



書評 生命科学・生命工学のための間違いから学ぶ実践統計解析： R・Pythonによるデータ処理事始め

著者：川瀬 雅也／松田 史生，編者：日本生物工学会

B5判・並製・モノクロ・本文204頁，近代科学社 Digital

価格：【印刷版基準価格】2,500円（税抜）

【電子版基準価格】2,000円（税抜）

一般的な数理統計学の専門書では抽象的な理論や数式を用いた説明が大半を占め、節目で具体的なデータを用いた練習問題を掲載していることが多いが、数学や統計学自体に興味薄い生物工学・生命科学の技術者、研究者は冒頭の理論の理解が進まず難解である。しかしながら、実験データのバラツキが大きい培養や生物反応に関するデータは統計解析なしに正確な議論ができない。研究室に配属されて、具体的な課題に直面してから、数理統計学の理解が必須であることに気づいた学会員の方も少なくないのではないだろうか。本書は一般的な数理統計学の専門書と逆の構成で、生物工学技術者が誰でも一度は経験しているような馴染みのある実験を題材に、数理的な説明を最小限に抑え、R言語やPython言語を使用した具体的な計算方法を示し、その解釈を主体に構成されている。統計学をツールとして、使いこなす必要がある本学会員の皆様にとっては、親しみやすい構成と思われる。全編会話形式の文章になっており、学生のAさん（学部・女性）、B君（大学院生・男性）らの話題提供、統計学に詳しいX教授への相談、X教授の解説という構成が取られており、単行本を読む感覚で統計解析を復習できる。大学の研究室でみられるような会話の中には、著者の川瀬先生、松田先生のお人柄が反映されたユーモアも含まれており、読者を飽きさせない工夫も随所にみられる。先生方の研究室での一場面を思い浮かべながら読むのも一興かもしれない。

第1部はR言語を使用した解説は基礎的な統計学に関する内容、第2部はPython言語を使用した解説はよりアドバンスな内容となっており、深層学習の基礎まで学べる。第3部では解釈に関する見直しの要点がまとめられている。R言語やPython言語で統計計算の方法やコードが記載されており、PCの前でソフトウェアを操作しながら、読み進めると基礎的な実践力が身につく内容となっている。一度、統計学の講義を受けたが十分に習得できていない方や、実験データの統計処理を定型的にはできているが十分理解できていないと思われる方に、おすすめの書籍である。統計的解釈や実験デザインについても、随所で強調されており、統計処理に慣れている方も一読すると、理解不足や誤りだけでなく、新たな発見に気付くきっかけになるかもしれない。近年、生物工学分野でも、R言語やPython言語を使用する場面は増えており、PCを使用したドライ解析に苦手意識を持っている方のドライ解析入門書や副教材としても良いと思われる。

本書の前書きにあるように、簡単なプログラミング技術と統計学の考え方ができる研究者や技術者を「データサイエンティスト」と呼ぶようになっている。バイオ技術と情報技術のダブルメジャーになる必要があると考えられることも多く、高いハードルを感じることもあるが、本書を一読して、統計学とプログラミングに親しみを持つことが「データサイエンティスト」への第一歩ではないだろうか。

小西 正朗（北見工業大学）