

ケンドール著 多変量解析

本書は Sir Maurice Kendall 著, “Multivariate Analysis”, 第2版, Charles Griffin(1980)の翻訳書である。ケンドール卿は、若い頃からロンドン大学教授などを勤め、老年になってから官庁統計や市場調査の実務にたずさわり、最近では国際統計協会の世界分働力調査の研究推進者として精力的に働いている。A. Stuart との共著の3部作: The Advanced Theory of Statistics は統計理論のバイブルとして広く世界の学者・研究者に利用されている。また、Time-Series, Rank Correlation Methods などの著書も名著とされている。

本書は、経営・工業・経済や応用科学の分野で多変量解析を必要とする実務家向けに、各手法のもつ背景や実用上の価値に深い考察を加えて書かれている。各章の構成は次のようになっている。

第1章 序説 第2章 主成分 第3章 分類とクラスター分け 第4章 因子分析 第5章 正準相関 第6章 分布理論 第7章 回帰分析の諸問題 第8章 関数関係 第9章 仮説検定 第10章 判別 第11章 多変量分類データ

内容は、判別分析、主成分分析、関数関係解析を特に重視し、一方、重回帰分析の基礎概念の拡張を試みている。まず序説では、多変量解析が関係する実際の問題を広く応用分野に対して例示し、多変量解析法が1次元手法の多次元への単純な拡張だけではない点を指摘しながら読者に対して多変量解析の特徴を概念的に理解させようとしている。また、諸手法の関連図を示して第2章以下へのよい導入部になっている。以下の章では、経済学、農業、教育、工業等の広い分野からの実際の例題を使用し、数式に偏らない手法の説明と解釈の方法を与えている。特に諸手法を適用するうえで誤りやすい点や、結果の解釈に深い示唆を与えてくれる。第2章では主成分分析の手法を簡潔に導いた後、測定単位の問題、行列の退化の問題のほか、相関行列から計算した固有値が1以上の成分のみ盲目的に採用する危険や固有ベクトルの係数(因子負荷量)の個々の値をあまり重視すべきでないことなど、実際上注意すべき点が指摘されている。

第3章では、クラスタリングにおける距離として簡明で計算量の少ないユークリッド距離を推めている。全体を通じ、ケンドール卿は過度の複雑化を避け、計算機を積極的に使用することも考慮して書いているようである。第6章分布理論では、一変量の理論を一般化して多変量へ進む道程を示し、両者の対応づけをし、主成分分析の固有値の間の有意差検定にふれている。第7章では、回帰分析の初等的知識を前提として実務上遭遇する幅広い問題について注意すべき点を指摘しており、きわめて有用である。たとえば、共分散行列の退化(共線性)の問題、擬変数使用上の注意、回帰係数を t 検定する問題、欠測値の推定法、段階回帰、残差の規則性、通常の回帰理論と自己回帰モデルとの関係に関する注意等が平易に述べられている。また、コンピュータ・プログラムを使用するためにその信頼度を吟味することも指摘している。第8章では、他書でほとんどふれられていない観測誤差を含む変数間の構造関係を扱っている。一般の読者にはなじみが薄いかも知れないが、回帰分析とともに重要なものであり、順を追って平易に説明されている。第9章では、尤度関数の導入、分散共分散の一様性が否定された場合の検定の進め方、変量間の相関により、ある変量の効果の有意性が隠れる場合の処理などを記述している。第10章では、有名な Fisher の Iris の例を使用し、誤分類確率の評価、母集団が3個以上の場合、判断の保留、分布型に依存しない判別法などがくわしく述べられている。第11章では、多次元分割表に並べられた個体数から成る多変量分類データについて、シンプソンのパラドクスなどを例に挙げながら解析上留意すべき点を説明している。

最後に、訳書では原著の誤解しやすいところやわかりにくい表現を脚注・補注で補い、引用の誤りを原論文にさかのぼって訂正するほか、記号法をわが国での慣用に統一するなど、読者のためにきめ細かい配慮がなされている。また、邦書を含む最近の文献リストが追加されており、学生、研究者、実務家にぜひ読んでほしい好著である。(石川 徳衛 三菱電機)