

新時代を支える 新しい統計科学を 実例を交えて リアルに学ぶ

理工学部数学科／統計科学研究室

酒折文武 准教授

Fumitake Sakaori

統計学とは非常に幅広い分野で活躍する学問である、ということを知っていただろうか？創薬、保険、金融、そしてスポーツマネジメントの世界でも、いまや業種の垣根を問わず必要とされている学問である。酒折先生は、そんな統計学を活かし、人々や社会へ貢献させていきたいと考える研究者の一人。インタビューでは、統計学とはどんな学問なのか、研究の具体例を交えながらわかりやすくご説明いただいた。



プロ野球も研究の題材に
データを統計的に分析して
多くの情報を引き出す

先生が専門とされているのは統計学。統計学をインターネットで調べると「集団現象を数量的に観察・把握し、統計データを処理・分析する方法に関する学問」とあるが、実際

どんなことを研究されているのだろうか？

「私が研究しているのは大きい枠組みでいえば統計学です。研究内容をいくつか列挙させていただと、『多変量解析』、『ノンパラメトリック法』、『関数データ解析法』、『統計的学習理論』、『統計的因果推論』などです」

専門的な言葉がずらりと並び「？」を浮かべていると、先生はすかさず具体的な例を挙げて、取り組んでいる研究について解説してくれた。

「わかりやすいものを紹介しましょう。いま興味を持って研究しているのがスポーツデータの解析です。あの企業から提供していただいた2008年のプロ野球のすべての行動デ

ータを元にして、さまざまな統計分析を行っています。たとえば、ダルビッシュ投手が時速150kmのストリートをどのコース、どの座標に投げて、バッターを空振りさせたかなど、全選手のすべての行動のデータがあります。そのデータを使って、選手がどのくらい勝利に貢献したか、各プレーの評価などを行うことがで

きます。その結果から選手の活躍度によって給料査定などの判断基準をつくることも可能になります。また、戦略や戦術の決定の他、トレードなどのチームマネージメントにも使えますのでしよう」

現在、スポーツ統計はアメリカをはじめとする海外では研究が進んでおり、実際メジャーリーグでは過去何年かのデータを元にして、次の年はこれくらい打つだろうと予測できるまでになっているとのこと。的率も非常に高いそうだ。

「日本ではスポーツデータの統計解

析がまだまだ未発達であり、統計の研究者が積極的に関与した例は少ないです。国際化社会で日本のスポーツが強くなるためには、これから統計専門家の関与が非常に重要になってくると思います」

時代のニーズに合わせて 統計学も変化している

さまざまな分野で物事を実証するために必要とされる統計学。時代が変われば必要とされる方法も変化していくため、時代のニーズに合わせて統計学は変わってきているという。「従来の統計データは人数が100人ほどいれば多かったのです。しかし、高度情報化社会の今はコンビニのPOSデータなどのように、データが膨大化しています。あれは、自動的にデータセンターに蓄積されていくため、データを取ろうと思わなくても貯まってくるのです。その一方で観測項目の数も爆発的に増えていますので、それに合わせた統計分析が必要になってきています」

膨大なデータを背景とした研究として、ゲノムのデータを扱うための方法も研究対象だ。

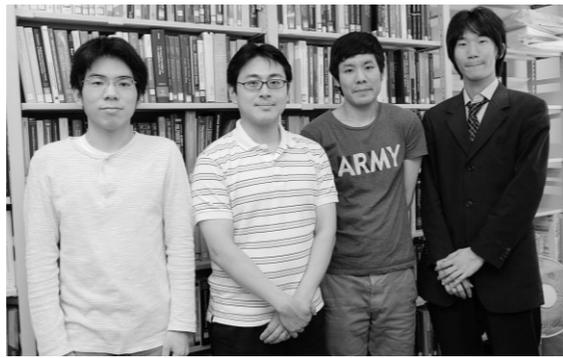
さかおり ふみたけ
1976年、東京都生まれ。1994年、中央大学附属高等学校卒業。1998年、中央大学理工学部数学科卒業。2000年、同大学院理工学研究科数学専攻博士課程前期課程修了。2003年同大学院理工学研究科数学専攻博士課程後期課程修了。2003年、(株)エスミ統計学学術顧問。2004年、立教大学社会学部助手。2007年、中央大学理工学部専任講師を経て2009年、中央大学理工学部准教授となる。専門は「統計科学」。特に「ノンパラメトリック法」「多変量解析法」「関数データ解析法」「統計的因果推論」「計算機統計学」「統計的学習理論」など。応用分野としては「ゲノム解析」「スポーツ解析」「画像処理」「パターン認識」など。著書に『実践ワークショップ Excel徹底活用 統計データ分析基礎編』(秀和システム・共著・2008)、『Excelで学ぶ理論と技術 多変量解析入門』(ソフトバンククリエイティブ・共著・2006)、『実践ワークショップ Excel徹底活用 推測統計入門』(秀和システム・共著・2006)など。論文多数。



先生が執筆した統計学の入門書など。統計学の普及も先生の仕事のひとつである。



実は先生も中央大学理工学部卒業生。先生としてはもちろん、大学の先輩としても皆さんの相談に答えてくれるはずだ。



基礎理論を学びながら、生きたデータの分析も行う

中央大学に赴任して3年目という酒折先生。研究室に所属する学生は、統計学をどのように学んでいるのかを聞いた。

「院生に関しては、自分が興味を持ったテーマを研究してもらい、それに対して時間をかけてとことんつきあう指導方法をとっています。学部生に関しては、いまは基本的な統計理論が書いてある本をじっくりと読んでもらい、基礎力を身につけているところですか。とはいえ、それだけではなく、先日は教育実習で授業に参加できない学生が多かったため、野球のデータを使って、それぞれ自由に分析してもらいました。みんな楽しんでやっていましたよ」

また、先生の研究室では理論だけでなく、生きたデータを解析してみることも重要と考えている。その一環としてデータ解析コンペティションというものにも参加しているそうです。

「昨年は200世帯分の1年間の食卓メニューのデータを分析するコンペでした。朝昼晩3食のメニューは

もちろん、材料まで1年分のデータが与えられて、そこからどんなことが見えるかを分析するのです。また、一昨年は中古車のオークションがテーマでした。車種、走行距離、車の排気量などの仕様、傷の有無などから、落札額がいくらになるかを予測しました」

そのコンペティションに参加する意味は、分析目標を見つけることはもちろん、生きているデータに触れることでデータの処理などの方法やその大変さを実際に体験してもらうこと。また、数学科はグループ学習の機会が少ないため、こうしたシチュエーションやプレゼンテーションの大切さを学んでもらうことを狙っている。

最後に、これから中央大学に入ってくる皆さんに、どんなことを期待しているのか、また、どんな風に大学生活を過ごしてほしいかを聞いた。

「今回は研究について分



趣味はマンドリン。オーケストラと一緒に演奏できるほどの腕前だとか。

「自分の研究でこれまで見えなかったことが、見えるようになる。そして問題解決につながっていくという流れが、いちばんのやりがいですね。また、統計学は社会的にも必要性が高まっている学問です。私の立場としては、統計学を正しく広めていくことがひとつの務めだと思っています」

「自分が研究でこれまで見えなかったことが、見えるようになる。そして問題解決につながっていくという流れが、いちばんのやりがいですね。また、統計学は社会的にも必要性が高まっている学問です。私の立場としては、統計学を正しく広めていくことがひとつの務めだと思っています」

大事です。統計的な概念の意味をきちんと理解し、さらに自分で実行できるだけの数学的素養も必要です。サッカーでもバスケットボールができないと活躍できないように、統計学でも基礎が大事なのです。

また、大学生になったら、今しかできないことをやってほしいですね。私も大学院生のときに1ヶ月以上の間1人で海外を回りました。それまで英語が嫌いでしたが、そうした経験を通じて、英語を話せるようになりたいと強く思うようになりました。それまでの考え方や積極性が変わったのです。教員になってからは1ヶ月以上も海外に行くことはなかなかできませんから、あの体験をしておいて本当に良かったと思っています」



先生が提案した因果分析の方法。薬の効果測定などで活用することができる。

「メジャーリーグでは送りバントはまったく意味がないとされています」

先生は、たとえば野球での送りバントの効果測るなど、スポーツでの応用を考えているという。

「メジャーリーグでは送りバントはまったく意味がないとされています」

先生は、たとえば野球での送りバントの効果測るなど、スポーツでの応用を考えているという。

「自分の研究でこれまで見えなかったことが、見えるようになる。そして問題解決につながっていくという流れが、いちばんのやりがいですね。また、統計学は社会的にも必要性が高まっている学問です。私の立場としては、統計学を正しく広めていくことがひとつの務めだと思っています」

「自分の研究でこれまで見えなかったことが、見えるようになる。そして問題解決につながっていくという流れが、いちばんのやりがいですね。また、統計学は社会的にも必要性が高まっている学問です。私の立場としては、統計学を正しく広めていくことがひとつの務めだと思っています」

「ゲノム解析では遺伝子についてもすごい数のデータを取るんです。人数よりも観測項目の方が多い。たとえば、複数の集団があったとして、塩基配列のどこに差があるのかを調べたいとしても、項目が多すぎて従来の統計解析ではうまくいかないのです。そこで、どのように検証すれば良いかをシミュレーションして、新たな方法論を導き出すのが私たちの仕事です。その方法が見つければ、あとは自由に実務側に使ってもらおうというわけです」

つまり、先生は結果を導き出すというよりは、導き出すための方法論を探し出すという立場だ。続けて先生は、統計的因果推論について紹介してくれた。

「統計的因果推論は、たとえば薬の服用の有無によって病状の改善に影響があるか、といった因果関係を統計的に調べる方法論です。具体的に言うと、熱のある人が解熱剤を飲んだとします。しかし、薬を飲んだから熱が下がったのかどうかはわからないですね。ひよっとして薬を飲

まなくても熱は下がったかもしれない。同じ人が薬を飲んだ場合と飲まなかった場合の両方の結果を知ることができません。かといって、薬を飲んだ人全体、飲まなかった人全体の病状の改善具合を比較するだけでは、たとえば、薬を飲んだ人は重症の人ばかりで、飲まなかった人は軽い薬の効果を測ったとは言えない。そこで、傾向スコアという方法を用いれば、背景にある要因を等しいとした状態で薬の効果を測ることができ

ます。つまり同じ重症度の人が薬を飲むと、どのような効果があるかを統計的に推測することができるようになります。そういった方法論は他の問題にも応用することが可能です」

先生は、たとえば野球での送りバントの効果測るなど、スポーツでの応用を考えているという。

「メジャーリーグでは送りバントはまったく意味がないとされています」

先生は、たとえば野球での送りバントの効果測るなど、スポーツでの応用を考えているという。