

氏名	金 鉉 廷
授与した学位	博 士
専攻分野の名称	理 学
学位授与番号	博甲第 1782 号
学位授与の日付	平成 10 年 3 月 25 日
学位授与の要件	自然科学研究科システム科学専攻 (学位規則第 4 条第 1 項該当)
学位論文題目	STATISTICAL ANALYSIS OF INCOMPLETE DATA AND STATISTICAL GRAPHICS 欠損データの統計解析と統計的グラフィクス
論文審査委員	教授 垂水 共之 教授 大竹 正徳 教授 田中 豊 教授 北村 吉朗 教授 藤井 弘章

### 学位論文内容の要旨

Part I では、欠損データの統計解析における 2 つのトピックスについて議論した。1 つは“無視できる欠損構造”をもつ場合に、多次元データの平均ベクトルと分散共分散行列を最尤推定し、その結果を用いた多変量解析（主成分分析など）において、個々の観測値が分析結果にどのように影響するか、影響の大きい観測値が存在するかどうかを分析する、影響分析（感度分析）の方法について論じた。もう 1 つは“無視できない欠損構造”を持つ場合に、欠損のおこるメカニズムを表すモデルパラメータ  $\psi$  を想定し、指数分布と正規分布の場合について欠損メカニズムのパラメータ  $\psi$  と他のパラメータとを推定する直接的方法と  $\psi$  を既知として他のパラメータを EM アルゴリズムを用いて推定する 2 つの方法について論じた。

データを図示する統計グラフィクスの中には、不注意に描かれ、誤認識されるものが少ない。Part II では、新聞にあらわれる統計グラフィクスを分析し、データのより良い表現法のために、グラフィクスの一般的な描き方の規則と、表現法の基本になる部分を整理して提示した。

## 論文審査結果の要旨

本論分では大きく2つの問題について論じている。一つはデータに欠損値を含むデータの分析法であり、もう一つは統計的グラフィックスの表示の問題である。

第1部では、欠損データの統計解析は、欠損の起き方がランダムで欠損構造が無視できる場合と、欠損の起き方に構造があり、欠損構造が無視できない場合とを取り扱っている。

欠損構造が無視できる場合に、分散共分散行列の最尤推定や、推定した分散共分散行列を用いて多変量解析を行った場合の解析結果に与える、影響の大きい観測値の集合を求める方法を示し、欠損値を含むデータの感度分析法を示した。

さらに欠損の起き方に構造があり、欠損構造を無視できない場合として、値が大きいほど欠損となる割合が高くなるモデルを想定した。この場合に、欠損構造のパラメータを含めてイテレーションで直接最尤推定を行う方法、EMアルゴリズムを使ってパラメータを推定する方法を求めた。これらの有効性を調べるために、数値実験を行い、これまで欠損値を含むデータを捨てて、欠損のない完全なデータだけで解析していた場合に起きていた偏りを無くすることができることを示した。また、2次微分が必要なものの、イテレーションによる推定では、欠損構造のパラメータ推定まで行えるなど、メリットが大きい。

第2部では、国内の主要な全国紙5紙と、一つの地方紙に現れる統計グラフィックスを分析し、誤認識されかねないグラフィックスの問題点を指摘し、その解決法を示した。さらにより良い表現法のために、統計グラフィックスの一般的な描き方の規則と、その表現法の基本部分について示した。

以上のように、本論文ではデータ解析上重要な欠損値の取り扱いと、グラフ表現に新しい観点を与えるものであり、学位審査会では、学位論文の内容、参考論文等を総合的に判断し、博士（理学）の学位に値するものと判断した。