

序

私たちは2021年に、『医療統計、データ解析しながらいつの間にか基本が身につく本～Stataを使ってやさしく解説』を上梓しました。おかげさまで多くの方にご好評をいただき、今なお多くの読者の方からご感想や励ましのお言葉を頂戴しております。

上記の書籍では、Stataを用いたやさしい解説を通して、初学者の方でも医療統計の基本的な解析手法を学べるよう工夫を凝らしました。読者の皆様が「気づけば統計解析の基礎体力が身についていた」と感じてくださるような構成をめざした本です。その点において、私たちはこの書籍が一定の役割を果たせたと自負しております。

しかし、現実の医療統計の現場においては、統計解析以上に重要な行程があります。それが「データクリーニング」です。

研究者は、統計解析を行う際、多様な形式・構造のデータと向き合い、それらを解析可能な形に整えるという作業を繰り返しています。実際に、統計解析に要する時間のうち多くを占めるのが、この前段階の「データをきれいにする」プロセスです。

これは地味でありながら、解析の信頼性を左右する、きわめて本質的な作業でもあります。

私たちは、前著を執筆した当時から、この「データクリーニングの基本から丁寧に学べる本」が必要であると感じてきました。そして今回、その思いを形にすることができました。

今回も前著と同様、Stataを中心に解説していますが、データクリーニングの基本的な考え方は、RやPythonなどほかの解析ツールでも共通です。そのため本書では、Stataに加えてRのコード例も羊土社ホームページから利用できるようにしました。どのツールを使う場合でも、実務に根ざした知識と丁寧な手順の積み重ねこそが、信頼に足る解析の礎となります。

本書が、読者の皆様にとってその一助となることを心から願っております。

2026年2月

道端伸明，麻生将太郎，藤雄木亨真