

楢円母集団における多変量分散分析の 次元検定に関する研究

学位論文内容の要旨

本論文は、多変量楢円分布のもとでの多変量分散分析の次元に関する検定問題について考察し、新たな検定方式を提案したものである。現代社会において、我々の周りには多種多様のデータが氾濫しており、それらを適切に解析し、データが発生した背景的要因や、データ間の相互関係、あるいは将来起こり得る事象等を推測、予測する必要性がしばしば生じる。データは通常、身体測定における、(身長、体重、胸囲、座高)のように、いくつかの変量の組として得られ、このような多変量データを解析する手法である、多変量解析が計算機の発展と普及にともない、その応用が著しく広がっている。

多変量解析の主要な目的は、i) 分類、ii) 構造の単純化、iii) 相互依存性の解析などがあげられ、これらを扱う手法としては、判別分析、主成分分析、正準相関分析などが代表的である。多変量解析において、データの分布や誤差の分布として、しばしば、正規分布が仮定される。正規性の仮定は、中心極限定理により、理論的に妥当であることが主張されているが、実際には、数学的扱いが容易になるという利便性に起因するものであり、その仮定が自然である状況はむしろ稀である。したがって、正規性を仮定しない非正規分布のもとでの、多変量解析の議論が重要となる。

データが複数の母集団から観測されたとき、母集団の比較を行う手法として、多変量分散分析があげられる。多変量分散分析の主要な話題としては、分散が各母集団で共通であるときの、平均に関する検定である。たとえば、各母集団の平均が全て等しい場合は、母集団間に差異がないと解釈できる。また、ある母集団のみ平均が異なり、他の母集団の平均が共通の場合は差異が1となる。このような母集団間の差異をあらわす指標を次元という。次元に関する、従来の研究は通常、正規母集団の下で考察され、検定方式としては尤度比 (Likelihood Ratio = LR) 型、Lawley-Hotelling (= LR) 型、Bartlett-Nanda-Pillai (= BNP) 型がしばしば用いられてきた。しかし、次元に関する検定は、複合仮説検定となり、局外母数に依存する。前述の3タイプの検定方式は、大標本の場合、漸近的にカイ二乗検定となることが示されているが、小標本の場合は、局外母数の影響を無視できない。実際、検定を行うときは、局外母数を推定量に置き換える必要があり、仮説を棄却する確率が有意水準より大きくなる可能性が生じる。これを回避するために Schott(1984) は、正規母集団の下で、局外母数に依存しない、検定方式を提案している。

このような背景より、本研究では、以下の2点を主な目的とした。

1. 楢円母集団を仮定した場合の次元検定に関する新たな検定方式の提案。
2. 新たな検定方式の性質、および有効性の考察。

1において、仮定した楕円分布は、正規分布の拡張として知られ、分布の頑強性に関する研究でも、しばしば用いられる。本論文では楕円分布の性質をまず議論した後、母集団分布として仮定し、Schottの方法を利用することにより、次元に関する新たな検定方式を導出した。楕円母集団の下での次元検定に関する先行研究としては、Muirhead and Waternaux(1980), Seo et al.(1995)などがあげられるが、これらは漸近理論の枠組みで行っており、小標本の場合の局外母数の影響を取り除いた、本論文で提案した検定方式は、従来の研究とは一線を画した新しい結果である。また、2における検定方式の性質とは、従来の検定方式を T 、本論文で提案した検定方式を T^* とすると、

$$\Pr\{T > c\} \leq \Pr\{T^* > c\} \quad (\forall \text{局外母数})$$

および、

$$\sup \Pr\{T > c\} = \Pr\{T^* > c\}$$

なる性質がある。これらが成り立つとき、それぞれ、帰無分布に関する上界、および上限であるといい、提案した検定方式が有効であることを意味している。

本論文は以下の7章より構成されている。第1章では、本論文の背景と目的、及び章構成について述べた。第2章では、本研究で扱う楕円分布の定義と性質を与えるとともに、正規分布との相違点、類似点について論じた。楕円分布は正規分布の有する性質を多く引き継ぐが、独立変量の同時分布が楕円分布にしたがわないため、無作為標本を仮定した議論が、正規母集団の場合と比較して複雑になる。第3,4章では、正規母集団の下での従来の研究結果を総括した。第3章では、多変量線形モデルの一般的な推定、検定問題についてまとめ、特に線形仮説検定における、LR検定、LH検定、BNP検定について論じた。これらの検定形式は第4章で述べた、多変量分散分析における平均ベクトルの一様性の検定に適用でき、さらに次元に関する検定問題の礎となる。第5,6章は本研究の中心部であり、第5章ではSchottの手法を楕円母集団において適用することにより、新たな検定方式 T^* を提案した。この検定方式を用いることにより、局外母数の影響を受けずに、次元に関する検定が可能となった。さらに、 T^* にもとづいた、有意水準 α の有意点 c_α^* が得られるならば、任意の局外母数に対して $\Pr\{T > c_\alpha^*\} \leq \alpha$ が成り立ち、最悪の場合でも $\Pr\{T > c_\alpha^*\}$ は有意水準以下になることが保証された。 $\Pr\{T > c\}$ と $\Pr\{T^* > c\}$ の差が非常に大きいならば、 T^* にもとづく検定では仮説を過大の受容する可能性が生じる。そこで第6章では、contaminated 正規母集団と多変量 t 母集団においては、LR型、LH型について、この差が0に近づくための条件をもとめ、 T^* は T の帰無分布に関して上限となることを示した。このことにより、多変量 t 母集団の下でも、 T^* を用いた仮説検定が有効であることを示した。第7章では本論文の結論をまとめるとともに、今後の課題を提起した。

学位論文審査の要旨

主査	教授	佐藤義治
副査	教授	宮腰政明
副査	教授	工藤峰一
副査	教授	栗原正仁
副査	教授	水田正弘

学位論文題名

楢円母集団における多変量分散分析の 次元検定に関する研究

多変量解析は様々な現象をある観点から複数個の属性(変量)に関して、それらの値を可能な尺度で測定されたデータに対して種々のモデルや仮説を通して現象の簡潔な表現や構造を解明するための一連の技法である。本論分で扱っている多変量分散分析は多変量解析の一手法であり、複数の母集団から得られる多次元データを用いて、それらの母集団の相違(平均的な)を明らかにする手法である。

多変量分散分析における次元とは、母集団間の平均の差異を表現するに必要な空間の最小の次元であり、その検定問題は、通常、多変量正規分布の仮定の下で考察されてきた。しかし、データに多変量正規分布が常に仮定できるとは限らず、非正規分布の下での議論も必要であることは言うまでもない。

次元に関する従来の検定方式は尤度比型(LR型)、Lawley-Hotelling型(LH型)、Bartlett-Nanda-Pillai型(BNP型)が代表的であり、これらの漸近分布はカイ二乗分布に従うことが知られている。しかし小標本の場合、これらの帰無仮説の下での分布(帰無分布と呼ばれる)は、局外母数に依存し、実際検定をおこなうときには局外母数を推定量で置き換える必要があり、この場合、帰無仮説を棄却する確率が有意水準を超える可能性を有している。

楢円母集団における次元検定も同様の性質を有しており、従来は標本数が大いとして、局外母数を含む項を0に収束させることにより、局外母数に依存しない検定方式が考察されてきた。

以上のような背景の下で、本論文は、楢円母集団における多変量分散分析の次元に関する検定について、帰無仮説の下での分布が局外母数に依存しない検定方式を考案し、かつその性質を論じたものである。

本論文の内容と成果はつぎの通りである。

最初に本研究で扱う楢円分布の定義と性質を与えるとともに、正規分布との相違点、類似点について論じている。楢円分布は正規分布の有する性質を多く引き継ぐが、独立変量の同時分布が楢円分布に従わないため、無作為標本を仮定した議論が、正規母集団の場合と比較して複雑になることを示唆している。

さらに多変量分散分析の次元検定を構成する上で基本的な概念である多変量線形モデルの一般的な推定、検定問題についてまとめ、特に線形仮説において、群内変動行列と群間変動行列に関する行列式方程式の解の関数として表される不変検定（LR 検定，LH 検定，BNP 検定）について論じている。

つぎに多変量分散分析における次元検定の特殊な場合である平均ベクトルの一様性の検定について論じている。この検定は、多変量線形モデルにおける線形仮説検定に帰着され、したがって、LR 型，LH 型，BNP 型検定を用いることができることを述べている。また、大標本理論の下で、正規母集団における次元検定の従来の研究を概略し、従来の検定方式はカイ二乗分布に従うことを述べている。

多変量正規母集団における Schott(1984) の研究成果を楕円母集団の下に拡張することにより、次元に関する新たな検定方式を提案している。この検定方式は、帰無分布が局外母数に依存せず、かつ、従来の検定方式の帰無分布に関する上界である性質を有している。したがって、この検定方式を用いることにより、局外母数の影響を受けることなく次元検定が可能になるばかりではなく、最悪の場合でも棄却確率が有意水準より大きくなることのないことが保証される。

楕円母集団の具体的なものとして挙げられる contaminated 正規母集団、および、多変量 t 母集団の下で、提案した新たな検定方式は、LR 型および LH 型について、従来の検定方式の帰無分布に関する上限となることを証明している。このことにより、過大に帰無仮説を受容することがないことが示されている。これらの成果により、楕円母集団の下での次元検定をおこなうとき、局外母数の影響をうけないずに、かつ、有意水準を保つ検定が可能となることを示した。

これを要するに、著者は非正規母集団として代表的な楕円母集団における多変量分散分析の次元検定に関して、漸近理論を用いずに局外母数の影響を受けない新しい検定方式を考案し、それに関連する種々の新知見を得たものであり、情報解析学、計算機統計学の発展に寄与するところ大なるものがある。

よって、著者は北海道大学博士（工学）の学位を授与される資格あるものと認める。