

## 「特集 統計環境におけるインターネットの利用」 について

中野純司<sup>1</sup> (オーガナイザー)

統計学の主要な目的は、現実のデータの性質を明らかにすることである。したがって、実際に得られるデータの質と量およびデータを処理するための計算機は、統計学の手法と理論に対して、常に密接な関係を持っている。例えば、計算機の演算能力の急速な増大があつてこそ、計算機集約型的手法(3次元画像を用いるデータ表現法やブートストラップ法など)が開発され、また、その理論が研究されるようになった。

ところが計算機の進歩は、単に数値計算のための“高性能そろばん”を実現することにとどまらず、大きな変革を現代社会全体にもたらしつつある。それは World Wide Web (WWW) を中心とするインターネットの普及である。

インターネットは、ルネサンス期のグーテンベルクの活版印刷術や現代のラジオとテレビに比較できるほどの、情報通信における革命を引き起こしつつある。グーテンベルクの活版印刷術は、書物をごく一部の限られた知識階級のものから、より広い層にまで利用可能にし、宗教改革や科学革命の原因の一つとなつた。ラジオとテレビは、情報の伝達速度と量を、片方向ではあるが飛躍的に増大させ、マス・メディアを形成し、現在の大衆社会の基礎のひとつとなっている。

そして、過去約 10 年の間にインターネットと WWW の急速な普及があつた。WWW は情報の発信も受信も安価で容易な、伝達の即時性を持つ世界規模の新しいマス・メディアである。これまでのマス・メディア(テレビ、ラジオ、新聞、本など)においては、情報の発信を行なうには多くの費用と労力が必要であつた。そのため、情報発信者は組織に限られ、それらが大きな権力を持つことにもなつた。ところが、WWW においては、大組織でも個人でも、情報発信にかかる費用はほとんど同じである。そして、すべての情報発信者は、まったく平等に扱われる。個人がこのような広範囲で強力な情報発信手段を持ったことはこれまでになく、社会のあり方を一変させる可能性が大きい。また、情報取得に関して、これまで図書館が持っていたような役割をインターネットが果たしつつある。

このような状況にあつては、科学も様々な変革と無縁ではいられない。特に現実との接点があるデータという形で本質的に重要な統計学においては、大規模な変化が起こる可能性がある。実際に、POS システムでは計算機およびネットワークを用いてデータが自動的に収集されるようになっており、そのため、これまでになかった多量のデータが蓄積されている。そして、その処理のための手法が、データマイニングの名のもとに開発されてきているのはひとつの例である。

統計学の分野において、インターネットや WWW の影響をもっとも直接的に被るのは、統計

---

<sup>1</sup>統計数理研究所：〒106-8569 東京都港区南麻布 4-6-7

システムやそれを利用する研究教育活動であろう。この数年の間に、WWW を含むインターネットを統計環境の中で利用する試みは、世界中で活発に行なわれてきた。それらは多岐にわたり、データの公開や教育の試みだけを見ても、非常に多くの場所で実現されている。日本国内に限っても、それらの状況を把握することは易しいことではない。また、WWW の開始が 1990 年であることを考えると、その影響がどのようなものであるかを明確に示すことは、現時点ではまだ不可能であろう。

本特集は、日本で現在行なわれているそのような試みの一部を集めたものであり、統計におけるインターネットの利用の可能性の一端を示すことを目的としたものである。ここで、採択されている 7 編の論文を概観してみよう。渡辺・末永・竹内・宿久・山口・浅野論文は、WWW を利用した統計教育環境作成の試みをまとめたものである。WWW による対話的なテキスト、統計学に関する絶版書籍の電子図書館、ソフトウェアの配布などの機能を持つ WWW サイトを構築している。佐藤・大瀧論文では、基礎的な汎用統計グラフィブラリを作成し、それを用いて、WWW 上で疫学的なデータをわかりやすく表示する例を紹介している。飯塚・森・垂水・田中論文では、主成分分析における変数選択手法を実行するためのシステムを WWW から利用できるようにする試みについて述べている。山本・垂水・佐藤論文では、WWW 上で対話的な 3 次元統計グラフを実現するために、VRML (Virtual Reality Modeling Language) を用いることを提案し、その具体的な方法と例を示す。佐藤論文は、季節調整のための手法 Decomp を表計算ソフトウェア EXCEL 上で利用できるようにしたものであり、WWW 上のシステム Web Decomp のもうひとつの利用法を示している。横内・柴田論文は、インターネットにおける統計データの標準形として DandD ルールと XML (eXtensible Markup Language) の利用を提案し、そのために実現したシステムを紹介している。藤原・中野・山本・小林論文は、インターネットと WWW の利用を最初から考慮して設計された新しい統計解析システムについて述べたものである。

これらの論文は、インターネット上の統計環境における新しい“道具”を整える試みと言える。そして、ほとんどのプロジェクトは終了したものではなく、より高い完成度を目指して、現在も研究が進行中である。われわれは、インターネットおよびこれらの新しい“道具”は、今後、統計学の理論や方向を徐々に変化させることになるのではないかと考えている。