

道標

文部省認定社会通信教育機関誌

平成 6 年

1994

NO. 439

10

〈特集〉 修了生からのアドバイス

Le Semeur *L'angélus* J. F. MILLET



臨床分野でのパソコンによる 統計解析について

高 木 廣 文

統計学の授業では、今でも電卓などによって、平均値を求めたり分散や相関係数などを学生に計算させているようである。これだけパソコンなどが普及しているのであるから、計算はパソコンでやればよいのに、電卓を使った手計算などは時代遅れもはなはだしい、との考えも多いものと思う。しかし、これまでの授業経験からいえば、統計学の理論のみをいくら教えても、各種の解析方法が簡単に身につくものではない。例題にしたがって実際に計算をやってみることが、最も学習効率のよい方法であることは確かだと思う。もっとも確実に統計学の手法を理解し、自分のものとするには、学んだ方法の計算を行うためのプログラムを作成することであろう。しかし、そのためにはプログラム言語を、少なくとも一つ以上知っていなければならないので、そのための教育もさらに必要になるのだが。

統計学を学ぶ場合に、各種の方法を着実に理解するために必要な手計算ではあるが、実際のデータを統計学的に処理する場合、電卓で計算を行うのは止めた方がよい。例えば、100人について10個の変数のデータがあるとしよう。変数間の全ての単相関係数を求めるとすると、少なくとも10個の平均値と10個の分散、45個

の共分散を求めねばならない。普通の間であれば、キーの押し間違いによるデータの入力ミスによる計算間違いを何度も繰り返したあげく、諦めるところであろう。または、自分では完璧に正しく計算できていると思っても、実際には間違いだらけの結果であつたりもする。

実際に、新薬の効果をそれまでの薬剤と比較する、患者のQOL（生活の質）の調査を行う、またはQOLを測定するための質問票を開発する、肺癌の手術後の生存率に影響する要因を解析する、というような場合、とても手計算では間に合わないし、人間は必ず間違えるので、機械に任せるに限る。

20年程前までは、ちょっとした調査データの集計でも、大学の大型計算機で、しかも集計や統計解析用のプログラムをFORTRAN言語などで自作しなければならなかった。バグやエラーの対策などをしてしていると、1日などあつというまに過ぎてしまい、単純集計だけで1週間もかかったこともあった。その後、大型計算機用の統計学プログラムパッケージが登場した。社会科学系のデータ解析用パッケージとしてSPSSが紹介されたと

きは、あまりに簡単にデータの集計や解析が行えるので、極めて衝撃的であった。その後、SASが紹介され、パソコンやワークステーション上で、広く使われるようになってきた。このような経緯から、現在では、もはや大型計算機を使う必要はないし、通常の集計や解析ならば、プログラムも自作する必要はなくなった。当時のことを考えると、実に便利な世の中になったはずである。

我々も、調査データの集計・解析のための統計学パッケージHALBAU（はるぼう）を開発してきた。より簡単に、誰でもパソコンさえあれば、簡単な集計から、より高度な解析が行えるようにとの考えからである。例えば、肺癌患者の予後の解析を行いたいのであれば、生命表解析や生存解析が簡単に行えるし、種々の要因の死亡に対するリスクなどの検計が容易に行える。また、患者のQOLなどの調査の集計や解析とともに、QOL測定のための質問紙による尺度構成を行うための信頼性解析なども、極めて容易に行うことができる。

このように、統計学の実践の場では、パソコンの利用は当然のことであるが、いくつかの問題もある。最も強調すべきことは、さまざまな統計学ソフトを簡単に使用できる分だけ、統計学の誤用を助長するかもしれないという点にあるだろう。どのような解析ソフトを用いるにせ

よ、通常はそれらに含まれる手法は極めて多数であり、あまりにも簡単に様々な分析ができるはずである。そのため、利用する側が必ずしも最適な方法を用いるという保証はない。極端な場合、統計学の理論など何も知らなくても、画面に表示されるメニューにそってキーを押していけば、何らかの解析結果が得られることだろう。しかし、その解析方法がそのデータに相応しいものであるかは、全くわからない。さらに、そのような状況で得られた結果の解釈が正しく行えるか、疑問だらけである。

医学データの解析結果は、ときによっては新薬の承認、疾病予防、健康維持、厚生行政の指針など、実際の世の中ですぐに普及する可能性をもっている場合も多い。したがって、あまりに誤解に満ちた解釈を、統計学の名の下に行う訳にはいかない。結局のところ、データ解析を行うためには、統計学の各手法の具体的な計算方法を覚える必要はない。しかし、解析に用いる方法のもつ意味を正しく把握し、なぜその方法がデータに適したものなのかを、正確に理解しておくことは常に必要である。欲をいえば、より幅広い一般教養および専門的知識を身に付け、統計学的分析の結果を正しく読み、誤りのない結論を導ける実力を身につけることが望ましいといえよう。

（文部省統計数理研究所 助教授）