

# はじめに

## (改訂の序)

「統計は難しい……」

「どの検定を使えばいいかわからない」

「平均値はわかるけれど、分散や標準偏差と言われると自信がない」

そんな悩みに応える1冊として、本書初版を2019年に発行しました。実はこれらはすべて、私自身が抱えていた悩み、そのものです。

私は大学院修了後、あるうことが製薬会社の統計解析部門に配属されました。学生時代は統計の講義で寝てばかり。試験直前に詰め込んでなんとかパスしても、終われば何も残っていない。T検定の使い方すら怪しかった私が、統計を仕事にするなど夢にも思っていませんでした。

しかし、仕事を始めて数年後、統計に関して、2つのことが理解できたときに、ブレイクスルーが起きました。

その2つというのは、

- 母集団と標本の関係
- 誤差の概念

です。

「なぜこのデータを解析するのか？」という目的をクリアにし、データには必ず「誤差」があるからこそ統計が必要なのだと気づいたとき、景色が一気に変わりました。この気づきから、本書初版を執筆し、同じ悩みをもつ多くの読者の方に、ご活用いただくことができました。

さて、初版の刊行から6年半が経過しました。この間、AIやデータサイエンスの技術は飛躍的に進歩し、誰もが簡単に高度な解析ツールを使える時代になりました。しかし、ツールがどれほど便利になっても、入力するデータの背景にある「母集団と標本」を理解し、出力された結果に含まれる「誤差」を正しく評価する力がなくては、解析結果を適切

に解釈することはできません。むしろツールの進化とともにその重要性は増しています。

そのため、この2つの概念をきちんと捉えることが、統計を学ぶうえで最も大切なことだと思っています。そこで、本書では「母集団と標本」の関係をまずは重点的に学べるようにしてあります。この「母集団と標本」の概念がわかるようになれば、自然と「誤差」について考えるようになりますし、95%信頼区間の重要性についても理解できるようになります。

初版で大切にしていた通り、改訂版でも極力、数式は排除して解説することを心がけました。また、改訂にあたっては、さらに「わかりやすさ」を追求し、図解や身近な例を交えて内容をブラッシュアップしました。加えて、昨今の研究や論文執筆の現場で議論になることが多い「P値」の扱いについても、新たなトピックを加筆しています。具体的には、「名目上のP値」の解釈や、「そもそも、いつも多重性を厳密に回避する必要があるのか？」といった、多くの方が現場で直面する疑問に答える内容としました。そのほか、回帰分析や相関についても項目を追加し、より実践でお役立ていただけるようアップデートしました。ぜひ本書で、統計学の本質をイメージできるようにしてください。

本書を改訂・刊行するにあたり、多方面にご尽力いただいた(株)羊土社の秋本佳子様、森悠美様に深く感謝いたします。また、本書を執筆するにあたって常に協力してくれた、家族に感謝いたします。

2026年3月

株式会社データシード  
吉田寛輝