

## 新刊紹介

坂元慶行・石黒真木夫・北川源四郎著 「情報量統計学」 共立出版, 1983年1月発行 236ページ, 3,500円

私の研究活動の本拠地である北米では統計数理研究所の赤池氏によって考案された情報量規準(AIC)が高い関心を呼んでいる。私の専門にしている心理計量という狭い分野に限ってみてもAICを用いた論文の数がここ数年急速に増え始めている。これは「役に立つ」ものは積極的に採り入れようとする北米のプラグマティックな関心からしてなるほどと思わせるものであるが、私がこう納得できるのもAICが実際に役に立つという実感が私の側にあるからである。今回研究休暇で一時帰国し、表題の本を手にした。この本は非常に読み易く、かつ内容がこれまでの統計書にはみられない斬新なこともあって興味深く、手にしてすぐ一気に読みあげてしまった。この本の紹介を二つ返事で引き受けたのはAICを使ってみた経験に根ざしているが、

読んでみて改めて引き受けてよかったと思った。

本書は従来統計的検定が果していた役割をAICに基づく最適モデル選択によって全面的に置き換えることを意図したものである。検定の基本が二つのモデルの比較にあることは論を待たないが、従来の検定には有意水準の任意性、一度に二つのモデル、しかも一方が他方の特別の場合であるようなモデルの比較しかできないといった欠点が存在する。これに対しAICによる方法では統一的にモデルの適合度を評価し、広い範囲のモデルの比較が一義的にできるという利点がある。もちろんこうした考え方は赤池の原著にもはっきり現われているが、本書はそれを統計学の専門家に対する問題提起としてだけでなく、広く統計学の使い手にひろめる役割を担っている。この目的のためにまず第一部でAICの理論的根拠が与えられ、続いて第二部でAICによるモデル選択の方法論が従来検定が用いられてきたような場面にいかに適用されるかが具体的な例を用いて解説されている。応用例は離散分布モデル、分割表モデル、正規分布モデル、回帰モデル、分散分析モデルの多岐にわたり、説明は明快で説得力に富んでいる。最後に第三部として応用のための計算プログラムが読者の利用に供されている。

本書を通読して次のような感想・疑問を持った。思いつくままに述べておく。一読者の素朴な反応として受けとめて欲しい。

(1) 著者達は「視点と構成」の中で個々の標本分布論を廃すると述べているが、モデル選択、すなわち検定のためのそれは廃止できても、AICによるモデル選択によって最適のモデルが定まった時点で、そのモデルによる予測値の安定性などを評価する上で再び標本分布論が必要となって来るのではないか。

(2) AICによる方法論は完全に検定論にとって代わることができるのだろうか。AICの計算はモデルを直接データに当てはめることを前提とする。ところがモデルの中には直接当てはめができる程特定化が十分でないものも存在する。例えばノンパラメトリック検定で仮説として採用されるようなモデルである。AICはこのような場合にも有力な武器になり得るであろうか。

(3) AICによるモデル選択が大変便利であるだけに、乱用の危険性も大きいと考えねばならない。例えばモデルAとBがあり、AICによる比較の結果モデルAの方が良いという結果になったとしよう。そうなった理由は様々なことが考えられ、極端な場合にはデー

タの中の一つのアウトライヤーのためにそうなった可能性も否定できない。このように様々な理由を一つの統計量の値から読みとることは原理的に不可能である。データ解析にはAICを機械的に用いてどちらのモデルが良いかを判定する以上のものがあるはずである。AICの適用はその根底にデータを見る目があってこそその威力を発揮するのである。

(4) 著者によって結論の表現に若干の食い違いが見られる。AICによるモデル選択の結果はデータと比較されたモデルの範囲に依存するが、これらのことを常に結論の中で明示する必要がある。AICはその理論的根拠を漸近理論に置いているが、現実のAICは有限な標本数 $n$ の関数である。もちろん $n$ が大きい程正確な判定ができることが期待される。これは $n$ が大きければそれだけ精度の高いモデルの推定が可能であると考えれば当然である。しかし現実のモデルが真のモデルであることは不可能に近いから、このことを論理的につきつめると $n$ をどんどん大きくしていったとき、AICによって選ばれる最適モデルのパラメータ数もどんどん大きくなってしまふことになる。これはAICの主旨に反することである。このことは実際のデータ解析の場面では大きな障害にはならないかのように見える。 $n$ を無制限に大きくすることは現実的に不可能だからである。しかしデータを所与とする場合はともかく、現実 $n$ をどう定めたら良いかを考えるとき、やはり問題になるだろう。

著者達は序で「複雑なデータにはデータ解析の手を加えられてその内包する構造をいきいきと語り始める瞬間がある。われわれはAICを使って実際のデータを解析しながら幾度となくこのみずみずしい発見のよるこびを味わった。われわれが本書のねらいとしたのは(中略)この感覚を伝えることである。」と述べているが、同様の経験を持つ評者にとって本書の出版はこの上ない喜びである。同時に本書によって啓発された読者が一人でも多くAICを用い、同じ経験を味わわれることを願ってやまない。最後にもう一つ付け加えるならばこのような本が早く英語で書かれて欲しいものである。(高根芳雄)