

## 科学研究費助成事業 研究成果報告書

平成 29 年 6 月 14 日現在

機関番号：32634

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26730020

研究課題名(和文)高次元枠組みにおける頑健な統計的仮説検定方式の研究

研究課題名(英文) Study on robust statistical hypothesis testing procedures in high-dimensional settings

研究代表者

西山 貴弘 (NISHIYAMA, Takahiro)

専修大学・経営学部・准教授

研究者番号：30516472

交付決定額(研究期間全体)：(直接経費) 2,200,000円

研究成果の概要(和文)：本研究では、高次元枠組みの下で、平均ベクトルや共分散構造に関する統計的仮説検定の理論および現実問題への応用についての研究を行った。特に、1. 高次元データに対する平均ベクトルに関する仮説検定、2. 非正規性の下で、高次元データに対する平均ベクトル間の多変量多重比較法、3. 高次元データに対する共分散構造に関する仮説検定、に対していくつかの研究成果が得られた。

研究成果の概要(英文)：In this study, we discussed the statistical hypothesis testing for mean vector and for covariance structure in high-dimensional settings. In particular, we derived some results concerning with following 3 topics. 1. Statistical hypothesis testing for mean vector in high-dimensional settings. 2. Multivariate multiple comparisons among mean vector for high-dimensional data under non-normality. 3. Statistical hypothesis testing for covariance structure in high-dimensional settings.

研究分野：統計学

キーワード：仮説検定 多重比較 高次元漸近理論 高次元データ

## 1. 研究開始当初の背景

近年の情報化の発展に伴って、様々なデータがインターネットなどを通じて簡単に手に入るようになった。特に最近では日々新しく生まれるデータがますます巨大化し、それらのデータが互いに融合し複雑化する「ビッグデータ時代」となっている。このビッグデータに関連して、最近では遺伝子データ、画像・音声データなどといった「高次元データ」の統計解析手法の研究が盛んに行われている。このようなデータに対する平均ベクトルに関する仮説検定問題は、1 標本問題、2 標本問題については多くの研究結果があるが、一般の  $k$  標本問題に対する研究はそれほど多くはなく、特に「多変量多重比較法」に関する研究は皆無であった。

平均ベクトル間の多変量多重比較法は、例えば医学・薬学の分野で臨床試験データ解析に用いられるなど、自然科学や社会科学を問わず様々な分野で用いられる非常に重要な統計手法である。この問題は、通常、同時信頼区間の形を用いて議論されるため、「平均ベクトル間の比較に関する同時信頼区間の構成」が重要な問題となり、総標本数が次元数よりも多いという大標本枠組みの下では様々な研究が行われている。しかしながら、それらの結果は次元数が総標本数よりも大きいという「高次元枠組み」の下では使用することができないため、高次元データに対して使用可能な新たな理論と方法論の開発が望まれていた。

また、多変量解析手法には共分散構造の仮定が必要なものがある。そのため、共分散構造に関する統計的仮説検定は重要な問題であり、近年、高次元データに対する理論研究も多く行われていた。

## 2. 研究の目的

本研究では、これまでの研究成果を踏まえ、多変量統計解析理論の分野である複数母集団における多重比較法および関連する問題の統計理論に焦点を絞り、高次元データに対する平均ベクトルや共分散構造に関する統計的仮説検定の理論および現実問題への応用について以下の研究を行った。

(1) 高次元データに対する新たな平均ベクトルに関する仮説検定方式の提案。

(2) 非正規性の下で、高次元枠組みにおける平均ベクトル間の多変量多重比較法に対する新たな手法の提案。

(3) 高次元データに対する新たな共分散構造に関する仮説検定方式の提案。

## 3. 研究の方法

研究の目的(1)~(3)に対して、以下の方法で研究を遂行した。

(1) 高次元枠組みの下で、平均ベクトルの同等性問題に対して、既存の検定統計量を組

合わせることによって、検出力のより高い新たな検定統計量を提案する。また、プロフィール分析に対して、デンプスター型統計量を提案し、高次元データに対する検定方式を提案する。

(2) 非正規性の下で、平均ベクトル間の多変量多重比較法に対して、検定統計量としてデンプスター型統計量を提案する。また、提案統計量の極限分布を導出し、多重検定方式を提案する。

(3) 正規性および非正規性の下で、高次元データに対して共分散構造が「ブロック対角構造」を持つかどうかの検定統計量として、フロベニウスノルムに基づく統計量を提案する。また、提案統計量の極限分布を導出し、この検定問題に対する検定方式を提案する。

## 4. 研究成果

研究の目的(1)~(3)に対して、期間内に以下の研究成果を得ることができた。

(1) まず、高次元ではあるが次元数が標本数よりも小さいという「大標本高次元枠組み」の下で、平均ベクトルの1標本問題を議論し、 $T^2$ 統計量とデンプスター統計量の重み付き和による新たな検定統計量を提案した。また、正規性の下で、提案統計量の近似分布を導出し、大標本高次元枠組みの下での新たな検定方式を提案した。さらに、検出力の観点から最適な重みの推定量を与え、前述の2つの統計量との検出力比較を行った結果、提案統計量の有用性を確認することができた。

次に、正規性の下で高次元データに対する平均ベクトル間のプロフィール分析を議論した。本研究では、等分散性を仮定しない場合(Behrens-Fisher問題)を考え、この問題に対してデンプスター型統計量を提案した。また、この検定統計量の近似分布を漸近展開近似によって与え、高次元枠組みの下での新たな検定方式を提案した。また、検出力を理論的に考察し、数値実験の結果と併せて、提案手法の有用性を確認することができた。

(2) 複数母集団の平均ベクトル間の多変量多重比較法について、高次元枠組みの下での検定方式について議論を行った。現実問題として高次元データを取り扱う際には、一般に正規性の仮定は厳しい制約となってしまうため、本研究では非正規性の下でデンプスター統計量に基づく最大値型統計量を提案した。実際に仮説検定を行う際には、この統計量の棄却限界値が必要となるが、この正確な値を求めることは正規性の下ですら困難である。そのため、ボンフェロニの不等式とマルチンゲール理論を用いることによって、この統計量の近似帰無分布の導出を行い、正規性が成り立たない場合における頑健性を示すことができた。

(3) 母集団の持つ共分散構造に関する仮説検定問題は重要な問題であり、球面性の検定や、対角構造かどうかの検定など、多くの研究がされている。本研究では、これらの検定問題を含む構造である「ブロック対角構造」に注目し、高次元枠組みの下でこの構造を持つかどうかの検定について議論を行った。まず、正規性の仮定の下で、フロベニウスノルムに基づく検定統計量を提案し、この統計量の近似帰無分布を導出した。さらに、近似精度と検出力を数値的に評価することによって、提案手法の有用性を確認することができた。次に、非正規性の下で新たな検定統計量を提案し、マルチンゲール理論を用いることによって、この統計量の近似分布を導出した。また、検出力を理論的に考察し、数値実験の結果と併せて、提案手法の有用性を確認することができた。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計8件)

Nobumichi Shutoh, Takahiro Nishiyama and Masashi Hyodo, Bartlett correction to the likelihood ratio test for MCAR with two-step monotone sample, *Statistica Neerlandica*, 査読有, 掲載確定, 2017.  
DOI: 10.1111/stan.12106

Masashi Hyodo and Takahiro Nishiyama, A one-sample location test based on weighted averaging of two test statistics when the dimension and the sample size are large, *Communications in Statistics - Theory and Methods*, 査読有, vol.46, 2017, 3526-3541.  
DOI: 10.1080/03610926.2015.1066812

Yuki Yamada, Masashi Hyodo and Takahiro Nishiyama, Testing block-diagonal covariance structure for high-dimensional data under non-normality, *Journal of Multivariate Analysis*, 査読有, vol.155, 2017, 305-316.  
DOI: 10.1016/j.jmva.2016.12.009

Mizuki Onozawa, Takahiro Nishiyama and Takashi Seo, On test statistics in profile analysis with high-dimensional data, *Communications in Statistics - Simulation and Computation*, 査読有, vol.45, 2016, 3716-3743.  
DOI: 10.1080/03610918.2014.95386

Masashi Hyodo, Nobumichi Shutoh, Takahiro Nishiyama and Tatjana Pavlenko, Testing block-diagonal covariance

structure for high-dimensional data, *Statistica Neerlandica*, 査読有, vol.69, 2015, 460-482.  
DOI: 10.1111/stan.12068

兵頭 昌, 渡邊 弘己, 西山 貴弘, 共分散行列の逆行列の不偏推定量の改良, *計算機統計学*, 査読有, vol.28, 2015, 3-17.  
[https://www.jstage.jst.go.jp/article/jscswabun/28/1/28\\_3/\\_article/-char/ja/](https://www.jstage.jst.go.jp/article/jscswabun/28/1/28_3/_article/-char/ja/)

西山 貴弘, 兵頭 昌, 高次元枠組みにおける平均ベクトル間の多変量多重比較法, *日本統計学会誌*, 査読有, vol.44, 2015, 341-362.  
[https://www.jstage.jst.go.jp/article/jjsj/44/2/44\\_341/\\_article/-char/ja/](https://www.jstage.jst.go.jp/article/jjsj/44/2/44_341/_article/-char/ja/)

Takahiro Nishiyama, Masashi Hyodo and Takashi Seo, Recent developments of multivariate multiple comparisons among mean vectors, *SUT Journal of Mathematics*, 査読有, vol.50, 2014, 247-270.  
[http://www.rs.tus.ac.jp/sutjmath/\\_userdata/50-2/06-SUTJM14-32-nishiyama-survey.pdf](http://www.rs.tus.ac.jp/sutjmath/_userdata/50-2/06-SUTJM14-32-nishiyama-survey.pdf)

[学会発表](計11件)

兵頭 昌, 西山 貴弘, Simultaneous testing of the mean vector and the covariance matrix for high-dimensional data, 科研費シンポジウム「統計科学における数理と理論の最先端について」, 2016年10月29日, 北海道大学(北海道)。

首藤 信通, 西山 貴弘, 兵頭 昌, Bartlett correction to the likelihood ratio test for MCAR with two-step monotone missing data, 科研費シンポジウム「統計科学における数理と理論の最先端について」, 2016年10月29日, 北海道大学(北海道)。

兵頭 昌, 西山 貴弘, 高次元かつ大標本における平均ベクトルと共分散行列の同時検定について, 日本数学会秋季総合分科会, 2016年9月17日, 関西大学(大阪府)。

西山 貴弘, 首藤 信通, 兵頭 昌, Modified likelihood ratio test for MCAR with two-step monotone missing data, 統計関連学会連合大会, 2016年9月6日, 金沢大学(石川県)。

Takahiro Nishiyama, Yuki Yamada and Masashi Hyodo, Test for covariance structure for high-dimensional data under non-normality, International Conference on Computational Statistics, 2016年8月25日, Auditorio Principe Felipe, Oviedo, Spain.

西山 貴弘, 山田 雄紀, 兵頭 昌, 高次元枠組みにおける共分散構造に関する検定について, 科研費シンポジウム「多様な分野における統計科学の新展開」, 2015年10月25日, 富山県民会館(富山県).

Takahiro Nishiyama, Masashi Hyodo and Nobumichi Shutoh, Testing independence among populations for high-dimensional data, Eighth International Workshop on Simulation, 2015年9月22日, University of Natural Resources and Life Sciences, Vienna, Austria.

西山 貴弘, 山田 雄紀, 兵頭 昌, Test for block-diagonal covariance structure in high dimension under non-normality, 統計関連学会連合大会, 2015年9月8日, 岡山大学(岡山県).

西山 貴弘, 兵頭 昌, 首藤 信通, Test for covariance structure in high-dimensional settings, Workshop on Statistical Methods for Large Complex Data, 2014年11月12日, 筑波大学(茨城県).

西山 貴弘, 兵頭 昌, 首藤 信通, Tatjana Pavlenko, Testing the block-diagonal covariance structure for high-dimensional data, 統計関連学会連合大会, 2014年9月16日, 東京大学(東京都).

Takahiro Nishiyama and Masashi Hyodo, Pairwise comparisons among mean vectors in high dimension under non-normality, International Conference on Computational Statistics, 2014年8月19日, Centre International De Conference Geneve, Geneva, Switzerland.

〔図書〕(計0件)

〔産業財産権〕

出願状況(計0件)

取得状況(計0件)

〔その他〕

ホームページ等

<http://reach.acc.senshu-u.ac.jp/Nornir/search.do?type=v01&uid=1208470>

## 6. 研究組織

### (1) 研究代表者

西山 貴弘 (NISHIYAMA, Takahiro)

専修大学・経営学部・准教授

研究者番号: 30516472

### (2) 研究分担者

なし

### (3) 連携研究者

なし

### (4) 研究協力者

兵頭 昌 (HYODO, Masashi)