

生物統計家の数学との関わり方について

土居正明（京都大学医学研究科社会健康医学系専攻医療統計学分野）

まず、私の経歴を大まかにまとめさせていただきます。農学部を卒業後、修士は数理科学研究科で表現論という純粋数学を学びました。その後、主に生物統計を学んだ上で MPH（Master of Public Health）を取得した後に企業に就職し、医薬品および医療機器の開発の統計解析職に 11 年と 2 か月従事しました。企業に在職中、理工学研究科数学専攻の博士課程に進学し、ベイズ統計に関する研究で博士（理学）を取得し、2018 年 6 月より現職となります。このように私が統計を本格的に勉強し始めたのは修士課程修了後であり、それなりに紆余曲折を経て、現在に至っています。生物統計の素晴らしい点の一つは、私のような人間を含めた様々な人が受け入れてもらえ、途中から参加できる、という柔軟性および多様性ではないかと感じています。さて一方で、人によっては非常に大きな参入障壁となりかねない、数学との関わり方について、1 人の「比較的数学の好きな生物統計家」として、少し考えてみたいと思います。

まず、大前提として、生物統計家にとって数学とは「対象とする生命現象の定量化もしくは判断など、各自の目的達成のために必要なレベルで用いる道具」かと思います。従いまして、他に学ぶべきものを差し置いて数学の勉強に追われすぎる、などのことは、多くの場合本末転倒かと思います。一方、特に周りの統計家に対して、「最低限の数学は理解しておいて欲しい」と思われる方も多いのではないのでしょうか？さて、この「生物統計家にとって、最低限の数学とは何か？」という問いも、なかなか難しいものだと思います。と申しますのも、先に述べました生物統計家の多様性を考えますと、なかなか一律に「ここまでは全員知っていて欲しい」という線を引くのは難しいと思うからです。要求を厳しくしすぎると、参入できる方をかなり制限することとなり、またその後の学習内容も大幅に偏ってしまいます。とはいえ昨今の生物統計で用いられている手法を支えている理論の難解さを踏まえますと、「このくらいは最低限」の「このくらい」のレベルを年々上げられている方も多いのでは、と思います。

少し具体的に考えてみましょう。たとえば、（独立同分布の場合の）大数の法則や中心極限定理について、①名前すら知らない②名前は知っているが仮定や結論は全く知らない③仮定はあまり把握していないが、結論は大体知っている（シミュレーションベース等でイメージがわく）④仮定と結論を正確に知っているが証明は知らない。ただし結果を適切に使うことはできる⑤証明まで正しく理解している、くらいに分けられるのではないかと思います。①②では統計家として少し残念な気がしますが、⑤を広く求めるのは現実的ではないでしょう。④も、大数の強法則の主張の理解に測度論の知識が必要になってきたり、など数学を専門的に勉強されていない方には難しい問題が出てきますので、現実的には③と④の間くらいを「最低限」と考えられる方が多いものではないでしょうか。

次に、「全員」からもう少し対象を絞り、「比較的数理に興味のある生物統計家」に限定し

ましょう。現状、ある程度理論寄りの研究論文を書かれる場合でさえ、⑤まで必要になる方はそこまで多くはないのではないのでしょうか。また、現在のところ③や⑤を目指している教材はある程度豊富にみられるように思います。一方で、③のレベルでは、論文を読んでも難し過ぎて実務に応用しづらかったり、研究テーマに設定できなかったり、などのことも多いかと思えます。従いまして、様々なテーマにおいて④のレベルに到達することができれば、(一定の前提のもとで)内容を理解できる論文の数が増え、実務に取り入れられる手法が増えたり、研究の幅も広がり、有益なのではないかと考えています。たとえば、生存時間解析の背後に控えているマルチンゲールや経験過程の理論などは、現状少なくとも日本語の書籍などは、③か⑤のレベルを目指すものがほとんどのように思います。このあたりについて④のレベルを目指す「コンパクト」な資料を増やしていくことができれば、実務や研究に用いられる道具の選択肢が増えるのではないかと思います。

さて、そもそも数学という学問は、「論理的に道筋をたどれば、誰にでも主張の真偽の判断が可能」なものであり、それは数学の大きな魅力のはずなのですが、その「論理的に道筋をたどる」という行為に時間や労力がかかりすぎることも多いように思います。従いまして、たくさん勉強すればするほど、その後得られた正しい主張と証明に対して、周りの人が理解できない、ということになりかねません。大変悩ましいところですので、「よい感じの飛ばし方」をしつつ、応用に困らない範囲で『理解』する方法を模索できれば、と思います。一方、学習者の心構えとして先日学生たちに話したのは、「数理的な側面の強い講義が、聞いたその場で理解できない、というのは得意な人でも大変よくあることなので、『自分は頭が悪い』などと思ったり、落ち込んだりする必要はありません。一方、理解できていない、というのは事実なので、歩みを止めず勉強し続けて欲しい」ということです。このような、「難しく、すぐには理解できないものとの向き合い方」や、「分からないときにどのように試行錯誤するか」という技術は、私自身が学生時代に数学の学習を通じて身につけた(辛抱強い師匠たちに時間をかけて養っていただいた)重要なスキルだと感じており、学生たちにもぜひ一定のところまでは身につけて卒業してもらいたい、と考えております。

やや話がそれましたが、本稿の結論といたしましては、「高度だが、現在用いられている手法の背後にある数理的テーマに対して、『厳密に全ての証明は理解できない』ものの、『適切に使える状態になれる』ことを支援するような『コンパクト』な教材」が増えることが、今後の日本の生物統計家の実務および研究を、より一層推し進めるのではないかと、いうものです。私自身も「くまなく証明をつけて、全て説明したい」という欲求を時に抑え、「ほどよい塩梅」を探していければ、と思います。周りの統計家の皆様と協力しつつ、自分にできることを検討していきたいと考えております。