

Paul R. Rosenbaum 著・阿部 貴行、岩崎 学 訳

『ローゼンバウム 統計的因果推論入門
— 観察研究とランダム化実験 —』

共立出版、2021年5月刊、A5、416ページ、本体4,800円+税

大学は卒業したが「因果推論」(Causal Inference)に全く関心がないと云う読者が果たしているだろうか？ 大学には理科系・文科系、実に様々な学問分野が存在するが、語学や体育はともかく、「因果関係」に全く関係のない授業はかなり少ないのではないだろうか？ さて、因果関係に関心があるとしたら「因果関係は統計学では証明できない」というよく聞く主張に対し、「因果関係は統計学でしか証明できない」というハーバード大学統計学科を創設した故モステラー (F. Mosteller) 教授の主張をどう評価するだろうか？ この言葉から本書の「まえがき」は始まっているのである。全体を通じてモステラー教授が主張したかった内容を、より分かりやすい言葉でなるべく多くの読者に解説することが本書の目標ではないかと思われる。

まずは本書の内容を要約しておこう。第1部「ランダム化実験」では統計的因果推論を巡る事項、Fisherの帰無仮説と検定のロジック、ランダム化実験の方法を巡る基本事項をまとめている。第1章ランダム化臨床試験、第2章構造、第3章ランダム化実験による因果推論、第4章非合理性とポリオ、という章からなっている。いずれもランダム化実験に関する具体的な例を用いた平易な説明であるが、特にR.A. Fisherによる正確検定の意味やシンプソンのパラドックスなどを丁寧に説明している。第II部「観察研究」は古典的なランダム化実験からより現代的な生物統計、臨床試験における統計的因果推論を巡る諸問題が内容である。第5章観察研究

と実験研究の間、第6章自然実験、第7章理論の精緻化、第8章準実験の手法、第9章バイアスの感度、第10章デザイン感度、第11章マッチングの技法、第12章気質によるバイアス、第13章インスツルメント、最後に第14章の結論、という章の構成である。

普段は生物統計学や臨床試験などとはあまり縁がない本評者にとっても、本書により統計的因果推論を巡る様々な論点について理解が深まり、重要な論点を学ぶことができた。ここで本書を一読した感想を幾つか述べておこう。

第1に本書の特徴として前半、後半を通じて随所に生物統計学の展開、臨床試験、統計的因果推論の重要性を巡り様々な具体例を踏まえた説明がなされている。例えばProCESS試験、1954年に40万人以上の米国幼児が参加したSALKポリオワクチンの臨床試験など、専門家には良く知られている事例ではあろうが、生物統計や薬効・薬害などの門外漢にもその重要性がかなり理解できると思われた。

第2に第I部の内容、および第5章「実験研究と観察研修」の話題、表5.1～表5.4を巡る基礎事項ぐらいの内容までは門外漢にとってもかなり分かりやすい。統計学では定番の2×2分割表を用いたシンプソンのパラドックスや誤った分割表の利用例については特に熟読する必要があるだろう。現代的な統計的因果推論はD. Rubinによる反実仮想(counter-factual)モデルの提唱に始まるとされているが、そのアイデアは実は仮説検定論を建設したJ. Neymanによる農事

試験まで遡るらしい。マッチング、傾向スコア、多重比較などを巡る議論はフィッシャーやネイマンに始まる近代的な統計学の発想を理解するためには是非にも学ぶべき内容と思われる。

これに対して第Ⅱ部の後半の内容は率直に述べるとそれほどすぐに理解できるとは言い難い。自然実験、準実験などは最近では経済学や政治学などの応用分野でも必須の考え方になりつつあるようであるが、感度分析やデザイン感度などの内容は著者の研究に基づくと思われるがやや非標準的と思われた。実務的にはランダム化比較実験の実施が可能でない場合が多いと思われるが、その場合にどうしたらよいか、本書からは必ずしもまだ明確ではなさそうとの印象である。

第3には、統計的因果推論の理論についてほとんど言葉による説明が多く、数式などによる正確な説明はわずかである。これは著者がより多くの読者を想定して本書を書いたので意図的だろう。多くの読者にとって読みやすいかもしれないが、その反面、「無視可能な処置の割り付け」(ignorable treatment assignment)など何を意味しているのかなど分かりづらい所も散見される。特に後半のインストゥルメント(Instruments)の説明は日本語の翻訳のせいかもしれないが、「exclusion restrictions」を巡る内容や説明を理解するのはかなり困難である。操作変数法(instrumental variables method)はもともとeconometrics(計量経済学)では多用されている方法であるが、元々の説明は外生性(exogeneity)などを含めてより明確に定義されているのではないだろうか。また弱操作変数(weak instruments)と強操作変数、さらには多操作変数(many instruments)の問題なども重要ではないかと思われる。操作変数の定義にもよるが、「ランダム化比較実験」におけるランダム割付、傾向スコアによる方法などもある種の操作変数法の利用と見なすことができるのではないかと、とは単なる感想である。

最後に述べた論点と関係するが、2021年ノーベル経済学賞はデービッド・カード、ヨシユア・アングリスト、グイド・インベンスの3名が受賞、「カード氏は最低賃金の引き上げが雇用に与える影響を調査、最低賃金を上げても雇用が減るとはかぎらないことを突き止め、アングリスト氏は所得と教育期間の関連性を検証、インベンス氏は経済現象を独自の手法で解析した」、とのことである。この中のG. Imbens氏は統計的因果推論を開拓したD. Rubin氏と共同で教科書「Causal Inference in Statistics, Social and Biomedical Sciences」(2015, Springer)を書いている。そういえば2019年ノーベル経済学賞はマイケル・クレマー、エステル・デュフロ、アビジット・バナジーの開発経済学者3氏、その授賞理由は「世界の貧困軽減に向けたフィールド実験」であるが、ここで実験とはむしろ「ランダム化比較実験」を意味していたことに気が付く読者もいるのではないだろうか。むしろノーベル経済学賞については色々な評価がありうるが、少なくとも多くの関係者が「ランダム化比較実験」や「統計的因果推論」に関係する応用研究を高く評価している。こうした学界での動きは近年の統計学、データサイエンスの興隆と密接に関係しているのである。

本書を一読して統計的因果推論について優れた研究業績を重ねている著者から生まれた書籍という感想を持った。さて、こうした本書で説明されている統計的発想を様々な学問・実務分野において「因果関係」を分析している研究者、政策担当者、あるいは実務家は日本でも十分に活躍しているのだろうか？ 統計データ・統計学・データサイエンスに関心のある読者が「ランダム化比較実験と統計的因果推論」などを考察するきっかけとなればと考え、本書の一読を推奨する。

評者：国友 直人・くにとも なおと
(統計数理研究所特任教授、東京大学名誉教授)