

統計を知らないと

村田 勝 敬

■ プロローグ

ある医学生は臨床実習中の外科医局で「術後乳癌患者 204 名の術前データから乳癌組織の診断式を作成し、新規乳癌患者でその式を確認すると正答率は 75%でした」と発表した。この結果を聞いたひとりの先生は「それは凄い。外科学雑誌に発表するか!」と賞賛した。暫くして、医局の別の先生が「ところで何人用いて検証したのか」と質問した。医学生は誇らしげに「4 人です」と答えた。その瞬間、医局は静まりかえった。

近年、内閣支持率とか政党支持率の動向がテレビのニュースで 1 月毎に流される。無作為抽出された 2,000 名（実際の回答者は 1,200 名前後）の率が一人歩きしているように感じる人も多いのではないか。すなわち、日本国民の 0.001%前後から算出される数値がなぜ国民の代表値となり得るのかと疑問を懐くのである。一方、高校や予備校の“統計を駆使した偏差値”判定に従って受験・合格した大学生の多くが「統計は難しくて判らない」と呟いている。親たちはその傍らで「大学に合格しても統計を知らないのか?」と失笑する。

■ 医学教育における統計

医学・歯学教育の在り方に関する調査研究協力者会議が開催され、『21 世紀における医学・歯学教育の改善方策について—学部教育の再構築のために—』が平成 13 年 3 月 27 日にまとめられた。この会議では『準備教育モデル・コア・カリキュラム—教育内容ガイドライン—』も提示された。ここでは良き医療人を目指す医学・歯学教育の前提として身につけておくべき基本的教養として、①物理現象と物質の科学、②生命現象の科学、③情報の科学、④人の行動と心理の 4 つが掲げられた。このうちの「情報の科学」の一般教育目標 GIO には、情報リテラシー（パソコンの取り扱いを学び、パソコンを利用してネットワークに接続することを学び、パソコン上の各種アプリケーションソフトの利用の仕方を学ぶ）、統計の基礎（確率論的なものを見方を理解し、確率変数とその分布、統計的推測の原理と方法を理解する）、統計手法の適用（医学生物学でよく遭遇する標本に、統計手法を適用するとき生じる問題点や統計パッ

テージの利用を含めた具体的な扱い方を習得する)が含まれる。そして、統計手法の適用における到達目標 SBOs では、①正規性の検定、②対応のない（対応のある）2 群間の平均値の差の検定、③Mann-Whitney の U 検定、④ χ^2 検定、⑤一元配置および二元配置分散分析、⑥Kruskal-Wallis 検定、⑦回帰分析と相関分析、などを実行できるようになることが求められている。

■ 教育専用の統計ソフト

1970 年代後半、大型計算機センターで STATPACK や SPSS などの統計ソフトを使用することができたが、電卓を叩いて計算する方が一般的であった。当時は、FORTRAN の全盛期であり、数値解析のプログラム書も幾つかあった。パーソナルコンピュータの出現後、BASIC による統計プログラム集が 1980 年代初頭より現れ、『マイコンによる医療統計処理』、『パソコン統計解析ハンドブック』、『多変量解析ハンドブック』などの本が相次いで出版された。その後、Macintosh や Windows に世の趨勢が移る中で SAS、Stat View、S-PLUS などの統計ソフトが台頭し、また優れた無料ソフト（例えば R や SPSS）も多数出現するようになった。

1999 年の日本計算機統計学会のホームページでオンラインパネルディスカッションが行われた。その中のパネラー A の基調提案の中の 1 つに「教育専用の統計ソフトを充実させよう」というのがあった。私はこの考えに賛成の意を表明したが、教育用ソフトに関する意見の中には「限られた時間内で、社会



に役に立たないソフトの使い方を教えて意味があるか」などの批判も飛び出した。

■ 統計がわかる？

教育用統計ソフトをどのように定義するかにより、その有用性を評価する尺度は異なる。例えば、統計に熟知している人にとって、2群のデータを入力して得られる有意差検定の結果に、①2群の平均値、標準偏差、サンプル数、②母分散の検定値 (F 値)、③ t 値と自由度、④確率 P 値が表示されておれば十分であろう。しかしながら、初心者が異なるデータを幾度となく入力し、これらの解析結果 (数値) を眺めても“比較”の概念を理解することは容易でない。また、対のデータを入力して得られる解析結果が相関係数とその確率 P 値だとしたら、同様に初心者にとって“関係”の概念は判らずに終わってしまうかもしれない。なぜなら、多くの教師は統計の意味を学生に図示しながら教えるが、実際に学生がコンピュータで解析する時には数値結果しか表示されないことが多いからである。

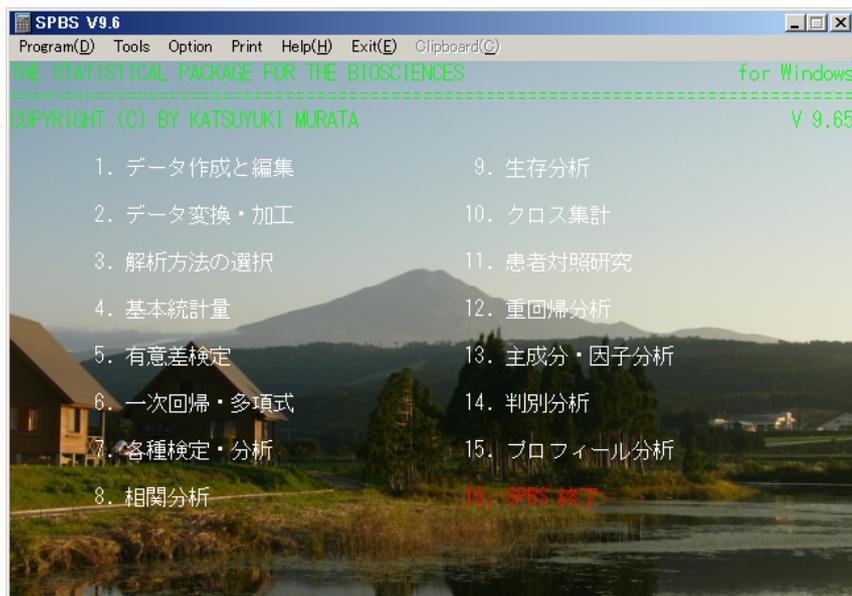
上述の“比較”の例で、統計が少し判ってくると学生は「有意差の有無」について述べることができるようになる。しかし、結論として「どちらの群が高いか」については (教師から質問されるまで) 言及しないことが多い。同様に、“関係”の例では、「有意な相関の有無」について述べることもできても、“正負”いずれの関係があるのか結論が出せない。その上、データが正規分布型であるか否かは全く議論の外である。初心者における統計の理解の向上に

資することが教育用ソフトの主たる目標であるならば、集めてきた数値データをもっと直感に訴える方法で提示する必要があるだろうし、使い方も簡便でなければならない。統計ソフトの使い方 (解析手順の記述) に多くの時間を割かねばならない状況は、どんなに評判の高いソフトであろうとも、統計の理解とはほど遠い。また、統計ソフトを上手く扱えることと、統計の意味を理解していることとは全く別次元である。

■ エピローグ

自然科学や経済学における証拠の多くは統計 (確率) に基づいて導出される。統計はデータのバラツキあるいは偶然 (random error) が許容できるかどうかを検証する道具であり、95%信頼区間値ないし (帰無仮説の生起) 確率 P 値で判断する。特に、我々が普段使い慣れている比率 (パーセント) は、分母の情報がない場合、“くせ者”となる。例えば、分子3、分母4の比率の点推定値は $3/4=75%$ 、その95%信頼区間は $19.4\sim99.4\%$ であり、信頼に足る範囲があまりに広すぎて、逆に信頼できないと考える。一方、無作為抽出調査の対象者1,200名の内閣支持率が75%であったならば、その信頼区間は $72.4\sim77.4\%$ となり、かなりの精度が保証されよう (もっとも、“無作為抽出”操作が完璧であるとの保証が得られた場合に限られる)。統計を馬鹿にする医学生は、根拠に基づく医療 (EBM) を行えず、思い付きで話す“ホラ吹き”医者として将来呼ばれるやもしれぬ。

「秋大生活のひろば」No. 143 (2013年6月刊)



SPBS ダウンロードアドレス

<http://www.med.akita-u.med.ac.jp/~eisei/link.html>