中から最も良い発表を決めるベストポスター賞を設け、研究者と大学院生からそれぞれ1名ずつの表彰者を選出しました。今回は初めてのオンライン開催ということで、過去の企画の踏襲ができない中、ゼロベースから企画立案することとなりました。反省点も多くありましたが、今回の試みがクロストークの将来の発展のためにとてもよい前例になることを期待したいと思います。来年のクロストークも楽しみにしていただければ幸いです。

(野村俊一)

統計研トピックス

統計数理研究所オープンハウスを開催

2020年10月27日にオープンハウス「データサイエンスが創る日本の未来」を開催しました。今年度は新型コロナウイルス感染症拡大防止のために、例年6月開催のところ時期を遅らせ、初めてオンライン形式での実施となりました。

公開講演会はYouTube Liveで配信し、伊藤聡副所長の司会のもと、始めに統計思考院長 川崎能典教授から「統計思考力育成事業による人材育成～統計思考院の取組紹介」と題した講演が行われ、公開講座等のリカレント教育や、共同研究スタートアップ等の産学連携、若手研究者の育成、DS教員の育成事業等の最近の活動が紹介されました。続いて国立天文台 先端技術センターの宮崎聡教授による「す



川崎教授の講演

ばる望遠鏡超広視野カメラによる最新天体観測とデータ解析の現場」と題した講演が行われ、宇宙の加速膨張の謎について、すばる望遠鏡の実際の観測画像を用いながら、最新の観測とデータ解析の現場の様子が紹介されました。最後に滋賀大学データサイエンス学部長の竹村彰通教授による「ウィズコロナ時代のデータサイエンス」と題した講演では、数理モデルの役割とPCR検査の精度、接触確認アプリの利用数と現状の限界について説明し、データの活用が不可欠とお話しされました。最後に椿所長より挨拶があり、講演会最初の挨拶動画が音声の不調のため配信されないトラブルは

ありながらも本講演会は107名の方にご視聴いただきました。

その他にも統数研の研究紹介として研究内容ポスターのPDFが88名分、内容を紹介する動画が18本公開されました。新型コロナ対応プロジェクトの動画や統数研と研究系の紹介をするミニレクチャー動画も公開し、多くの視聴がありました。

統計よろず相談室はZoomで9件の相談を受け、総研大統計科学専攻の大学院説明会はZoomで行い、21名が参加しました。
(広報室)

オープンハウス連携イベント「データサイエンスの現在と未来を解剖する! ~企業向け産学連携シンポジウム~」の開催

2020年10月26日に、統計数理研究所オープンハウス2020の連携イベントとして、シンポジウム「データサイエンスの現在と未来を解剖する! ~企業向け産学連携シンポジウム~」を、統数研と連携を深めてきた大学、企業のデータサイエンスの専門家を迎えて開催しました。当日は、北村浩三特命URAが司会を務め、椿広計所長の開会挨拶の後、中央大

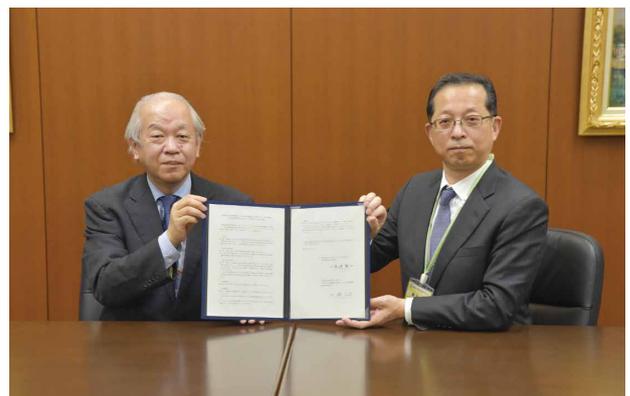
学 AI・データサイエンスセンター所長、理工学部経営システム工学科の樋口知之教授による「汽水型人材育成：製造業における人材育成の課題と加速」、パナソニック株式会社水野洋氏による「企業からデータサイエンスへの期待と要望」、さらに、椿所長による「統計科学の産学連携に果たすべき役割と統計数理研究所」の各講演が続きました。講演後は休憩を挟み、樋口教授がモデレータを務めるパネル討論を行いました。話題はコトビジネス、オープンイノベーション、エコシステム、マテリアルズ・インフォマティクス等から成長戦略まで広がり、データサイエンスを駆使した産学連携の経験知・成果・可能性等について、人材育成の観点を含めた率直な意見交換が行われました。本シンポジウムは新型コロナウイルス感染拡大防止のためオンラインで開催し、配信したYouTube Liveは394名の視聴があり大変盛況なイベントになりました。開催に当たり、ご講演者の皆様、ご後援としてご協力いただきました各機関・団体の皆様に深く御礼申し上げます。
(URAステーション)



パネル討論での椿所長

総務省統計研究研修所と統計数理研究所との連携に係る協定の締結

2020年12月1日(火)、総務省統計研究研修所より、平池栄一所長、伊達隆之次長、長尾伸一統括教授、高月由美子管理課長、塚尾聖調整係長が統計数理研究所を訪れ、連携に係る協定の締結式を行いました。本協定の目的は、総務省統計研究研修所と統計数理研究所が、統計技術の研究及び統計人材の育成を通じて公的統計の発展に努め、我が国における妥当な意思決定及び合理的な行動の変化に資するため、円滑に連携協力を図ることです。締結式では、平池所長と椿所長のご挨拶に続き、協定書への署名、記念写真の撮影を行いました。その後、施設の見学として、



左:椿所長、右:平池所長

オンサイト施設(データサイエンス共同利用基盤施設内)、赤池ゲストハウス、セミナー室、統計思考院を順次ご覧いただきました。平池所長をはじめ統計研究研修所の皆様は、担当

者の説明に熱心に聞き入っておられました。

(URAステーション)

Report

統計数理研究所子ども見学デー2020を開催

2020年12月26日～2021年1月11日に統計数理研究所子ども見学デー2020を、立川市の事業協力を得て、オンラインで開催しました。

例年は会場で体験していた10万個の白色と黒色のBB



弾の中から黒玉の数を調べる「白玉・黒玉数当て実験」と研究者による実験の解説を、統数研 YouTube チャンネルで動画公開しました。実験の動画と稲垣佑典客員准教授による「BB弾サンプリング実験動画の解説」は合わせて700回近くの視聴回数があり、多くの方にご覧いただきました。

「スーパーコンピュータ見学ツアー」動画は900回以上の視聴回数があり、「じゃんけんゲーム」Web版や、「日本人の国民性調査ミニクイズ」、統計用語に関する「謎解きチャレンジ」もたくさんの方に体験していただきました。

本イベントは冬休みの期間中に合わせて開催し、遠方の方にも統計学に触れていただく機会になりました。(広報室)

総合研究大学院大学複合科学研究科統計科学専攻関係

Report

大学院説明会について

2020年10月27日(火)に、2020年度第2回大学院説明会を開催しました。

新型コロナウイルス感染症の影響によりオンライン開催と

なったため、例年実施している教員との面談は行わず、カリキュラムの説明と入試ガイダンスのみの説明会となり、参加者は21名でした。(総務課・大学院係)

Information

●一般講座

一般社会人・学生を対象に、下記の公開講座をオンラインにて開催いたします。

① Rによる極値統計学

日時:5月25日(火)10時～16時45分(予定)
講師:西郷 達彦(山梨大学)
申込受付:4月5日(月)10時～4月12日(月)10時
受講料:5,000円
定員:100名(応募者多数の場合は抽選)
講義レベル:中級

極値統計学は偶然によっておこる現象の測定値の最大を推測する学問であり、土木工学・品質工学・保険・金融など多くの分野でリスク管理等に応用されている。この講義で

は極値統計学の基盤となる数理的な背景から始め、複数のデータ取得デザインに対応した推定法と、Rのパッケージを用いた実際の解析まで解説する。

受講者に期待する予備知識やレベル:

学部教養レベルの確率統計とRの初歩を知っていることが望ましい。確率統計は主要な確率変数と分布の性質および推定としてモーメント法と最尤法の導入部分ぐらいの知識が必要であり、Rは単回帰とそのプロット程度ができるとうい。

参考書:

- 1.Rによる極値統計学 西郷 達彦・有本 彰雄 オーム社 2020
- 2.極値統計学 高橋 倫也・志村 隆彰 近代科学社 2016

お
知
ら
せ

B 統計学を哲学する：統計はなぜ科学的な根拠になるのか

日時：7月29日(木) 10時～16時45分(予定)
 講師：大塚 淳(京都大学)
 申込受付：6月14日(月)10時～6月21日(月)10時
 受講料：5,000円
 定員：100名(応募者多数の場合は抽選)
 講義レベル：初級

統計学はなぜ、科学的根拠たりうるのだろうか。単にルーチンな統計処理を行うのではなく、その結果を正しく解釈し、結論を導くためには、統計学の数理的基盤のみならず、その哲学的前提を正しく理解する必要がある。本講座では、統計学を存在論的／認識論的観点から捉え直すことで、推測統計の基本的枠組みやp値などの基本概念を「腑に落とす」ことを目的とする。

受講者に期待する予備知識やレベル：

統計学の初歩的な知識があることが望ましい。

参考書：

●大塚淳(2020)『統計学を哲学する』(名古屋大学出版会)

●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「系統的教育コース・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)

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
いのちを支える対策推進センター 代理理事 清水 康之	行政における統計データの利活用の推進に関する研究	R2.8.3～ R3.3.31	5,000,000	樫 広計 名誉教授
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 彰博	数理知能表現による深層構造学習モデルの革新	R2.11.1～ R3.3.31	15,600,000	数理・推論研究系 福水 健次 教授

(研究推進課・研究推進係)

寄附金の受入れ

受入決定年月日	寄附者	寄附金額	担当教員	寄附目的
R2.12.9	公益財団法人カンオ科学振興財団 理事長 櫻尾 隆司	1,000,000	坂田 綾香	研究テーマ「ベイズ推定に基づく適応的グループテストのアルゴリズム開発:検体検査の効率化と誤り訂正」に対する研究助成

(財務課・予算・決算第二係)

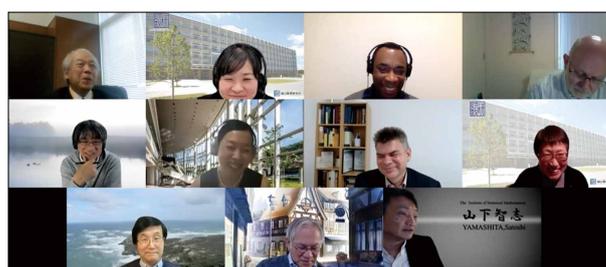
会議開催報告

国際アドバイザーボード：オンライン会議の開催

2020年12月18日、統計数理研究所国際アドバイザーボードのオンライン会議が開催されました。本来ならば、各委員を統数研に招へいしての開催を計画していましたが、コロナ禍のため、初の試みとしてオンラインで開催することになりました。

各地域で時差があったものの、シンガポール国立大学のYing Chen准教授、アメリカペンシルバニア州立大学のDonald Richards特別教授、オーストラリアマッコーリー大学のPavel Shevchenko教授、オーストラリア国立大学のAlan Welsh教授の4名の出席が得られました。

統数研からは、樫所長、伊藤副所長、山下副所長、宮里副所長、松井(知)教授、栗木教授が出席。樫所長が研究所全体、栗木教授が数理・推論研究系、そして松井教



授がリスク解析戦略研究センターについての活動状況の紹介を行いました。2時間半強にわたる本オンライン会議では、現在の活動状況の紹介の質疑応答、討論のみならず、今後統数研が注力すべき分野や取り組むべき新しい研究分野、国際連携の推進の仕方、SDGsやCOVID-19のデータ研究

解析について等について、国際アドバイザリーボードの各委員から、各委員の経験や海外での例を交え、率直な意見や助言を得ることができ、貴重な実り多い機会となりました。

(国際連携推進室)

刊行物

統計数理 第68巻 第2号

特集「Institutional Researchと統計科学」

「特集 Institutional Researchと統計科学」について

本多 啓介	195
日本におけるIRの動向：経営IR、教学IRから研究IRの誕生と推移 [総合報告]	
山田 礼子	197
学術文献DBにおける著者識別のためのトピックモデリングの利用とその性能比較 [総合報告]	
藤野 友和、濱田 ひろか	209
トピックモデルを用いた研究動向の分析 [原著論文]	
武井 美緒、藤野 友和、中野 純司	219
大規模大学における研究分野の研究実績の可視化 [原著論文]	
船山 貴光、山本 義郎、藤野 友和	233
学術分野における論文および統計学論文の引用状況について [研究ノート]	
張 菱軒、潘 建興、中野 純司	247
学術文献DBを用いた共著分析によるIoT研究における異分野融合の国際比較 [原著論文]	
水上 祐治、中野 純司	265
<hr/>	
グループ正則化に基づく順序ロジットモデルにおける隣接クラスの統合 [原著論文]	
永沼 瑞穂、吉川 剛平、川野 秀一	287

Annals of the Institute of Statistical Mathematics Volume 72, Number 6 (December 2020)

Natalie Neumeyer, Leonie Selk and Charles Tillier

Semi-parametric transformation boundary regression models	1287
Elena Hadjicosta and Donald Richards	
Integral transform methods in goodness-of-fit testing, II: the Wishart distributions	1317
Byungsoo Kim and Sangyeol Lee	
Robust estimation for general integer-valued time series models	1371
Abhijit Mandal	
An optimal test for the additive model with discrete or categorical predictors	1397
Paul Doukhan, Ieva Grublytė, Denys Pommeret and Laurence Reboul	
Comparing the marginal densities of two strictly stationary linear processes	1419
Yan Cui, Qi Li and Fukang Zhu	
Flexible bivariate Poisson integer-valued GARCH model	1449
D. Benelmadani, K. Benhenni and S. Louhichi	
The reproducing kernel Hilbert space approach in nonparametric regression problems with correlated observations	1479
Mineaki Ohishi, Hirokazu Yanagihara and Shuichi Kawano	
Equivalence between adaptive Lasso and generalized ridge estimators in linear regression with orthogonal explanatory variables after optimizing regularization parameters	1501

(メディア開発室)



通勤時間に広がる世界

Stephen Wu

データ科学研究系

二年ほど前から、天気の良い日には私は必ず立川駅から歩いて統数研に行くことが習慣になっている。そして、この通勤時間にはいつもポッドキャストを聞いている。今年はネイチャーポッドキャストにはまっていた(図1)。特に研究のためということではなく、ネイチャーポッドキャストは科学について様々な情報が流れており、面白い話題がいっぱいあるからだ。もちろん、統数研の池田思朗教授が活躍された史上初ブラックホールの撮影成功も話題になっていた。一日の仕事を始める前に、綺麗な緑と建物の中をゆっくりと散歩しながら、世界中の研究者たちの色々なテーマについての議論を聞くことは、いつも私に研究への活力をくれる(図2)。

2020年は人類歴史において深い傷を負った年と言えるだろう。ネイチャーポッドキャストでも3月20日からコロナポッドという特別なシリーズが始まった。いつも最新のコロナ情報を流す中で、特に5月15日のコロナポッドは深く印象に残った。テーマは「誤情報のパンデミックとの戦い」だった。世界中の人々がコロナの大流行により苦しんでいるようで、実際は誤情報の流行によって苦しんでいるというものだ。

「誤情報」は今、我々にとって一番の敵と言えるかもしれない。「科学」を名乗った検証していない情報や科学を否定するような情報を拡散され、コロナ対応策についても様々な意見が社会から聞こえてくる。このような「戦い」の中心にいる一人の人物はアメリカのトランプ前大統領である。ネイチャーポッドキャストは政治的なモチベーションはないが、科学者それぞれのスタンスはあるため、政治的な話は特に回避しない。例えば、10月7日と11月3日の「Trump vs. Biden: what's at stake for

science?」と「Stick to the science: when science gets political」の放送では割と客観的に科学と政治の関係性を議論した。一部の主流メディアに「反科学」というラベルをつけられたトランプ前大統領は自分に都合の良い仮説を真実として発信し、科学に大きいダメージを与えたとされる。さらに、コロナの出現は彼の「活躍」の起爆剤になり、誤情報のパンデミックは科学世界に注目を集める話題となった。

私の認識では、科学のコアは観測可能性と再現性である。ある現象や問題を説明するために、仮説を立て、観測データを集め、再現実験を行い、自然法則を発見する。特定の仮説に対して、完璧な観測データと再現実験は必ず存在するとは限らない。統計科学は不完全な観測データや有限の実験資源の中で最も論理的な結論を出すための、想像と現実世界を繋ぐ橋である。しかし、使い方によっては、悪用されることも十分ありえる。今の時代はインターネットの発達で、情報拡散の量と速度は格段に進歩した一方で、誰でも発信するために情報の質は保証できないと言える。厳密に検証された科学情報より、適当な検証で正しそうな誤情報の方が拡散しやすいと言われている。このような誤情報パンデミックは統計モデルで研究されている。この興味深い研究は、今まで個人の研究世界だけで満足していた自分には重要な反省となった。科学者の役目として、研究だけではなく、教育と社会貢献もとても大事なことであると考えたからだ。そのような面においても日本の統計界を牽引してきた統計数理研究所に所属していることをいかして、より広く見つめる視野を忘れずに研究を進めていきたいと思っている。

www.nature.com/nature/article?type=nature-podcast

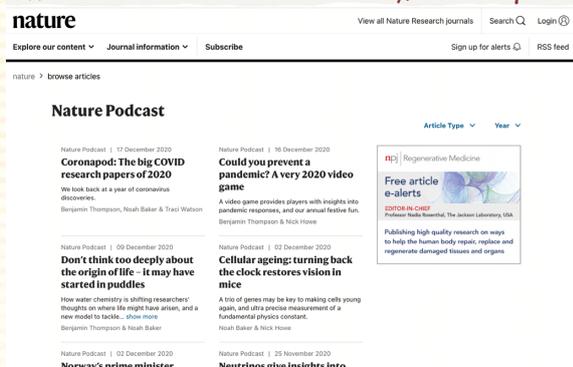


図1. ネイチャーポッドキャストのホームページ



図2. 立川駅から歩いて統数研に行く私