

論文審査の要旨

博士の専攻分野の名称	博 士 (理 学)	氏名	佐 藤 倫 治
学位授与の要件	学位規則第 4 条第 ①・② 項該当		
論文題目			
C_p type criterion for model selection in the generalized estimating equation method (一般化推定方程式における C_p 型モデル選択規準)			
論文審査担当者			
主 査	教 授	若木 宏文	
審査委員	教 授	井上 昭彦	
審査委員	教 授	柳原 宏和	
〔論文審査の要旨〕			
<p>一般化推定方程式法とは、同一の個体に関して、興味ある変数（目的変数）とそれに関連があると考えられる複数の変数（説明変数）を、特定の期間繰り返し測定することによって得られるデータから、目的変数と説明変数の関連の度合いを推定するため、Liang & Zeger (1986) によって提案された推定手法の一つである。</p> <p>個体 i の、第 j 時点における目的変数の値を Y_{ij} とする。一般化推定方程式法では、Y_{ij} の周辺分布に対して、ロジスティック回帰や、ポアソン回帰などの一般化線形モデルを仮定することにより、目的変数の平均と分散への説明変数の影響を記述し、一方、(Y_{i1}, \dots, Y_{im}) の相関構造については、作業量相関行列と呼ばれる単純な構造を指定して、それらを基にした未知母数に関する非線形方程式を解くことで未知母数の推定量を得る。</p> <p>候補の説明変数から、目的変数に影響を与える説明変数を選択する代表的な方法に、独立性を仮定した擬似尤度に基づく情報量規準である QIC があるが、未知母数の推定では目的変数間の相関構造を仮定しているので、相関構造を反映した変数選択規準が望ましい。</p> <p>Inatsu and Imori (2013) は、相関構造を考慮した予測平均 2 乗誤差に基づく変数選択規準を提案し、従来の QIC より優れた変数選択規準であることを示した。彼らは、作業用相関行列に含まれる相関パラメータと尺度パラメータについては既知と仮定していたが、応用上はいずれのパラメータも未知である場合が重要である。</p> <p>本論文では、第 3 節において、Inatsu and Imori (2013) が提案している変数選択規準の漸近バイアスを、相関パラメータと尺度パラメータが未知で推定する場合に評価した。申請者は、相関パラメータと尺度パラメータの推定量がいくつかの条件を満たすとき、Inatsu and Imori の変数選択規準を、漸近バイアスの修正なしに利用できることを明らかにした。一般化推定方程式によるパラメータ推定は R や SAS などの統計解析ソフトで実用化されているが、相関パラメータと尺度パラメータの推定量としてはモーメント推定量を用いる場合が多い。申請者はモーメント推定量を用いる場合、Inatsu and Imori の変数選択規準は修正なしに利用できることを確認している</p>			

(Theorem 1). また、この選択規準は、変数だけでなく作業用相関行列の選択にも利用できることを指摘している。

ガンマ分布をベースとした一般化線形モデルは正の連続データを解析する場合に良く用いられるモデルである。申請者は提案した変数選択規準の有限データでの性能をチェックするため、ガンマ分布と対数リンクを用いた一般化推定方程式モデルに対する数値実験を行っている。一般化推定方程式では、観測時点ごとの周辺分布に対して一般化線形モデルを仮定するが、個体に関する目的変数全体の分布形は仮定していない。本論文では、コピュラ法を用いて擬似データを発生している。その結果、次のような特徴が見て取れた。

- (1) 目的変数に影響のある変数のみを用いたモデルの選択確率が最大となり、その選択確率は個体数の増加とともに大きくなる。
- (2) 影響を与える変数をひとつでも含まないモデルの選択確率は個体数が大きくすると 0 となる。
- (3) 作業用相関行列が真の相関構造と一致している方がリスクは小さくなる。
- (4) モデルに取り込む変数だけでなく、作業用相関行列も選択規準を用いてデータごとに選択した場合、作業用相関行列を真の相関行列に固定して変数のみ選択した場合よりもリスクが小さくなる。

(1)~(3) は、変数選択規準の漸近性質から予想されることであったが、(4) の結果は予想を超えており、非常に興味深い発見である。

第 3 節では、漸近バイアスの評価において個体数は増加するが観測時点数は固定した漸近枠組みを採用しているが、第 4 節では個体数 (n) と観測時点数 (m) が共に増加する場合を扱っている。Xie and Yang はいくつかの条件を仮定して、線形予測子の回帰係数ベクトルの推定量の漸近正規性を証明しているが、申請者はさらに、回帰係数ベクトルの推定量の確率展開を導出し、それらを用いて変数選択規準の漸近バイアスが m^4/n のオーダーであることを証明した (Theorem 2)。また、一般化推定方程式法において、作業用相関行列が真の相関構造に一致している場合には、回帰係数ベクトルの推定量の漸近効率が低いことが知られているが、本論文で扱っている変数選択規準の漸近バイアスも小さくなり、 m^2/n のオーダーであることが示された (Theorem 3)。

第 4 節では観測時点数 m が大きい場合について、第 3 節と同様の数値実験を行っている。実験結果から $n = 400, m = 20$ であっても真の変数の組み合わせの選択確率が十分に高いことが確認された。

以上、審査の結果、本論文の著者は博士（理学）の学位を授与される十分な資格があるものと認める。

公表論文

- (1) A C_p type criterion for model selection in the GEE method when both scale and correlation parameters are unknown.
T. Sato, Y. Inatsu, to appear in *Hiroshima Mathematical Journal*, **49**(2), 2019.
- (2) Asymptotic bias of C_p type criterion for model selection in the GEE when the sample size and the cluster sizes are large.
T. Sato, to appear in *Hiroshima Mathematical Journal*, **49**(2), 2019.