

## データベースについて一言

齊藤昌宏\*

本会の1982年度夏季セミナーと第25回シンポジウムの記録を拝見して、たいへん参考になりました。特にデータベースとコンピュータに対する考え方の多様さは、門外漢でありながら利用はしたい小生にとって、面白く読ませてもらいました。

ところで先日、同様に門外漢でありながら利用はしたい人（これは小生がそう思っているだけかも知れませんが）とコンピュータの利用とデータの取り方について議論しました。要は大型コンピュータ派とマイコン派の議論です。大型コンピュータは大量のデータを高速処理できる、一通りの統計処理はすべて組み込まれているのでプログラムを作る必要がないなどの利点があります。そのかわりデータの取り方は処理システムを予想して、規格化して取り、結果を見てから理由は考えることとなります。これに対してマイコン派は少量データの低速処理で、プログラムは自分で作ることとなります。しかし、データの取り方は自由で、一段階ごとにデータを処理しながら考えることができます。簡単にいえば結果をにらみながらデータを取るか、データをにらみながら結果をだすかということになるでしょう。これは研究者、研究対象によって使い分けることであってどちらが良いという問題ではないようです。

しかしデータベースの場合には違うのではないのでしょうか。データベースにはマイコン派的考え方の入る余地はないとみてよいと思います。人によってデータを取る背景は異なっています。同じ森林を調査しても人によって出てくるデータは異なることが多いのではないかと思えます。またデータベースを作るためにデータを取る人はいません。普通は研究に必要なデータを取って当面必要な解析を終えたデータを公開するのではないのでしょうか。さらにいえば自分が研究を始めるときに必要にして充分な項目がそろっているデータの集積にであう経験はほとんどないと考えてよいでしょう。データベースの利用のしかたとしては、自分が研究を始めるときにこれまでのデータの集積あたりをつけて研究の方向を絞るとき。反対に自分の取ったデータから手作りで何かを見つけたならば、そのときデータベースの中で利用できるものを用いて、他でも同じことがいえるかどうか調べるときが大部分でしょう。つまり自分の研究の周辺を埋めるための利用のしかたが大半ではないのでしょうか。

このように考えると、データベースを作るのはかなり簡単になります。どんな研究にも必要にして充分なデータはない。しかし多様なデータの集積にしておけばその中から利用できるものを使えばよいわけです。文献の引用と同じで、データベースを利用し、データをいじった後の結果はすべて利用した人の責任でよいわけです。人のデータが使えないのであれば使わなければよいし、人に使っても

\* 新潟大学農学部

らいたくないデータはデータベースに入れなければよいわけです。データベースにデータを提供する場合にも、肉を食べたあとの骨を出すといった程度の考え方で提供すれば気が楽です。もちろん骨からでも調味料を工夫すれば立派なスープがとれます。極端な話データベースの最低条件としてどんな項目が必要で、どの程度の精度が必要かなどの議論は無用であり、データであればなんでも出せばよろしい。ファイルの形式も異なっても当分はしかたがないのではないのでしょうか。ただそれぞれのデータの概要一覧を一冊まとめておけばすむことです。要はデータベースを作るという合意があればあとは枝葉のことです。林業統計研究会でデータベースを作るならば、会員各自が利用してもらってもよいデータを提供し、物理的にデータベースを保管する場所と、それをどう配布するかの技術的問題を議論すればよいのではないかと思います。