

AICが生み出したブレイクスルー

○田村 義保 (統計数理研究所 / (独) 統計センター)

Breakthrough by AIC

*Yoshiyasu Tamura (The Institute of Statistical Mathematics & National Statistics Center)

Abstract— AIC is an information criterion derived by Dr. Hirotugu Akaike in 1973. His research is the beginning of research on criterion for model selection. In many areas of science and technology such as geophysics, life science, economics and engineering, the AIC is used. For this research, Akaike has received Asahi Award, the Kyoto Prize, the purple medal. AIC is worthy to call it a treasure.

Index terms— information criterion, statistical modeling, prediction, K-L divergence

1 推薦対象および推薦理由の概要

本稿では「赤池情報量規準 (Akaike Information Criterion: AIC)」をコトづくりの至宝に相応しい対象として推薦し、その理由について考察する。赤池は統計科学において、モデル選択という基本概念を世界で初めて導入し、統計科学、数理統計学の分野の発展に貢献しただけでなく、生命科学、地球科学、経済学、金融工学、文学、心理学、医学、薬学等の広範囲の分野のデータ解析の高度化に貢献し、さらに、火力発電所ボイラーの制御、船舶のオートパイロット等の実社会でも広く用いることから分かるように、統計学の発展とその利用の拡大に対して大きく寄与しており、コトづくりの至宝として十分な価値を持つと考えられる。

さて、「コトづくり」とは何だろうか? 「モノからコトへのシフト」ということを耳にすることは多い。しかし、本事業の『趣旨および推薦・選出方法』にもあるように、その定義はあいまいで、異なる認識が存在している。ここでは、本事業が選出対象として示している「コトづくり」の概念に従い、AICが至宝にふさわしいことを説明したい。

本事業の「コトづくり」

本事業ではコトづくりを「価値創造のプロセスに焦点をあてた有形無形を問わない総合的な取り組み」として考える。従って、物理的なモノを含んでいても構わない。また、方法論や計算方式など、概念的な事象も含まれる。

火力発電所の制御だけでなく一般的に工業プロセスの制御はAICを用いなくても可能である。しかし、工業プロセスを多変量時系列過程で表現できるプロセスと考えて、操作変数、被操作変数としてどのような変数を用いるべきであるかを、測定される多くの時系列データから選択することにより、工業プロセスの安定した動作を可能にし、生産効率、エネルギー効率を改良できることは、まさしく、工業プロセスの価値を高めることであり、新たな価値を創造したことになる。また、遺伝系統樹においても、長い間用いられてきた、形態の類似性からの親近度ではなく、ゲノムの類似性をAICで測る方法の方が優れていることは言うまでもないことである。AICによるモデル選択が、遺伝情報を扱う際の新たな価値を生み出したことになる。

データから法則、真実を探し出すのが統計科学の役

割の一つと考える。データ中心科学、データ駆動型科学という言葉をよく聞くようになっている。データがあり、それに対して、モデルを考える。AICは、どのようなモデルが良いかを選ぶための規準である。適切な分析を行い、データの持つ情報を引出し、データの価値を高めるために、統計学が提供する「コト」がAICである。

2 情報量規準の創造

ここでは、AICについて概説する。AICを提案した論文としては国際学会のプローシーディングス¹⁾とIEEEの論文²⁾がある。AICは確率密度関数間の距離の尺度であるK-L divergence ($g(x)$: 真の確率密度関数, $f(x)$: モデルの確率密度関数とする)

$$\begin{aligned} I(g; f) &= \int_{-\infty}^{\infty} \log \frac{g(x)}{f(x)} g(x) dx \\ &= \int_{-\infty}^{\infty} g(x) \log g(x) dx - \int_{-\infty}^{\infty} g(x) \log f(x) dx \end{aligned}$$

の第2項である平均対数尤度をデータ解析で中心的な役割を果たす最大対数尤度を用いて推定しようとしたものである。2項目(マイナス符号を含める)の $2n$ 倍(n はデータ解析に用いるデータサイズ)として、次の量を導いた。

$$AIC = -2 \sum_{i=1}^n \log f(x_i | \hat{\theta}) + 2k$$

ここで、 $\hat{\theta}$ がモデルに含まれるパラメータの最尤推定値であり、 k がパラメータ数である。複数のモデルの中で、どれを選択すればよいかの方法を与えたことになる。K-L divergenceの意味で、予測の観点でのモデル選択規準である。AICによって選ばれるモデルは「比較した中で最良」モデルであることを注意しておく。

AICの発表とほぼ同時に Mallowsは C_p ³⁾を提案した。また、SchwarzはBayesモデルの事後確率最大化の考えに基づいてBIC⁴⁾を提案した。赤池自身も経験Bayesモデルのための規準ABIC⁵⁾を提案している。このように、AICの提案は「情報量規準」のその後の発展の端緒となった。統計科学の応用分野ではモデリングの考え方の定着と発展に貢献しているが、詳細は後述する。

3 AICの誕生

統計数理研究所は「実問題を解決するために統計科学の理論・方法を構築・発展させる」という研究スタイル・理念を持っている。AICも秩父セメントとの共同研究の成果である。セメントキルンの制御のための時系列モデルを構築し、どのような変数を用いるか、また、構築した多変量自己回帰モデルのラグをどのようにするかを決めることが必要であった。FPE⁶⁾を提案し、この問題を解決している。この貢献に対して1972年に石川賞(日本科学技術連盟)を受賞している。FPEは自己回帰モデルのラグ決定のための規準であった。これに対してAICは自己回帰モデルだけでなく、すべての統計モデルのモデル選択のための規準である。AIC誕生の背景はプラント制御のためのモデル構築という実問題の解決のためであった。しかし、視点を広げると、AIC以前は統計科学では検定により、モデルを選んでいたと言っても過言ではないと考える。重回帰分析の変数選択においては、係数のt検定によっていた。モデル全体については、F検定(分散分析)が用いられていた。検定では、積極的なモデル選択が実際はできないことは自明であるが、AIC誕生により、モデル選択が可能になった。

情報量規準導入の効果については京都賞受賞を記念して出版された室田・土谷編集の「赤池情報量規準」¹⁴⁾に赤池自らの解説がある。また、AICの一致性やBICとの関係などについては、北川や甘利の解説があるので、これらを参照して欲しい。同書第II編第4章には下平による遺伝系統樹でのAIC利用の解説がある。AICの推定誤差についての説明があるが、AICの推定誤差を考えることはモデル選択においてより適切な判断ができることにつながり、データが生み出す価値をより高めることを可能にすることになる。このような新しいAICの使い方がAICの「コトづくり」の道具・方法としての価値を高めているものと考えられる。

4 AICが変えた統計科学の基礎研究及び応用研究

では、AICがどのように統計科学の分野およびそれを用いる諸科学の研究をどのように変えたであろうか。すでに、 C_p やBICについては述べたがAICを改良するという数理科学の分野の研究としては竹内によるTIC⁷⁾やSugiuraによる $cAIC$ ⁸⁾がある。新しい情報量規準としては渡辺により提案されたWAIC⁹⁾やKonishiとKitagawaによるGIC¹⁰⁾等がある。情報量規準について詳しく知りたい場合は上記の小西・北川の書籍を読むことを勧めたい。このように、AIC誕生により、多くの情報量規準に関する研究が統計科学の分野で行われ、統計科学の発展に貢献している。

統計モデルを用いて研究する科学の諸分野への貢献については実例を上げるときがないので、赤池のAICについての論文の引用統計を示すことで行う。2論文での累積引用数は30,000件を超えて現在でも急速に増え続けており、広範囲の分野で使われている。

Fig. 1及びFig. 3はWoSからの結果である。Fig. 2は北川による「赤池統計学の世界」¹¹⁾の図をそのまま使わせていただいた。

上述したようにAICが拓いた情報量規準の概念は、まだまだ、発展している。スパースモデルについては

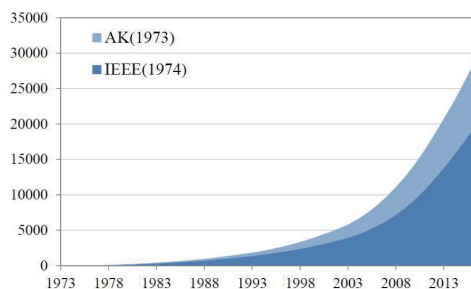


Fig. 1: Cumulative citation by 2017



Fig. 2: Example of fields in which AIC is used

L1正則であるLASSOが用いられるが、この場合の情報量規準も活発に研究されている。Fig.1では2017年までの引用数を示しているが、2008年までの引用数は約12,000であった。9年間で約18,000の引用があったことになる。累積件数でなく、1年ごとの引用数が指数関数的に増加している。推薦理由のところでAICが使われている学術分野の多様性について述べたが、図2にあるように、自然科学、社会科学、生命科学などのいわゆる理系の分野だけでなく、文学や心理学のような文系の分野でも用いられている。データ解析環境「R」の多くのパッケージでAICが実装されている。モデリングで微分方程式や差分方程式が用いられ、微分計算や逆行列計算、固有値演算が行われている。AICはモデル選択規準として、モデルにおける微分や行列と同等のレベルで用いられるようになってきているものと考えられる。日本発の理論や手法で優れた「コト」であるものは多くあると思うが、それらの中に間違いなく入ることが可能な「コト」の一つであると考えられる。

AICの活用は、これから、ますます盛んになっていくものと考えられる。「根拠に基づく」という概念のひろがり関係している。EBPM(根拠に基づく政策立案)に関して多くの閣議決定がなされている。EBM(根拠に基づく医療)の重要性が言及されるようになってから久しい。根拠となるデータを活かすも無駄にするのも、データをどのように分析するかである。どのようなモデルを考えるかである。もっとも良いモデルを選択する必要がある。そのために用いられるのがAICであり、「コト」としてのAICの重要性は増すばかりであると考えられる。

5 主たる貢献者

赤池情報量規準AIC提案・創造の主たる貢献者は赤池弘次氏である。赤池氏は、この貢献により、次の賞を

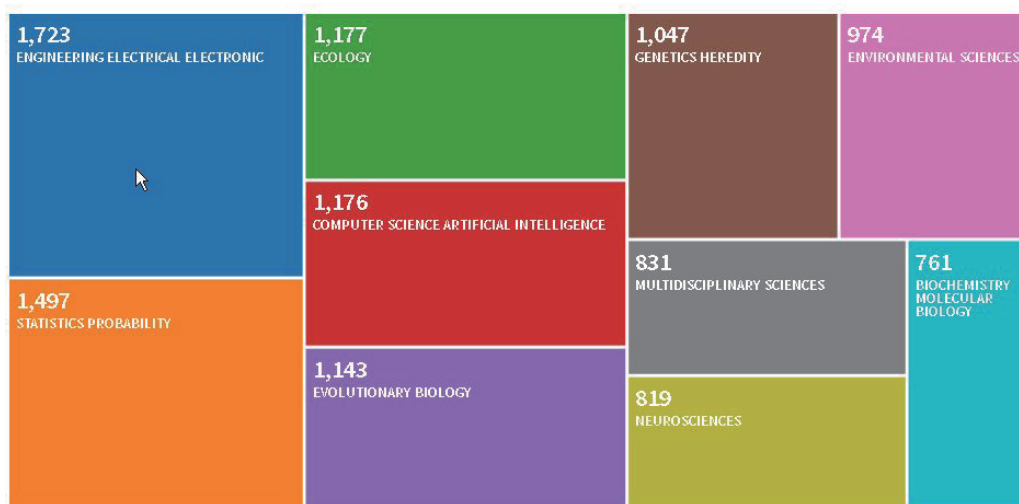


Fig. 3: Cited research area in IEEE

受賞 (受章) している。括弧内は受賞理由である。1989年：朝日賞 (統計学の研究—特にAICの理論と応用)，1989年：紫綬褒章 (統計学，特に時系列解析とその応用の研究)，1996年：第1回日本統計学会賞 (統計学の理論とその実用化の研究)，2006年：京都賞 (情報量規準AICの提唱による統計科学・モデリングへの多大な貢献)。AICの普及に貢献した者としては，赤池の共同研究者や統計数理研究所の研究者がいる。一部ではあるが，京都賞の受賞を記念にまとめられた総研大ジャーナル¹²⁾にその記載がある。赤池が中心になって開発した時系列解析プログラムパッケージTIMSACシリーズ¹³⁾(Rのパッケージも有る)が時系列解析利用の発展，AIC利用普及に貢献したことも間違いない。

6 価値の観点による考察

- i. 先導力 赤池情報量規準は統計科学において最初に提案された情報量規準である。その引用の拡がりを見ても，統計科学の理論・応用の発展を先導している。単にモデル選択規準を与えたに止まらず，モデリングの重要性を明らかにし，その後の統計的モデリングの潮流を形成した意味での貢献は大きい。
- ii. 規範力 繰り返しになるが，多くの科学分野で広く用いられており，定着している。統計解析のための言語Rの多くのプログラムでAICが用いられており，デファクトスタンダードとなっている。
- iii. 意味力 統計的データ解析への取り組み，モデリングの実行に大きく影響を与え，まさしく，統計科学とその周辺へのブレイクスルーとなった。
- iv. 解決力 モデル選択をデータに根差して，客観的に行う方法を創出している。多くの分野の問題解決を可能にした。火力発電所の統計的制御を可能にしたことで，莫大な量の重油が節約できていることは，二酸化炭素削減にも貢献している。

7 まとめ

コトづくりの至宝として赤池弘次氏による情報量規準AICの創造を，統計的データ解析にブレイクスルーをもたらしたこと，その後の情報量規準の研究の発展につながったこと，データ解析におけるデファクトスタンダードとなっていることを根拠に推薦する。至宝登録を契機に情報量規準そのものの研究とその利用が活性化して，EBPMやPDCAが根付き，これらによ

り，データの価値を高めることに寄与していくことを願う。

参考文献

- 1) H. Akaike, "Information theory and an extension of the maximum likelihood principle", Proc. 2nd International Symposium on Information Theory (B. N. Petrov and F. Csaki eds.) Akademiai Kiado, Budapest, (1973) 267-281. (B. N. Petrov and F. Csaki eds.) Akademiai Kiado, Budapest, 267-281, (1973). (Reproduced in Breakthroughs in Statistics, Vol. I Foundations and Basic Theory, S. Kotz and N. L. Johnson eds., Springer-Verlag, New York, 610-624, (1992).)
- 2) H. Akaike, "A new look at the statistical model identification", IEEE Trans. Automat. Contrl, AC-19, No. 6, 716-723, (1974).
- 3) C. L. Mallows, "Some Comments on Cp," Technometrics, Vol. 15, pp. 661-675., (1973)
- 4) G. Schwarz, "Estimating the dimension of a model", Ann. Statist., Vol.6, 462-464, (1978).
- 5) H. Akaike, "Likelihood and the Bayes procedure. Bayesian Statistics", J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds., Valencia, Spain; University Press, 143-166, (discussion 185-203), (1980).
- 6) H. Akaike, "Statistical predictor identification", Ann. Inst. Statist. Math., Vol.22, 203-217, (1970).
- 7) 竹内啓, 情報統計量の分布とモデルの適切さの規準, 数理科学, No.153, 12-18, (1976)
- 8) N. Sugiura, "Further analysis of the data by Akaike's Information criterion and the finite corrections", Commun. Statist.-Theor. Meth., Vol.7, No.1, 13-26, (1978)
- 9) S. Watanabe, "Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory", Journal of Machine Learning Research, Vol.11 3571-3594, (2010).
- 10) 小西貞則, 北川源四郎, 情報量規準, 朝倉書店, (2004).
- 11) www.ism.ac.jp/akaikememorial/pdf/Akaike_Statistical_World.pdf
- 12) 赤池統計学の世界, 総研大ジャーナル, No12, (2007).
- 13) Y. Tamura, "APPLICATION OF TIMSAC", Proceeding of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach, 251-271, (1994).
- 14) 室田一雄, 土谷隆 (編), 赤池情報量規準 AIC —モデリング・予測・知識発見—, 共立出版, (2007).