

統計数理研究所

文部科学省 大学共同利用機関

概要 2003

The Institute of Statistical Mathematics

目次

はじめに	1
統計数理の多様な世界	予測と知識発見 2, 3 不確実性とリスクの管理 4 データ設計と調査 5, 6 計算推論 7 統計基礎数理 8 統計リソース（ハードウェア・ソフトウェア・コンテンツ） 9 開発した主なプログラム 10
研究紹介	予測の視点からの柔軟なモデリング 11 統計学と機械学習の融合とゲノム科学への応用 12 統計的手法による音声・話者認識 13 複雑な図形の空間配置を測る 14 哺乳類の進化 15 試験問題プールの設計 16
研究系・附属施設	研究系 統計基礎研究系 17 調査実験解析研究系 17 予測制御研究系 18 領域統計研究系 18 附属施設 統計計算開発センター 19 統計科学情報センター 19
国際協力	交流協定締結研究機関 20 国際シンポジウム（平成14年度） 20 国際共同研究（平成14年度） 20 外国人研究員（平成14年度） 20
共同利用	共同利用 21 共同利用研究課題一覧（平成15年度） 22, 23
大学院教育	大学院組織 24 講座の概要 24 教育研究の特色 25 修了要件および学位の種類 25 入学定員および在学生数 25 入学者の出身大学 25 学位授与数 25 修了生等の現在 25
研究成果の普及	公開講座 26 統計数理セミナー 27 公開講演会 27 統計相談 27 平成14年度文部科学省大学等地域開放特別事業 統計数理研究所「数理のめがね」 27
決算	歳出決算 28 外部資金受入状況 28 科学研究費補助金 28
建物・設備	敷地・建物 29 計算資源 29 図書・資料 29
組織	機構図 30 定員 30 所員 31 評議員 32 運営協議員 32 共同利用委員 32 名誉所員・名誉教授 32
沿革	33

表紙デザイン解説

緑の多い季節になり、本研究所の隣りにある有栖宮記念公園を歩いてみると緑色のドームの中にいるみたいです。ふと見上げると葉を透過してやわらかな光が降り注ぎます。萌える緑を本研究所と重ねあわせてイメージし、デザインしました。（蛸名優子）



来年4月には国立大学および大学共同利用機関は法人化されることになりました。この機会に大学共同利用機関は4つの研究機構に再編され、統計数理研究所は情報・システム研究機構に設置される研究所として新たな活動を始める予定です。典型的な横断型基幹科学技術としての統計数理を担う当研究所は新たな発展の契機を得たことになります。

このような外的要因の変化とともに、統計数理の世界は今や大きな変革期にさしかかっています。環境、生命、経済、災害などの様々な社会的問題は複雑かつダイナミックな現象のモデリングの重要性を明らかにしています。また情報化社会の進展は統計数理を取り巻くデータ環境を一変させ、大量データからの知識発見という重要な問題を提起しています。統計数理の役割が今後ますます重要になることは間違いないでしょう。

このような状況を考慮して、統計数理研究所では予測と知識発見、不確実性とリスクの管理、データ設計と調査、計算推論、基礎数理、統計リソースという6つのプロジェクト研究を設定して1年間活動をしてきました。本概要ではこのプロジェクト研究の紹介を中心に研究所の活動をご理解いただくことを目的に編集いたしました。

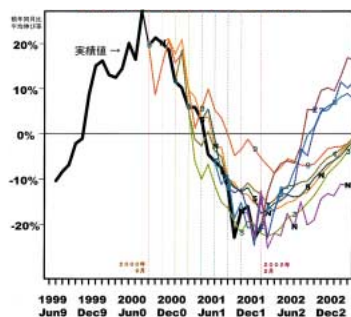
統計数理研究所は、統計数理的世界的中核としての活動をととして、今後とも社会に貢献する活動を推進して行きたいと考えています。関係各位のご理解とご支援をお願い申し上げます。

統計数理研究所長
北川 源四郎

統計数理の多様な世界

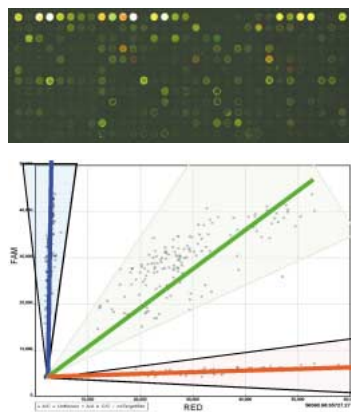
予測と知識発見

不規則で複雑な変動を示す現象の予測と制御は様々な分野で重要な課題となっています。本研究所が開発した統計の理論とプログラムパッケージは金融経済時系列データの解析、天文データの画像処理、地震データの処理、船舶の自動制御、車両の振動制御などに適用されています。近年は、生命、地球および宇宙科学において生じる大規模データから有意義な知識と情報とを発見する研究に取り組んでいます。



多次元ARモデルを通して探る日本経済

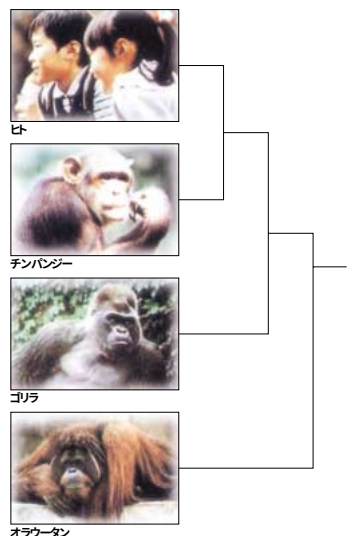
この図はある11個の経済変数に多次元ARモデルを当てはめ、長期予測を行った結果の中から、特に機械受注（前年同月比ベース）を描いたものです。これを見ると2000年9月以降の月について、足下までのデータに基づいたモデルから長期予測値をはじき出すと、その回復軌道（を予測したパス）が、月を追うに従って実績値に近づいていることがわかりますが、これはまさに、この時期の日本経済の悪化と回復を早くから捉えているといえましょう。



ゲノムデータの解析

ヒトゲノムは凡そ30億塩基からなっていますが、一人一人比較すると異なる部分が多く存在します。これを遺伝子多型と呼び、単一の塩基の違いによる多型をSNP（Single-Nucleotide Polymorphism）といいます。ヒトゲノム全体では約300万程度のSNPがあるといわれています。疾患や、薬の副作用発現性、有効性などに対して関連性が高いSNPの発見のために活発に研究されています。左の図はSNPの発見データとSNPタイピングのために開発された統計方法による一結果です。

提供：
財団法人癌研究会ゲノム情報センター



DNAによる進化過程の推移

生物進化の過程で起こるDNAの変化は確率的な現象です。従って、いろいろな生物のDNAを統計的に比較解析することにより、進化の系統樹を推定することができます。

形態の比較からは、チンパンジーとゴリラとが近い親戚だと考えられていましたが、DNAの解析からヒトとチンパンジーとが最も近い親戚であることがわかってきました。

情報量規準による画像処理

複数のパラボラアンテナから構成される電波干渉計のデータを計算処理して画像を構成し、COガスから放射される電波のドップラー効果を解析することによって、例えば、遠い宇宙空間における回転するガスの存在が分かります。原始惑星系円盤が見えているのです。

このような観測をさらに波長の短い領域にまで広げようとすると、パラボラ上空の電波の経路上の水蒸気の変動に起因するデータの乱れを補正する必要があります。そのための「大気ゆらぎモデル」も開発されています。



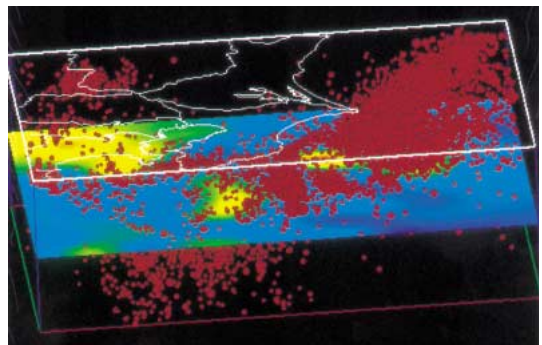
写真提供：国立天文台

地震の統計的解析

マグニチュードMの地震発生頻度は地震が小さくなるにつれて指数的に増加します。実際 10^{-bM} に比例しますが、この係数bの値は場所によって異なります。

写真には深さ100kmまでの関東直下の微小地震に対して、赤池ベイズ情報量規準ABICに基づき大規模ベイズ平滑化で求めたb値の三次元分布が示されています。

係数bが大きくなるにつれて寒色から暖色に変わっていきます。



データ提供：独立行政法人 防災科学技術研究所

船舶の自動操舵

東京商船大学との共同研究で開発された、統計的な予測・制御の方法によって設計されたオート・パイロットは実船の制御に生かされています。現在は更に新しい研究が進められています。



写真提供：東京商船大学

高速車両の振動の制御

大型車両（高速道路上のバス）のサスペンションとタイヤ部に関するバネ・マス・ダンパ系の統計モデルを作成し、双線形制御理論を用いて振動の抑制を行い、優れた振動抑制効果があることを実証しています。



写真提供：日産および日産ディーゼル

不確実性とリスクの管理

現代社会は、環境変動、技術進歩、経済システムの変容、社会の変動などの挑戦を受けて、不確実な未来に対して合理的に対処し、リスクを科学的に評価、管理することが求められています。本研究所では、環境の合理的なモニタリング、環境データの解析、金融のリスクの研究に取り組んでいます。



写真提供：
豊橋技術科学大学エコロジー工学系・藤江研究室

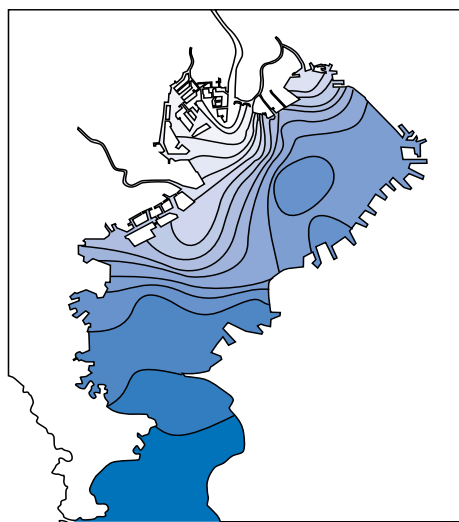
環境リスク管理

サンプリング点の決定、サンプリング時期の決定、解析方法の開発に関して統計科学が環境科学に寄与しています。

環境データの分析

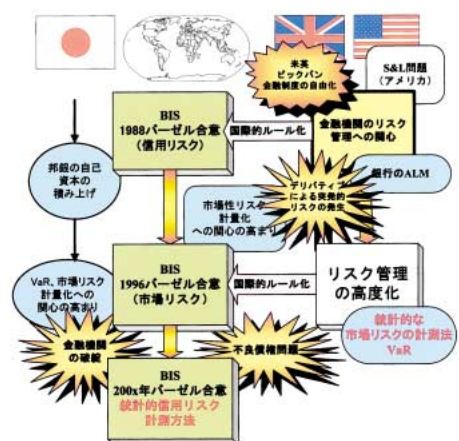
東京湾の水質の特性と長期変動を明らかにするため、時空間季節変動調整法による分析を行っています。図は東京湾内の塩分濃度を表します。分析の結果、東京湾内の恒常的な水の流れが明らかになりました。

その他にも、ダイオキシン類、浮遊粒子状物質、ヒートアイランド、地球温暖化等の問題にも取り組んでいます。



金融におけるリスク管理

市場のリスクは、市場価格の変動に起因する変動リスクと、売却に何らかの障害あることによって起きる流動性リスクに分類されます。変動リスクに関しては時系列モデルや確率過程などを用いて、市場リスクの統計学的測定モデルが実用化されています。金利変動リスクのモデル化やデリバティブのモデル化に取り組んでいます。



金融リスク計量化手法に関する社会的背景

データ設計と調査

科学的な調査は、データの獲得の段階において正確な設計が必要です。本研究所は1953年（昭和28年）から「日本人の国民性」の調査に取り組み、我が国の社会調査技術をリードしてきました。今日ではその対象を、インターネット調査、鯨類資源の調査、森林資源の調査、環境汚染の調査、さらには文化の計量にまで広げて研究に取り組んでいます。

日本人の国民性の分析

日本人の国民性調査は、5年ごとに行っている調査で、平成10年に第10次になりました。同じ質問を継続調査して、人々の意見や考え方にも時とともに変わる面と、時代の影響をほとんど受けない面があることなど、日本人の“心”の動きを調査結果から追跡しています。

これからも時代の変化に適応した社会調査法と解析法の開発を進めるとともに、国際比較調査を進展させることを目指しています。



国民性の国際比較

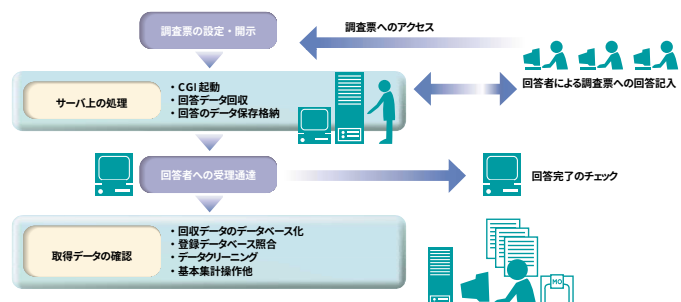
異なる国で得られた調査データは、標本抽出法の相違や翻訳などの問題があり、単純に比較はできませんが、私たちは、調査データの国際比較可能性を追究して、「連鎖的比較法」を開発してきました。これは、言語や民族の比較的近い国々を比較の輪で順につないでいき、個々の国を全体の中に位置づける方法です。

質問文の翻訳は、単純な逐語訳ではなく、バック・トランスレイションの繰り返しによって、慎重に各国の言語に変換していきます。この過程で、既に得られている調査の結果が次々と有効になっていきます。



新しい調査法

伝統的な調査法とインターネット調査との比較を実証的にを行っています。





鯨の調査と資源管理

鯨類資源の保護と管理を目指してフィードバック型の管理方式が開発され、現在この方式の運用に向けて各資源への適用条件のための試験が続けられています。また、さらに情報を得るために、毎年南極海においてモニタリング調査が行われています。

現在では、年齢や成熟状態などの個体の生物学的情報の収集解析に加えて、遺伝形質を用いて南極海に回遊してくる繁殖集団を同定する研究や汚染物質の蓄積状況などの環境情報についても収集・解析がすすめられています。



森林資源の調査と管理

個々の樹木は種の特徴や周りの環境などによって成長過程が大きく異なることを確率的な現象としてとらえ、統計的に解析することによって森林の動的変化を把握し、望ましい森林管理を実現できるような調査・分析が行われています。



写真提供：勉成社

文化の計量

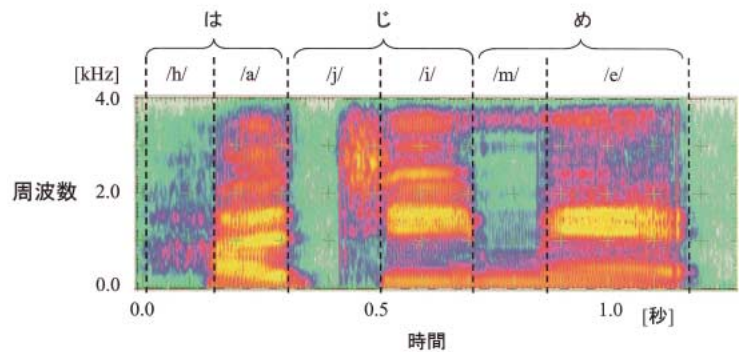
古文書には執筆時期、成立過程などに疑問が持たれている作品が数多くあります。このような問題を解決するために、文の長さ、語彙量、単語・品詞の出現頻度などの情報をデータベース化して、統計的方法によって作品の著者や時代などの推定の研究が進められています。その一つ、「源氏物語」の研究が、言語学や国文学の専門家とで大規模な共同研究が展開されています。

計算推論

統計モデルとコンピュータの発達によって、直接見たり観測したりすることが出来ない現象から得られる間接的データから、直接見えない情報を計算によって取り出す事が出来るようになりました。本研究所では、マルコフチェインモンテカルロ法、モンテカルロフィルター、逆問題の解法、最適化アルゴリズム、遺伝的アルゴリズム、線形計算アルゴリズムなどの研究に取り組んでいます。

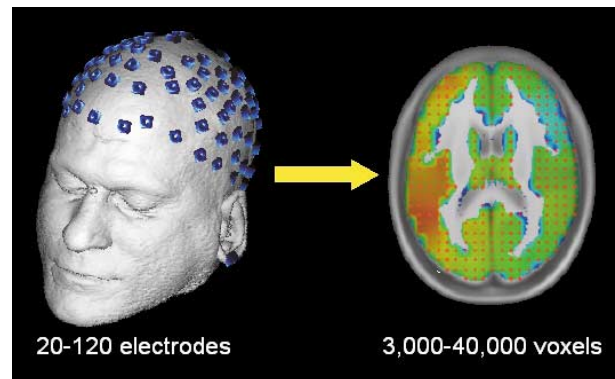
音声認識

ある女性が「はじめ」と発生した音声を周波数分析した様子：黄・赤・桃・青の順にその周波数成分が多く含まれています。各音によって周波数成分の含まれ方に特徴があります。音声認識ではこの特徴を利用します。



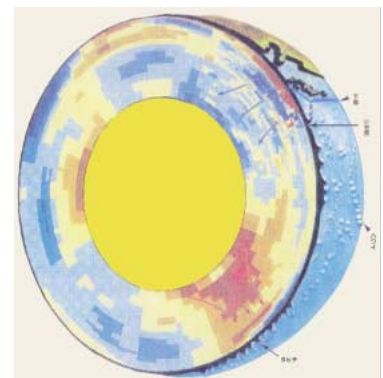
脳波のダイナミック逆問題

脳の外科的開頭手術をすることなしに脳波データを解析することによって脳の機能異常（アルツハイマー病など）の早期発見と発病のリスク軽減を目指す研究が進められています。頭皮上で計測される数十チャンネルの脳波データを用いて脳内の電氣的信号の発生源とそのダイナミクスを推定する方法が超大規模状態空間モデルとそのカルマンフィルターの平行計算アルゴリズムの開発によって実用化されつつあります。



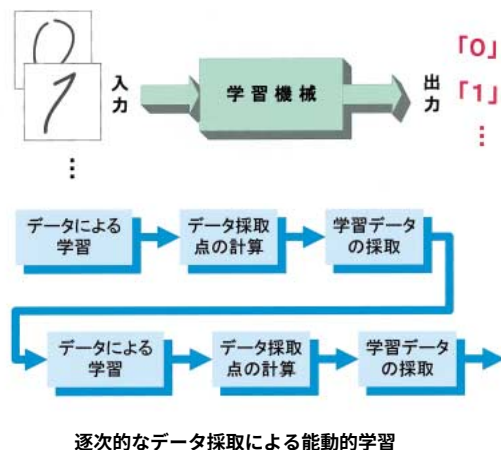
地球内部構造の推測

脳内の構造を知るためのX線断層写真と同様の方法を用いて、地球の内部構造を推測するための地震波による地震トモグラフィに取り組みました。統計モデルや最適化法が役立ちます。



統計基礎数理

統計学の基礎は、多変量の分布のモデリングとその同時分布の確率の計算です。本研究所では解析的な計算にはなじまない複雑な分布に関わる確率の計算法に取り組んでいます。さらに近年では、構造が十分には把握できない現象を、大量のデータを活用して自己組織的あるいは学習によって予測を可能にする方法の研究が活発に行われています。



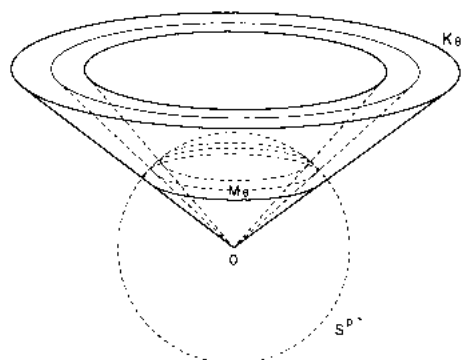
機械学習

サポートベクターマシン、ニューラルネットワーク、罰金付きロジスティックリグレーションマシンなど、データからの学習機能をもつ予測推論機械が研究されています。とくに汎化（帰納）能力の高いマシンの研究が活発に行われています。

$$\begin{aligned}
 & t^6 p^8 (1 - 3 p^{23} t^6 + 3 p^{24} t^6 - 3 p^{21} t^6 - 6 p^{20} t^7 - p^{11} t^3 - p^{10} t^4 \\
 & + 3 p^{27} t^{11} - 4 p^{26} t^7 - p^{24} t^7 + 5 p^{27} t^7 + p^{14} t^4 - p^{13} t^2 + p^5 t^4 \\
 & + 10 p^{32} t^8 - 5 p^{11} t^8 + p^{20} t^8 - p^{23} t^5 - 2 p^{14} t^3 - p^{10} t^4 - 3 p^4 \\
 & + p^{17} t^4 - 2 p^{15} t^4 - 3 p^{21} t^5 + 2 p^{22} t^3 + p^{19} t^4 + 2 p^{20} t^3 - 10 \\
 & + 20 p^{20} t^{10} + 6 p^{37} t^{10} - p^{41} t^{11} + 7 p^{42} t^{11} - 21 p^{43} t^{11} - p^{14} t^4 \\
 & - p^{36} t^{10} - 15 p^{40} t^{10} - p^{26} t^9 + 35 p^{44} t^{11} - 10 p^{38} t^9 - 10 p^{39} \\
 & + 3 p^{29} t^7 - p^8 t^4 - 4 p^{28} t^7 - p^{24} t^8 + p^{19} t^3 - p^{20} t^3 - p^{38} t^8 \\
 & + 4 p^{36} t^8 - 2 p^{20} t^6 - p^{36} t^6 + 3 p^{31} t^6 + 21 p^{46} t^{11} - 2 p^{27} t^6 + \\
 & - p^{42} t^{10} - 3 p^{32} t^7 - 5 p^{34} t^8 + p^{33} t^6 + p^{48} t^{11} - 7 p^{47} t^{11} - p^2 \\
 & + 6 p^{41} t^{10} - 5 p^{60} t^9 + p^{41} t^9 - 3 p^{32} t^6) / (1 - p^{21} t^7 - t - p^{23} t^6)
 \end{aligned}$$

パタンのモデリングと確率の計算

DNAのパタンにある種の構造がある場合、特定のパタンが生じる確率を計算するには実行不可能な程に膨大な計算量を要します。空間的な構造のモデリングを母関数の間の関係に表現することによって、確率の計算が実用的に可能となることが示されました。



幾何学による確率の計算

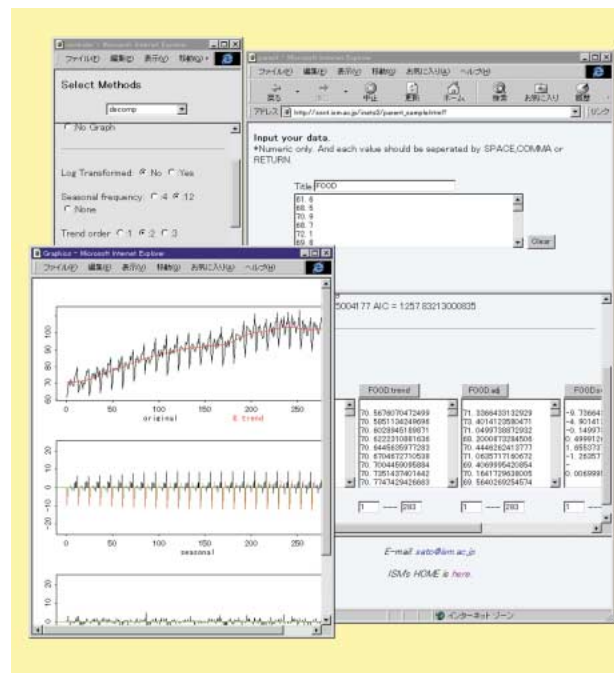
画像信号処理、地震統計学、計量経済などで求められる変化点の検出のための確率の計算は一般に困難な問題です。近年この問題に対処するためにチューブ法と呼ばれる方法に大きな進展がありました。射影追跡法、多次元配列データ、順序の制約があるデータの確率計算の方法としてチューブ法に着目して研究に取り組んでいます。

統計リソース

研究成果を、一般のユーザーが使いやすいようにソフトウェアの形に仕上げることに努めています。これまで開発したものの中で最も良く知られたものはTIMSACです。TIMSACシリーズの一部のプログラムのソースコードをFORTRAN77に完全準拠するように改編作業は終了しました。Windows上のQuick_Winアプリケーション用、Linux用を配布しています。WebCATDAPやCATDAP for Windowsも配布しています。その他のソフトウェアについてもユーザーインターフェースの改良などを行い、より使いやすくする計画があります。また、最新の統計科学の理論的成果を実用化するための新しいプログラムの開発も行っています。プログラム提供については統計計算開発センター(e-mail:cdsc@ism.ac.jp, FAX:03-5421-8796)にお問い合わせ下さい。

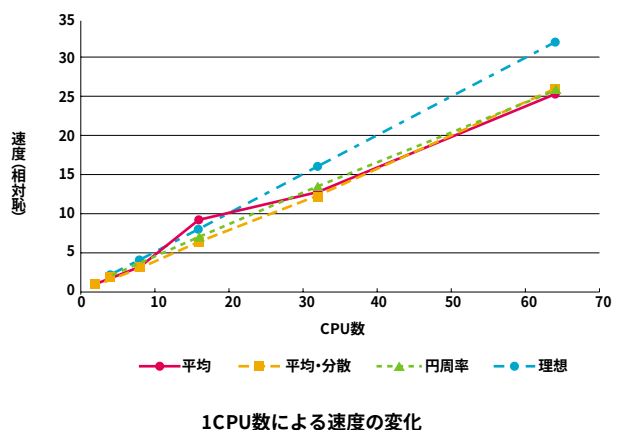
Webアプリケーション

予測制御研究系の佐藤整尚助教授が公開しているWebDECOMP使用時の画面のハードコピーです。季節調整や時系列解析をWWW上で行うことができます。Excelのアドイン化されたDecompも開発されています。



ブートストラップシステム

構成は、100台のパソコンを100BASE-TXによって接続したものに、各CPUに物理乱数発生ボードを装着したものです。OSはLinuxで、並列化のためにMPIを用いています。FORTRANやC言語から呼び出して使えます。



開発した主なプログラム

研究論文だけでなく、プログラムの形でも研究成果を公開しています。

プログラム名	利用分野＜事例＞	提供先機関名	
TIMSAC 時系列データの解析・予測・制御のための総合的プログラムパッケージ	・脳波分析 ・経済変動の分析 ・工業プロセスの最適制御 ・船舶のオートパイロットへの適用 ・地震データの解析	京都大学 東京大学 大分医科大学 九州大学 米国商務省	高エネルギー加速器研究機構 社団法人漁業情報サービスセンター 東京電力福島原子力発電所 サッポロビール株式会社 東京都老人医療センター 等
BAYSEA 季節変動・週変動・日変動等の周期的変動を含むデータを解析するためのプログラム	・経済時系列データの季節調整	東京大学 筑波大学 横浜市立大学 日本銀行	通商産業省 社団法人中央調査社 経済企画庁 米国センサス局 等
CATDAP カテゴリカルな目的変数に対する最適な説明変数を自動的に選択するためのプログラム	・多次元クロス表の分析 ・データマイニング	京都大学 日本女子大学 名古屋大学 東京女子大学	農林水産省 国立療養所南福岡病院 花王株式会社東京研究所 読売新聞社 等
NOLLS1 非線形最小二乗法のプログラム（関数群の二乗和を最小にするパラメータの値を数値的に求める）	・原子炉材料解析 ・プラント機器設計 ・新薬の薬動力学解析 ・呼吸器系の音波による内部解析 ・X線分光学におけるスペクトル解析	千葉大学 京都大学 名古屋大学 電力中央研究所 独協大学	日本IBM株式会社 東京大学海洋研究所 東京都環境科学研究所 東北大学電気通信研究所 UCLA 等
QUANT 数量化理論のプログラム質的データの多変量解析予測・判別・分類・要因分析	・青少年の行動調査分析 ・臨床医学データの分析 ・選挙予測 ・広告効果分析 ・教育心理等のデータ解析	東京大学 東京工業大学 筑波大学 兵庫教育大学 建設省	社団法人新情報センター 環境数理研究所 電通 朝日新聞社 読売新聞社 等
DALL 最尤法によるモデルあてはめのためのDavinson法による対数尤度最大化のプログラム	・医学データ解析 ・非定常多次元時系列データ解析 ・最尤法が必要な全分野	国立天文台 大分医科大学 米国国立電波天文台	等
ARdock TIMSACによるシステム解析を対話的に行えるようにしたプログラム	・プラント解析 ・システム解析 ・生体情報解析	大分医科大学 明治大学	等
TIMSAC for Windows TIMSAC72の一変量ARモデル、多変量ARモデルをMS-WINDOWS上で動作するようにしたプログラム	・脳波分析 ・生体活動の分析 ・商品売上予測 ・株価予測 ・地震データの解析	富士総合研究所 三菱総合研究所 明治生命 住宅金融公庫 住友生命 東京学芸大学	安田信託銀行 日本開発銀行 日経データ 和光経済研究所 一橋大学 九州大学 等
CATDAP for Windows カテゴリカルな目的変数に対する最適な説明変数を自動的に選択するためのプログラムのWindowsバージョン	・多次元クロス表の分析	京都大学	慶應義塾大学 等
DLL of TIMSAC TIMSACをDLL化したもの Linuxのshared libraryもある	・時系列解析	京都大学	東京大学 等

予測の視点からの柔軟なモデリング

— 金利期間構造の推定を例に —

1. モデリングと予測の視点

予測という営為を、「現在までの情報に基づいて将来に対して合理的な判断を行うこと」捉えれば、統計科学における予測は、まさに人間の知的活動の一般化に相当する。統計的モデリングとは、現在までの情報がデータによって与えられたとき、そのデータが生成される構造を確率モデルで表現することである。モデルとは所詮虚構であり、我々は通常複数のシナリオの中から尤もらしいものはどれかを知りたいと思う。このとき、「同じ確率構造から再びデータが生成されたときに、その新しいデータに対する予測性の優れたモデルを良いモデルとする」、というのが統計科学における予測の視点である。

2. 柔軟なモデリングとその評価

統計数理研究所では、赤池弘次元所長による情報量規準の提案以来、様々な分野で予測の視点からのモデリングを推進している。赤池情報量規準（AIC）は、理論的には最尤法と呼ばれる手法で推定された統計モデルの比較を念頭に導入されたものであったが、その後計算機の飛躍的な能力向上と相俟ってモデルも複雑化を遂げ、モデル推定が必ずしも最尤法によらない時代が到来した。そのひとつが、罰則付き最尤法を適用する場合である。典型的には次節の例に述べる非線形回帰のように、データを表現する仕組みが柔軟性に富む場合に、過度の柔軟性に対する罰則項を導入することで、データの精度に見合った複雑さのモデルを決定するという解決法である。このとき、罰則の

重さをどう合理的に決めるかが問題となる。このような柔軟なモデリングの評価に役立つのが、一般化情報量規準GICである。

3. 金利期間構造の推定

金利の期間構造とは、例えば一定の金額を現在から一定期間預けたらどのぐらいの利回りになるかを示したものである。一般には国債利回りが重要であり、日々取引される国債の価格 p から利回り曲線（イールドカーブ）が導かれる。その前提となるのが次のモデルである。

$$p = c\delta(t_1) + \dots + c\delta(t_{L-1}) + (c+R)\delta(t_L) + \epsilon$$

利付き国債は、償還期限まで半年ごとに c だけの利払いがあり、償還時（ L 回目）には元金も含めて $c+R$ だけの支払いが約束されている。これらは全て将来発生するキャッシュフローなので、その発生までの時間に応じて割り引いて考えられる。こうして割り引かれた償還金額の合計として、現在の債券価格が決まっていると考えるのが上のモデルである。各時点に於いて割引率を定義する関数 $\delta(t)$ がわかれば、利回り曲線はその逆算できる。通例 $\delta(t)$ は区分的に滑らかな非線形関数の集まりで近似されるが、関数型の柔軟さを制御することに失敗すると、解釈に困る利回り曲線が推定されることがしばしばある。特に、将来のある時点からほんの少し先までの金利を表すフォワードレートと呼ばれる指標を安定的に推定することは従来困難とされてきた。罰則付き最尤法で非線形回帰を行い、情報量規

準で罰則項の大きさを適切に決めたモデルの安定度を示しているのが図1である。図1は、リサンプリングで標本を複製する方法でフォワードレートの安定度を調べたものであるが、実験結果から信頼度の高いモデルが得られていることがわかる。これに対して図2は罰則項なしに推定した結果に基づく数値実験であるが、満期の長い区間で推定された曲線が不安定であることがわかる。

4. 予測システムの構築と知識発見

仮に「当たるか当たらないか」といった卑近な意味での予測が分析の第一目的でなくとも、統計的モデリングには予測の視点が不可欠である。また、予測の視点から継続的にモデルの改良を続ける過程にこそ、個別科学における発見の機会が内包されている。統計数理研究所では、研究プロジェクトの重要分野として「予測と発見」というカテゴリーを設けているが、ここに紹介した研究は、研究課題「予測システムの構築と知識発見」からの一例である。我々の研究グループでは、柔軟なモデリングとモダンな情報量規準の利用により、大規模なデータセットからの情報抽出（巨大データベースからの知識発見）を支援するシステムを構築するための研究を行っている。

このように学習と質問を繰り返して行っていく統計的方法の研究が進められており、実際に汎化能力が高い機械を作ることが可能なことが実験的に確かめられている。

（文責・予測制御研究系 川崎能典）

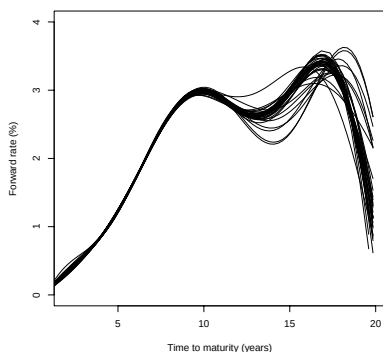


図1. 罰則付き最尤法と一般化情報量規準で推定されたフォワードレートに基づく数値実験

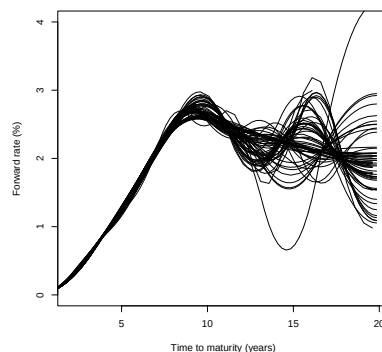


図2. 罰則なしの最尤法で推定されたフォワードレートに基づく数値実験

いずれも「統計数理」第50巻第2号掲載の川崎・安道論文から引用。

統計学と機械学習の融合とゲノム科学への応用

〔統計理論の課題〕

データを普遍的な立場から科学する方法は、統計学の基本課題である。特定の、例えば医学統計の分野で集中的に研究された方法論の一つ‘生存解析’は、医学の臨床データに関するばかりでなく、例えば、経済統計の金融データにも適用される。これはデータを普遍的な視点から研究するスタイルによって初めて可能になることで、ここに統計数理の多様な世界が実現される。このように統計数理は、その諸々の組み込まれた詳細を数理的に抽象化し、その視点から様々な分野に還元する方法論が展開される。

この基礎概念は、20世紀の初頭に、データのもつ‘情報’という視点から数理的な体系がほぼ完成されていたが、最近になってその有効性に問題が生じている。その原因はデータの質の多様化と量の巨大化に伴う急激な変質にある。つまり従来のデータに対する普遍的な立場の範囲を超えた形式が出現しているのである。例えば、データの形式がウェブページの集合だとすると、これはテキスト、画像、音楽、動画の複合体となり互いにリンクで結ばれている。また診察データが複雑な関数の形式で提供されることも日常になった。

万能な科学なんて存在しない！というわけだ。そのような背景から、‘データマイニング’は、統計学のアンチテーゼとして現れた。私たちの研究グループは、統計学の流れと異なる分野から出現したデータの科学的方法について研究を進める一方で、それによって得られた知見を融合して、統計学の基礎理論の再構築を目指している。図1で示されるよう左二つの球はかなり共有するものがすでにあ

る。データは一つなのだから、それぞれの球に住み分けることは良くないことだ。

〔機械学習〕

人工知能の分野で最近有力な進展を持っている‘機械学習’の基礎概念は、統計学においても刺激的である。彼らは、ある学習機械が、与えられた特定のデータを例題として捉え、その例題から学ぶことによって普遍的なデータの理解を得るプロセスに着目する。これをヒントとしてデータの科学のための新しい方法を提案する。ここで‘学習機械’とは特別な機械を指すのではなく、単純なルールによって記述されるプログラムの一つと考えても良いし、仮想的な‘脳’の機能の一つと考えても良い。

例えば、簡単な識別問題を考えよう。図2の一番左にあるように、黄色の線で与えられる正解の境界に、学習前の識別境界があったとする。これからブースティング法による学習アルゴリズムが薄緑の枠内で行われている。注目すべきは、それぞれのステップではかなり正解にずれた結果をだしているにもかかわらず最終的には重み付け多数決によってほぼ正解の結論に達している。ここで橙色の線の太さが重みを表す。間違っても、間違いいから学べばよいのだ！という考えは、従来の統計理論にはなかったことだ。

〔統計学と機械学習の融合の試み〕

データのもつ‘情報’という考えは統計理論の最初の段階にすでに確立されていたのだが、当時‘学習’という視点は生み出されなかった。我々人間が経験するすべての事柄は、自己の脳を通して認識される。これを学習と捉えることは、

データを普遍的な立場から科学することをも包含することになる。つまり我々の脳が認知することによって、はじめてデータは観測されるわけだ。このような有力な接近に対峙して、データの持つ‘情報’を中心に組み立てられた統計理論の再検討は重要である。またこの成果が人工知能の分野に刺激されたなら、すばらしいことだ。

現時点で注目している方法論は、教師なしデータに対するガウシアンミクスチュア、独立成分分析、グラフィカルモデル、教師ありデータに対するブースティング・マシンとサポートベクター・マシンなどが挙げられる。

〔ゲノム科学への応用〕

ゲノム科学の最近の目覚ましい進展によって、いわゆる‘ゲノムデータ’は膨大な規模で蓄積されている。現時点では一連のゲノムプロジェクトが一段落しつつあり、ポストゲノムとして新しい方向を求めてさらに大きく展開されている。データ普遍性の拡大として、このゲノムデータを対象にすることは大きな可能性があると考ええる。つまり、図1の右の赤球がこの研究プロジェクトのトレーニング場になっている。

このような見通しのもとで、昨年度より、財団法人癌研究会のゲノムセンターとの共同研究が開始されている。遺伝子多様性解析、特に塩基多型を探るSNPのデータや、マイクロアレイデータについて、上述のデータ解析の方法論の適用に向け、さまざまな観点から研究が進行中である。

(江口 真透)

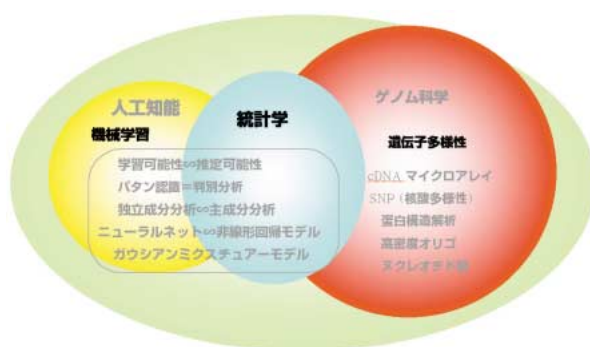


図1. 研究関連図

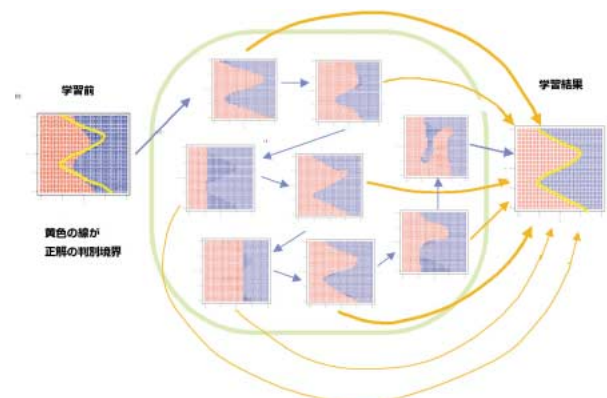


図2. ブースティングによる学習の過程

音声は、人間が用いるコミュニケーション手段の中で最も基本的なものである。近い将来、インターネットなどの情報ネットワークにパソコンや携帯電話に限らず、いろいろな電化機器からアクセスでき、場所や時間にあまり囚われない生活ができる社会を迎えようとしている。そのような社会では音声は手軽なアクセス手段として、またそのアクセスコントロールのための本人確認手段として有望である。アクセス手段として用いる場合、音声の内容を理解する必要があるが、その技術が音声認識である。また音声に含まれる個人性情報から本人確認を行う技術が話者認識である。ここでは、音声・話者認識技術の基礎と、そこでの統計的モデリングの役割について紹介する。

【音声・話者認識の基本的なしくみ】

音声は物理的には人間の発声器官で生成される空気の粗密波であり、文字や話者の目印や区切りが入っているわけではない。ある話者が「当たり前 /atarimae/」と発声した時の音声を細かく区切って、各時刻での波のパワーと、どの周波数成分がどのくらい含まれるかを調べた結果（色の濃淡で表示）を図1の上下部に示す。横軸は時間を、上部の縦軸はパワー、下部は周波数を表す。図1下部において濃い線、つまりその周波数成分が多く含まれている部分に注目すれば、例えば/a/音と/i/音ではっきり違っていることがわかる。また、/t/音は濃い線がはっきり出ずに不規則な模様となる。このように、周波数成分の情報には音を区別する特徴がよく表れる。また、その模様は話者によっても特徴がある。そのために、音声・話者認識では周波数成分の情報を一般的に利用する。

音声認識ではまず、音声を細かく（20～30ミリ秒）分けて周波数分析を行い、周波数成分の情報を10～40次元のベクト

ル時系列として抽出する。予め/a/や/i/などの短い音（母音や子音など）ごとに作成しておいたモデルを（意味のある単語の列を表すように）適当に並べ替えてみて、上記入力音声のベクトル時系列に最もよく適合するモデルの並びを探す。そして、その並びに対応する単語の列を認識結果とする。

話者認識では音声認識と同じように入力音声から周波数成分の情報を抽出した後、予め作成しておいた本人のモデルと、入力音声との類似度を調べ、その値に基づいて本当に本人かどうかを判定する。

【音声・話者認識の難しさと統計モデルの利用】

図1では/a/音が3回ほど出現するが、それぞれ時間的な長さが異なり、また前後の音に影響されて濃い線が傾きを持っている様子が見える。話者や発声内容が異なると、それらのばらつきの程度は大きくなる。音声のモデリングでは、各音（もしくは各話者）の周波数成分の特徴を要領よく表しつつ、時間的な伸縮や周波数成分のばらつきを吸収することが重要である。そこで統計的モデリングが行われる。

音声認識における各音のモデリングでは、時間的な伸縮が可能な隠れマルコフモデル（hidden Markov model; HMM）が一般的に用いられる。HMMの各状態は周波数成分のばらつきを吸収できるように混合ガウス分布などに対応させる。不特定話者を対象とする場合、数百から数千人の話者が発声した計数十時間分の音声データから、HMMのパラメータを学習する。

話者認識では特にキーワードを決めずに、どのような言葉からでも本人確認したい場合、各話者のモデリングでは、周波数成分の静的な特徴が表せる混合ガウス分

布モデルが一般的に用いられる。混合ガウス分布モデルのパラメータは、本人が発声した数十秒、数分の音声データから学習する。

【次のステップへ】

現在、統計的手法の利用によって、高い精度で音声・話者認識ができるようになった。しかし、実用上はまだ、いくつかの課題が残されている。例えば、周囲雑音や話者の違いによる学習とテストデータのミスマッチ、話し方（発声速度や声の大きさ）の違い、言い誤りや言い忘れ、通信状況の悪化などが挙げられる。特に話者認識では声の経年変化も問題となる。本研究ではそれらの課題を取り扱う、新たな統計的手法の利用について検討を進めている。

以下では音声認識において、話者の違いによる学習とテストデータのミスマッチを認識時にオンラインで軽減する、音声モデルの話者適応化法の研究例について簡単に紹介する。

【音声モデルの話者適応化】

オンラインの話者適応化では、まず入力音声自体を用いて各音のモデルをその話者へ適応化してから、その適応化後のモデルで認識を行う。その際、入力音声の発声内容は未知なので、音声中のどの区間をどの音のモデルの適応化に用いるか、その対応が不確かであることが問題となる。そのために我々が提案した方法では、発声内容の候補を正解が含まれるように複数選び、各候補について適応化を行う。次いで各適応化後のモデルと入力音声が最もよく適合するモデルを最終的に選択する。ここでは正解の発声内容に対しては、適応化が最適にできると考えるわけである。この方法により認識誤率が3割程度削減された。

（松井 知子）

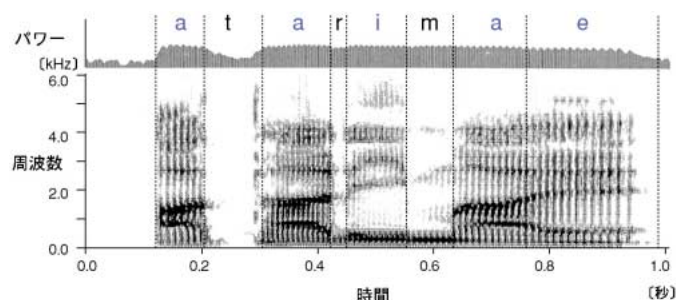


図1. ある話者が「当たり前 /atarimae/」と発声した時の音声パワー（上部）と周波数成分の情報（下部）

複雑な図形の空間配置を測る

— ヒト表皮神経線維分布の定量的診断への応用 —

〔問題の起こり〕

平面の有限領域に多数のいろいろな形状の図形たちが互いに重ならずランダムに散布されているとする。そのような空間配置データが与えられたとき、図形たちの配置の特徴をどのように表現するのか、また個々の図形の占有区域をどのように定めればよいのかなど、さまざまな問題が生まれてくる。散布される対象が点状のものであれば、これまで「空間統計学」において種々の統計的方法が提案され、応用されてきたのだが、複雑な図形の空間配置に対する空間統計学はまだ研究が進んでいなくて、将来性のある分野である。

〔ボロノイ分割を使う方法〕

点状のものが空間に散布されているとき、点配置の特徴をボロノイ分割の統計的性質によって捉えるという方法はきわめて有効であり、散布された母点の占有区域は各母点のボロノイ多角形によって客観的に定められる。ボロノイ多角形とは、散布された母点の一つに注目するとき、他の母点との垂直二等分線で区切られる（注目する母点を含む）半平面の交わりとして得られる凸多角形であり、平

面領域は各母点に割り当てられたボロノイ多角形で分割され、それをボロノイ分割とよぶ。

〔複雑図形の空間配置のボロノイ分割〕

そこで、上の考え方を形状が複雑な多数の図形が散布された場合に拡張できないだろうか。それが実は可能である。

一つの図形は線分や曲線だけから構成されるとする（図1に青色で示した曲線分たちの配置を参照）。そのとき、各図形に対して割り当てられた区域が図1に黒線で示されていて、それらが平面領域を分割している。

この図を作成した方法は、次の通りである。まず、各図形は近似的に多数の点からなると見なせるので、図形を小さな間隔でとった点の集まりで置き換える。次に、これらの点を母点として、上記のボロノイ分割を実行する。そして、一つの図形に所属する母点のボロノイ多角形を連結する。この操作の結果、図1のように一つの図形に一つの区域が割り当てられる。図形がある面領域である場合は、その周を多数の点で置き換えて、同様な操作をすればよい。このようにして各図形に割り当てられる区域を一般化ボロノ

イ領域とよぶ。

〔ヒト表皮神経線維分布への応用〕

われわれの皮膚は「痛み」、「痒み」を感じるが、それらは表皮にある神経線維で知覚される。アトピー性皮膚炎患者と健常者では、この神経線維の空間分布に違いがあると考えられるが、これまでそれらの違いを定量的に調べることは困難であった。図1と図2の太線はそれぞれ健常者とアトピー性皮膚炎患者の神経線維の顕微鏡観察から得られた分布である。これらに、一般化ボロノイ領域による分割を適用した結果が、それぞれの図に示されている。これによって、個々の神経線維に占有区域を客観的に割り当てることが可能になり、例えば占有区域の面積分布がアトピー性皮膚炎と健常皮膚と統計的な差が現れることが定量的に示された。

この研究は、今山修平氏（国立病院九州医療センター）および本多久夫氏（兵庫大学）との共同研究の一部である。

（種村 正美）

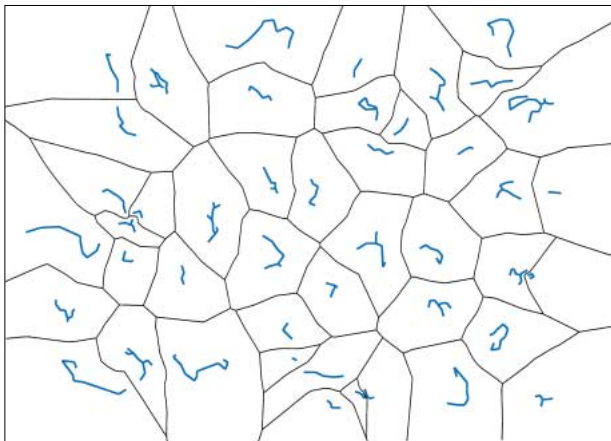


図1. 健常者表皮の神経線維分布の観測例。青の太線が神経線維で、黒の細線で示された区域が一般化ボロノイ領域。

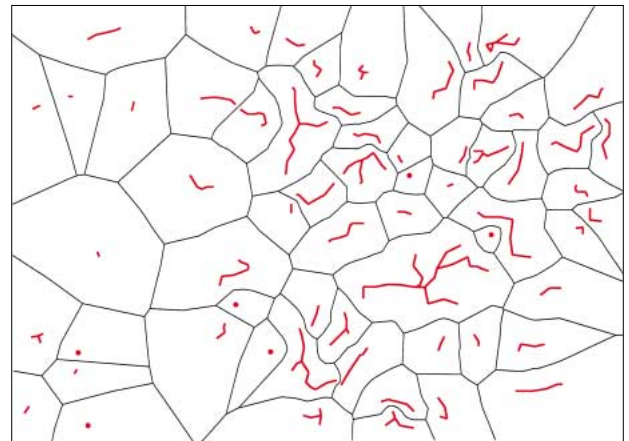


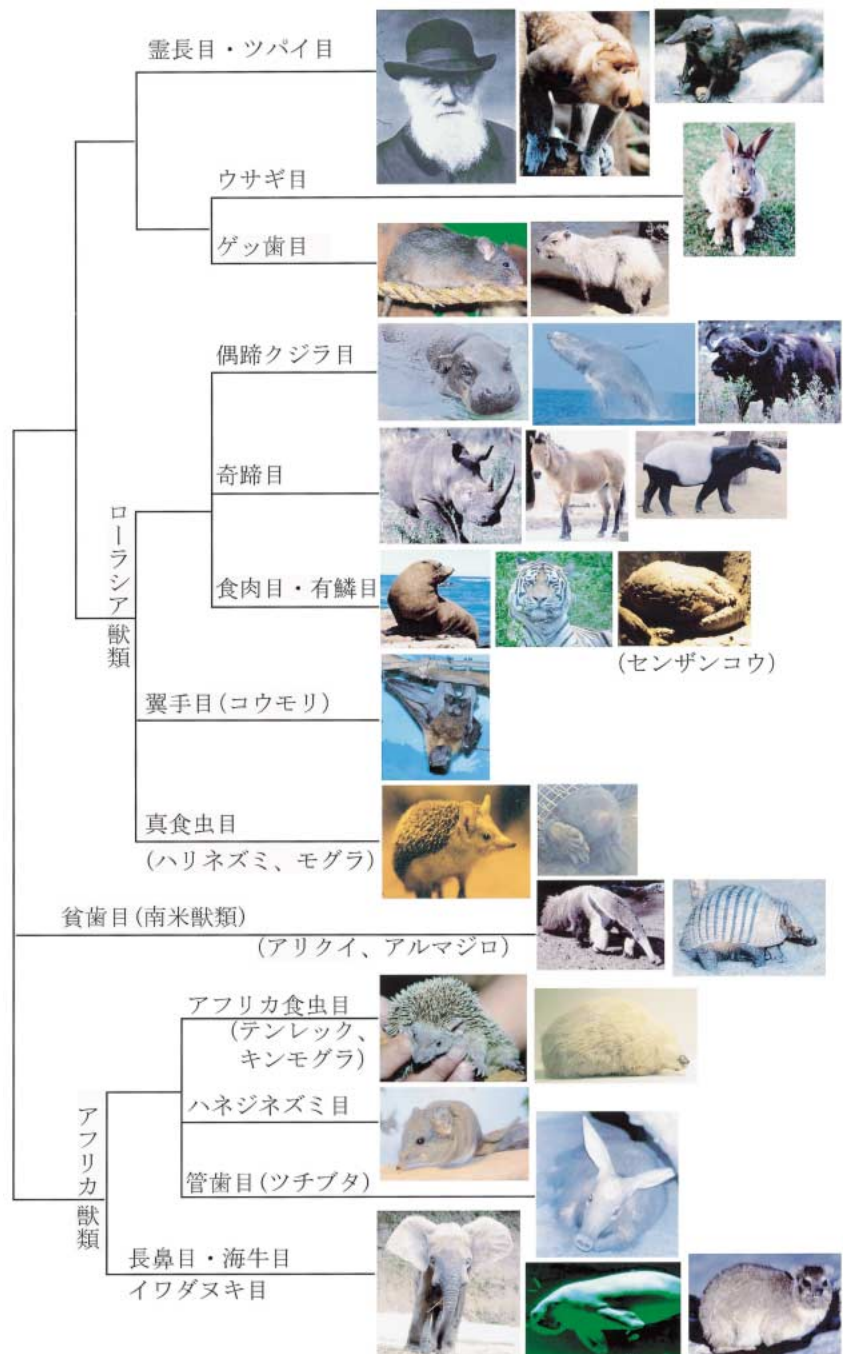
図2. アトピー性皮膚炎患者表皮の神経線維分布の観測例。赤の太線が神経線維で、黒の細線で示された区域が一般化ボロノイ領域。

哺乳類の進化

近年のゲノムプロジェクトなどの成果として、いろいろな生物の遺伝子配列データが急速な勢いで蓄積している。それぞれの生物のもつ遺伝子は長い進化の歴史の産物であり、これらのデータを解析することによって生物進化の系統樹を推定することができる。進化の過程で起こる遺伝子配列の変化には、ランダムな変動がつきものであり、データ解析には統計的な方法が必要となる。本研究では、生物学の実際問題の解決をはかりながら、そのような方法の開発を行ってきた。

生物学の実際問題としては、雌が胎盤をもつ哺乳類のグループである真獣類の進化を扱ってきたが、その成果として図のように真獣類進化の全体像が明らかになってきた。例えば真食虫目のハリネズミとアフリカ食虫目のテンレック（マダガスカル島にだけ生息する）とは、形態的に非常によく似ていて、これまで同じグループに分類されてきたものであるが、ハリネズミはモグラに近縁であるのに対して、実はテンレックはゾウなどアフリカ起源の動物に近縁であることが分かってきた。またアフリカとアジアに生息する有鱗目のセンザンコウと南米に生息する貧歯目のアルマジロとは非常によく似ているが、これもまた系統的には近くないことが明らかになってきた。つまりこれらの類似性は、似たような生活環境に独立に適応したためのものであり、収斂進化と呼ばれるものであり、遺伝子データ解析の結果、収斂進化がこれまで考えられていた以上に頻繁に起こっていることが明らかになってきたのである。

(長谷川 政美)



試験問題プールの設計

〔設計科学の流れ〕

最近の科学研究には、目的を明確にした上で、その目的に則した方法を設計して、効率的に研究する大きな流れがある。天才的な研究者のひらめきに頼るのではなく、多くの優れた研究者の英知を結集して作り上げていく研究である。研究の設計を明確にすることは、研究に伴う論理と仮定を明確にして、結果として信頼性の高い研究成果を得る。ここで合理的な設計とは統計学の論理と理論に即応した設計のことに他ならない。試験の問題を作製する場合も、明確な目的と入念な設計が必要である。

〔試験問題の作製〕

試験は、学習者が自らの学習の到達度を確認するために、また第三者が学習者が確かに習得したことを確認するために行われる。試験問題は学習の到達度を測定する道具であり、実験のための測定装置の製作に似ている。しかしその作製は見かけ以上に遥に困難な仕事であるが、これまでは試験委員にすべてを委嘱していた。だから、問題の質は専ら委員の熱意と能力に委ねられていた。その結果として、試験にはいわゆる傾向と対策がつきものであったし、いわゆる難問・奇問も決して珍しくはなかった。学習者の能力を測定することを目的に、入念に設計された手続きを経て作製されてはこなかったのである。

先ず求められる知識内容を明示的に決める必要がある。たとえば、現在準備が進められている医歯学生の臨床実習前の試験では、臨床実習に必要な教育内容（コアカリキュラム）が定められている。学校教育では指導要領が対応する。内容

が明示化されていないと試験委員によって解釈が異なることになる。更に、実際に出題する試験問題が目的とする内容に則していなければならない。そのためには、数多くの問題のプール（問題集）を作製しておいて、その中から少数の問題を無作為に選んで出題することが原則となる。以上をまとめて試験の枠組みを図示したのが図1である。予め多くの試験問題を作製しておくことは大変な労力である。だから様々な変法も考えられる。しかし何らかの無作為な選択は必要である。この手順を踏まないから従来の試験では予期しない傾向が生じてしまっていたのである。実際のでしかも偏りのない無作為選択についての研究は今後大きな進展が期待できる。

試験問題プールは、教育内容に則すると共に、良い問題が集められなければならない。ここで良い問題とは、学習した学生には解けるが、学習しない学生には解き難い（これを識別度が高いという）問題である。

〔試験問題の試験〕

普通には試験というと受験者を試すという一面のみに目が向けられる。しかし他面では試験問題の良さが試される試験でもある。単に試験問題を眺めていても良さを判定するのは困難で、試験を通じて結果を統計解析しないと分からない。出題者が考える問題の困難度と受験者が感じる困難度とはしばしば異なる。設問が高度な知識を問う場合に正答率が低いのは当然だが、設問の文章が拙くても正答率は低くなってしまふ。後者の場合には識別度が低くなる。図2は項目反応モデルを用いて問題の困難度、識別度

を推定した例である。図中に傾斜の緩い線があることが見て取れる。これらの問題は再考の余地がある。要するに、試験問題の良さは実証データにより吟味される。

ところで、近年同じ問題をすべての受験者が解く試験ではなくて、受験者ごとに異なる試験問題を解く方式が普及してきた。そうすると、いよいよ多くの数の試験問題が必要になる。実際、医歯学共用試験では、2年間に7千題が作製され更に蓄積が進んでいる。

〔高次元母数モデル〕

望ましい試験を構築するためには、良い試験問題プールが必要である。受験者の数も多いので、推定する母数の次元は極めて高くなる。今日では広範囲の分野で高次元母数モデルが用いられている。そこで開発された、ベイズ推論、修正尤度法などは近年理論統計学の進歩が最も著しい分野である。そこでの理論的展開における発想と必要な計算量は、従来の統計解析とは時代を画している。

高次元の母数が頻繁に現れるようになった理由は多くある。しかし最も大きな理由は、科学研究が組織的に設計されて遂行されるようになったことに求められる。試験問題プールを作製して望ましい試験を構築する試みは、焼き畑農業から脱却して定住型の農耕に移行した経緯に準えることができる。試験問題プールの作製と利用は統計理論により支えられてきた。逆に、現場からは高次元母数モデルにおける斬新な推論方法を開発するよう求められている。

（柳本 武美）

1) 求められる知識内容

2) 明示された試験の内容

3) 試験問題プール

4) 出題する試験問題

図1. 論理的に設計された試験の枠組み



図2. 項目反応理論による解析例
（東工大・前川教授提供）

研究系

統計基礎研究系

統計科学を科学として基礎づける独立した統計理論の研究と、技術としての統計的方法に明確な理論的根拠を与える基礎的研究を行っています。4つの研究部門は互いに有機的に連携し、不確実性のもとでの統計科学の基礎にかかわるさまざまな問題を発見し解決するための統計理論の構築とその適切な適用方法の研究に取り組んでいます。

目的	研究部門	研究内容
統計の基礎理論の研究	推測決定理論	データから事象の本質を統計的に推定するための推測理論、最適行為を決定するための決定理論の研究を行う。
	基礎概念	統計学を構成する基礎概念の研究を行う。
	応用確率論	不確実な現象のモデル化とその解析、確率過程論とその統計理論への応用の研究を行う。
	確率・分布理論（客員）	統計に関する確率理論並びに分布理論等の研究を行う。

調査実験解析研究系

個別諸科学における具体的な数理的問題を素材に、データの獲得や分析の方法を数理的に定式化し、実際の統計数理的方法を研究・開発することを課題としています。したがって、接点となる研究対象分野は自然・人文・社会科学のさまざまな領域にわたりますが、目的は、統計数理を用いた個別科学の研究ではなく、個別科学に想を得た統計数理の研究にあります。

目的	研究部門	研究内容
データ獲得法とデータ解析法の研究	標本調査	社会調査法（標本調査法を含む）と推測法の研究を行う。
	空間事象	形状に関する統計的解析など空間事象の研究を行う。
	多次元解析	計量データ、および質的データなどを含む非計量データの多次元解析の研究を行う。
	パターン解析	探索的方法論を含む分類解析、パターン認識の研究を行う。
	系列事象	突発現象や構造変化現象など時間および時空間データのモデル化を統計解析法の研究を行う。
	実験計画（客員）	実験の最適計画法、推定理論の研究を行う。

予測制御研究系

多数の要因が複雑に関連し合いながら時間とともに変動するダイナミックな現象の予測・制御・最適化およびそれを実現するための計算手法に関する研究を行っています。このため特に時系列模型、確率場模型、非数値的模型、大規模システムなどの統計モデルの設計・解析・運用およびそれに基づく統計的推論法の理論とその実用化に関する研究を行っています。また、これらの基礎となる計算技術と計算システムの開発をも併せて行っています。

目的	研究部門	研究内容
統計モデルによる予測、制御および最適化ならびにこれらに関する計算手法の研究	システム解析	相互に依存しながら時間的に変動する確率的対象を動的システムとして表現、同定、状態推定するための理論、方法及び応用の研究を行う。
	予測理論	統計的予測の理論と応用に関する研究を行う。
	制御理論	統計的予測の理論と応用に関する研究を行う。
	数値的最適化	複雑なシステムの最適化の方法である数値的最適化の研究を行う。
	統計計算システム	統計計算のための計算システム、アルゴリズム、シミュレーションの研究を行う。
	非数値的情報処理研究部門	データから必ずしも数値的でない「質」や「構造」などの形の情報を取り出すこと、及び非数値的データから情報を抽出する方法を研究する。めた技術としての情報処理に限らず自然現象としての情報処理の研究も行う。
	大規模システム（客員）	複雑かつダイナミックな大規模システムの推測に関する研究を行う。

領域統計研究系

自然科学、人文社会科学等の種々の領域における実証的分析を通じて、各領域に適した統計的方法に関する研究を行っています。たとえば生物、医学統計、結晶の対称性、意識の国際比較、文章の計量分析、考古学データの計量分析などの研究が、国内外の大学、研究機関等の研究者と共同ですすめられています。

目的	研究部門	研究内容
自然科学の領域、人文社会科学の領域ならびに両領域にまたがる分野の統計的研究	自然科学領域	自然科学領域の統計的研究を行う。
	人文社会科学領域	人文科学・社会科学領域の統計的研究を行う。
	複合領域（客員）	自然科学領域および人文社会科学領域にまたがる分野の統計的研究を行う。

附属施設

統計計算開発センター

統計計算の基礎的研究とそれを共有記憶型並列計算機、分散記憶型並列計算機、計算機クラスター上で実現することに関する研究を広く行っています。物理乱数発生のための機構についても研究しています。物理乱数を所内外からインターネットを通してオンデマンドでダウンロードするためのシステムを研究・開発し、公開に向けたテスト中です。グラフィック・ユーザー・インターフェース(GUI)を利用したクライアントサーバー型の統計解析システムについての研究・開発も行っています。共同利用、所内利用のための計算システムとLAN、WANに関する企画・整備・管理・運用も行っています。

目的		内容
統計計算に関する研究および計算システムとネットワークに関する企画・整備・運用・管理	研究開発第1部	統計計算に関する研究を行うとともに、計算システムに関する企画、整備、管理及び運用にあたる。
	研究開発第2部	統計計算に関する研究を行うとともに、計算機ネットワークに関する企画、整備、管理及び運用にあたる。

統計科学情報センター

統計科学に関する幅広い研究を行うとともに、外国人客員教官部門を持つ研究施設として国際交流を図り、統計科学に関する研究交流一般の推進、学術情報の提供、統計科学の教育普及に当たっています。統計科学では国内唯一の独立した国立研究機関として統計科学に関する図書・資料を幅広く収集しており、インターネットによる検索が可能となっています。また、本研究所の研究成果を出版して国内外の研究者に伝える一方、衛星通信による研究教育活動の公開、インターネットによる研究情報の提供、公開講座やセミナーの開催等により統計科学の成果を社会に還元することにも力を注いでいます。

目的		内容
統計科学に関する研究および国内的、国際的研究交流の促進、研究成果の出版、公開、統計科学の教育・普及	統計科学研究交流部	統計科学に関する研究を行うとともに、統計科学に関する国際・国内研究交流の推進をはかる。
	統計科学教育情報部	統計科学に関する研究を行うとともに、統計科学に関する研究指導及び普及、学術情報並びに図書・資料等の収集、管理提供及び公開に当たる。
	統計科学研究部 (外国人客員)	外国人客員教官による統計科学の理論と応用に関する研究を行う。

交流協定締結研究機関

アメリカ合衆国センサス局	アメリカ合衆国（ワシントン）	1988.7.27～
数学センター財団	オランダ王国（アムステルダム）	1989.5.10～
キューバ神経科学研究センター	キューバ（ハバナ）	1999.12.27～2004.12.26
統計学研究協会	ニュージーランド（ウェリントン）	2001.10. ～2006.10.
ソウル大学複雑系統計研究所	大韓民国（ソウル）	2002.10.17～

国際シンポジウム（平成14年度）

ISM Symposium:環境科学と統計科学の新たな融合—環境科学は統計科学に何を期待するか— 統計科学は環境科学に何ができるか—	場所：統計数理研究所
期間：2002年8月19日(月)～20日(火)	
マルコフ連鎖モンテカルロ法の発展と応用計量分析	場所：統計数理研究所
期間：2002年12月19日(木)～20日(金)	
ISM Symposium Statistics, Combinatorics and Geometry	場所：統計数理研究所
期間：2003年3月20日(木)～22日(土)	
ISMシンポジウム：インターネット調査の現状を検証する — 調査法としての評価方法と標準化をどう考えるか—	場所：統計数理研究所
協賛：(社)日本マーケティング・リサーチ協会、日本行動計量学会、日本分類学会	
期間：2003年3月25日(火)～26日(水)	

国際共同研究（平成14年度）

情報伝達関連分子種の分子系統樹解析に関する共同研究	ロックフェラー大学（アメリカ）	橋本 哲男
対称錐上の線形計画問題の解法とその周辺	ジョージア工科大学（アメリカ）	土谷 隆
（日米科学協力事業共同研究）		
データマイニングのための局所尤度法	ソウル国立大学、Keimyung大学（韓国）	江口 真透
（日本学術振興会「日韓科学協力事業」）		
選択的観測における統計推測について	ウォーリック大学（イギリス）	江口 真透
データ科学に基づく社会調査法に関する研究	中国人民大学、華東政法学院、江科学院 雲南大学（中国）	村上 征勝、鄭 躍軍
マングローブの進化について	上海復旦大学（中国）	曹 纓
分子進化に関する共同研究	中国中山大学（中国）	長谷川政美、鍾 揚
分子系統に関する共同研究	上海生物信息技術研究センター（中国）	曹 纓

外国人研究員（平成14年度）

Peng HUI<中華人民共和国>（中南大学）：発電プラントの排出NO _x の非線形予測と脱硝システムの最適制御
Wenwei REN<中華人民共和国>（復旦大学生物多様性研究所）：ヒト進化の分子系統学的解析
Jeffrey Kojiro LUM<アメリカ合衆国>（東京女子医科大学）：ヒト進化の分子系統学的解析
Dorothee HUCHON<フランス>（東京工業大学）：げっ歯類の分子系統学的解析
Michel-Marie DEZA<フランス>（フランス国立科学研究センター）：組み合わせ論の統計数理への応用
朴 承根<大韓民国>（韓国統計庁）：意識の国際比較方法の調査研究
岩田 貴樹<日本>（University of Liverpool）：潮汐と地震発生との相関を調べる新しい統計的手法
Andreas GALKA<ドイツ>（キール大学）：脳機能の非線形ダイナミクス解析
Pedro VALDES-SOSA<キューバ>（キューバ国立神経科学センター）：脳機能の非線形ダイナミクス解析
Arnaud DOUCET<フランス>（メルボルン大学）：モンテカルロ法の統計科学への応用
Renato D.C.MONTEIRO<ブラジル>（ジョージア工科大学）：対称錐上の線形計画問題の解法とその周辺
Samuel BURER<アメリカ合衆国>（アイオワ大学）：対称錐上の線形計画問題の解法とその周辺
Jimbo Henri CLAVER<カメルーン>（なし）：ファイナンス研究における非線形時系列解析手法の利用
Byeong PARK<大韓民国>（ソウル国立大学）：セミパラメトリックの新展開
Marc LAVIELLE<フランス>（パリ大学）：脳機能の非線形ダイナミクス解析
程 維虎<中華人民共和国>（北京工業大学応用数学院）：生物統計学における理論と実際
Gasper KOREN<スロベニア共和国>（リュブリャナ大学）：Web調査の方法論に関する研究（Methodology of Web surveys）
Stephane Jean-philippe SENECA<フランス>（なし）：複雑なモデル（非線形性・非ガウス性・長期記憶性等を有するモデル）のための微分幾何的方法及び確率的シミュレーション法の研究
Anthony J. HAYTER<連合王国>（ジョージア工科大学）：逐次数値積分に基づく高次元正規分布の確率計算
Juan Carlos JIMENEZ-SOBRINO<キューバ>（キューバ国立サイバネティクス数学・物理学研究所）：確率微分方程式の推定の研究
Elena PANTELEeva<ロシア>（モスクワ州立教育大学）：ボロノイ・ラティスの研究
Vasja VEHOVAR<スロベニア共和国>（リュブリャナ大学）：Web調査の方法論に関する研究（Methodology of Web surveys）
Mathieu DUTOIR<フランス>（パリ高等師範学校）：結晶群の出現頻度
Emmanuel GUERRE<フランス>（パリ大学（第6校））：オークションの統計理論
Hsien-kuei HWANG<中華民国>（台湾統計科学研究所）：計算アルゴリズムの確率論的解析

共同利用

大学等に所属する研究者が、研究所の施設を利用したり、研究所において統計に関する数理及びその応用の研究を行い、学術研究の発展に資することを目的としています。

共同利用の専門分野

共同利用は次のような専門分野に分類されています。この表は、申請者が主な研究領域の欄を参照して、適切な共同利用を申請していただくための参考資料です。

専門分野	主な研究領域
理論・方法論	1. 基礎理論関係 統計推測理論、応用確率論、データ解析の基礎研究、統計学の歴史、統計教育
	2. 計算と最適化 最適化、大規模線形計算、超高次元数値積分、インポートランスサンプリング、乱数、ニューラルネット、推論計算
	3. 時系列 時系列のモデリングと解析・予測・制御に関する研究
	4. 調査理論 統計データの取得法と解析法、統計的実証の理論と方法に関する研究
応用分野	5. 理工学関係 理学、工学における統計科学の実践
	6. 宇宙・地球科学 地球・惑星科学、宇宙科学に関連するデータの統計的解析
	7. 生物・医学 生物学・医学・薬学・農学・生活科学等の生命科学分野における統計的実証研究
	8. 人文・社会科学 人文・社会科学における統計的情報の活用法の研究
	9. 環境科学 環境と生態学に関する統計学的研究
	10. その他 上記以外の研究領域

(2002年の概要より)

専門分野		平成10年度	平成11年度	平成12年度	平成13年度	平成14年度
理論・方法論	1. 基礎理論関係	20件	26件	24件	21件	22件
	2. 計算と最適化	17	10	7	9	8
	3. 時系列	17	19	21	20	16
	4. 調査理論	4	4	3	4	2
応用分野	5. 理工学関係	14	11	18	11	10
	6. 宇宙・地球科学	6	4	4	4	8
	7. 生物・医学	32	28	27	23	26
	8. 人文・社会科学	19	13	12	15	13
	9. 環境科学	8	4	8	9	6
	10. その他			1	4	3
計		137	119	125	120	114

共同利用採択課題一覧（平成15年度）

平成15年6月1日現在

共同利用登録（9件）

分野	研究課題名	利用登録者（所属）
基礎理論研究系	学校教育における統計教育について	伊藤 一郎（東京学芸大学・教育学部）
計算と最適化	数値的最適化についての研究 (1)計算理論の力学系的基礎付け (2)学習・記憶の力学系モデル 豊かな時間構造を持つ動的な生物情報処理の理論研究	水野 眞治（東京工業大学・大学院社会理工学研究科） 佐藤 譲（理化学研究所・脳科学総合研究センター） 藤本 仰一（東京大学・大学院総合文化研究科）
時系列	カオス時系列におけるフラクタル次元の推定	川口 淳（九州大学・大学院数理学府）
生物・医学	Beyond Analysis of Variance データ解析用ソフトウェアの開発 アミノ酸サイト別進化速度の比較分子生物学	広津 千尋（明星大学・理工学部） 吉岡 耕一（東京医科歯科大学・ 大学院保健衛生学研究科） 和田 康彦（佐賀大学・農学部）
人文・社会科学	論文文の文章構造の研究－論文文を支える文型の確定を目指して－	村田 年（慶應義塾大学・国際センター）

共同利用研究 1（23件）

分野	研究課題名	利用登録者（所属）
基礎理論研究系	非可換確率空間における分布論に関する研究 非線形モデルによる分光スペクトルの解釈方法の検討	吉田 裕亮（お茶の水女子大学・理学部） 貝原巳樹雄（一関工業高等専門学校）
時系列	臨床における経時データの解析 脳機能の非線形ダイナミクス解析 金融リスクの制御と資産最適配分に関する研究 ガスコンバインドサイクル発電プラント排煙脱硝制御改善に関する研究 重症脳性麻痺児の脳波活動と脳幹下部網様体活動との相互応答 計量ファイナンスにおける離散モデリングとその応用 超高頻度の円ドルレートを用いた為替レートの変動性に関する研究 高頻度金融データのデータベース整理と計算機解析 非対称分布を用いた非ガウス型時系列モデル	石黒真木夫（統計数理研究所・予測制御研究系） 尾崎 統（統計数理研究所・予測制御研究系） 尾崎 統（統計数理研究所・予測制御研究系） 尾崎 統（統計数理研究所・予測制御研究系） 小川 昭之（大分医科大学・医学部） 川崎 能典（統計数理研究所・予測制御研究系） 須齋 正幸（長崎大学・経済学部） 田中美栄子（宮崎大学・工学部） 永原 裕一（明治大学・政治経済学部）
理工学関係	複雑系の相転移の数値的研究 統計論的トンネル速度理論の開発	小野 豆郎（日本女子大学・理学部） 武次 徹也（お茶の水女子大学・ 大学院人間文化研究科）
宇宙・地球科学	応力変化に起因する地震活動の時間変動 ランダムな不均質を含んだ地下構造の地震波によるイメージング	岩田 貴樹（防災科学技術研究所） 西澤 修（産業技術総合研究所）
生物・医学	比較ゲノム学による真核生物の系統進化の解明 喘息発作の環境因子によるモデル 時系列解析法による動物の高精度行動解析 含菌性嚢胞と良性腫瘍との鑑別点についての研究	橋本 哲男（筑波大学・生物科学系） 清水 悟（東京女子医科大学・医学部） 花井 一光（京都府立医科大学・医学部） 池島 厚（日本大学・松戸歯学部）
人文・社会科学	生産関数の特定化に関する統計的推測	川崎 能典（統計数理研究所・予測制御研究系）
環境科学	東京湾の水質の長期的な変動に関する研究 地球環境のリモートセンシングと統計的手法 ダイオキシン類組成解析による挙動並びに発生源の検討	柏木 宣久（統計数理研究所・調査実験解析研究系） 柏木 宣久（統計数理研究所・調査実験解析研究系） 佐々木裕子（東京都環境科学研究所・分析研究部）

共同利用研究 2（53件）

分野	研究課題名	利用登録者（所属）
基礎理論研究系	乱数生成法とその検定の研究 無限分解可能過程に関連する諸問題 システム故障の統計的推測 漸近展開の研究 電子社会における統計的手法のニーズ把握のための基礎研究 統計データ公開システムの開発・利用に関する研究 スキャン統計量に関する確率分布と統計的応用 データ指向型統計事例データベースの構築に関する研究 確率解析による統計学の研究 独立成分分析に関する理論とその応用 混雑されたイルカ数への拡張Zero-Inflated Poisson回帰モデルの適用 スピンをもつ粒子の一次元ランダムバックギングの解析的研究 非線形構造探索のための統計的モデリング 平面・球面・空間における円・球・ロッドのランダムな配置の研究	高嶋 恵三（岡山理科大学・理学部） 石川 保志（愛媛大学・理学部） 二ツ谷昌夫（弘前大学・理工学部） 吉田 朋広（東京大学・大学院数理学府研究科） 村上 征勝（統計数理研究所・領域統計研究系） 宿久 洋（鹿児島大学・理学部） 平野 勝臣（統計数理研究所・統計基礎研究系） 山本 義郎（多摩大学・経営情報学部） 阪本 雄二（広島国際大学・人間環境学部） 南 美穂子（統計数理研究所・統計基礎研究系） 南 美穂子（統計数理研究所・統計基礎研究系） 伊藤 栄明（統計数理研究所・領域統計研究系） 小西 貞則（九州大学・大学院数理学府研究科） 磯川 幸直（鹿児島大学・教育学部）
計算と最適化	次世代統計システムJaspの開発 局所モーメント法に関する研究	小林 郁典（徳島文理大学・工学部） 寒河江雅彦（岐阜大学・工学部）
時系列	粒子型フィルタによる時系列解析の動的識別問題への応用 POSデータにおける価格反応分析	生駒 哲一（九州工業大学・工学部） 近藤 文代（筑波大学・社会工学系）

（次頁につづく）

分野	研究課題名	利用登録者（所属）
調査理論	個票データの開示におけるリスクの評価と官庁統計データの公開への応用	佐井 至道（岡山商科大学・商学部）
	女性の自立意識についての構造分析	高倉 節子（長崎純心大学大学院）
理工学関係	不完全情報下における制御系設計に関する研究	宮里 義彦（統計数理研究所・予測制御研究系）
宇宙・地球科学	超大規模人工衛星データセットからの情報の組織的抽出・統合と可視化技術の応用（2）	樋口 知之（統計数理研究所・予測制御研究系）
	プラズマ粒子速度データの混合分布モデルによる分析	中村 永友（札幌学院大学・経済学部）
生物・医学	看護職における疲労度解析	中山 晃志（大分県立看護科学大学・人間科学講座）
	ブナ林の遺伝構造とその解析のための空間統計学	島谷健一郎（統計数理研究所・調査実験解析研究系）
	多年生林床草本の空間的個体群動態解析	島谷健一郎（統計数理研究所・調査実験解析研究系）
	不正咬合患者に見られる時系列生体情報の統計数理解析の検討	相馬 邦道（東京医科歯科大学・大学院歯学総合研究科）
	表皮ランゲルハンス細胞の空間配置モデル	窪田 泰夫（香川医科大学・医学部）
	植物の形態形成形質評価における光計測による簡便化とその利用	平田 豊（東京農工大学・大学院農学研究科）
	分子系統樹推定の諸問題	長谷川政美（統計数理研究所・予測制御研究系）
	皮膚病変の（コンピュータによる自動診断のための）数値化と増殖モデル作成	今山 修平（国立病院九州医療センター・皮膚科）
	ヒトの縦断的成長に関する研究	正法寺孝雄（福山大学・人間文化学部）
	水産資源に対する観察データ解析のための統計推測	北門 利英（東京水産大学・水産学部）
	コウホート分析による歯科疾患調査データの国際比較	中村 隆（統計数理研究所・調査実験解析研究系）
	類洞の立体構造に影響を及ぼす肝細胞の形状---その統計的解析	種村 正美（統計数理研究所・調査実験解析研究系）
	安全と効果の評価のための設計と基盤	柳本 武美（統計数理研究所・領域統計研究系）
	21世紀型保健医療指標の開発（2）	大野ゆう子（大阪大学・医学部）
	生物統計学における科学的証拠生成のための研究方法の改善	藤井 良宣（宮崎大学・教育文化学部）
人文・社会科学	カオス理論による糖尿病血糖値時系列データの短期予測システムの構築と検証	有田清三郎（関西医科大学・数学教室）
	看護における物語と対話の効果	柴田 義貞（長崎大学・医学部）
	英語コーパスを用いた意味類型論の考察	高橋 薫（豊田工業高等専門学校）
	文化計量のための基礎研究	村上 征勝（統計数理研究所・領域統計研究系）
	犯罪統計データのコウホート分析	中村 隆（統計数理研究所・調査実験解析研究系）
	スポーツ・レクリエーション参加のコウホート分析	中村 隆（統計数理研究所・調査実験解析研究系）
	心理学と社会調査における文章型データと選択肢設問の統計解析	土井 聖陽（大阪樟蔭女子大学・短期大学部）
	日中国民性データの統計的解析	鄭 躍軍（統計数理研究所・領域統計研究系）
	千葉県の縄文貝塚の時期別モデルの作成	植木 武（共立女子短期大学・生活科学科）
	CBTにおける項目プール構築と管理のための統計的方法	村木 英治（東北大学・情報教育学研究部）
環境科学	点過程モデルと分子遺伝マーカーによる天然林施業の検討	島谷健一郎（統計数理研究所・調査実験解析研究系）
	日本列島の気候傾度における森林群集の空間構造に関する比較解析	久保田康裕（鹿児島大学・教育学部）
	リモートセンシングデータの解析とその植物生態学への応用	西井 龍映（広島大学・総合科学部）
	温暖化防止に向けた循環型森林資源管理を通じた炭素収支モデルの構築	吉本 敦（東北大学・大学院環境科学研究科）
その他	オンライン・ラーニングを目指した統計科学コンテンツ開発に関する基礎的研究	金藤 浩司（統計数理研究所・統計科学情報センター）

共同利用研究 3（1件）

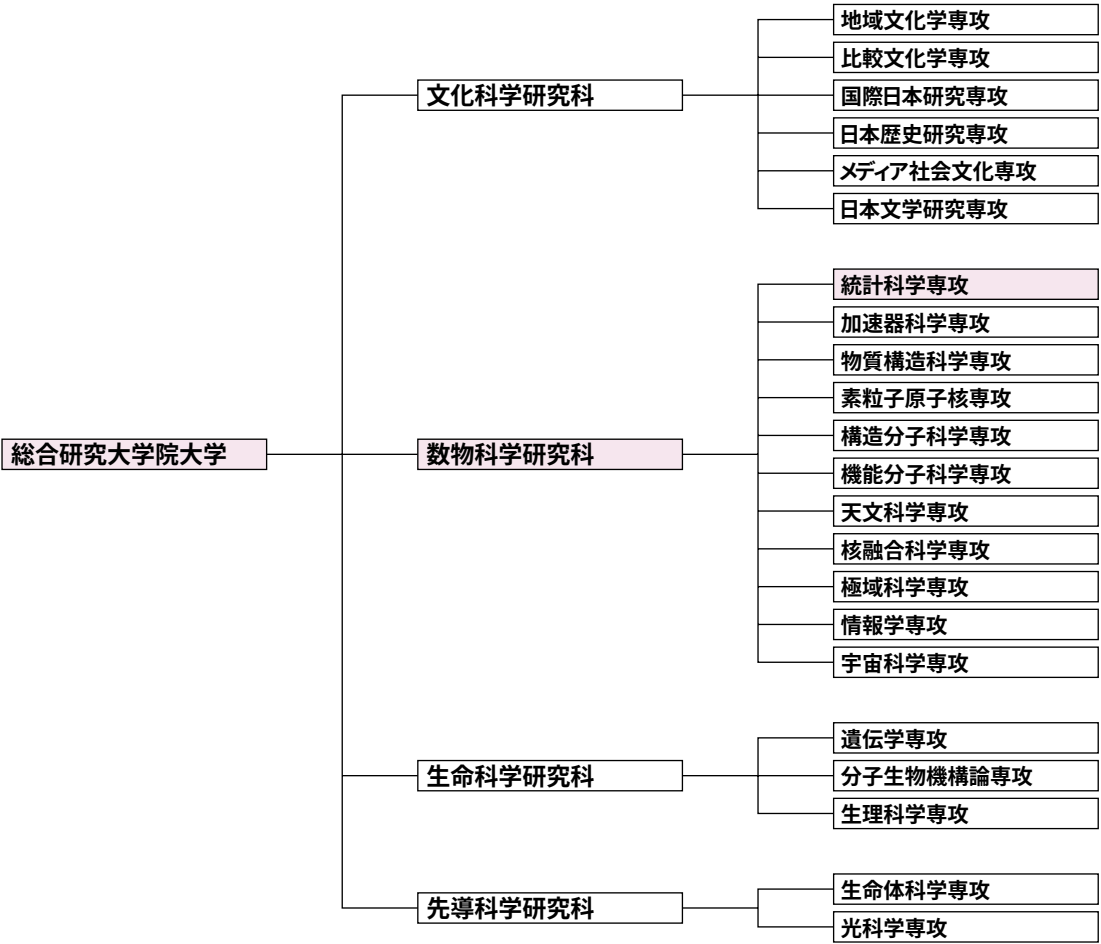
分野	研究課題名	利用登録者（所属）
環境科学	水環境問題に於ける統計的認識手法の確立と実務での検証	岩瀬 晃盛（広島大学・大学院工学研究科）

共同研究集会（7件）

分野	研究課題名	利用登録者（所属）
基礎理論研究系	極値理論の工学への応用	高橋 倫也（神戸商船大学・商船学部）
計算と最適化	最適化：モデリングとアルゴリズム	田邊 國士（統計数理研究所・予測制御研究系）
	動的システムの情報論III	佐藤 譲（理化学研究所・脳科学総合研究センター）
理工学関係	21世紀の診断工学とその周辺	伊良皆啓治（東京大学・大学院医学系研究科）
	乱流の統計理論とその応用	岡崎 卓（統計数理研究所・統計科学情報センター）
環境科学	環境統計データ解析の理論と実際	清水 邦夫（慶應義塾大学・理工学部）
その他	統計サマーセミナー	桜井 裕仁（北海道大学・大学院工学研究科）

大学院組織

統計数理研究所は、昭和63年10月に開学した学部を持たない大学院だけの大学、総合研究大学院大学（神奈川県三浦郡葉山町）の基盤機関の一つとして、創設時から数物科学研究科・統計科学専攻を設置し、平成元年4月から学生を受け入れて、博士後期課程の教育研究を本研究所で行っています。



講座の概要

本専攻は、科学的仮説の構築・検証等、合理的な推論あるいは予測の実現を目的とするデータの有効利用に関する教育研究を行います。統計基礎、調査実験解析、予測制御、領域統計の4講座で構成され、統計科学の新分野を生み出すような創造性豊かな研究者を養成することを目的としています。

講座名	内 容
統計基礎	統計科学の基礎概念（ランダムネス、尤度、エントロピー等）、統計的推論、並びにその基礎となる確率、確率分布、確率過程等の確率機構に関する教育研究を行います。
調査実験解析	標本調査法、実施計画等のデータ収集に係る諸方法、並びに空間事象、多次元解析、パターン解析、点過程・時空間事象等の統計解析に係る諸方法に関する教育研究を行います。
予測制御	統計モデル構成とデータに基づくモデル評価を組織的に用いて統計的推論法及びそれに基づく予測・制御・管理の理論と応用、並びに統計的推論をコンピュータによって実行するための数値・非数値計算及び最適化の理論とアルゴリズムに関する教育研究を行います。
領域統計	治療効果・生活習慣の健康要因評価及び安全評価に関した生物統計手法、理工学等の分野における空間的、かつ、動的な扱いを要する現象の解析法、並びに人間行動の測定と分析のための統計的方法等に関する教育研究を行います。

教育研究の特色

＊理論から応用までの多分野にわたる研究者により、統計科学全般についての教育、研究が行われています。 ＊統計数理研究所においては大学共同利用機関として、他の大学・研究機関とのユニークな共同研究が活発に行われ、それらに参加することができます。 ＊優れたハードとソフトの環境が準備されています。 ＊国の内外からの多くの訪問者と研究交流を行うことができます。 ＊総合研究大学院大学では東京工業大学大学院、お茶の水女子大学大学院と学生交流協定を結んでおり、必要であれば単位を取得することができます。

修了要件および学位の種類

本専攻の修了要件は、大学院に3年以上在学し、10単位以上を修得し、かつ、必要な研究指導を受けたうえ、本大学院の行う博士論文の審査及び最終試験に合格することとしています。その場合には、博士（学術）の学位が授与されます。なお、優れた研究成績を上げた者の在学年限については、弾力的な取扱いがされることがあります。

入学定員及び在学生数

平成15年4月1日現在

入学年度	平成15年度	平成14年度	平成13年度	平成12年度	平成11年度	平成10年度
定 員	4	4	4	4	4	4
現 員	6②	5(1)④	7①	2①	3②	1①

() は国費留学生で内数、○は有職者で内数

入学者の出身大学

国公立：東北大学2・福島大学1・筑波大学5・埼玉大学1・千葉大学1・東京大学11・東京学芸大学1・東京工業大学3・東京商船大学1・お茶の水女子大学1・一橋大学1・名古屋大学2・大阪大学2・島根大学2・九州大学2・北陸先端科学技術大学1・大阪市立大学1
私 立：慶應義塾大学1・中央大学4・東京理科大学2・東洋大学1・日本大学2・法政大学4・早稲田大学4・京都産業大学1
外 国：中国科学院1・中国科学技術大学1・東北工科大学1・吉林師範大学1・香港科技大学1・Colorado州立大学1・Hawaii大学1・Campinas大学(ブラジル)1・Dhaka大学(バングラディシュ)2・Jahangirnagar大学(バングラディシュ)2・Rajshahi大学(バングラディシュ)1

学位授与数

平成3年度	博士（学術）2名	平成9年度	〃（〃）1名
平成4年度	〃（〃）1名	平成10年度	〃（〃）4名（論文博士1名を含む）
平成5年度	〃（〃）2名	平成11年度	〃（〃）6名
平成6年度	〃（〃）7名	平成12年度	〃（〃）5名
平成7年度	〃（〃）6名（論文博士1名を含む）	平成13年度	〃（〃）5名
平成8年度	〃（〃）3名（論文博士1名を含む）	平成14年度	〃（〃）4名

修了生等の現在

国公立大学等：東京大学助教授・筑波大学助教授・埼玉大学助教授・山梨医科大学助教授・電気通信大学助教授・神戸商科大学助教授・九州工業大学講師・筑波大学講師・北海道大学助手・琉球大学助手・東京水産大学助手・広島大学助手・東京工業大学助手・理化学研究所研究員・統計数理研究所研究員
私立大学等：旭川大学教授・札幌学院大学教授・明治大学助教授・札幌学院大学助教授・城西大学講師・札幌学院大学講師・青山学院大学助手
外国の大学等：Jahangirnagar大学教授・Deakin大学講師・Massey大学研究員・Otago大学研究員・Wurzburg大学研究員・Hong Kong Baptist大学講師
民間企業等：日立製作所中央研究所研究員・新昭和・公立学校教諭・NTTコミュニケーション科学研究所研究員・財団法人統計情報研究開発センター・東京都精神医学研究所学振特別研究員・みづほ信託銀行年金運用研究所研究員・ニッセイ基礎研究所主席主任研究員

公開講座

沿革

統計数理研究所における社会人教育は、研究所設立時（昭和19年）に附置された文部省科学研究補助技術員養成所数値計算第一期養成所に始まります。

昭和22年には、当時の統計行政組織の改善や不足していた統計職員の養成機関として、中核となる統計技術職員や統計技術教育者を養成するために、附属統計技術員養成所が開設され、本格的な社会人教育が始まりました。

その後、社会情勢の変化に伴い、当初の目的であった優秀な統計技術員を養成し、社会に供給するということから、しだいに一般社会人に対する統計教育に重点が移り、公開講座が開講されるようになりました。また、統計的方法が普及し、様々な分野に応用されるに至り、より広範で高度な統計学の教育の必要性が叫ばれ、その要請にこたえるべく講義内容も豊かになっていきました。

昭和40年代に入ると、講座数は年間に6～8講座となり、大阪、岡山、福岡などの地方でも講座が開かれるようになりました。

昭和60年、本研究所の大学共同利用機関への改組転換に当たり、附属統計技術員養成所は廃止されることになりました。しかし、公開講座に対する社会的要求は強く、統計科学情報センターが中心になり、年間3～4講座を開講して現在に至っております。

講座の内容

昭和44年度から平成14年度までに開設した講座数は延べ185、受講生総数は15,373人にのぼり、その内容は基礎から応用まで多岐にわたっています。これまでに開講された講座の主な内容は次のとおりです。

年度	種類	講座名	年月	受講者数(人)
平成10年度	統計数理概論	統計解析入門	7月	108
	統計数理要論A	数理計画法	10月	98
	統計数理要論B	基礎からの情報幾何	11月	78
	統計数理要論C	わかる多変量解析	12月	114
平成11年度	統計数理概論	統計解析入門	7月	109
	統計数理特論	幾何学的構造の統計	12月	23
	統計数理要論A	サンプリング入門と調査データの分析法	12月	99
	統計数理要論B	とっても楽しい医学統計	2月	81
平成12年度	統計数理要論A	情報量規準の世界	7月	49
	統計数理概論A	統計分析の基礎	7月	106
	統計数理概論B	統計学概論	10月	69
	統計数理要論B	統計学・ニューラルネットワーク・学習，その最前線	12月	66
平成13年度	統計数理概論	統計学概論	7月	93
	統計数理要論A	資源管理のための統計分析	10～11月	16
	統計数理要論B	数量化法入門	11月	49
	統計数理要論C	統計理論に関わる最近の情報通信の理論と実際	12月	15
平成14年度	統計数理概論	統計学概論	7月	95
	統計数理要論A	モンテカルロフィルタ入門と実戦的応用	10月	71
	統計数理要論B	統計学・ニューラルネットワーク・学習，その最前線 part2	11月	64
	統計数理特論	計算科学と統計科学の接点 —並列計算，グリッド，専用計算機，乱数—	3月	38

統計数理セミナー

毎週水曜日、午後1時半から約1時間、所内教官、及び国内、海外からの所外の方々によるセミナーを統計数理研究所本館2階講堂で開催しています。このセミナーは時には衛星通信ネットワーク（SCS）を利用して、他の大学等と結んで遠隔地間で同時開催をすることもあります。セミナーの聴講は自由です。このセミナーに関する開催予定表及び関連する情報は統計数理研究所のホームページで扱っています。<http://www.ism.ac.jp/>

公開講演会

毎年、教育文化週間（11月1日～7日）に、本研究所の活動の一端を紹介し、統計科学の普及をめざして開催しています。特定のテーマのもとに、数名の講師が、統計科学の先端的話題について分かりやすく講演します。この3年間には次のようなテーマが取上げられました。「インターネット調査とそれを巡る諸問題」、「これからの教育と統計」、「統計科学に期待するもの」。今年度も11月5日に開催を予定しております。公開講演会の聴講は自由です。

統計相談

本研究所では、研究成果の社会還元に積極的に取り組み、統計科学情報センターを窓口として、一般社会人や研究者等からの統計科学に関する相談に随時応じています。相談の内容は、基本的なものから専門的なものまで多岐にわたり、約半数が民間からの相談で、残りを公的機関、大学の教員、学生が占めています。教官が直接対応する専門的な相談は年間約20件あり、その内の4割程度が学会などでの具体的成果として、社会に還元されています。

平成14年度文部科学省大学等地域開放特別事業 統計数理研究所「数理のめがね」

小中学生を対象にした平成14年度大学等地域開放特別事業「数理のめがね」を平成14年11月2日に実施し、当日は小学生を中心に20余名の参加者がありました。前半では、東京農工大学小谷善行教授による「コンピュータは素晴らしいおもちゃ」の講義があり、コンピュータは、優れた道具であると共に、新しい発想力を培う優れた「おもちゃ」であり、子供達は、じゃんけんや相撲などのコンピュータゲームで楽しみました。

後半では、統計数理研究所長谷川政美教授による子供向け「生物の進化」の話があり、クジラの祖先は、ウシ、カバ、ラクダ、ブタなどの仲間の偶蹄類であって、その中でも特にカバに近い仲間であることが最近のDNAの研究で分かったことなどの説明に対し、子供達は、驚いた様子で、付き添いの保護者からは、さわやかな秋の1日を有意義に過ごすことが出来たと好評でした。

第2回目として3月15日にも実施し、当日は、小学生2名とその保護者2名計4名の参加者があり、昨年11月2日に同事業「数理のめがね」第1弾を実施し、その時に使用したソフトに改良を加え、対戦型ゲームソフト（戦略方法を自分で組むことが出来る。）とし、家族的な雰囲気の中、子どもたちと本研究所教官が対戦するなどして、じゃんけんゲームを楽しみました。



歳出決算

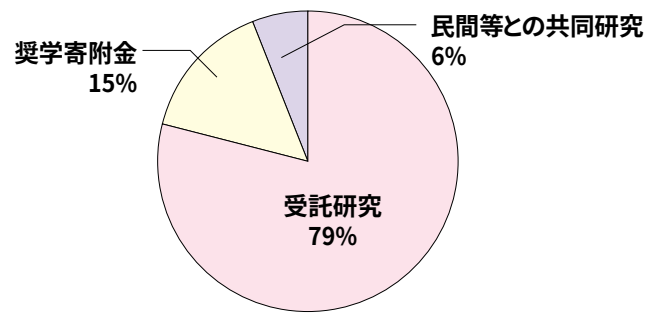
平成14年度
単位：千円

区	分	歳 出 額
人 件 費		806,149
物 件 費		1,022,489
施 設 整 備 費		8,660
合 計		1,837,298

外部資金受入状況

平成14年度
単位：千円

区	分	件 数	受入金額
受 託 研 究		7	19,977
奨 学 寄 附 金		8	3,900
民間等との共同研究等		2	1,500
合 計		17	25,377



科学研究費補助金

平成14年度
単位：千円

研 究 種 目	件 数	交付金額	研 究 種 目	件 数	交付金額
基盤研究 (A)	2	23,500	特定領域研究	1	3,300
基盤研究 (B)	7	22,800	萌芽研究	1	1,200
基盤研究 (C)	14	13,800	若手研究 (A)	1	5,900
			若手研究 (B)	3	3,597
			特別研究員奨励費	4	3,700
			合 計	33	77,797

敷地・建物

平成15年4月1日現在

- 敷地面積 ————— 5,032m²
- 建物面積（延べ面積） — 6,286m²

建物名称	構造階数	延べ面積
庁舎	R3	4,818m ²
情報統計研究棟	R3	1,024m ²
電子計算機棟	R2	368m ²
体育場等	S1	76m ²

計算資源

平成15年4月1日現在

並列ベクトル計算機HITACHI SR8000（20ノード、主記憶160GB）と管理用計算機HP 9000V2250（16CPU、主記憶12GB）を中心とした統計科学スーパーコンピュータシステムを平成11年1月に導入し大規模統計的データ解析のために用いています。なお、平成16年1月により高性能の新システムに更新する予定です。また、SGI2800（64CPU、主記憶51.5GB）を中心としたシステムに、計算統計学支援システムを平成12年1月に更新し統計科学のための並列処理の研究も引き続き行っています。さらに、PCI物理乱数発生ボードを搭載したLinuxパーソナルコンピュータSGI1200、100台のクラスタであるブートストラップシステム、Onyx3000を中心とした高精度高速グラフィックスシステムをそれぞれ平成13年1月、平成13年3月に導入しました。

従来の10Base5によるイーサネットに加え、1000Base-SXを幹線とし、100Base-TXを支線に持つ所内情報網を平成10年12月に敷設しました。研究室に配置されたワークステーション、パーソナルコンピュータ、統計科学スーパーコンピュータシステム、計算統計学支援システム、ブートストラップシステム、高精度高速グラフィックスシステム等が接続されています。この所内情報網を利用した分散処理が可能であり、計算資源、統計データの有効利用が行われています。また、研究室に配置されたワークステーション、パーソナルコンピュータを用いて統計科学スーパーコンピュータシステム、計算統計学支援システム上で動作するプログラムの開発が行われています。国内外の研究者との共同研究を活発にするため及び電子メールの交換等のためにSINETによってインターネットと接続されています。通信速度は平成11年度途中に1.5Mbpsになり、平成14年7月からは100Mbpsになっています。また、平成15年10月からはスーパーSINETに接続予定です。公衆回線からもターミナルサーバーを通して利用可能です。

図書・資料

平成15年4月1日現在

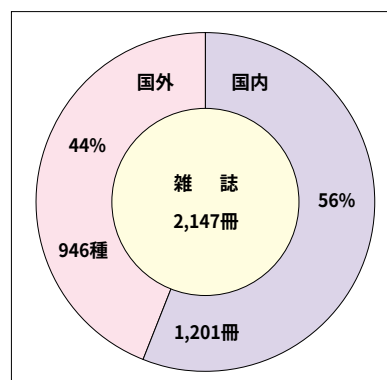
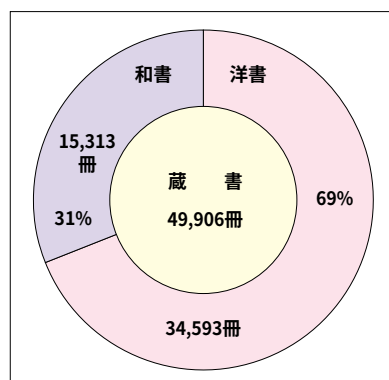
本研究所の広範な研究分野を反映して、統計学、数学、計算機科学、情報科学に関わる内外の主要学術誌を多数備えています。収蔵図書はこれらの分野に加えて人文・社会科学から生物、医学、理工学の広範な領域にわたっています。

また本研究所が刊行する欧文誌「Annals of the Institute of Statistical Mathematics」（Kluwer Academic Publishersから発行）、和文誌「統計数理」、「日本人の国民性の研究」など調査研究のための「統計数理研究所研究リポート」、「Computer Science Monographs」、共同利用における共同研究のための「共同研究リポート」、「Research

Memorandum」、「計算技術報告」、「研究教育活動報告」および内外からの寄贈による資料も備えています。

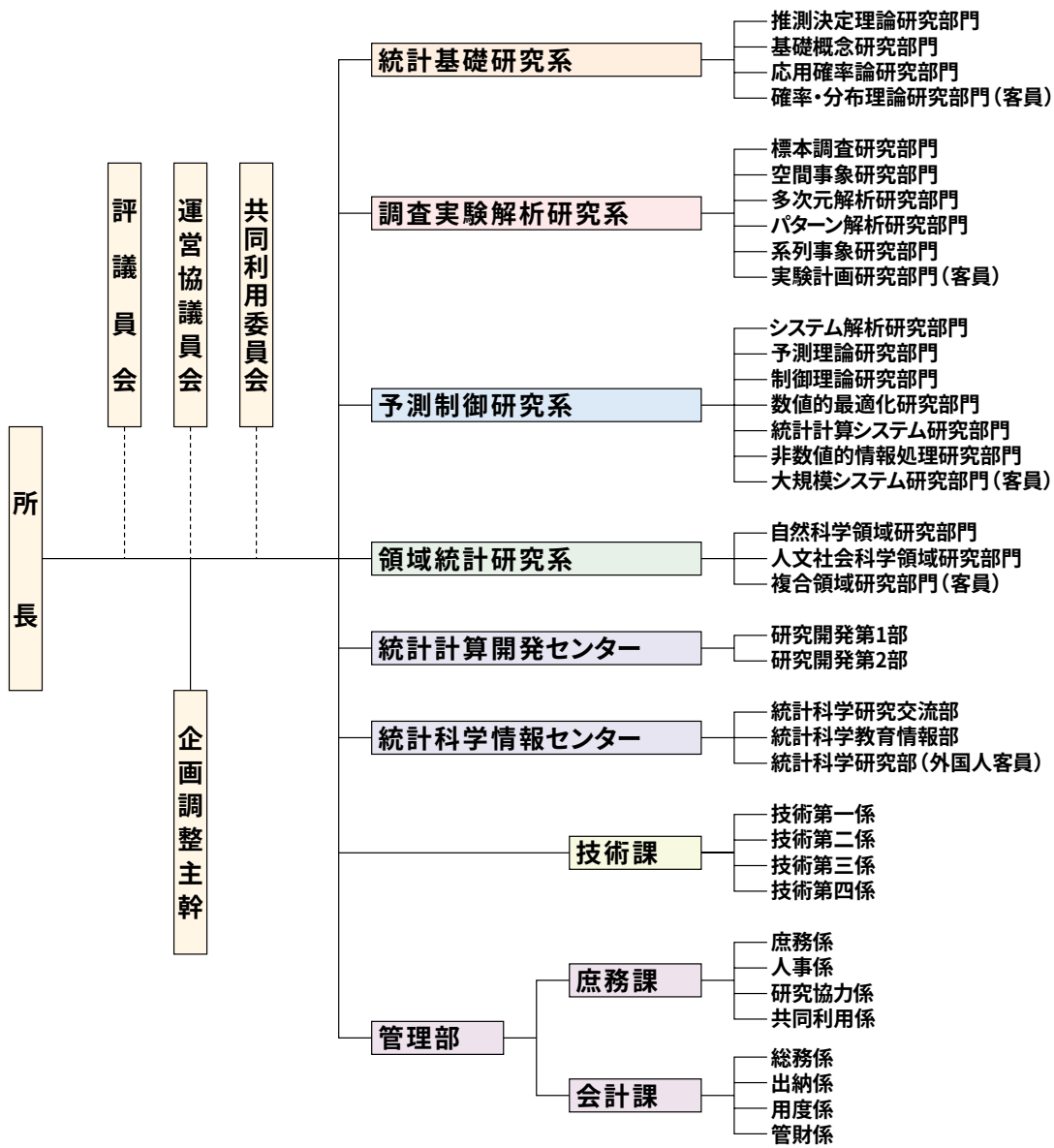
あらゆる分野の研究者の需要に応えるため、図書・資料を整理し、Webから検索出来るようになっています。また文献の問い合わせと複写サービスも行っています。

（問い合わせ先：03-5421-8719）



機構図

平成15年4月1日現在



定員

平成15年4月1日現在

区	分	所 長	教 授	助教授	助 手	事務官	技 官	合 計
所	長	1						1
統 計 基 礎 研 究 系			3 (1)	3 (1)	2			8 (2)
調 査 実 験 解 析 研 究 系			5 (1)	5 (1)	4			14 (2)
予 測 制 御 研 究 系			6 (1)	5 (1)	5			16 (2)
領 域 統 計 研 究 系			3 (2)	2 (2)	2			7 (4)
附 属 研 究 施 設			4 ①	4 ①	1			9 ②
管 理 部						17	1	18
技 術 課							11	11
計		1	21 (5) ①	19 (5) ①	14	17	12	84 (10) ②

※ ()内は客員で外数、○内は外国人客員で外数

所 員

平成15年4月1日現在

所 長 北川源四郎
企画調整主幹(副所長)(併) 田邊 國士

■統計基礎研究系

研究主幹(併) 平野 勝臣

推測決定理論研究部門

教 授 平野 勝臣
助教授 栗木 哲
助 手 西山 陽一

基礎概念研究部門

教 授 松縄 規
助教授 南 美穂子

応用確率論研究部門

教 授 江口 真透
助教授 福水 健次
助 手 志村 隆彰

確率・分布理論研究部門(客員)

教 授(併) 高橋 倫也
(神戸商船大学)
助教授(併) 石川 保志
(愛媛大学)

■調査実験解析研究系

研究主幹(併) 坂元 慶行

標本調査研究部門

教 授 坂元 慶行
助 手 前田 忠彦

空間事象研究部門

教 授 種村 正美
助教授 丸山 直昌
助 手 島谷健一郎

多次元解析研究部門

教 授 中村 隆
助教授 池田 思朗

パターン解析研究部門

教 授 大隅 昇
助教授 土屋 隆裕
助 手 清水 信夫

系列事象研究部門

教 授 尾形 良彦
助教授 柏木 宣久

実験計画研究部門(客員)

教 授(併) 前原 闊
(琉球大学)
助教授(併) 村田 泰章
(産業技術総合研究所)

■予測制御研究系

研究主幹(併) 長谷川政美

システム解析研究部門

教 授 樋口 知之

助 手 川崎 能典

予測理論研究部門

教 授 長谷川政美
助教授 足立 淳
助 手 曹 纓

制御理論研究部門

教 授 尾崎 統
助教授 宮里 義彦
助 手 上野 玄太

数値的最適化研究部門

教 授 田邊 國士
助教授 伊庭 幸人
助 手 上田 澄江

統計計算システム研究部門

教 授 土谷 隆
助教授 佐藤 整尚
助 手 染谷 博司

非数値の情報処理研究部門

教 授 石黒真木夫
助教授 瀧澤 由美

大規模システム研究部門(客員)

客員助教授 矢船 明史
(北里研究所)
助教授(併) 篠田 浩一
(東京工業大学)

■領域統計研究系

研究主幹(併) 伊藤 栄明

自然科学領域研究部門

教 授 柳本 武美
教 授 伊藤 栄明
助教授 藤澤 洋徳
助 手 大西 俊郎

人文社会科学領域研究部門

教 授 村上 征勝
助教授 吉野 諒三
助 手 鄭 躍軍

複合領域研究部門(客員)

教 授(併) 藤江 幸一
(豊橋技術科学大学)
教 授(併) 栗原 考次
(岡山大学)
教 授(併) 大野 優子
(大阪大学)
客員助教授 泰地真弘人
(理化学研究所)

■統計計算開発センター

センター長(併) 田村 義保

研究開発第1部

教 授 田村 義保

助教授 松井 知子

助 手 荒畑恵美子

研究開発第2部

教 授 中野 純司
助教授 伊藤 聡

■統計科学情報センター

センター長(併) 石黒真木夫

統計科学研究交流部

教 授 馬場 康維
助教授 金藤 浩司

統計科学教育情報部

教 授 岡崎 卓
助教授 山下 智志

統計科学研究部(外国人客員)

客員教授 DOLBILIN Nikolai
Petrovich
(H.15.4.1~15.6.30)
客員助教授 BOSCH-BAYARD
Jorge Francisco
(H.15.4.1~15.12.31)

■管理部

部 長 濱崎 豊

庶務課

課 長 河野 克俊
課長補佐 新川 昇
課長補佐 斉藤 仁
庶務係長 須藤 文雄
人事係長 服部 斎
研究協力係長 関 祐子
共同利用係長 細田 明

会計課

課 長 平松 昌弥
課長補佐 永野 謙一
(移転推進室長 斉藤 仁)
総務係長 高野 米孝
出納係長 瀬見千恵子
用度係長 長谷川清徳
管財係長 山本 雅久

■技術課

課長(併)教授 種村 正美
技術第一係長 桂 康一
技術第二係長 太田 岭子
技術第三係長 清水 恵子
技術第四係長 寺尾 節子

評 議 員

平成15年6月1日現在

相澤 益男	東京工業大学長	佐和 隆光	京都大学経済研究所長
甘利 俊一	理化学研究所脳科学総合研究センター長	末松 安晴	国立情報学研究所長
安西祐一郎	慶應義塾大学塾長	高橋陽一郎	京都大学数理解析研究所長
伊藤 彰彦	(財)日本統計協会理事長	田中 直毅	21世紀政策研究所理事長
甲斐 睦朗	独立行政法人国立国語研究所長	藤田 文子	津田塾大学学芸学部教授
勝木 元也	岡崎国立共同研究機構基礎生物学研究所長	松野 陽一	国文学研究資料館長
北原 保雄	筑波大学長	丸山 瑛一	理化学研究所フロンティア研究システム長
木村 孟	大学評価・学位授与機構長	吉村 功	東京理科大学工学部第一部教授
小池 唯夫	パシフィック野球連盟会長	渡邊 興亞	国立極地研究所長
小平 桂一	総合研究大学院大学長		

運営協議員

平成15年6月1日現在

所外委員		所内委員	
及川 昭文	総合研究大学院大学教育研究情報資料センター教授	石黒真木夫	教授（統計科学情報センター長）
小西 貞則	九州大学大学院数理学研究院教授	伊藤 栄明	教授（領域統計研究系研究主幹）
佐藤 俊哉	京都大学大学院医学研究科教授	尾形 良彦	教授（調査実験解析研究系）
竹内 正弘	北里大学大学院薬学研究科教授	坂元 慶行	教授（調査実験解析研究系研究主幹）
竹村 彰通	東京大学教授大学院情報理工学系研究科	田邊 國士	教授（企画調整主幹）
田中 勝人	一橋大学大学院経済学研究科教授	種村 正美	教授（技術課長）
林 文	東洋英和女学院大学人間科学部教授	田村 義保	教授（統計計算開発センター長）
廣津 千尋	明星大学理工学部教授	土谷 隆	教授（予測制御研究系）
水本 好彦	国立天文台天文学データ解析計算センター長	長谷川政美	教授（予測制御研究系研究主幹）
宮野 悟	東京大学医科学研究所附属ヒトゲノム解析センター教授	樋口 知之	教授（予測制御研究系）
		平野 勝臣	教授（統計基礎研究系研究主幹）

共同利用委員会委員

平成15年6月1日現在

所外委員		所内委員	
内田 直	早稲田大学スポーツ科学部教授	平野 勝臣	教授（統計基礎研究系主幹）
鎌倉 稔成	中央大学理工学部教授	中村 隆	教授（調査実験解析研究系）
岸野 洋久	東京大学大学院農学生命科学研究科教授	中野 純司	教授（統計計算開発センター）
小島 宏	国立社会保障・人口問題研究所国際関係部国際関係部長	伊庭 幸人	助教授（予測制御研究系）
清水 邦夫	慶應義塾大学理工学部教授		

名誉所員・名誉教授

平成15年6月1日現在

名誉所員		名誉教授	
青山博次郎		鈴木 達三	
松下嘉米男		赤池 弘次	
西平 重喜		駒澤 勉	
藤本 熙		鈴木義一郎	
		清水 良一	

昭和19年 6 月	昭和18年12月の学術研究会議の建議に基づき「確率に関する数理およびその応用の研究を掌り並びにその研究の連絡、統一及び促進を図る」ことを目的として、文部省直轄の研究所として創設される。
昭和22年 4 月	附属統計技術員養成所を開設。
5 月	第1研究部（基礎理論）、第2研究部（自然科学に関する統計理論）、第3研究部（社会科学に関する統計理論）に分化。
昭和24年 6 月	文部省設置法の制定により、所轄機関となる。
昭和30年 9 月	第1研究部（基礎理論）、第2研究部（自然・社会科学理論）、第3研究部（オペレーションズ・リサーチ・統計解析理論）に改組されるとともに、9研究室および研究指導普及室の編成からなる研究室制度が採用される。
昭和44年10月	新庁舎を建設。
昭和46年 4 月	第4研究部（情報科学理論）を設置。
昭和48年 4 月	第5研究部（予測・制御理論）を設置。
昭和50年10月	第6研究部（行動に関する統計理論）を設置。
昭和54年11月	情報研究棟を建設。
昭和60年 4 月	国立学校設置法施行令の改正により、「大学における学術研究の発展に資するための大学の共同利用機関として、統計に関する数理およびその応用の研究を行い、かつ、大学の教員その他の者でこれと同一の研究に従事する者に利用させることを目的とする」国立大学共同利用機関に改組・転換される。それにともない6研究部が4研究系（統計基礎、調査実験解析、予測制御、領域統計）へと組織替えが行われ、統計データ解析センタおよび統計教育・情報センターが設置され、附属統計技術員養成所は廃止される。
昭和63年10月	総合研究大学院大学 数物科学研究科 統計科学専攻を設置。
平成元年 6 月	国立学校設置法の改正により、大学共同利用機関となる。
平成 5 年 4 月	企画調整主幹制を設置。
平成 9 年 4 月	附属施設である統計データ解析センターが統計計算開発センターに、統計教育・情報センターが統計科学情報センターに転換された。





統計数理研究所

〒106-8569

東京都南麻布 4 丁目 6 番 7 号

電話 03-3446-1501 (代表)

FAX 03-3443-3552 (庶務課)

URL <http://www.ism.ac.jp/>