

統計学, コンピュータ, そして統計教育

日本の大学における統計教育は今まで文科系では主に経済学部(時には経営学部)また理科系においては数学科をはじめ多くの学科で教えられています。それにともない統計学者の多くはそういう学部、学科に属している事が多いようです。さて統計学というものの性質上それを応用する部門とともにあり、常に現実の世界との接点を持っている事が望ましいと言えますが実際にはなかなか難しいようです。どうしても自分の好きな方向や、学術(?)論文の書き易い方向へと流れがちとなります。そのような傾向の極にあったのが実は1960年代の数理統計学でした。不必要的抽象化と並び現実には全く存在し得ないような条件の下での理論が数多く展開されていたのもこの時代でした。しかしながら数学的に意味のある結果を無理に出そうとすればするほど、その理論が無意味な抽象化に走るのは世の常です。ただし、そのような事になってしまった一つの理由は、世の中が学問の世界のように簡単ではないからです。実際各分野で使われてきている多くの統計手法の有効性(無効性?)を数学的に証明したり、問題に即した新しい手法を開発し、その性質までも理論的に論ずる事は少なくとも現在の数理統計学、確率論、さらには数学のレヴェルをもってしては殆ど不可能な事が多いからです。それ故理論的に何か良いとされる結論を出すための無意味な抽象化が行われ、そして全く役に立たない無味乾燥な体系が残されてきました。しかしながら幸か不幸か、役に立たない統計学は統計手法としてはもはや存在できない訳で、その60年代の過度なる抽象化に対する反省が70年代に入り現実に役に立つ統計学への回帰としてアメリカを中心に始まりました。

この回帰は現実からの要請(特に医学、薬学か

らの要請が強かったが)もさる事ながらコンピュータの飛躍的な発展に伴うモンテカルロ法の普及も見逃す事ができません。上でも述べたように数学的に解く事の出来ない問題であってもコンピュータによる数値実験を繰り返し行う事によりある程度の事が示されます。70年代に入り盛んになってきたノン-パラメトリック統計学(基本となる分布を仮定しない、実際には正規分布を仮定せずにいろいろな推測を行う方法)や医学統計の世界で盛んになってきたセンサーされたデータ(癌患者の治療効果等を分析する際、観察期間の途中で患者が他の病院に移ったり、当該の病気以外の原因で死亡したりする事により完全な形のデータが得られない事)を扱うための手法等がその代表的な例といえます。勿論数学的な議論を完全に排除するといった事ではありませんが、それ以上にコンピュータによる数値実験が重きをなしてきています。我々にとっていちばん大切なものは空虚な理論ではなく現実に役に立つ手法であるという、プラグマチックな考えが支配的になってきています。このように書いてくると統計学は今完全にアカデミズムの世界から離脱し実学の世界へと邁進しているかの印象を与えてしまいますが、実は反対に健全な方向にきていると筆者は感じています。それはひとたび答が判れば、そこに到達すべく多くの(理論的)手法が作られてくるからです。先が見えていれば、どのような道具を使えば目的地にたどり着けるかを考える事は比較的容易となるでしょう。その結果例えれば、今まで非常に難解かつ抽象的で誰もが敬遠していたような確率論の一分野も統計理論の中に応用される事と成り、ひいてはその分野にも大きな影響を与える事となるのです。

一橋大学経済学部教授 高橋一

さてコンピュータを理論研究の道具として見る事と並んで、それをデータ解析により積極的に活用していく方法も過去20年間に大きく進歩してきています。最近多く出回っている各種の統計学ソフトがそれです。ロータス1-2-3から始まって、SAS, S, Minitab, TSP, Shazam, … 等枚挙に暇がないでしょう。またパソコンが会社、学校、そして各家庭にまで普及し始めそれらの統計ソフトが簡単に利用され始めているようです。特に多変量分析、時系列分析なるものはひと昔前までは大学ないしは大学院の上級統計学の授業で論じられる一教科にすぎなかつたのですが、今では大学や研究所における研究者は勿論、証券会社、銀行に勤め始めた文化系出身の若い社員までがそれらを使い金融分析を行っています。そうなると若干気になるのがそれらの統計手法が果たして正しく使われているか否かでしょう。これは別にビジネスの世界だけの問題ではありません。実際にデータを取って研究する全ての科学においても問題となってきます。そのなかでも最も切実なのは医学の分野にあるかも知れません。少々古い話になりますが、1978年にアメリカのミシガン大学統計学部で当大学に提出された過去10年間(?)程のドクター論文に使われていた統計学手法で誤った使い方をされたものの割合を調査した事がありました。結果は惨憺たるもので確か、正確な数字は発表されなかったと記憶しています。同様の調査は数年後にハーヴァード大学のモステラー教授を中心となってNew England Journal of Medicineというアメリカでは最も権威のある医学雑誌の一つに掲載された論文について行っています。こちらの方は当該雑誌に調査結果が載っていますから、興味のある方はご覧になって下さい。ただ問題は

今は統計ソフトが10年前よりも数段使いやすくなっていると考へられる事です。統計学教育に今要求されるものは、数々の統計ソフトを使いこなす技術面をいかに効率的に教えるかと同時に一つ一つの技法の限界、適用範囲をどの様に整理してユーザーに示すかにあります。しかしながら後者については非常に重要であるにも拘らず、現在の所殆ど不可能と言えるくらい難しいと言えます。例えば時系列分析のソフトを使うには数時間の講習を受ければ十分でしょう。しかし現在の教育方法では時系列分析が何かを理解するには少なくとも数年間、統計学の(退屈な)授業を受ける必要があるのです。非常に優れた統計エキスパートシステムができればこの様な問題はすぐに解決するでしょうが、見通しは決して甘くはありません。それよりは大学における統計教育を今大きく代える事がいちばん必要なかも知れません。車の動く理屈は知らなくとも、車を正しく運転できる様な方法が果たして統計学にも存在するものかを今考える必要があるようです。

