

▼ CONTENTS

- 02 研究室訪問
持橋 大地／土屋 隆裕
- 06 シンポジウム報告
[Seminar on Application of Mathematical Modeling in Forest Resource Management]
共同開催報告
[Joint International Symposium By Japan and Czech Republic Data Acquisition,
Statistical Modeling and Decision-Making Toward Better Forestry] 共同開催報告
[Training on Introduction to Statistical Analysis in "R" for Forest Resource Management]
共同開催報告
- 07 研究教育活動
2013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(2013年11月～2014年1月)
2013年公開講演会「統計学が切り拓く脳科学、脳工学の未来」
千葉大学大学院 特別講演「脳・神経系の研究への新しい試み」
文部科学省委託事業：数学協働プログラム「安心・安全・快適な社会インフラ維持への数理科学の適用」
- 10 統数研トピックス
SC13に出展／大学共同利用機関シンポジウム2013開催
2013年度富士通SS研合同分科会での展示
新スーパーコンピュータシステム導入
情報・システム研究機構永年勤続者表彰式
- 12 総合研究大学院大学複合科学研究科統計科学専攻関係
- 13 お知らせ
公開講座
- 14 共同利用
平成25年度共同利用公募追加採択課題
- 14 外部資金・研究員等の受入れ
共同研究の受入れ／受託研究の受入れ／外来研究員の受入れ
- 15 人事
- 16 所外誌掲載論文等
- 16 刊行物
Research Memorandum (2013.11～2014.1)
統計数理／Annals of the Institute of Statistical Mathematics
- 18 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 ニュース



生きたことばをモデル化する

統計科学の研究



持橋 大地

数理・推論研究系
学習推論グループ准教授

「いつれの御時にか、女御更衣あまたさぶらひける中にいとやむごとなき際にはあらぬが、すぐれてときめきたまふありけり」で始まる『源氏物語』。平安時代中期（11世紀）に成立した宮廷の恋愛物語は、世界最古の長編小説として有名だが、江戸時代の人々にすら極めて難解なことばの集合で、多くの注釈書が発刊されていた。

ことばは時代と共に変化し、それを正しく理解することは失われた意味を復元する作業にも似る。冒頭にかかげた源氏の一節も、教えてくれる人がいなければ、それぞれの単語がどこで切れるのかすら判然としない。

言語を確率的にとらえ、計算式によって解析（処理）するなら現代語も古典も、未知の言語すらも容易に理解できるようになる——このように考えて「計算言語学」の旗をかかげる統計科学者がいる。

言語の差異構造を統計的に理解する

持橋の専門分野は、「統計的自然言語処理」と「統計的機械学習」。それは簡単に表現するなら、言語を数学と統計の力を使ってモデル化しようとする科学である。「言語には様々な意味で人により、地域により、時代によってバリエーションがあり、同時にそれらには共通性がある。そうした言語の差異の構造を統計的手法によって理解したい」と持橋は言う。

「計算言語学は言語のもつ統計的な性質に着目し、機械翻訳、構文や意味の解析、感情情報の処理といった様々な応用を可能にする分野です」。持橋は言語の文字列に適切な確率を与えて統計モデル化する研究に取り組んできた。まず、単語の出現が話題や文脈によって異なること、意味的に関係のある単語群が同時に現れる傾向があることに注目し、話題の「文脈」を時系列としてモデル化することを試みた。大量の新聞記事から細かい話題を統計的に抽出する統計モデルを利用し、「話題の変化」を自動的に見つける研究に打ち込んだという。ここから、持橋は「単語に次の単語が続く確率、また、文字に次の文字が続く確率をすべてモデル化することができれば、言語の形態を鮮明に印象づけることができ

順位	単語	頻度	出現確率
1	の	1266	0.055005
2	。	1120	0.048662
3	、	988	0.042927
4	た	951	0.041319
5	て	884	0.038408
18	ジョバンニ	189	0.008212
34	カムパネルラ	101	0.004388
104	風	26	0.001130
104	天の川	26	0.001130
482	橙	5	0.000217
482	ポート	5	0.000217
482	ステーション	5	0.000217
1307	燈火	1	0.000043
1307	天蚕	1	0.000043
1307	鶴嘴	1	0.000043

表1. 宮沢賢治『銀河鉄道の夜』での単語の頻度とその確率。「た」「の」「て」のような高頻度の単語の一方で、低い頻度の単語が同時に大量にあることがわかる。

言語を言語学者の主観から解き放ちたい

るのではないだろうか」と考えた。

「教師なし形態素解析」で文章を読み解く

たとえば、宮沢賢治 (1896-1933年) の代表作『銀河鉄道の夜』を例に考えてみよう。賢治の没後に発表されたこの作品は、「けれどもほんたうのさいはひは一體何だらう」など、主人公の華麗な台詞と造語にあふれている。

この作品を構成する単語に分解し、統計的な観点から考察すると、ひんぱんに現れる単語と、そうではない単語があることが容易に見てとれる (表1)。単語のそのような出現頻度の実態は、「ポアソン=ディリクレ過程」という統計モデルで説明できる。持橋によれば、「それは簡単に言うと “rich gets richer” つまり、いったん頻度が高くなった単語はどんどん頻度が高くなるという現象のモデル化で説明される」という。具体的には「た」「の」「て」のような高頻度の単語がある一方、低い頻度の単語が同時に大量にあることを指摘できる。『銀河鉄道の夜』では「橙」「ステーション」「天蚕」などの単語が該当する。そして、確率と順位の積がほぼ一定になる Zipf の法則が成り立っていることが右下がりの直線から読みとれる (図1)。

このような方法から生み出された統計モデルを使えば、未知の言語や古文、話し言葉などの単語分割が可能になるというのが、統計学者としての持橋の考え方だ。持橋はそれを「教師なし形態素解析」の手法と呼んでいる。

『源氏物語』に話を戻す。難解な読みも上記の統計モデルを用いて解析すると、図2のように、意味を持つ単語の区分が明らかになり、容易に読み進めることができるようになる。これは持橋が学会で発表した具体例であり、多くの研究者に驚きを与えた。

持橋は、自然言語処理と数学の接点をテーマにしたエッセーで、次のように書いている。「言語を統計的にとらえることによって、複雑で膨大な言語現象を計算機で自動的にモデル化できるとともに、

規則ではとらえきれない曖昧性や例外、文脈依存性を数学的に適切に扱うことが可能になる」。

ことばの不思議な「構造」を探り続ける

「自分では統計学者というよりは、言語学者のつもり」。そのように自認する持橋の研究姿勢は、統計学を道具にして言語の世界の不思議に迫ろうとする営みである。

その一方で、「言語を言語学者の主観から解き放ちたい」とも、持橋は語る。「源氏物語を例にした単語分割の研究は確かに私の代表的な研究成果ですが、基本は言語構造を統計的にとらえたいということ」。学者の経験と主観によって生み出された仮説を積み重ねる従来の言語学は、「生きたことば」の構造解明には対応できなくなっている、というのが持橋の主張だ。

ところで、持橋は少年時代から大の音楽好きだった。高校では合唱部の指揮者。「自分のパートをきわめるだけではなく、曲の全体像を理解したいと、いつも思っていた。一方で楽譜も一種の記号であると考えていた。今から振り返ると、音が作り出す数理に惹かれていたのだと思う」。音楽の中に、数学的な言語の構造を見出そうとする姿勢が、当時から育まれていたようだ。

文系で入学した大学の3年次に理転してコンピュータサイエンスを専攻。奈良県の大学院卒業後、NTTの研究所を経て2011年に統数研入り。統計的機械学習のバイブルともいえる教科書『パターン認識と機械学習』の訳者の一人でもある。

「言葉の配列の中に隠れているもの、つまりリズムや反復などの確率を、統計的手法を使って表に出すのが私の仕事。普通の生活の中で日々接している言葉を、どうやって機械に乗せることができるだろうか、といつも考えている」と語る。

時代とともに海のように広がっていく言語現象。広大さに対峙するかのように、統計研究の立場からの問いかけが続く。(広報室)

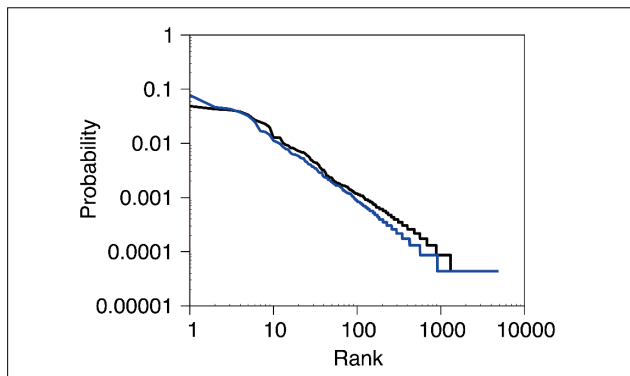


図1. 単語の確率と順位を両対数でプロットしたもの。図1のデータでの単語の確率と順位を両対数でプロットした (黒線)。確率と順位の積がほぼ一定になる Zipf の法則が成り立っていることが、右下がりの直線から読みとれる。青線は、ポアソン=ディリクレ過程 $PD(2, 0.8)$ からランダムに生成した、同じ長さのテキストでの単語の確率と順位を表す。自然言語のもつ性質が、この確率過程によってほぼ再現できていることがわかる。

しばしは夢かとのみたどられしを、やうやう思ひしづまるにしも、さむべき方なくたへがたきは、いかにすべきわざにかとも、問ひあはすべき人だになきを、忍びては参りたまひなんや。若宮の、いとおぼつかなく、露けき中に過ぐしたまふも、心苦しう思さるを、とく参りたまへ」など、はかばかしうも、のたまはせやらず、むせかへらせたまひつつ、かつは人も心弱く見たてまつらむと、思しつつまぬにしもあらぬ御気色の……



階層ベイズモデルによる学習

しばしは夢かとのみたどられしを、やうやう思ひしづまるにしも、さむべき方なくたへがたきは、いかにすべきわざにかとも、問ひあはすべき人だになきを、忍びては参りたまひなんや。若宮の、いとおぼつかなく、露けき中に過ぐしたまふも、心苦しう思さるを、とく参りたまへ」など、はかばかしうも、のたまはせやらず、むせかへらせたまひつつ、かつは人も心弱く見たてまつらむと、思しつつまぬにしもあらぬ御気色の……

図2. 「教師なし形態素解析」による『源氏物語』の学習。古文や話し言葉、未知の言語などでは単語が何なのか分からないが、階層ベイズモデルにより学習することで、文字列だけから何が「単語」であるか知ることができる。

質の高いデータを求めて 調査方法を研究



土屋 隆裕

データ科学研究系
データ設計グループ准教授

昨年10月9日、「読解力、数的思考力ともに日本は1位」というニュースが各紙朝刊1面で報じられた。経済協力開発機構（OECD）が世界24か国・地域で行った「国際成人力調査」（PIAAC）の結果である。このPIAACでNational Sampling ManagerおよびNational Data Managerとして日本の調査設計やデータの取りまとめに当たったのが土屋である。日本人のIT活用力は紙の調査に答えた人も含めると高順位ではなかったが、質問にパソコンで答えた人の中では世界トップだった。

全国的に関心が高い「全国学力・学習状況調査」では文部科学省専門家会議の一員として平成19年度の初回調査から設計・分析に加わり、抽出調査となったときの設計も行った。この延長線上で「横浜市学力・学習状況調査」の分析も行い、最近では体力調査にも関わる。統計数理研究所では珍しい教育学部（東大）出身で、教育関係の調査に関わることが多いが、他にもさまざまな調査に携わっている。研究所が昭和28年（1953年）から実施している世界的に珍しい継続調査「日本人の国民性調査」では、平成6年（1994年）の入所直後から分析や広報などいくつかの作業を担当している。さらには東京多摩地域住民意識調査を立案して実施しているほか、内閣府や財務省、国税庁、経済産業省、特許庁、国土交通省、総務省、環境省、法務省など国の多くの重要な調査に関わっている。

調査票の設計、間接質問法、調査不能バイアスの調整などに取り組む

多くの調査に携わる中で気がついたことがあるという。「最初のデータ、回答がいかにか頼りないものか、深く関われば関わるほど分かってきた。統計調査はどういう仕掛けで、いかに質の高いデータを採るかという最初の設計がもっとも肝心。良質のデータさえ得られれば、統計的な分析方法はむしろ可能な限りシンプルの方が世

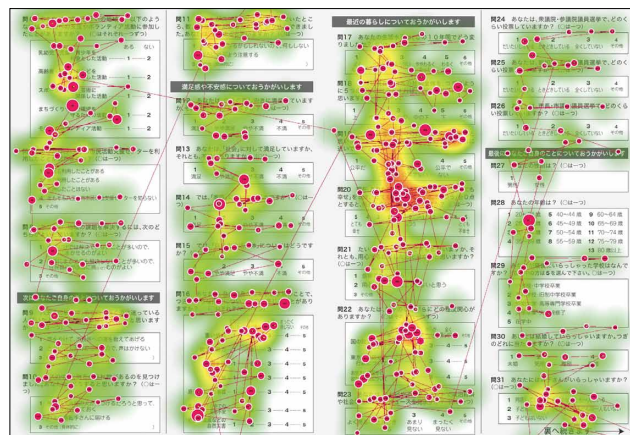


図1. 自記式調査票の視線追跡結果。回答者が調査票のどこを見ているのか調べることで、調査票の見やすさや質問項目の分かりやすさなどを向上させ、回答の質を高める方法を研究している。

今後の研究目標は調査統計の研究を通じ 人は何のためにどう生きるべきかを追求していくこと

の中へのインパクトは大きく、役に立つ調査になる」。質の高いデータを採るために調査方法の研究を行っている。

例えばプライベートな経験の有無など、なかなか素直には答えづらい事柄を調べる方法として「間接質問法」を研究している。集団を2つに分け、一方の集団には、プライベートなことを含む5つの選択肢のうち経験した項目の数のみを答えてもらい、別集団にはそれを含まない4つの選択肢の中で経験項目数を答えてもらう。その差から経験者の割合を推定するというものだ。この方法は当初うまくいかなかったが、土屋が成功例を報告した論文を出してから、いま世界的に関心が高まっている。

最近は回答者自身が調査票に記入する自記式の調査が非常に多いが、回答者は調査票をきちんと見て読んでいるのだろうか。そんな問題意識から、人の視線は調査票のどこに行きやすいか、視線追跡装置を使って調べることに取り組みは始めている。

調査結果からより正確な全体像を得るため、「調査不能バイアスの補正法」も研究課題の一つだ。調査には意識が高く協力的な人が回答することが多く、全体としては偏ってしまう。そこで、調査の中で聞いた年齢、職業、学歴などいくつかの項目を手がかりに国勢調査等を参考として日本全体の姿に戻し、その上で結果を出すなどというものである。さらに、回答者が調査に積極的に協力したかどうかなどの情報も取り入れ、補正の効果をより高めようとしている。

肩の力を抜いて楽しみながら考えたものの方が使えるものになる

中学生のころ、まだ珍しかったパソコンをお年玉貯金をはたいて購入し、ゲームなどをつくって遊んだ。将来はその道にと思ったこともあったが、「自分は何のために生まれてきたのか、人生の意味は何か」と悩み、答えを哲学や心理学に求めて文系へ進んだ。統計研へ入って中学生のころの遊びが仕事につながった。国民性調査の継続的な結果がよく分かるホームページをつくり、理解しや

すいグラフの考案、膨大な調査結果のデータをひと目で見られるようにする可視化の取り組み。国民性調査の結果を広く国民に知ってもらいたいと、調査結果を印刷したトランプや、調査結果と歴史年表を入れた物差しをつくった。「難しく考えるよりも、ときには肩の力を抜いて楽しみながら考えた方が、自分でも驚く発想が出てきたり、使えるものになることが多い」。子どものころに培った柔軟な思考が生きている。

学力調査のデコボコチャートで子どもたちの学力の底上げを期待

「楽しんで考えたこと」の1つが、全国学力・学習状況調査の結果報告のために考案した学校ごとの円形チャートである。円の真ん中に点線で全国平均値のマルが描かれる。チャートは学力、生活習慣、学習習慣、関心、自尊感情、規範意識などの項目に分かれ、各学校の得点が平均値を上回れば点線から外へ出て、下回れば点線より低くなって空白ができ、デコボコになる。レーダーチャートの変形でローズチャートとも呼ばれるが、土屋が提案したこのチャートはレーダーチャートよりも分かりやすいとして、実際に学校現場で生かされるよう全国3万の小中学校へ配布されている。数字中心の報告書では分かりにくかったことが、感覚的にひと目で判断できるようになった。全国の学校でこのチャートの活用が進めば、子どもたちの学力の底上げが期待できるかもしれない。

今後の研究目標は短期的には「より質の高いデータ設計の方法を研究すること」だが、長期的には「調査統計の研究を通じ人は何のためにどう生きるべきかを追求していくこと」と語る。「『大学』に『物に本末あり、事に終始あり』とあるように、学問にも本・末があって、統計学などいまの多くの学問は末学にすぎない、人間としての生き方につながる学問こそ本学といわれる。末学の先に本学を求めるようになって初めて自分のしていることが本物といえるのではないか」。人間としての学問を追求する研究者である。(広報室)



図2. 国民性調査の結果を知ってもらうために作られたPRグッズのトランプや物差し。

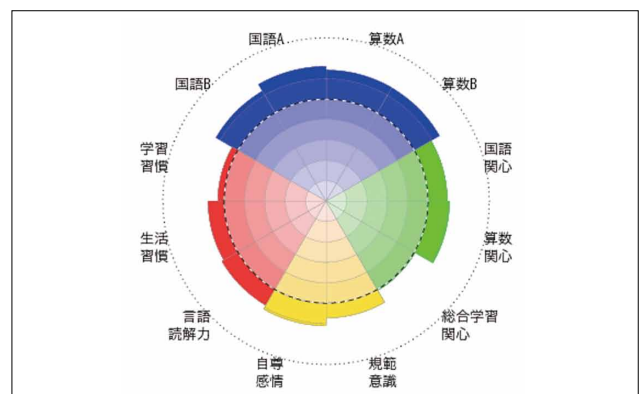


図3. 一般的なレーダーチャートに代えて、学力調査で用いられているチャートの例。

「Seminar on Application of Mathematical Modeling in Forest Resource Management」共同開催報告

平成25年11月18日(月)に、マダガスカル・アンタナナリボのUniverité d'Antananarivoにおいて、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、「Seminar on Application of Mathematical Modeling in Forest Resource Management」を開催しました。このセミナーは、現在様々な自然環境問題に直面しているマダガスカルにおける効率的・効果的な土地利用・森林資源管理の実施に向けて、それらの問題に対する数理モデルの応用、最新の知見・情報の共有および、今後の国際共同研究の萌芽を探る目的として開催しました。セミナーの前半は、木島氏(琉球大学)が様々な攪乱とそれに伴う被害軽減に向けた空間的なシミュレーションモデルの応用について紹介し、後半は、吉本が森林資源管理における最適化フレームワークについて動的計画法による林分レベルの間伐スケジューリング最適化モデル、線形計画法、整数計画法による森林ランドスケープレベルの伐採スケジューリング最適化モデルについて基本的な考え方から最新の応用例まで紹介しました。セミナーの最後には、参加学生からの質問や疑問について事例を交えて答え、議論しました。19日には、エコツーリズムを導入し始めたMandrakaにある演習林の視察を行いました。20日にはAntsirabeという町まで景観および、水源管理に関する視察に行きました。土壌流出や土砂崩れがひどい地域が多

くみられ、こういったリスクの軽減がこの地域における1つの重要課題であることが分かりました。(吉本 敦)



セミナーの様子



ランドスケープ・水源涵養機能 視察

「Joint International Symposium By Japan and Czech Republic Data Acquisition, Statistical Modeling and Decision-Making Toward Better Forestry」共同開催報告

平成25年10月8-9日の二日間に渡り、チェコ・プラハのCzech University of Life Sciences Pragueにおいて、統計数理研究所リスク解析戦略研究センター、Czech

University of Life Sciences Prague, Japan Society of Forest Planningの共催により、「Joint International Symposium By Japan and Czech Republic Data



ワークショップの様子



Acquisition, Statistical Modeling and Decision-Making Toward Better Forestry」を開催しました。今回のシンポジウムは、今後のより効果的な森林資源管理に向けて、データ収集、統計分析、モデリング、意思決定サポートシステムについて最新の研究成果を議論することを目的にしました。シンポジウムでは、5つのセッション

ンにおいて最新の技術を駆使した3次元データ収集と計測についての報告、収集したデータをもとにした樹木の成長モデリングとモデル選択の研究成果の報告、そして、意思決定サポートシステムについて最適化モデルを用いた様々な応用研究の報告が行われました。また、ワークショップではインフォマティクスと様々な可視化ソフトウェアの紹介がありました。それぞれの発表について活発な議論が展開され、参

加者にとって、最新の知見、技術、ノウハウを今後の森林資源管理に効果的に導入、活用、反映させるための課題と解決方法について検討する貴重な機会となりました。尚、ワークショップの運営は、Czech University of Life Sciences Prague の Róbert Marušák 氏、ISM の Peter Surovy 氏らが中心に行いました。(吉本 敦)

「Training on Introduction to Statistical Analysis in "R" for Forest Resource Management」共同開催報告

平成25年9月25-27日の三日間に渡り、カンボジア・プノンペンの Forest and Wildlife Training Centerにおいて、統計数理研究所リスク解析戦略研究センター、Institute of Forest and Wildlife Research and Development の 共催により、「Training on Introduction to Statistical Analysis in "R" for Forest Resource Management」を開催しました。今回のトレーニングワークショップは、前年に引き続き、カンボジアにおける森林資源管理の実施に関わる人材育成を念頭に、統計解析ソフトRを習得し、参加者が実

際に収集・蓄積しているデータに対してRを用いた統計解析ができるようになることを目的にしました。今回は2年目でありましたが、Rを新たに学びたいという参加者が多く、前年度とは異なる参加者を対象に、1日目はRのイントロダクション、インストール作業、データの入力、データハンドリングについて講義を行いました。2日目は、前日のレビューを行った後、引き続き、分析の基礎となるデータハンドリングについて講義を行い、その後、Rでのグラフィックス作成を行いました。3日目は、簡単な回帰分析について講義し、与えられたデータに対して、回帰分析により推定されたパラメータを用いた樹木の成長予測を行いました。皆、非常に熱心にワークショップに参加しており、休憩時間中にも多くの質問があり、講師陣は休憩も十分に取れない状況でありました。3日目のワークショップ終了後には、今回のワークショップ参加者に対して、修了証の授与が行われました。参加者は今回のワークショップについて、非常に、満足しており、すでに来年度の開催を楽しみにしている様子でした。ワークショップの運営は、Institute of Forest and Wildlife Research and Development Forestry Administration、Cambodia の Acting Director、Sokh Heng 氏らが中心に行いました。

(吉本 敦)



ワークショップの様子

研究教育活動

2013年10月—12月の公開講座実施状況

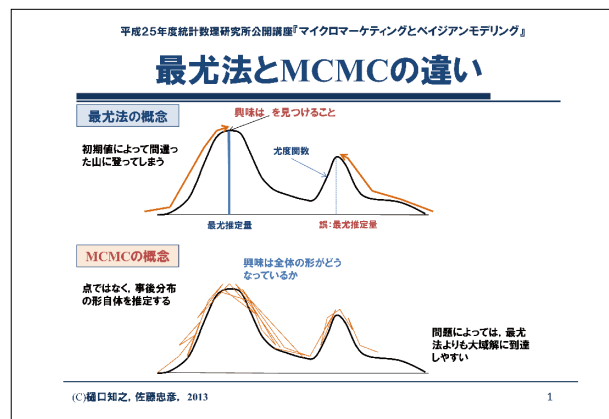
平成25年度後期公開講座は10月3日(木)と4日(金)の「マイクロマーケティングとベイジアンモデリング」で始まりました。筑波大学の佐藤忠彦准教授と当研究所の樋口知之所長が、ベイジアンモデリング、特に状態空間モデルに焦点を当て、ビッグデータを用いたモデリング技法とその推定技術について講義をしました。会場は満席でしたが、更に、先着順のため申込のできなかった受講希望者が多数あり、同

一内容の講座を平成26年2月4日(火)と5日(水)に追加開催することになりました。

後期2回目は11月22日(金)と25日(月)の「離散最適化」で、講師は東京工業大学の松井知己教授、東京大学の岩田覚教授、大阪大学の梅谷俊治准教授、そして成蹊大学の池上敦子教授でした。受講生の数は40名の定員を大幅に超え、68名となりました。

12月19日(木)と20日(金)には「確率的トピックモデル」が行われました。1日目は、当研究所の持橋大地准教授が独りで講義をしました。2日目の講義は、NTTコミュニケーション科学基礎研究所の石黒勝彦氏が午前を、同じく岩田具治氏が午後を担当しました。96名の受講生の中には、宮城、京都、大阪、兵庫、福岡などの遠方からの参加者が含まれていました。

(情報資源室)



公開講座「マイクロマーケティングとベイジアンモデリング」資料より

統計数理セミナー実施報告(2013年11月～2014年1月)

毎週水曜16時から所内研究教育職員及び外部の方が1人40分ずつ、1日に2人の講演を行っています。11月から2014年1月のセミナーは以下の通り行われました。

日程	氏名	タイトル
2013年11月13日	志村 隆彰	極小値データの特徴
11月13日	斎藤 正也	2009年インフルエンザ動向調査に見られる感染流行への外部地域の影響
11月20日	中村 隆	鶴岡言語調査データのコウホート分析(III)
11月20日	吉田 亮	4Dライブセルイメージングのデータ解析について
11月27日	持橋 大地	音象徴の教師なし学習
11月27日	北野 利一	極値統計学 再訪 ～ Rule of Thirds の視点から
12月 4日	上野 玄太	数値予報モデルへのデータ同化
12月 4日	藤澤 洋徳	経験推定可能でアフィン不変性をもつダイバージェンス
12月11日	深谷 肇一	多変量時系列モデルを用いた生物季節の長期変化の統計的推測
12月11日	黒木 学	Effect restoration and measurement bias in causal inference
12月18日	金藤 浩司	環境統計学で用いられる母幾何平均について
12月18日	塩田さやか	統計的手法に基づく音声合成のための聴覚的特徴を用いたクロスリンガル話者適応
2014年 1月 8日	栗木 哲	チューブの体積を最小にするフーリエ・多項式最適実験計画
1月 8日	吉野 諒三	アジア・太平洋価値観国際比較調査の概要
1月15日	Robert C. Griffiths	Multivariate Jacobi Polynomials on the Dirichlet and their reproducing kernels
1月15日	小林 景	データ空間上の距離とその変形
1月29日	伊藤 聡	クリンチとエリミネーションの数理
1月29日	逸見 昌之	共変量シフトと欠測データ問題

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。 <http://www.ism.ac.jp/>

(メディア開発室)

2013年公開講演会「統計学が切り拓く脳科学、脳工学の未来」

脳機能を巡る最新の研究成果を一般に紹介する2013年統計数理研究所公開講演会「統計学が切り拓く脳科学、脳工学の未来」が、11月7日午後1時半から統数研2階の大会議室で開催されました。公開講演会は毎年、教育・文化週間(11月1日～7日)の関連行事として企画。5回目となる今回は学生、近隣の方々、企業や医療関係の方など約100人が参加し、脳科学、脳工学研究の最前線に触れながら、その背後でいかに統計学が重要な役割を果たしているかについて理解を深めていました。

プログラムは以下の通り。

○主催者あいさつ 樋口知之(統計数理研究所所長)

1. 脳信号解析における統計学の役割
三分一史和(統計数理研究所 准教授)
2. 脳からの信号で機械を操作する
神作憲司(国立障害者リハビリテーションセンター研究所 脳機能系障害研究部脳神経科学研究室 室長)
3. ヒト脳活動を精細に測るためのデータ統合技術
山下宙人((株)国際電気通信基礎技術研究所 脳情

報解析研究所 計算脳イメージング研究室 室長)

司会 中野純司(統計数理研究所 統計科学技術センター)

約8,600億もの神経細胞が活動する人間の脳。そこでは無数の信号情報がやりとりされ、それを統合する精巧なシステムが日夜、稼働しています。トップバッターの三分一准教授は、様々な方法で行われる脳信号の計測の手法について紹介。ヒトの脳波計測は80年以上前から行われ、近年、脳内の血流変化から活動する部位を調べる機能的磁気共鳴画像法(fMRI)が広く普及しています。脳信号からヒトの意志や思考を読み取るなど、工学的な応用研究が飛躍的に進んでいることを解説しました。

続いて登壇した神作室長は、脳からの信号で機械を操作するための、ブレイン-マシン・インターフェイス(BMI)という技術に注目しました。脳からの信号でワープロや家電を操作する環境制御システムが開発中です。脳波信号を計測しそれを解析することで、電気製品などに指令を送る夢のシステム。会場内からも実用開発への期待の声が上がりました。

統数研の旧広尾時代に総合大学院大学に通学した経験のある山下室長は「脳活動データの計測は、脳を科学するための第一歩です。計測の精度を向上させることは脳科学の新しい発見につながります」と指摘し、脳活動を可視化するいくつかの方法を紹介しました。統計手法を用いることで複数の脳計測データが統合され、脳内ネットワーク情報処理の研究が飛躍的に進んでいます。統計科学が脳研究の可能性をたくましく切り拓いていることが、あらためて理解されました。

ネット検索で講演会を知ったという立川市の医師(35歳)



は「たまたま休日だったので参加した。脳波を解析する側の個人としての能力差をどのように考慮すべきか。神経内科を専攻しているが、あらためて自分の疑問点を整理することができた。」と話しました。近隣に配布されたチラシを見て日野市から参加した男性(76歳)は「これまでは都心まで行かないと、最新の科学研究の講演に触れることができなかった。統数研のおかげで学習機会が広がった感じがする。質問も受ける時間を作ってくれると有り難かったが…」と要望。また、統数研ホームページを見て参加した認知科学の女性研究者は「データ解析から脳機能の解析に迫る手法が、良く理解できた。レジュメにいくつかの図表が欲しかった。こんどは土、日曜にも開催していただけると、元学生だった主婦なども参加できるのでは」と、今後の公開講演会の充実に期待を寄せました。

樋口所長は「立川に移転して5年目。地域の一般の人々の問題意識の向上に応えられるような統数研像をさらに追求していきたい」と話していました。(広報室)

千葉大学大学院 特別講演「脳・神経系の研究への新しい試み」

2013年10月23日、千葉大学・大学院工学研究科において、瀧澤由美准教授と深澤敦司元特任教授とによる特別講演が開催されました。これは国際神経学会 Best Paper Awardの受賞(2012年7月)を機に須貝康雄教授の企画により行われました。

瀧澤准教授は、生体の神経系は自然の原理(物理法則)と生命の原理(有機的生体反応)の両面を満たすとし、実空間と脳内皮質の相似的対応(脳内地図)のしくみの研究成果を述べ、またその応用について話しました。

また、深澤博士は個(ニューロン)と集合(ニューロン群)の關係に着目し、電気・物性・通信理論の基礎から神経系の動作原理を解明し、上記研究の基礎を構築したことを講演しました。

物性、化学、通信、電気専門の教授(教員7名)と、電

気・通信系OB3名、院生約20名が参加し、世界的な研究成果であるとの高い評価を得ました。(瀧澤由美)



会場風景(瀧澤由美准教授による講演)

文部科学省委託事業:数学協働プログラム「安心・安全・快適な社会インフラ維持への数理科学の適用」

統計数理研究所は文部科学省委託事業として数学協働プログラムを推進しています。数学協働プログラムには、ワークショップ・スタディグループ・作業グループという三つの枠組みがあります。本集会はスタディグループの一つとして行われました。文部科学省からも視察に来られています。

スタディグループは、企業などと大学などが一緒に短期集中で議論を行い、何らかの成果を得ようという試みです。今回は、鉄道総合技術研究所と統計数理研究所が一緒になって行いました。前半は11月25日-26日に行われました。後半は1月27日-28日の予定です。

写真は、前半初日午前、統計数理研究所のセミナー室で、公開で行った際のものです。20名程度の参加者がいました。そこでは三つの課題「直交格子法流体解析における統

計的補正モデル」「車輪・レール間接触問題の数値解析手法について」「振動加速度波形からの乗客アンケートの結果推定」が提示されました。公開講演会の後に、3つのグループに分かれて、それぞれで議論を行いました。（藤澤洋徳）

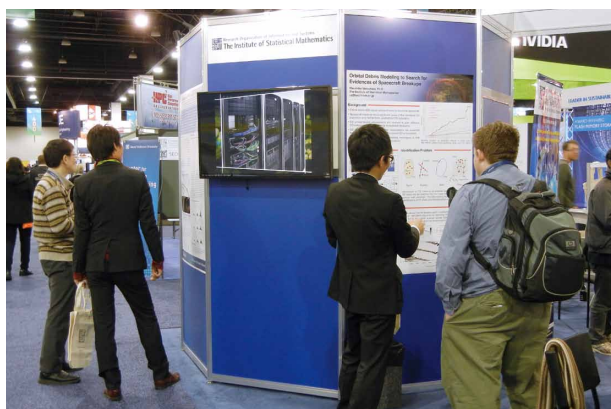


統数研トピックス

SC13に出展

アメリカ合衆国・デンバーで11月18~21日の4日間開催されたSC13に出展してきました。この展示会は、スーパーコンピュータに関する要素技術の開発や大規模計算による科学研究を行っている多くの企業や研究所が参加しているもので、今年で25周年目をむかえます。

来年度に統計数理研究所のスーパーコンピュータの刷新・新規導入を控えており、その中のひとつに、HPCIシステムに計算資源を提供するデータ同化スーパーコンピュータシステムがあります。ノード当たり512GBという巨大なメモリは類例を見ないものです。中野純司統計科学技術センター



長らが開発を進めている新しいRの並列化環境であるRhpiの紹介に対しても、すぐに使ってみたいので入手やインストールの方法を教えてほしいといった質問が相次ぎました。ハードとソフトの両面から大規模データ処理を支援する環境の整備とその広報が必要であると感じられました。また、例年配布している国民性調査トランプも展示1日目にほとんど無くなるほどの人気がありました。SC14では、実施中の第13次調査結果も取り込みながら、統数研が調査に果たす役割を発信していきたいと思います。（斎藤正也）



大学共同利用機関シンポジウム2013開催

11月16日に東京国際フォーラムにおいて大学共同利用機関シンポジウム2013「万物は流転する—因果と時間—」が

開催され、①高エネルギー加速器研究機構多田将助教による「ニュートリノ振動実験—この世で最も小さな粒子を探るこ

の世で最も大きな実験—」、②統計数理研究所間野修平准教授による「時間をさかのぼる推理—ゲノム解析から明らかになる私たちの起源—」、③人間文化研究機構国文学研究資料館武井協三名誉教授による「『曾根崎心中』因果応報—作者近松はなぜ残酷になるのか—」、④自然科学研究機構核融合科学研究所伊藤公孝教授による「法則と予測—過去・現在・予測・そして未来—」の4つの講演が行われました。

本研究所の展示コーナーにおいては『人との繋がり、コミュニケーション力の重要性』をコンセプトに、所長研究紹介コーナー、研究所紹介DVDビデオ、研究所の研究紹介ポスターを展開し、沢山の来場者が足を止めてくださいました。来場者は、約260名ほどあり盛況のうちに終了しました。統



計数理研究所が一体何を研究しているところなのか、コンパクトにお知らせすることは大変難しいものであることを実感しました。
(広報室)

2013年度富士通SS研合同分科会での展示

2013年10月24日(木)に、ホテルオークラ神戸に於いて開催された富士通SS研合同分科会においてデータサイエンティスト育成事業についての展示を行いました。富士通SS研は(<https://www.sskn.gr.jp/MAINSITE/aboutus/index.html>)にあるように、大学、研究所などの科学技術分野のコンピュータ利用機関を主体とした研究会(任意団体)であり、合同分科会は一年に一回、開催されています。

教育関係の分科会もあることやビッグデータ解析等の研究開発を行っている富士通関係者も多く参加されることから、データサイエンティスト事業の紹介をすることが会員への

情報提供になると考え、丸山副所長の協力を仰ぎ、展示に参加させていただきました。当日は文科省からの委託事業を丸山副所長に協力して勢力的に推進している藤田さんにブースでの説明を行っていただきました。分科会参加者は127名だったそうですが、その中の多くの方に、統数研の展示を見に来ていただけた、展示(全部で8展示)のアンケートでは第一位の評価を得ることができています。これからも、統計数理研究所はデータサイエンティスト育成のための諸事業に積極的に取り組んでいく予定です。
(田村義保)

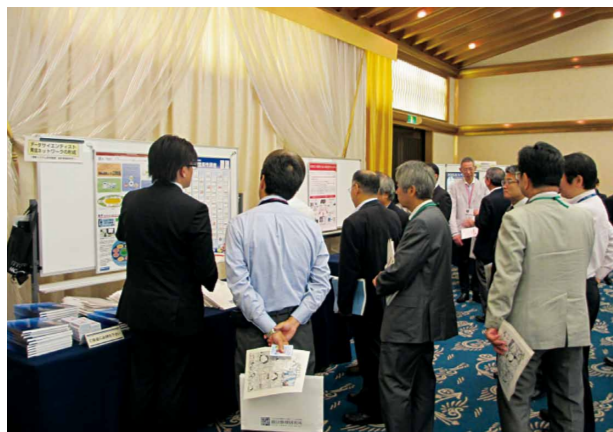


写真1 会場風景

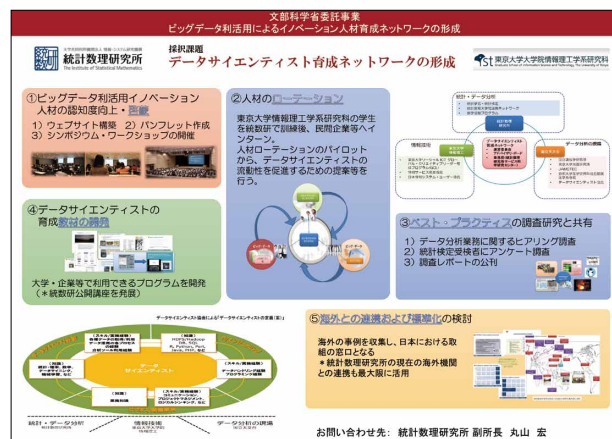


図1 当日のポスター原稿

新スーパーコンピュータシステム導入

統計数理研究所においては、平成26年4月1日から、共用クラウド計算システム及びデータ同化スーパーコンピュータシステムの2式のスーパーコンピュータシステムを稼働させ、共同利用のために公開いたします。

前者は平成23年から27年の5年計画で実施中の特別事業「統計数理クラウドを用いた共同研究による新学術分野の

創成」で必要な計算機器のための経費が平成24年度補正で措置していただけたことにより導入可能となったシステムです。標記事業を推進するために活用いたします。使いやすく高度な統計数理研究環境をクラウド上で提供する予定です。多くの方が共同利用していただければ幸いです。

後者は、昨年度から一般利用が始まった神戸に設置され

ている京コンピュータを中心とするHPCI事業の計算資源の一つとして提供するために平成24年度補正予算を用いて導入しました。世界最大の64TBの共有記憶空間を持つ分散共有型計算機2セットを中心とした構成です。HPCI事業を通じての利用だけでなく、統計数理研究所の共同利用者也利用可能です。大容量の共有記憶ということで大規模データ解析だけでなく、ゲノム解析や社会科学的なシミュレーション、相関分析にも適していると考えています。これらのシステムはデンバーで開催されたSC13でも紹介させていただいております。

(田村義保)

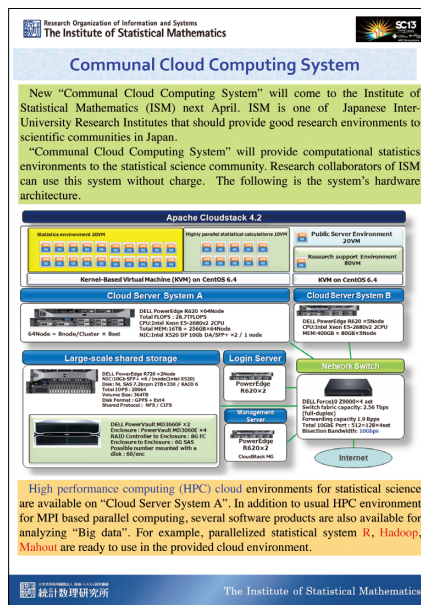


図1 SC13での共用クラウド計算システムのポスター

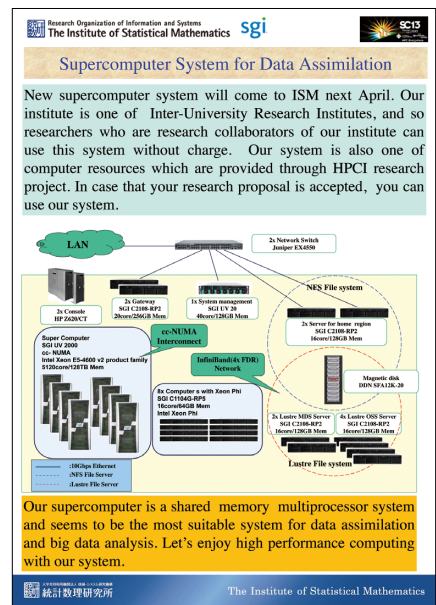


図2 SC13でのデータ同化スーパーコンピュータシステムのポスター

情報・システム研究機構永年勤続者表彰式

平成25年12月4日(水)に会議室にて、情報・システム研究機構永年勤続者表彰式が行われました。被表彰者は極地研・統数研統合事務部共通事務センター 塩原研一氏、統計科学技術センター 中村和博氏で、樋口所長から表彰状が授与され、併せて記念品が贈呈されました。

(企画グループ・人事担当)



総合研究大学院大学複合科学研究科統計科学専攻関係

大学院説明会の実施

平成25年11月8日(金)に会議室1において、平成25年度第2回大学院説明会を開催し、16名の参加者がありました。内容は、今年度の入試ガイダンス、5年の課程と後期3年の課程のカリキュラムの説明、在学生による研究テーマの紹介、修了後の進路紹介などについての説明を行い、説明会終了後には、希望者に対して教員との面談を行いました。

(企画グループ・研究支援担当)



●公開講座

一般社会人・学生を対象に、公開講座を開催します。

【受講料、受付方法変更のお知らせ】

諸般の事情により、平成26年度は受講料と受付方法を変更いたします(受付方法の詳細についてはホームページをご覧ください)。

変更によりご不便をおかけする面もあるかと存じますが、ご理解をお願いいたします。

① 統計学概論

日時:5月13日(火)～5月16日(金) 10時～16時(20時間)
講師:山下智志・小林景・小山慎介・野間久史(統計数理研究所)
申込受付:3月27日(木)10時～4月9日(水)17時
受講料:20,000円
定員:100名(申込多数の場合は抽選)
講義レベル:初級

統計学の入門編として基礎的な講義です。

これからデータ分析・モデリングを行う初心者に必要な統計学の知識を、網羅的に紹介します。基本数学的による統計学の記述だけでなく、簡易なデータ分析例を適宜示すことにより、直感的な理解を深めることを目的としています。

高校数学の微積分と初歩的な線形代数の知識があることを前提としています。

- ・統計のための基礎数学
- ・記述統計、確率分布
- ・統計的推定
- ・統計的検定
- ・回帰分析と統計モデル
- ・重回帰分析とモデル選択
- ・多変量解析の概要と用例
- ・時系列モデルの初歩

② 非定常時系列解析

日時:6月5日(木)～6月6日(金) 10時～16時(10時間)
講師:川崎能典(統計数理研究所)、姜興起(帯広畜産大学)
申込受付:4月28日(月)10時～5月7日(水)17時
受講料:10,000円
定員:70名(申込多数の場合は抽選)
講義レベル:中級

本講座では、平均が非定常な(すなわち時間とともに変動する)時系列に対する統計的モデリングの方法を解説する。取り扱う内容は以下の通り。[1] 定常時系列解析の復習 [2] 状態空間モデル [3] トレンドの推定と季節調整 [4] 循環変動を含む時系列の季節調整 [5] 局所定常ARモデル [6] 時変係数AR (VAR) モデル

学部卒業程度の統計学・統計解析の知識(例えば

線形回帰モデルに関する推測統計的な議論になじみがあるレベル)を前提に講義を行う。

参考書:

- 北川源四郎「時系列解析入門」岩波書店
- 姜興起「ベイズ統計データ解析」共立出版

③ 統計モデルによるロバストパラメータ設計

日時:6月26日(木) 10時～16時(5時間)
講師:河村敏彦(統計数理研究所)
申込受付:5月19日(月)10時～5月28日(水)17時
受講料:5,000円
定員:40名(申込多数の場合は抽選)
講義レベル:中級

ロバストパラメータ設計は、創始者である故田口玄一博士が半世紀かけて体系化した品質改善のための技術方法論である。これは使用環境条件などの誤差因子に対してロバスト(頑健)になるように制御因子を設計することにより、特性や機能性のばらつきを低減する方法として知られている。本セミナーでは、統計的モデリングアプローチによる最適化を述べ、SN比解析との類似点および相違点などの観点から実データを用いて解説する。また、統計解析ソフトJMP上で動作するアドインを用いて、データ解析例を紹介する。なお、本セミナーは、基本的な実験計画法やタグチメソッド、多変量解析法などの入門レベルの知識があることが望ましい。

教科書:

- 河村敏彦・高橋武則(2013)『統計モデルによるロバストパラメータ設計』日科技連出版社

参考書:

- 河村敏彦(2011)『ロバストパラメータ設計』日科技連出版社

④ 標本調査データの分析

日時:7月14日(月)～7月16日(水) 10時～16時(15時間)
講師:土屋隆裕(統計数理研究所)
申込受付:6月9日(月)10時～6月18日(水)17時
受講料:15,000円
定員:30名(申込多数の場合は抽選)
講義レベル:中級

標本調査データの分析にあたっては、標本の抽出デザインを考慮することが必要です。当講座では、層や集落、不等抽出確率などを用いて抽出された標本データの分析法について、理論の解説と実際の演習をとおして理解を深めます。想定する受講者は、社会調査など統計調査に関心のある大学生・大学院生や、実際に調査データを扱う実務者の方々と、統計学と標本調査の用語について基本的な知識を前提にします。受講者が持参したパソコン上で、RまたはSPSSを用いた演習を行います。その詳細は申込み後の案内をご覧ください。

参考書:

- 土屋隆裕(2009)『概説 標本調査法』朝倉書店

E 動的幾何学ソフトウェアGeoGebraの使い方と数学教育における活用

日時:8月1日(金) 10時～16時(5時間)
 講師:丸山直昌(統計数理研究所)、濱田龍義(福岡大学)
 申込受付:6月23日(月) 10時～7月2日(水) 17時
 受講料:5,000円
 定員:40名(申込多数の場合は抽選)
 講義レベル:初級

無償で利用できる動的幾何学ソフトウェアGeoGebraを受講者が持参するパソコン上で動かしながら実習を行う。題材は中学校、高等学校での数学の問題を中心に選ぶ。高等学校の新しい指導要領に統計が入ったことを踏まえ、統計の問題も扱う。このような実習を通して、GeoGebraが中等教育の現場で極めて有用なソフトウェアであることを認識して貰い、現場での利用上の問題点・注意点も話題とする。GeoGebraは現在Windows、Mac、Linux、FreeBSDなど種々の環境下で動かせるが、それぞれの動作環境下での違いの解説も行う。さらにGeoGebraで描いた図を取り込んだ文書作成について、解説と実習を行う。

F 多変量解析法[社会調査士資格E科目対応(予定)]

日時:9月9日(火)～9月12日(金) 10時～16時(20時間)

講師:馬場康維・清水信夫(統計数理研究所)、大森拓哉(多摩大学)
 申込受付:7月22日(火) 10時～7月30日(水) 17時
 受講料:20,000円
 定員:100名(申込多数の場合は抽選)
 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、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。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。

資料を配布する。

詳細は、以下のホームページをご覧ください。

<http://www.ism.ac.jp/lectures/kouza.html>

(情報資源室)

共同利用

平成25年度共同利用公募追加採択課題

【共同利用登録】2件

分野	研究課題名	研究代表者(所属)
f2	大規模ショック時の日経平均先物の価格変動について	阿部 千晶(同志社大学 商学部・学部学生)
g3	融合プロジェクト研究におけるNGSのデータ解析と結果の検証	高田 豊行(大学共同利用機関法人 国立遺伝学研究所)

【一般研究1】3件

分野	研究課題名	研究代表者(所属)
a2	多段変調装置による管内圧力波のベイジアンモデリングと信号解析	小山 慎介(統計数理研究所 モデリング研究系・助教)
d7	日本人の意識調査のコウホート分析	中村 隆(統計数理研究所 データ科学研究系・教授)
d7	都市尺度としての人口ポテンシャルの再検討	赤枝 尚樹(関西大学 社会学部・助教)

(企画グループ・研究支援担当)

外部資金・研究員等の受入れ

共同研究の受入れ

受入年月日	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
H25.4.1	株式会社地球快速適応インスティテュート 取締役副所長 田中 栄司	分子設計のためのケモインフォマティクス技術の開発	H25.4.1～H26.3.31	3,000,000	モデリング研究系 吉田 亮 准教授
H25.6.3	株式会社日立製作所 中央研究所 所長 長我部 信行	社会現象のシミュレーション技術の研究	H25.6.3～H26.3.31	4,000,000	モデリング研究系 丸山 宏 教授

受入年月日	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
H25.8.1	ムサノ機器株式会社 代表取締役 阿部 正治	電磁波による測位方式の研究	H25.8.1～ H26.3.31	1,980,000	モデリング研究系 瀧澤 由美 准教授
H25.11.13	パナソニック株式会社 所長 岡村 和男	i-Vectorを用いた高精度な音識 別アルゴリズムの評価、及び独 自アルゴリズムの研究	H25.11.13～ H26.3.31	1,507,000	モデリング研究系 松井 知子 教授

(企画グループ・研究支援担当)

受託研究の受入れ

受入年月日	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
H25.4.1	国立大学法人筑波大学分任 契約担当役研究担当副学長 三明 康郎	(i) 気候変動リスクの評価の基 盤となる確率予測情報の創出 b. 気候変動予測データの統計 学的解析手法の開発	H25.4.1～ H26.3.31	15,038,250	モデリング研究系 上野 玄太 准教授
H25.11.15	横浜市教育委員会 教育長 岡田 優子	横浜市学力・学習状況結果分 析チャートの作成委託	H25.11.15～ H26.3.31	2,155,335	データ科学研究系 土屋 隆裕 准教授

(企画グループ・研究支援担当)

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	特命教授又は客員研究員の称号付与
藤原 海世	東京工業大学大学院・修士 課程	データサイエンティスト育成ネッ トワークの形成	H25.11.14～ H26.3.31	丸山 宏 教授	
柳本 武美	中央大学 理工学部・客員 教授	データの情報量に対応する事前 分布の情報量	H25.12.1～ H26.3.31	山下 智志 リスク解析戦略 研究センター長	
辻川 美沙貴	パナソニック株式会社 R&D 本部、クラウドソリューションセ ンター・社員	音識別要素技術に関する研究 開発	H25.11.25～ H26.3.31	松井 知子 教授	
郭 一村	北京大学・地球宇宙科学研 究科・博士後期課程	地震活動フェーズの自動客観識別 における隠れマルコフモデルの開発	H25.12.16～ H26.1.30	庄 建倉 准教授	
王 婷	ニュージーランドオタゴ大学・ ポストドクター	地震活動フェーズの自動客観識別 における隠れマルコフモデルの開発	H25.12.16～ H26.1.31	庄 建倉 准教授	
Somayah Danafar	ダレモレ人工知能研究所 (ス イス)・博士後期課程	カーネル法によるセミパラメトリッ ク変数選択法の研究	H25.11.12～ H25.12.11	福水 健次 統計的機械学習 研究センター長	
片山 翔太	大阪大学大学院 基礎工学 研究科 システム創成専攻 数理科学領域 (狩野研究室) 日本学術振興会特別研究員PD	高次元データにおけるロバストな スパース推定	H25.9.17～ H26.3.31	藤澤 洋徳 教授	
Donald ST.P Richards	ペンシルバニア州立大学 統 計学部・教授	高次FKG不等式の研究	H26.1.27～ H26.2.1	栗木 哲 教授	
Mercedes T. Richards	ペンシルバニア州立大学・教 授	計算天文学、とくに相互作用を持っ た2体問題に関するデータ解析	H26.1.27～ H26.2.1	栗木 哲 教授	
河口 朋子	株式会社 筑波総合研究所・ 主任研究員	自殺の地域統計作成のための 政府統計を用いたデータベース の構築とその利活用	H26.1.1～ H26.3.31	椿 広計 教授	
Elisa Varini	国家研究評議会 応用数学と 情報技術研究所・研究者	地震活動フェーズの自動客観識別 における隠れマルコフモデルの開発	H26.1.22～ H26.3.14	庄 建倉 准教授	
Ilia Negri	ベルガモ大学・准教授	確率過程の統計的推測	H25.12.23～ H25.12.27	西山 陽一 准教授	
周 仕勇	北京大学 地球空間科学学 院・教授	地震活動フェーズの自動客観識別 における隠れマルコフモデルの開発	H26.1.11～ H26.1.20	庄 建倉 准教授	

(企画グループ・研究支援担当)

人 事

平成25年10月31日転出者(事務職員)

異動内容	氏 名	新職名等	旧職名等
辞 職	織田 知子	東京農工大学	極地研・統数研統合事務部企画グループ (統数研担当) チームリーダー(研究支援担当)

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
グリフィス ロバート チャールズ Griffiths Robert Charles	Department of Statistics, University of Oxford, Emeritus Professor 連合王国	数理・推論研究系 統計基礎数理グル ープ	客員 教授	数理集団遺伝学にお ける確率過程の研究	H26.1.8～ H26.3.28	間野 修平 准教授
リウ シュアンツェ Liu Shuangzhe	Faculty of Information Sciences and Engineering, University of Canberra, Associate Professor オーストラリア連邦	データ科学研究系 構造探索グループ	客員 准教授	環境データ解析のた めの影響診断法の研 究	H26.1.10～ H26.2.5	金藤 浩司 教授

(企画グループ・人事担当)

所外誌掲載論文等

本研究所の教員、研究員、総研大(統計科学専攻)大学院生によって発表された論文等を前号に引き続き紹介します。

- 足立 幸子, 池上 敦子, 宇野 毅明, 作業者の操作観点に基づく訪問介護勤務表作成支援システムの利点と課題, 日本経営工学会論文誌, 64(3), 488-497, 2013
- Fujiki, Y., Yokota, S., Okada, Y., Oku, Y., Tamura, Y., Ishiguro, M. and Miwakeichi, F., Standardization of Size, Shape and Internal Structure of Spinal Cord Images: Comparison of Three Transformation Methods, *PLOS ONE*, 8(11):e76415, 2013.11
- Jia, K., Zhou, S., Zhuang, J. and Jiang, C., Possibility of the Independence between the 2013 Lushan Earthquake and the 2008 Wenchuan Earthquake on Longmen Shan Fault, Sichuan, China, *Seismological Research Letters*, 85, 60-67, doi:10.1785/0220130115, 2014.01
- Jiang, C., Zhuang, J., Long, F., Han, L. and Guo, L., Statistical analysis of ETAS parameters in the early stage of the 2013 Lushan MS7.0 earthquake sequence, *Acta Seismologica Sinica*, 35, 661-669, doi:10.3969/j.issn.0253-3782.2013.05.005, 2013.09
- Kohana, M., Okamoto, S. and Ikegami, A., A Meta-heuristic Approach for Dynamic Data Allocation on a Multiple Web Server System, *IEICE TRANSACTIONS on Information and Systems*, E96-D(12), 2645-2653, 2013
- Kumazawa, T. and Ogata, Y., Quantitative description of induced seismic activity before and after the 2011 Tohoku-Oki earthquake by nonstationary ETAS models, *Journal of Geophysical Research*, 118(12), 6165-6182, doi:10.1002/2013JB010259, 2013.12
- Ogata, Y., A prospect of earthquake prediction research, *Statistical Science*, 28(4), 521-541, doi:10.1214/13-STS439, 2013.11
- Takada, T., Ebata, T., Noguchi, H., Keane, T., Adams, D., Narita, T., Shin-I, T., Fujisawa, H., Toyoda, A., Abe, K., Obata, Y., Sakaki, Y., Moriwaki, K., Fujiyama, A., Kohara, Y. and Shiroishi, T., The ancestor of extant Japanese fancy mice contributed to the mosaic genomes of classical inbred strains, *Genome Research*, 23, 1329-1338, doi:10.1101/gr.156497.113, 2013.08
- 田辺 隆人, 池上 敦子, 「最適化モデリング」, 応用数理ハンドブック(日本応用数理学会(監)/薩摩 順吉, 大石 進一, 杉原 正顕(編)), 朝倉書店, 東京, 2013.11
- Yamashita, S. and Yoshida, T., A collateralized loan's loss under a quadratic Gaussian default intensity process, *Quantitative Finance*, 13(12), 1935-1946, doi:10.1080/14697688.2012.762459, 2013.12
- Yamashita, S. and Yoshida, T., Analytical solutions for the expected loss of a collateralized loan: a square root intensity process negatively correlated with collateral value, *Journal of Credit Risk*, 9(2), 27-44, 2013.06
- 吉野 諒三, 調査機関紹介 統計数理研究所の歴史と展望, よろん 日本世論調査協会報, 112, 40-43, 2013.10
- Yoshino, R., Trust of nations on cultural manifold analysis (CULMAN) - sense of trust in our longitudinal and cross-national surveys of national character-. in Dryakhlov, N. I., Ishikawa, A., Kupreychenko, A. B., Sasaki, M., Toshchenko, Zh. T. and Shadrikov, V. D. (eds.), *Trust in society, business and organization*, National Research University, Moscow, 213-256, 2013.10
- 吉野 諒三, 最新 心理学事典「調査法」の項, 平凡社, 2013.12
- Zhuang, J., Werner, M. J. and Harte, D.S., Stability of earthquake clustering models: Criticality and branching ratios, *Physical Review E*, 88(6), 062109, doi:10.1103/PhysRevE.88.062109, 2013.12

(情報資源室)

刊行物

Research Memorandum(2013.11~2014.1)

No.1181: 石綿 元, 高輝度放射光屈折コントラスト撮影法における多波回折の適用

No.1182: Komori, O., Eguchi, S. and Copas, J., Generalized t-statistic for two-group classification

(メディア開発室)

統計数理 第61巻 第2号

特集「環境リスクと統計解析 —データ基盤構築と解析—」

「特集 環境リスクと統計解析 —データ基盤構築と解析—」について

金藤 浩司、吉本 敦	179
森林炭素動態シミュレーションシステムを用いた気候変動が森林炭素吸収量に及ぼす影響評価の試行 [研究ノート]	
光田 靖、鹿又 秀聡、松本 光朗	181
離散データに対する回帰モデルによる冠雪害の解析 [研究ノート]	
加茂 憲一、嘉戸 昭夫、吉本 敦	189
林分施業法に基づく持続的・順応的森林管理とデータ基盤 [総合報告]	
尾張 敏章	201
ビッグデータ時代の環境科学 —生物多様性分野におけるデータベース統合、横断利用の現状と課題— [総合報告]	
大澤 剛士、神保 宇嗣	217
鉛はんだ代替におけるリスクトレードオフ評価のための用量反応関係の導出 [原著論文]	
竹下 潤一、蒲生 昌志	233
東京電力福島第一原子力発電所近隣地域における空間放射線量率と直下土壌の放射能汚染度との関連について [研究ノート]	
大瀧 慈、大谷 敬子、今中 哲二、遠藤 暁、星 正治	247
地震音波データ同化システムの開発 —双子実験による検証— [原著論文]	
長尾 大道、樋口 知之	257
ゼロの多いデータの解析：負の2項回帰モデルによる傾向の過大推定 [原著論文]	
南 美穂子、Cleridy E. Lennert-Cody	271
環境科学における方向統計学の利用 [総合報告]	
清水 邦夫、王 敏真	289
人工林長期継続調査データを利用した林分成長モデルのパラメータ推定 [研究ノート]	
光田 靖、細田 和男、家原 敏郎	307

(メディア開発室)

Annals of the Institute of Statistical Mathematics

Volume 66, Number 1 (February 2014)

Debdeep Pati and David B. Dunson

Bayesian nonparametric regression with varying residual density	1
---	---

M. Rauf Ahmad

A U -statistic approach for a high-dimensional two-sample mean testing problem under non-normality and Behrens-Fisher setting	33
---	----

Kengo Kamatani

Local consistency of Markov chain Monte Carlo methods	63
---	----

Peter J. Brockwell, Jens-Peter Kreiss and Tobias Niebuhr

Bootstrapping continuous-time autoregressive processes	75
--	----

Masaaki Fujii

Momentum-space approach to asymptotic expansion for stochastic filtering	93
--	----

Sanying Feng and Liugen Xue

Bias-corrected statistical inference for partially linear varying coefficient errors-in-variables models with restricted condition	121
--	-----

Vyacheslav A. Vasiliev

A truncated estimation method with guaranteed accuracy	141
--	-----

Weihua Zhao, Riquan Zhang, Jicai Liu and Yazhao Lv

Robust and efficient variable selection for semiparametric partially linear varying coefficient model based on modal regression	165
---	-----

Jan Beran, Dieter Schell and Milan Stehlík

The harmonic moment tail index estimator: Asymptotic distribution and robustness	193
--	-----

(メディア開発室)



統計解析システムSとR私記

中野 純司

モデリング研究系

統計解析システムRは統計コミュニティの標準システムとなっており統計科学に多大な貢献をしています。多くの分野のデータ解析に利用されていますし、ビッグデータ解析のツールとしても注目されており商用のいわゆる“Analytics”ソフトウェアの中にもRを利用しているものがあります。Rはもともとアメリカのベル研究所で開発された統計解析システムSの互換ソフトウェアとして作成されました。したがってRの成功は(部分的には)Sの成功と言っても良いでしょう。S自身も統計コミュニティの中ではある程度の成功を収めました。歴史的にはRのオリジナルとして記憶されそうです。

わたしが最初にSに出会ったのは、1987年9月に東京でISI総会が開かれたときに慶應義塾大学の渋谷政昭、柴田里程氏がSの作者であるJohn Chambers、Richard Becker氏を招聘し、統計数理研究所で講演会を開いたときでした。当時私は徳島大学に勤務しておりSが稼働するワークステーションを手に入れる事は不可能だと思っていたのですが、たまたま研究室でApolloワークステーションを買っても良いということになりました。Sは学術関係者にはソースコードで配布されていたので、ちょうどUNIXのソースコードをほしかった総合科学部の数学の先生たちと一緒に、非常に分厚い英語のATTとの契約書を全部日本語に翻訳して事務と交渉し、手に入れることができました。ただ、OSやコンパイラの知識不足でなかなか稼働させられずに困っていたところ、ちょうど同じシステムを所有していた山口大学の刈谷丈治氏と知り合うことができ、そのおかげでApolloワークステーション上でSを動かすことができました。ヒストグラムが出たときの感動は忘れることができません。それを利用していくつかのデータ解析を行いました。しばらくして大学を移動したため、その後数年はSを利用することができませんでした。ただ、Sの商用化権を取った会社がS-PLUSという形でWindows上でも動くようにしたので、それを買って利用するようになりました。その頃、Rというものがあるということは聞いていましたが、手は出ませんでした。

2002年にニュージーランドのオークランド大学統計学科を訪問する機会がありました。その時にRの作成者のひとりであるRoss Ihaka氏と彼の学生でRのグラフィックスライブラリGridの作者であるPaul Murrell氏とゆっくり話をする機会がありました。当時はまだRもそれほど有名ではなく、Rに対する大学の評価も高くなかったようで、かれらも苦労している

様子でした。それで応援も兼ねて、日本でもRをより有名にしようとして2003年12月に二人を日本に招き研究会を行いました。ちょうどその頃、中間栄治氏がRの中で日本語を利用できるようにしたためもあり、筑波大学の岡田昌史氏が管理していたRjpWikiの上で疫学や生態学関係の方を中心にRが盛り上がっていました。わたしは自分たちで統計ソフトウェアを作っていたこともあり、Rに関しては好意的ではありましたが、あまりコミットする気はありませんでした。ただその後、いくつかの研究集会などでRのコアグループの人々と知り合う機会が多くあり、かれらの親切さ、寛大さ、情熱などに感銘を受けました。そこでコアグループやその周辺の人を毎年一人または二人日本に招き、最新の情報を教えてもらうとともにかれらに日本の情報を知ってもらうということを計画しました。2005年から始めたその研究集会はほぼ毎年続いており11、12月頃に2日ないし3日かけて開いています。そのうち1日は日本のRユーザーに発表してもらっています。これまでにほとんどのRコアグループのメンバーに参加してもらいました。例えば今年度はWindowsバージョン作成の中心となっているカナダのDuncan Murdoch氏に来てもらい、かれの作成しているパッケージでの日本語利用などについて議論しました。

Rは優れた開発環境を提供することにより世界中の統計研究者を巻き込んだ巨大なシステムになっています。典型的なボトムアップ体制のため今後どのようになるかはコアグループでさえわからないようです。わたしもボトムから少しでもSやRを生んだ計算機統計学に貢献していきたいと考えています。



Rの最初の正式版CD (コアメンバーのサイン入り)