

氏名	植木 優夫
授与した学位	博士
専攻分野の名称	環境学
学位授与番号	博甲第3674号
学位授与の日付	平成20年 3月25日
学位授与の要件	環境学研究科生命環境学専攻 (学位規則第5条第1項該当)
学位論文の題目	Studies on statistical modeling with bootstrap and information criteria (ブートストラップ法と情報量規準を用いた統計的モデリングに関する研究)
論文審査委員	教授 垂水 共之 教授 栗原 考次 教授 梶原 毅

### 学位論文内容の要旨

データに基づいて統計的推論を行う際、事前に統計モデルを仮定することが多々ある。この統計モデルがデータの発生構造を十分表現できるものであれば、データ解析を非常に円滑に進めることができる。統計モデリングは統計解析において最初に行う基本的な手順であるが、早急に解決しなければならない問題が多く残っている。本論文では、ブートストラップ法と情報量規準を用いた統計的モデリングに関連する近年の発展について述べ、私の研究成果と併せて報告した。特に以下の3点について研究を行った：(i) 予測、(ii) 信頼区間、(iii) 情報量規準。これらは現実問題において広く一般的に利用されている。またこれらには共通する部分が多くあり、互いに密接に関連している。

まず本論文での(i) 予測では、パラメトリックな確率分布モデルに基づいたものを考察した。特に次の2通りのアプローチを議論した。ひとつは予測分布を考えるもので、確率分布の推定を主眼とし、真の将来の確率分布との近さをカルバック・ライブラー情報量などの確率分布間の距離を利用して評価するものであり、もうひとつは予測限界を求めるもので、将来のデータが発生しうる予測区間を構築するものである。上記の2つは、過去のデータを用いて将来の挙動を捉えるという意味で類似しているが、実は大きな差があることが知られている。まず予測分布に関しては理論的な考察を行なった。次に予測限界の構成に関しては Ueki and Fueda (2007)で提案した手法を中心に述べた。またTOPIX, S&P500, 香港ハンセン指数, JASDAQ, という現実の4指数の株価データを用いて予測限界を計算し、その有用性について議論した。

次に(ii)の信頼区間では、あるスカラー値統計量に関する $100\alpha\%$ 信頼区間を求めることを考えた。これまで多くの高精度な信頼区間構成法が提案されている。本論文ではUeki and Fueda (2007)の手法を応用し、2次の精度を有するように補正する方法を提案した。さらに導いた方法を正規信頼区間と符号付対数尤度比信頼区間に適用した。特に補正正規信頼区間はブートストラップ-t法と同じく2次の精度となり、一方で計算コストは同程度となった。補正正規信頼区間はブートストラップ-t法よりも安定したパフォーマンスを示した。

最後に(iii)については、情報量規準を用いたモデル選択、特に赤池情報量規準 (AIC) のひとつの一般化であるKonishi and Kitagawa (1996)の一般化情報量規準 (GIC) について考察した。本論文では、経験分布にあまりにも近すぎる統計モデルを想定し、解析的な計算に基づく情報量規準はそのモデル評価を失敗する可能性があることを示した。そのモデルを「過剰に近いモデル」と呼び、その問題に対処するためのGICの改善法を提案した。

## 論文審査結果の要旨

本論文は、環境データの解析にも必要となる数理統計学の一分野：母集団分布の推定に関する研究を行っており、ブートストラップ法と情報量規準を用いて関連する複数の成果をあげている。

一つ目は分位点の推定に関するものである。未知の母集団分布に対して、その上側5%点などの分位点を推定するために最もよく知られた簡便な方法は誤差が多いことが知られており、いくつかの改良法が提案されている。しかし既存の改良法は目の子での繰り返し計算を必要とし、計算量が膨大となっていた。本論文では、ブートストラップ法を利用して既存の改良法と漸近的に同じ精度が得られ、しかも計算量が遥かに少ない方法を開発した。二つ目は、信頼区間に関するものである。信頼区間の構築には推定量の分布が必要であり、広く用いられる方法としては推定量の分布を正規分布などで近似する方法とブートストラップを用いて推定する方法がある。しかしながら前者は信頼区間が真の値を含む確率の誤差が大きく、後者はしばしば信頼区間が長くなりすぎる。本論文では一つ目の成果を応用することで真の値を含む確率を保ちつつ区間が長くなりすぎることを防ぐことが出来た。三つ目は、統計モデル選択に用いられる情報量規準の改良である。この分野は赤池のAIC、小西・北川のGICなど日本の研究者による業績が世界的に認められているが、ある状況の下でGICによるモデル選択が良くないことを調べ、その改良法を開発した。

本論文の一つ目の成果は国際誌Biometrika(2007)に掲載され、二つ目の成果はThe 9th Japan-China Symposium on Statistics(2007)、三つ目の成果は応用統計学会(2006)にてそれぞれ発表した。

以上のように、本研究は数理統計学の理論の発展に寄与するところが大きく、博士(環境学)の学位に値すると判断する。