

情報量統計学

共立出版 1983年 定価3500円 236頁

AIC (Akaike Information Criterion). —それは、国内・外を問わず、多くの統計研究者の間で注目されている画期的な理論である。提唱者は、統計数理研究所の赤池弘次氏。かつて筆者も、赤池氏ご自身が“熱っぽく”説かれたAIC理論を拝聴して敬服したものが、それから10年。統計の活用という面において、この理論がいくつもの大きな花を咲かせてきた。

本書は、日常赤池氏の近くにあつて、AICの奥義を体得された気鋭の研究者たちが、AICの上に独自の分野をきり拓いた統計学実践のための手引書である。

本書は従来、推定や検定、あるいはデータの記述法としてとり扱われてきた統計学の諸問題を、モデルの構成と情報量による評価という視点から見直そうとの一貫した態度で書かれている。

その第1部は理論編、第2部が応用編、第3部はプログラム編となっているが、第2部の応用編が圧巻である。応用編は、次のように組まれている。

- 第5章 離散型確率分布モデル
- 第6章 分割表解析モデル
- 第7章 正規分布モデル
- 第8章 回帰モデル
- 第9章 分散分析モデル

われわれが、データにもとづいてヒストグラムを書こうとするとき、階級の数をいくらにしたらよいかで、戸惑うことが多い。ヒストグラムを書く場合、階級の数をデータ数に比べて少なくしすぎると、データから得られるはずの情報を捨て去ることになってしまう。逆に、階級の数を多くすると、こんどはデータの個々の値を、そのままながめるのと同じようなことになって不都合である。そこで、最適な階級数をもつヒストグラムを求めることが必要になる。この問題を、AICが解決してくれる。第5章には、そのようなことが、詳細に述べられていて、興味ぶかい。

ところで、統計数理研究所では継続的に「日本人の国民性調査」を実施していることを、ご存じの方が多いと思うが、第6章では、前回行なった1978年全国調査のなかの1つ、

「男と女のどちらに生まれかわりたいか」

を例示して、解説している。すなわち、

男が男に、男が女に、女が男に、女が女に生まれたい、としている人が、それぞれ何名いるかといった2次元分割表と、回答者の年齢（30歳未満と30歳以上に分類）をも考慮にいたした3次元分割表とで、どちらがより有効であるか、などについて比較している。この章あたり、AICの有用性が躍如としている。

第7章正規分布モデルでは、データへの正規分布のあてはめを述べている。ただし、これは正規分布に限らず一般の分布へのあてはめという、重要な問題であると捕えることができる。この章もおもしろいが、ORマンにとって、たいへん参考になるのは、第8章回帰モデルであろう。ORマンは、しばしばデータを方眼紙にプロットしたとき、これに多項式をあてはめることを考えるが、その次数をどうしたらよいか。1次式でよいのか2次式がよいのか、といったことで悩むことも少なくない。これについて、次数をいろいろ変えてみて、AICの値が最小になるものを選べばよい、という提案は優れている。これはいわば、次数選択の問題へのAICの適用である。これに対し、第9章分散分析モデルは、目的変数の動きを説明する必要最小限の説明変数の組を見出そうとするものであって、変数選択問題とよばれるものである。この問題もAIC最小化の方策によって解決されることが、第9章で明確に述べられている。

以上、応用編について概観してみた。

われわれが、統計を活用しようとして、テキストなどを開いてみても、とかくモヤモヤした霧がたちこめているような気分させられることが多いが、この本を読むと、AICがその霧を一気に晴らしてくれそうな気がしてくる。AICのもつ魅力と、それをたくみにとりいれた著者の統計的センスのよさによるものといえよう。

(牧野都治 東京理科大学)