



# 計算機科学とその応用

## 手塚 集

学位: 工学博士(東京大学)

専門分野: 計算統計、一様分布論

コンピューターの処理速度の大幅な向上と記憶媒体の大容量化により、不確実性を含んだより現実的なモデル構築のための大量なデータの収集と解析、およびそのモデルを用いた大規模な高精度シミュレーションが可能になっています。そこで主要な役割を果たしているのが、計算機科学の重要な応用分野である「計算統計」です。この分野は、今日、従来の科学技術分野だけでなく、金融・経済といった社会科学や政策決定の分野においても重要な役割を果たすようになっています。

統計学を大きく2つに分けるとすると「モデリング」と「計算(シミュレーション)」になります。確率的な考え方に基いてモデリングをするということが「統計学」に入ってきたのは100年以上も前のイギリスでのこととなりますが、「計算」に確率的な考え方が本格的に導入されたのはそれよりかなり遅れて今から約60年前、第2次世界大戦中のことです。それは統計学ではなく、物理の世界での話でした。アメリカで原子爆弾の開発が行われていたときです。ここでは、確率的な考え方を使って計算する方法を「モンテカルロ法」と呼んでいました。

コンピューターが誕生したのも同じ時期ですが、その普及と進歩に伴って、当然、統計学にもコンピューターが使われるようになっていきます。特に90年代に入るとコンピューターの処理能力、記憶容量が飛躍的に進歩し、またその一方でコンピューターの値段はどんどん安くなり、われわれの日常生活でも当たり前のように使われるようになっています。今では、1人1台コンピューターを持っていると言っても大げさではありません。また、同じく90年代にはインターネットが出現し、世界中から実にさまざまな種類の情報を瞬時に手にすることもできるようになりました。そしてそれらの情報を大量に、自分のコンピューターに保存することもできます。

そういう背景から、大量データの高速処理の技術として統計学に対する期待が実に広範な分野において高まっています。もともと、統計学とは関係の深い医学や遺伝学の分野では、2003年にヒトゲノムプロジェクトが終了し巨大な塩基配列のデータベースが構築されました。現在そのデータベースにインターネットをとおして容易にアクセスすることもできます。地球環境にかかわるさまざまなデータが通信衛星から時々刻々送られてきて、気象予測、地球温暖化の研究などに統計学が使われています。ビジネスの世界でもコンビニのPOSデータが全国から集められ解析され販売・需要予測などに使われています。金融の世界ではこの10年でグローバル化が格段に進み、実に多彩で複雑な金融商品がコンピューター上で作られ、世界中で売買され、巨額のお金が動いています。このような背景から「計算統計」は現在、急速にその重要性を増しています。

それらさまざまな種類のデータを解析する道具として統計学がつかわれているわけですが、例として、統計量の計算の中でも基本中の基本といえる期待値の計算を考えましょう。かつては、正規分布や線形性というような条件をモデルに課していたために、多くの場合解析的あるいは電卓レベルで期待値を計算することができました。逆に言えば、統計量の計算が負担にならないようなモデリングが行われていたということもできます。しかし、今ではコンピューターの高速化と大容量化によりモデリングにも大きな自由度が与えられています。非正規モデル、非線形モデル、非定常モデルなど、大規模な数値計算を前提としたモデリングがさまざまに行われています。

以上に述べたような背景から「計算統計」は現在、急速にその応用分野を広げています。私は、「計算統計」の理論研究にとどまらず、計算機科学の新しい応用分野の開拓や若手研究者の教育・育成に力をいれています。

