

▼ CONTENTS

- 02 響き合う人とデーター 統数研プロジェクト紹介
第32回「時空間データ解析プロジェクト」
- 06 研究教育活動
2022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(2022年11月～2023年1月)
2022年度 ROIS クロストーク開催報告
- 07 統数研トピックス
情報・システム研究機構「第4期戦略的研究プロジェクト」における統計学とプラズマ／
核融合科学の協調的發展を目指した研究者交流と討論会の報告
第1回「自己点検アドバイザーボード」を開催
有竹俊光特任助教と日野英逸教授が2021年度IBISML研究会賞を受賞
津山高校での出張講義を実施
模擬授業とインタビュー動画の公開
- 09 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
- 10 お知らせ
2023年度公開講座(一般講座)
- 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
- 12 会議開催報告
令和4年度第4回運営会議の開催
- 13 刊行物
統計数理研究所調査研究レポート
統計数理／Annals of the Institute of Statistical Mathematics
- 14 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

時空間データの解析手法を 社会課題の解決に寄与



▲村上大輔助教

幅広い時空間データの統計モデリング により各種要因の影響を分析

近年、センサー観測技術の発達により、さまざまな「時空間データ」が収集されるようになった。時空間データとは、位置情報と時刻情報を併せ持つデータのことで、例えば、人工衛星による観測では、気象はもちろん地表や地中の温度、森林・水域・都市域といった地表面の状態などが分かる。アメリカのGPSや日本の「みちびき」などのGNSS（全球測位衛星システム）からは、人やモノの正確な位置情報も取得できるようになった。

また、地図情報などのオープンデータ化も進みつつある。自由に地図を編集・利用できる「オープンストリートマップ」や、国土交通省が推進する3D都市モデル



▲松井知子教授

「PLATEAU（プラトー）」などはその一例だ。

さらには、即時的にデータ更新がなされる態勢も整ってきた。新型コロナウイルス感染症に関して言えば、厚生労働省や自治体によって、新規陽性者数や重症者数、死亡者数などのデータが日々更新されている。

統数研のリスク解析戦略研究センターに誕生した「時空間データ解析プロジェクト」では、こうした時空間データを基に、統計モデルを開発して解析し、社会リスクの要因を解明したり、今後の動向を予測したりする研究を幅広く手がける。これまで統計的機械学習研究センターの「都市インテリジェンス研究プロジェクト」で進めてきたテーマの一部を独立・発展させ

るべく設けられたプロジェクトだ。

スタート時点のメンバーは、都市インテリジェンス研究プロジェクトからリーダーの松井知子教授と村上大輔助教、そして「環境情報に対する統計解析手法開発プロジェクト」から船渡川伊久子准教授の3人がいずれも兼務で着任。時空間データ解析プロジェクトのリーダーは村上が務める。

時空間データに統計モデルを適用する意義とは何か。コロナの陽性者数の要因分析を例に挙げると、国や自治体などが状況に応じた政策を検討するとき、観測された生のデータだけでは役立てにくい。なぜなら、陽性者数がどのような要因の影響で日々上下しているかは、陽性者数データだけではわからないからだ。

また、自治体による報告の遅れや入力ミス、検査を受けていない無症状感染者などによって、観測される陽性者数は上下してしまい、必ずしも正確とは言えない点もまた問題である。

「統計モデルを当てはめることで、まず、密集や通勤、政策の影響といった説明可能な要因と、入力ミスなどのノイズをきちんと分ける。さらにそのうえで、説明可能な要因について解析を行えば、政策検討に役立つ知見を得ることができます」と、村上は説明する。

密集・密接によってどれだけ陽性者が

開発し、

新型コロナウイルス感染症や熱中症といった社会リスクへの対策を検討するうえで、人々の移動や場所による温度変化など、時空間データに基づいて要因を分析するニーズが高まっている。2022年8月、統計数理研究所のリスク解析戦略研究センターに誕生した「時空間データ解析プロジェクト」の研究内容の概要を紹介する。

増えたか。あるいは、通勤の抑制やマスクの着用でどれだけ減ったか。それらを定量的に評価することで初めて、有効な「次の一手」を考えるよりどころとなるのだ。

熱中症のリスク対策に役立つ「熱波」の発生を予測

時空間データ解析プロジェクトが、都市インテリジェンス研究プロジェクトの成果を踏まえ、これから取り組むテーマの一つが「熱波の予測」だ。

ここ数年、大きな社会問題となっている地球温暖化や都市のヒートアイランド現象は、時空間データ解析が求められる事例の一つ。機械学習による統計モデルの開発を得意とする都市インテリジェンス研究プロジェクトではこれまで、都市の気候などに関する時空間モデリングの方法について検討してきた。2013年から18年までの5年間、時空間モデリングに関する国際ワークショップを開催した実績もある。

熱波の予測には、気象庁の地域気象観測システム「アメダス」の観測データと、地球観測衛星から送られてくる地球表面温度の画像データを利用する。両方のデータを使うのは、それぞれに一長一短があるからだ。

というのは、アメダスの観測局は東京

都内に8カ所しかない。このため、気象観測データは1時間間隔ではあるものの、カバーできていないエリアが多く存在する。一方、衛星から取得された地表面の熱画像データは、1日あたり4時点のみであるが空間解像度は高い。その反面、雲で遮られるとデータが取得できず、曇りの日などまったく得られないこともある。

そこで両者を組み合わせ、日々の気温の空間分布を予測する。「さらに最新の統計モデルを使って解析する

ことで、都心部は高温になる確率が高い、沿岸部は温度のばらつきが大きいといったさまざまな特徴が見えてきます。そして、これを効果的に計算する手法が重要となります」と村上は話す（図1）。

時空間データ解析プロジェクトでは、都市インテリジェンス研究プロジェクトと連携して、この時空間モデルの効率的な解析・計算手法について研究を進める。

気温観測と衛星観測を用いた地表面温度の分析

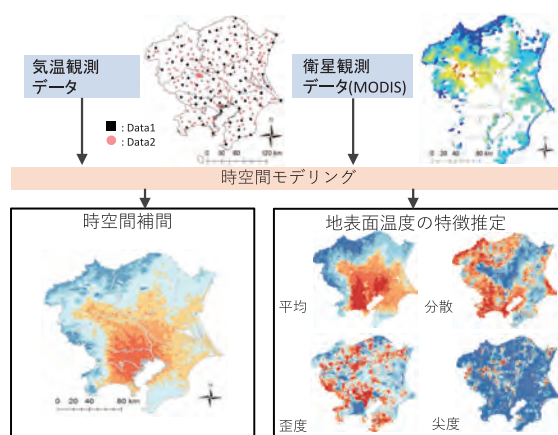


図1：上左図に示すように気温観測データは地点数が限られている。一方で上右図に示すように衛星観測データは、空間解像度は高いが、雲などにより観測値が得られる地点・時点は制限される。そこで両方を考慮した時空間モデリングを行った。極端な熱波などもとらえるために、Tukey g-and-hモデルを用いた。それにより地表面温度の時空間補間やその特徴（平均、分散、歪度、尖度）などを面的に推定した。

暑さに関連する“つぶやき”を気温予測の補正に活用

都市インテリジェンス研究プロジェクトではまた、Twitterデータから「ジメジメ」「ムシムシ」「熱中症」「だるい」「暑い」「水分補給」といった暑さに関連するキーワードを含むツイート（Heat-tweet）を抽出し、それがどんな気温のときにどんな場所で発生しやすいかを調べた。その結果を気温予測の補正に活用するのだ。

ただ、こうしたHeat-tweetはあくまでも個人の感覚に基づくものであり、本当に気温と連動しているかどうかは分からない。そこで村上は、Heat-tweetの起こり方をモデル化し、8月1カ月間の日変動と平均的な1日の時間変動の観測値と比較した(図2)。

その結果、Heat-tweetの起こりやすさは、気温の日変動とはおおむね一致。しかし、1日のうちで見ると、気温が最も高くなるのは午後2〜3時ごろであるのに対し、Heat-tweetは午前10時ごろに一番多くなっており、時間変動ではズレが生じていた。

Heat-tweetが起こりやすいのは、気温が高い時間帯ではなく、気温が急上昇する時間帯だ。「1時間前に比べてどれだけ温度が上昇したか」を表すグラフは、Heat-tweetの発生数のグラフとピークの位置がほぼ一致し、急に暑くなったときにつぶやきが増えることが分かった。

この結果を踏まえ、気象観測データを補間したところ、補間をしないデータだけ

を使った予測よりも実際に近い予測ができたという(図3)。

「東京都心の地表はコンクリートで覆われているので、郊外よりも温度が高くなります。しかし、少ない観測点のデータから予測した地表温度の分布図では、郊外の北側の温度が高くなってしまいました。これにHeat-tweetのデータを補間すると、都心ほど高く、妥当な結果となりました。実際には、この補間を効率的に計算する手法が重要となります」(村上)。

標準的な統計解析では、気象観測データとTwitterのようにまったく異なる複数種類のデータを一つのモデルで扱うことは少ない。気温データだけを使うのであれば、「正規分布」を仮定する基本的な統計モデルで十分であるし、Twitterデータだけでも特定の分布を想定すれば足りる。だが、二つのデータを組み合わせて予測しようとする、難易度が急上昇する。

都市インテリジェンス研究プロジェクトでは、気温データとTwitterデータの背後

に共通の空間分布パターンがあると仮定したうえで、機械学習でよく使われるガウス過程を応用することで気温の時空間予測を行った。

気温と暑さ関連Twitterの空間分布をできる限り正確に予測するようにモデルのパラメータを調整することで「気温は近い場所では類似する」「短時間で大きくは変わらない」「土面より舗装道路やコンクリート面のほうが高温になる」といった各種要因の影響度を推定した。それにより予測精度を高めている。

時空間データ解析プロジェクトでは、この研究成果を受けて、これらをさらに効率的に計算する手法を研究開発していく。

こうした手法は、さまざまな分野への応用も期待できる。プロジェクトでは例えば、万引などの犯罪がどういった場所で起こりやすいかを推定する「犯罪分析」や、住宅地価がどのように決まるかを推定する「不動産価格の要因分析」などにも着手している。

Heat-tweetと気温のプロット

Heat-tweetの起こりやすさ

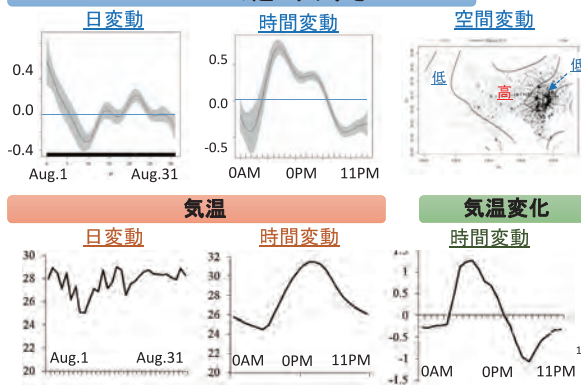


図2：Heat-tweetの起こりやすさをプロットした上の図と、気温と気温変化をプロットした下の図を比較すると、日別では気温がHeat-tweetに相関していることがわかる。一方で、時間別で見ると気温よりもむしろ気温変化がHeat-tweetと強く相関していることが確認できる。

Heat-tweetの具体的な活用例

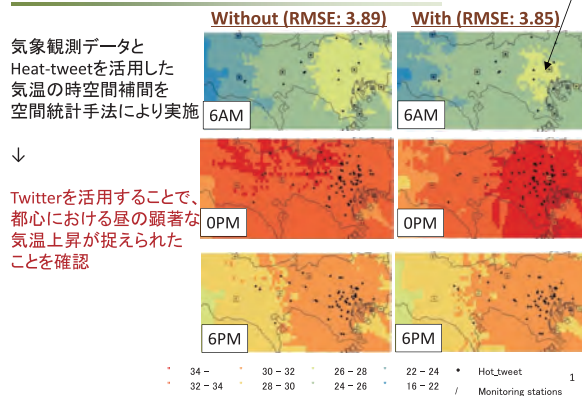


図3：気象観測データとHeat-tweetを活用した気温の時空間補間を空間統計手法により実施した。Twitterを活用することで、都心における昼の顕著な気温上昇が捉えられた。例えば中段左は気温観測データだけを使った気温分布の予測結果で、右はHeat-tweetも考慮した結果だが、後者の方が都心付近の気温上昇が捉えられていることが分かる。

従来手法との計算時間の比較

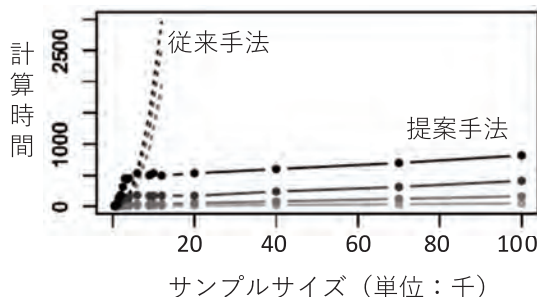


図4：従来手法ではサンプルサイズの増加に伴い計算時間が急増する。一方で、提案した大域的な近似手法を行うと、計算時間の増加が緩やかになり、標本数が10万の場合でも数分程度で推定が終わる。

計算負荷の低減に効果絶大な新手法をフリーソフトで提供

時空間データを解析するとき、最も苦労するのはコンピュータの計算負荷をいかに低減するかだという。「大規模な時空間データを柔軟に統計モデリングすることは、計算コストの観点で容易ではありません。例えば住宅地価の場合でも、データ数が膨大であるため、計算負荷が大きくなってしまいます。精度と計算効率の両立が重要な課題です」(村上)。

熱波の予測の際にもみられた「近隣の値は類似する」という性質は空間相関と呼ばれ、空間データの最も基礎的な性質とされる。しかし、空間相関を加味すると、計算コストは標本数の3乗のオーダーになることが知られている。例えばデータ数が2倍になると計算コストは2の3乗、つまり8倍である。これは、空間相関関係を記述する標本数×標本数の行列の逆行列を評価しなければならぬからだ。そのため標本数が大

き場合は、途中でコンピュータがフリーズして回らなくなってしまうこともよくあるという(図4)。

このため、主要な空間成分のみを用いて空間相関パターンを近似することで、負荷の軽減を図る。村上が提案した近似手法を導入したところ、計算時間の増加が緩やかになり、標本

数が10万の場合でも、数分程度で推定が完了することが確認できた(図5)。

村上は2017年に、開発したアルゴリズムを統計分析フリーソフト「R」のパッケージ「spmoran (エスピーモラン)」として公開。現在も拡張した手法を開発する

たびにアップデートを欠かさず、公開講座などを通じて活用方法の普及にも尽力している。すでに、環境分野や都市計画分野の解析などに世界中で広く活用されつつある。

「都市の抱える問題は、感染症のパンデミックや熱中症といった健康被害から災害、事故、渋滞、住宅まで非常に多岐にわたっています。プロジェクトで開発した統計モデルを研究者や実務者に提供することで、そうした問題の解決に寄与していければ」と村上は抱負を語る。

村上と以前から共同研究を進めてきた松井は「統数研は伝統的に時系列解析ですぐれた成果を出してきましたが、近年はやや手薄になっています。その意味でも、村上先生との連携を強化できたらと思っています」と話す。統数研全体が、若きプロジェクトリーダーの活躍に期待し、応援している。(広報室)

住宅地価要因分析への応用例

各種要因が住宅価格に及ぼす影響を場所毎に推定

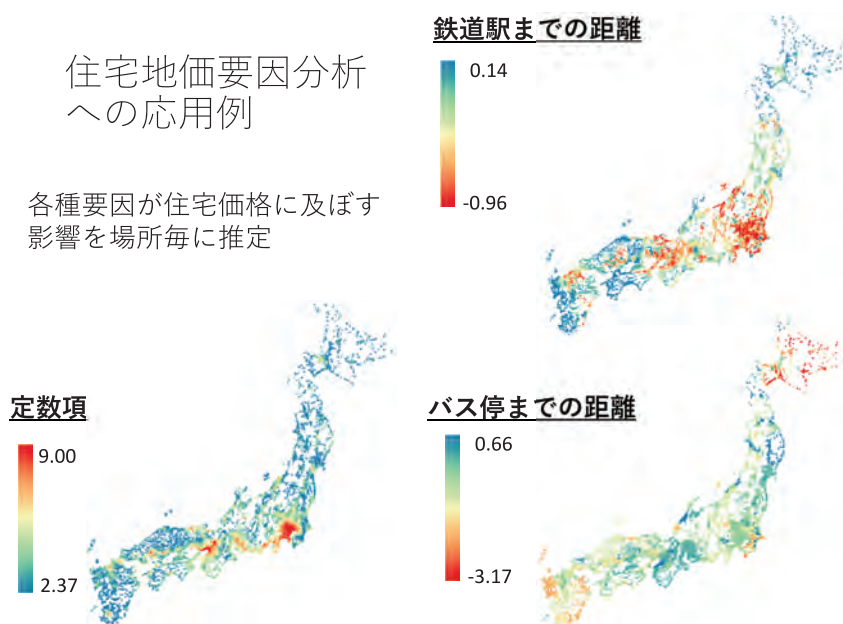


図5：提案手法をわが国の住宅地価の要因分析に応用した。この図の定数項の推定結果からは、東京や大阪で住宅地価が高いことが分かる。駅までの距離の影響の推定結果(右上)からは、関東・近畿の広域では駅に近いほど地価が高いという傾向が強いことが分かる。バス停距離の影響の推定結果(右下)からは、北海道の鉄道が少ないエリアでバスの影響力が強いことが確認できる。

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。

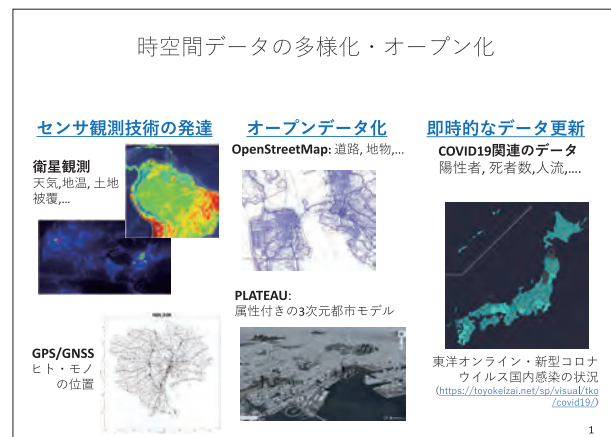
2022年10月-12月の公開講座実施状況

10月21日(金)に、当研究所の村上大輔講師による「Rによる時空間モデリング入門」の講義がオンラインにて行われました。社会・経済、気象、生態などの各分野で位置座標つき(地理)空間データが利用できるようになっており、それらの利活用は研究・ビジネスの両方で、今後より重要となっていくことが予想されます。それら空間データや時空間データのための統計手法とそのRによる実装方法について講義し、空間データを扱うための基礎事項からはじめて、(時)空間データを統計的にモデリングする方法を、Rを用いた実例を交えて解説しました。

12月5日(月)~6日(火)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「Rによる時空間モデリング入門」資料より

統計数理セミナー実施報告(2022年11月~2023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。2022年11月~2023年1月のセミナーは下記の通り行われました。新型コロナウイルス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
2022年11月 2日	Le Thanh Tam	Optimal transport and its applications on machine learning
11月 2日	今泉 允聡	非スパースな高次元漸近論の理論と応用
11月 9日	大久保 祐作	ベイズモデルを用いたマクロ進化の種間比較生物学研究とBSDSモデル
11月 9日	上野 玄太	極値時系列の状態空間モデル
11月16日	逸見 昌之	診断研究のメタアナリシスにおける要約ROC曲線について
11月30日	本武 陽一	材料パターン情報学のための各種手法開発について
11月30日	持橋 大地	Nonparametric Bayesian Deep Visualization
12月 7日	加藤 昇吾	多次元トラス上のデータのためのコピュラモデル
12月 7日	中野 慎也	ISSの極端紫外光画像からの電離圏構造推定
12月14日	Petrillo Giuseppe	Including stress relaxation in point-process model for seismic occurrence
12月14日	船渡川 伊久子	COVID-19と健康関連指標の推移
12月21日	菅澤 翔之助	ベイズ的予測統合による柔軟な空間予測
2023年 1月11日	野間 久史	ネットワークメタアナリシスにおける外れ値の検出と影響力解析
1月18日	三分一 史和	円筒プラズマ乱流データの理解に向けた統計的時系列解析の適用性
1月25日	伊庭 幸人	事後分布からのサンプルを利用したブートストラップ法とデータ空間での漸近論
1月25日	彭 鴻	2001年から2021年までの日本におけるMj3.0からMj7.4の本震に対する前震の研究

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)

2022年度 ROISクロストーク開催報告

2022年12月19日に、情報・システム研究機構主催の「ROISクロストーク」がオンライン開催されました。これは長年「若手クロストーク」として開催されてきた研究会です。

昨年はそれに「+ベテラン」が加わりましたが、今年は年齢の枠を取り払ったシンプルな名称に変更し、分野と世代を超えた交流の促進を目指しました。



Spatial Chat 上のポスター発表風景

参加者は73人で、教員や大学院生など22件（23名）の発表がありました。

当日はSpatial Chatを用いてライブのポスター発表を行い、参加者それぞれの専門領域を生かした活発な議論が行われました。このポスター発表の中で最も盛り上がった発表に、後日機構長からベストポスター賞が贈られました。ポスター発表の後には、参加者から御提案のあったテーマに沿ってフリーディスカッションが行われ、世代や分野の枠を超えて活発な交流が図られました。

当日は反省点も色々ありましたが、オンライン開催ならではの長所を生かした広範囲の交流が行えたのは非常に良かったと思っています。来年のクロストークも楽しみにして頂ければ幸いです。

(清水 信夫)

統計研トピックス

情報・システム研究機構「第4期戦略的研究プロジェクト」における統計学とプラズマ／核融合科学の協調的發展を目指した研究者交流と討論会の報告

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、12月12日から13日にかけて、量子科学技術研究開発機構（QST）の六ヶ所研究所・核融合科学研究所（NIFS）六ヶ所研究センターおよび八戸市公民館で施設見学、共同研究者の研究紹介および討論会を開催しました。12月12日はQST六ヶ所研究所 石井康友次長のご案内で、フランスに建設中の国際熱核融合実験炉（ITER）遠隔実験センターや核融合炉シミュレーション用のスーパーコンピュータなどの設備の見学を行い、さらに、世界と日本の核融合研究の現状や将来の見通し、核融合炉の設計から材料開発まで網羅的に貴重なレクチャーを頂きました。また、ITERの実験開始後には、ITERで取得される大規模なデータが六ヶ所に転送されてくることとなり、核融合コミュニティのみならず、このプロジェクトをはじめとする幅広い分野の研究者に活用してほしいというコメントがありました。その後、NIFS六ヶ所研究センター 横山雅之教授のオーガナイズで参加者による研究紹介と討論会を開催し、13日は場所を八戸市公民館に移し継続しました。

本プロジェクトは多様な統計数理・データ駆動手法の実データへの適用を行い、手法の改良や社会実データへの実装を促進することで、統計学とプラズマ／核融合科学の双方のコミュニティの協調的發展を目指すものです。今回の企画は8月に統計数理研究所で開催されたキックオフ集會に引き続くもので、本プロジェクトの代表である統計思考院 菊地和平特任助教が精力的に準備を進めておりました。しかし、その最中、菊地さんは突然帰らぬ人となりました。多大な喪失感とともに一時は開催中止を検討しましたが、菊地さんの遺志を継ぎプロジェクトの推進をすべく、代



国際熱核融合実験炉（ITER）遠隔実験センター（QST六ヶ所研究所サイト内）の見学

表を三分一准教授に交代し、菊地さんとともに準備を進めていた横山教授のご尽力の下で予定通り開催する運びとなりました。

今回は菊地さんの呼びかけに応じて、三分一准教授、横山教授、石井次長の他、当研究所から矢野恵佑准教授、田中未来准教授、本武陽一特任助教、東京大学今泉允聡准教授、京都大学大学院で核融合研究に取り組んでいる森下侑哉さん、安齋亮慶さん、田原康祐さんが参加しました。2日間の会合を通して、菊地さんへの追悼や感謝、生前のエピソードも数多く聞かれ、改めて菊地さ

んの遺志を継いでいくことを心に強く誓いました。

(三分一 史和／横山 雅之)



菊地さんを囲んでの集合写真

第1回「自己点検アドバイザリーボード」を開催

統計数理研究所が中核機関となる「統計エキスパート人材育成コンソーシアム」では、2021年度から5年間の事業期間の初期における活動状況について自己点検を実施することにより、大学統計教員育成研修などのコンソーシアム活動の更なる改善を図ることとしています。

この点検を効果的・客観的に行うため、2022年10月26



第1回自己点検アドバイザリーボードの開催の様子(オンライン開催)

日(水)に、統計関連学会連合会員などの外部有識者から構成される「自己点検アドバイザリーボード」の第1回会合を開催しました。

アドバイザリーボードは、大学統計教員育成センター・千野雅人センター長の進行のもと、椿広計所長のごあいさつにより開会しました。そして、ボードメンバーである(株)NTTデータ数理システム・小木しのぶ取締役、実践女子大学・竹内光悦教授、東海大学・山本義郎教授から、コンソーシアムの活動や自己点検の進め方などについて、評価や助言を頂きました。

2023年2月には、これらの助言やコンソーシアム参画機関・研修生に対して実施したアンケート調査の結果を踏まえ、コンソーシアム活動の課題と今後の改善の方向・方策を取りまとめた「自己点検報告書」を作成し、HPなどで公開します。

(大学統計教員育成センター)

有竹俊光特任助教と日野英逸教授が2021年度IBISML研究会賞を受賞

有竹俊光特任助教と日野英逸教授が、電子情報通信学会第一種研究会IBISMLにおいて、2021年度IBISML研究会賞を受賞しました。IBISML研究会では、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、IBISML研究会賞を授与しています。今回の受賞は2022年1月に開催された第44回IBISML研究会にて発表した「変数の拡張に対する最適輸送を用いたドメイン適応」に対するものです。2022年11月に開催された第25回情報論的学習理論ワークショップでの表彰式において賞状が授与されました。

(日野 英逸)



津山高校での出張講義を実施

2022年11月8日(火)、岡山県立津山高等学校の1年次生(241名)を対象に川崎能典副所長が出張講義を行いました。

津山高等学校の生徒には5年前に来訪した際も川崎教授が統計学やデータサイエンティストについて生徒からの事前質問に答える形で講演しており、本年は統計・データサイエンス分野に係る基礎的な概念・スキルについて伺いたいという講義依頼があり、実施されました。

「統計・データサイエンス概論 シミュレーションを使った推測とパターン発見」と題した講義では、始めに研究者になるためのステップや川崎教授のキャリアなどが紹介されました。次にwebアプリを使用した3つの実験が行われ、生徒たちは「ランダムの中に生まれる法則がある」ことを体験し、統計学の考え方とそれがどう意思決定の役に立つのか学びました。最後には健全な科学の発展のために研究者には守るべき倫理があることが説明され、「知・好・楽(楽しめることを探そう)」のメッセージも伝えられました。

生徒からは講義内容やキャリアに関する多くの熱心な感想があり、講義後日にも質問が寄せられました。

本講義はスーパーサイエンスハイスクール指定校である津山高校の人材育成のための設定科目「十六夜プロジェクト(iP)」の一環として行われました。

(広報室)



模擬授業とインタビュー動画の公開

2022年の夏から公開している子供向けページ「すーりん、けいとの『統数研の部屋』」に東京学芸大学 先端教育人材育成推進機構の山下雅代准教授の模擬授業とインタビュー動画を掲載しました。模擬授業は「身近な“問題”に活かす統計の使い方」と題し、感性を扱った課題について、問題発見から原因追及、問題解決までのプロセスをわ

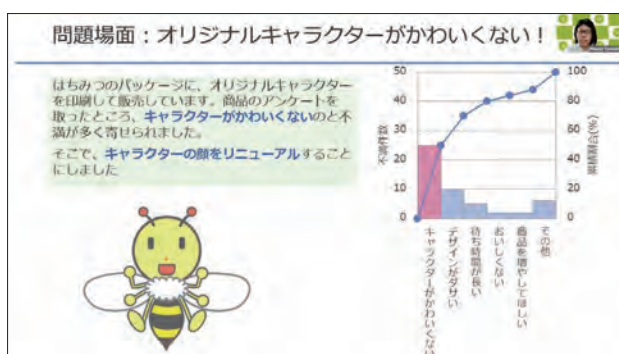
かりやすく説明しています。「蜂蜜のパッケージに載せるオリジナルキャラクターの顔をかわいくリニューアルする」という課題に対し、「目がかわいくない」という問題発見をし、原因追及では「かわいい目の大きさや位置、形」を探り、問題解決では改善後の検証と理由付けを行うという流れで、身近に感じる素材を扱うことで子供の興味を引く内容になっています。

本ページには他にも統数研の紹介や子ども見学デー2020、2021(オンライン開催)で公開された研究者インタビューと講演、白玉・黒玉数当て実験と解説、謎解きチャレンジの動画などを掲載しています。

すーりん、けいとの「統数研の部屋」

<https://www.ism.ac.jp/events/oyako/index.html>

(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

大学院説明会について

2022年11月11日(金)に、2022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48名でした。

(総務課・大学院係)

●2023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）を開催いたします。(健)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。

①A 行列の固有値という数学：統計学と生態学からの再入門 (健)

日時: 6月13日(火) 10時～16時(5時間)
 講師: 島谷健一郎(統数研)
 申込受付: 5月8日(月)10時～5月15日(月)10時
 受講料: 5,000円
 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

個体生態学において、生残・成長・繁殖を表す行列の固有値が個体群成長率を表す話を中心に据える。余裕のある範囲で、主成分分析、微分方程式の解の安定性、さらに、マルコフ連鎖モンテカルロ法(MCMC)と個体群行列モデルの間で共通するもの、について解説する。

受講者に期待する予備知識やレベル:

大学1-2年生レベルの線形代数の中の固有値・固有ベクトル

教科書:
 資料を配布します。

①B 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— (健)

日時: 7月4日(火) 10時～16時(5時間)
 講師: 酒井憲司(東京農工大学農学研究院)
 申込受付: 5月29日(月)10時～6月5日(月)10時
 受講料: 5,000円
 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

カオス理論の農学諸現象への応用を試みるのがカオス農学である。近年、トラクタ事故のメカニズムの非線形動力学の解明が進み、カオスを含む多彩な挙動に起因することがわかってきた。本講座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する。そのうえで、この理論の社会実装として世界的に農業の持続性への潜在的脅威とされている農作業事故へのアクティブ安全対策として展開を示す。

1. 農耕の起源と科学
2. トラクタの非線形動力学
3. 仮想ドライブによる事故メカニズムの解明
4. モーション型ドライビングシミュレータによる安全な危険体験
5. 内発的リスクマネジメント

受講者に期待する予備知識やレベル:

農学研究者、農業従事者、農業機械・自動車関連産業、保険会社、研究機関、行政機関など、何らかの形で農学と関わる人を主な対象とする。トラクタの車両力学的内容にも触れるが、それらの専門的知識がなくても、農業や農業関連の経験や知識があれば十分に役立つように講義を構成する。

教科書:
 資料を配布します。

①C 多変量解析法 (健)

日時: 8月3日(木)～4日(金)、8月7日(月)～8日(火)
 10時～16時45分(23時間)
 講師: 馬場康雄・清水信夫(統計数理研究所) 今泉忠(多摩大学)
 申込受付: 6月19日(月)10時～6月26日(月)10時
 受講料: 20,000円
 会場: オンライン
 定員: 100名(応募者多数の場合は抽選)
 講義レベル: 初級

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によって把握される。すなわち多次元のデータによって現象が表現される。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ある。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する。また、クラスター分析、分散構造分析などの解説をする。

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する。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うが、理解をたやすくする点から、大学初級程度の線形代数と微分積分の知識があることが望ましい。

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「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
グラクソ・スミスクライン株式会社 代表取締役社長 ポール・リレット	ベリムマブ治療と標準療法を受けたルー ブス腎炎患者における日本の実臨床下 での有効性評価:MOONLIGHT研究	R4.12.2～ R5.12.31	5,148,000	データ科学研究系 野間 久史 准教授
国立研究開発法人新エネルギー・ 産業技術総合開発機構 理事長 石塚 博昭	革新的セラミック材料設計のための材 料パターン情報学の創成	R4.8.1～ R5.3.31	0	統計的機械学習研究 センター 本武 陽一 特任助教
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	全球とミクロを繋ぐ要約に基づく災害解 析基盤の開発	R4.10.1～ R5.3.31	1,300,000	データ科学研究系 村上 大輔 助教
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	フェイゾンエンジニアリングのための統計 的機械学習	R4.10.1～ R5.3.31	4,680,000	データ科学研究系 吉田 亮 教授
一般社団法人いのちを支える自殺 対策推進センター 代表理事 清水 康之	ポストコロナの自殺対策に資する統計等 のミクロデータ利活用推進に関する研 究	R4.11.1～ R5.3.31	2,602,000	椿 広計 名誉教授

(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	称号付与
清水 邦夫	慶應義塾大学・名誉教授	統計エキスパート人材育成プロジェクト	2022.10.6～ 2023.3.31	川崎 能典 教授	特命教授
Evgeny Spodarev	Ulm University, Professor	裾の重い確率場のイクスカーション集合 の研究	2022.10.6～ 2022.10.10	栗木 哲 教授	
Wei Zhang	Northeast Normal University, PhD Student	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	2022.10.11～ 2023.10.10	福水 健次 センター長	
貝淵 響	みずほ第一フィナンシャルテク ノロジー株式会社・アナリスト	条件付バリューアットリスクのバックテスト 法に関する研究	2022.10.6～ 2023.3.31	川崎 能典 教授	
津田 悠人	東京都市大学・博士後期課 程学生	確率モデルを用いた地盤物性の空間分 布推定	2022.11.14～ 2022.11.25	Stephen Wu 准教授	
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	2022.11.14～ 2022.11.25	Stephen Wu 准教授	
Michael Choi	National University of Singapore/Yale-NUS College, Assistant Professor	相互作用粒子系に対するランドスケープ 修正	2022.12.1～ 2022.12.15	鎌谷 研吾 准教授	
村瀬 博典	愛知製鋼株式会社・チーム 長	異常検知のメタ学習に関する研究	2022.11.3～ 2023.3.31	福水 健次 センター長	
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	2022.12.8～ 2022.12.13	庄 建倉 准教授	
李 永波	Institute of Geophysics, China Earthquake Administration, Graduate Student	ベイズ型重力インバージョンにおけるモン テカルロ法の応用	2022.12.20～ 2023.12.19	庄 建倉 准教授	
Hsien-Kuei Hwang	Academia Sinica, Institute of Statistical Science, Distinguished Research Fellow	積分型適合度検定統計量における組み 合わせ構造の解明	2022.12.9～ 2022.12.17	栗木 哲 教授	
Fatima Jenina Tolentino Arellano	大阪大学大学院工学研究 科・博士後期課程学生	低温プラズマ科学	2023.1.23～ 2023.2.3	吉田 亮 教授	

氏 名	職 名	研究題目	研究期間	受入担当研究教育職員	称号付与
Patrick Asante	BC Ministry of Forests Lands and Natural Resources Operations, Manager & Senior Researcher	森林資源管理における生物保護区域設定に関わる現地実情分析	2022.12.8～2022.12.13	吉本 敦 教授	
金川 平志郎	Newcastle University, Research Associate	カーネル法による機械学習手法の評価	2022.12.19～2022.12.23	福水 健次 センター長	
奥田 悠介	株式会社システムサポート ICT事業部・社員	RadonPy 高分子物性データベースの開発	2022.12.12～2023.3.31	吉田 亮 教授	
沖田 大佑	株式会社システムサポート ICT事業部・社員	RadonPy 高分子物性データベースの開発	2022.12.12～2023.3.31	吉田 亮 教授	
石垣 利洪	順天堂大学大学院 医学研究科修士課程 データサイエンス専攻・大学院生	メッシュデータを用いた集計データの統計分析	2023.2.1～2023.3.31	川崎 能典 教授	

(研究推進課・共同利用係)

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
ワン カルロス ヒメネス・ソブリーノ Juan Carlos Jiménez-Sobrino	Senior Professor	Dept.Interdisciplinary Mathematics Institute of Cybernetics, Mathematics and Physics, Cuba	客員教授	ニューロン発火パターンの動力的解析と確率微分方程式モデリング	R4.10.10～R4.12.11	三分一 史和 准教授

(総務課・人事係)

外国人客員紹介



● Juan Carlos Jiménez-Sobrino 客員教授

Returning to ISM after two years of covid world isolation is gratifying. This time, for two months, I will join Prof. F. Miwakeichi in conducting a new challenge research project on the simulation of plasma models and their identification from actual data. As in our previous collaborative works, we hope to achieve novel results in this area as well. Through this ISM News, I would like to send my most cordial greetings to the ISM colleagues and friends.

(総務課・人事係)

会議開催報告

令和4年度第4回運営会議の開催

2022年12月23日(金)に、オンライン形式で令和4年度第4回の運営会議が開催されました。

はじめに、研究教育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、研究所の主な年間活動報告があり、情報・システム研究機構関連として、機構シンポジウムやクロストークの報告がありました。

その後、他分野と広く融合できる統計科学の強みを生かした研究所の組織改革、今後の公募人事企画のあり方、科学研究費助成事業への対応、国際的な共同利用・共同研究への取組み、研究所の年間イベントに基づいたアウトリーチ活動などについて、活発な意見交換が行われました。

(統数研事務課)

Report 統計数理研究所調査研究レポート

No.122: Kiyohisa Shibai, Cross-National Survey on Nuclear Disarmament Issues — Japan, Hiroshima, Nagasaki, and the U.S. 2022 Web Survey — (2022.11)

(メディア開発室)

Report 統計数理 第70巻 第2号

特集「データ同化の方法」

「特集 データ同化の方法」について

上野 玄太	129
雲解像粒子フィルタを用いた積乱雲の非ガウス性に関する研究 [研究詳解]	
川畑 拓矢、上野 玄太	133
データ同化を用いたオーロラ活動指数の推定 [原著論文]	
三好 由純、上野 玄太、山本 凌大、町田 忍、能勢 正仁、塩田 大幸、中村 紗都子	153
アンサンブルカルマンフィルタにおける変数局所化を利用した気象場と大気濃度場の同時データ同化 [原著論文]	
関山 剛、梶野 瑞王	165
大気解析のための変分法データ同化における背景誤差共分散行列の根の定式化 [総合報告]	
石橋 俊之	181
共役ベクトルとBFGS公式を用いた解析誤差共分散行列の導出 [研究ノート]	
丹羽 洋介、藤井 陽介	195
変分法データ同化システムにおけるBFGS公式を利用したアンサンブルメンバー生成について [原著論文]	
藤井 陽介、吉田 拓馬、久保 勇太郎	209
アンサンブルを用いた変分法データ同化と計数データのための拡張 [研究ノート]	
中野 慎也	235
シグネチャ法入門 [研究詳解]	
杉浦 望実	251

(メディア開発室)

Report Annals of the Institute of Statistical Mathematics
Volume 74, Number 6 (December 2022)

Bofei Xiao, Bo Lei, Wei Lan and Bin Guo

A blockwise network autoregressive model with application for fraud detection	1043
---	------

Hammou El Barmi

On comparing competing risks using the ratio of their cumulative incidence functions	1067
--	------

Michael Kohler, Adam Krzyżak and Benjamin Walter

On the rate of convergence of image classifiers based on convolutional neural networks	1085
--	------

Ke Yu and Shan Luo

A sequential feature selection procedure for high-dimensional Cox proportional hazards model	1109
--	------

Chih-Hao Chang, Hsin-Cheng Huang and Ching-Kang Ing

Inference of random effects for linear mixed-effects models with a fixed number of clusters	1143
---	------

Holger Dette and Martin Kroll

Asymptotic equivalence for nonparametric regression with dependent errors: Gauss-Markov processes	1163
---	------

Toshiaki Tsukurimichi, Yu Inatsu, Vo Nguyen Le Duy and Ichiro Takeuchi

Conditional selective inference for robust regression and outlier detection using piecewise-linear homotopy continuation	1197
--	------

(メディア開発室)



トモグラフィと研究

矢野 恵佑

数理・推論研究系

トモグラフィ (tomography) とは、直接見るできない対象物を多方向の切断面の描像から復元する逆解析技術のことで、物理探査や医療診断等で用いられている。対象物を取り囲むように配置されたセンサの観測から対象物の描像を描くのである。トモグラフィの具体例としてX線CT (Computerized Tomography) やMRI (Magnetic Resonance Imaging) を想像するとわかりやすいだろう。地震学では、地震波の各観測点での伝播をもとに震源の情報や地下の構造に迫っていく地震波トモグラフィが行われ、地球の内部構造を明らかにしてきた。地震波トモグラフィの先駆的な研究に、ベイズ統計で知られるHarold JeffreysとKeith Edward BullenのまとめたJeffreys-Bullenの表がある。遠地地震波の伝播に関してはJeffreys-Bullenの表でそれなりの精度で近似できることが知られている。脇道にそれるが、Harold Jeffreysは統計学者でもあり、地球物理学者でもある。Bradley Efron and Trevor HastieのComputer Age Statistical Inference: Algorithms, Evidence, and Data Scienceには「part-time statistician, working from his day job as the world's premier geophysicist of the inter-war period」と書かれている。

トモグラフィにおいて対象物をどの程度正確に描けるかは、センサがどの程度対象物を囲っているかに強く依存する。球の中心に小球があり、球面にセンサが位置している状況を想像してみよう。もし、上半球面にしかセンサが配置されていないならば、小球の上半分はどれだけ詳しくわかったとしても小球の下半分については全くわからないままである。した

がって、センサの位置をこちらが指定できる場合は対象を一樣に囲む形にするのが望ましい (CT スキャンやMRI を考えてみると良い)。一方、物理的な事情からセンサをそのように配置できない場合もある。そのような場合は決定精度が著しく良い状況を除いて、解を一つに定めることが本質的に難しくなる場合がある。トモグラフィをベイズ的に行うとそういった不定性が視覚的に明らかになるため、ベイズ的トモグラフィは昔から根強い人気がある。

「研究する」というのもまたトモグラフィだと思うことがしばしばある。いくつかの切り口を試してみても対象の「感じ」がようやく分かり研究が進むのは、いかにもトモグラフィだと思う。単一の切り口だけでは行き詰まることはしばしばある (もちろん、単一の切り口だけで解明できるならそれに越した事はない)。あるいは、単一の切り口のみで開発されたものが別の視点を取り入れることで改良されることがある。時には、自分の専門外の知識を取り入れることが問題解決の糸口になるだろう。大学・大学院時代に韓太舜先生の「情報理論における情報スペクトル方法」・小林昭七先生の「接続の微分幾何とゲージ理論」・藤重悟先生の「Submodular functions and optimization」といった本を輪読したことを今でも覚えている。非常に難解で苦戦した (苦しかった記憶しかない) が、直接引き出しが増えることに繋がる事もあれば、ああいう考え方があ・ああいう世界があるということを認識しておくことが柔軟な発想に繋がる事もある (あるいは、そういう血肉になっていると信じていものである)。色々な分野の色々な考え方を知り、研究の解像度をもっともっと上げていきたい。



図1: Bradley Efron and Trevor Hastie, Computer Age Statistical Inferenceの日本語訳



図2: モンレアーレのドゥオーモ (パレルモ)。至る所に工夫が凝らしており、360度細かく見ても見すぎるという事はない。