

科学研究費助成事業 研究成果報告書

平成 26 年 5 月 29 日現在

機関番号：32660

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500360

研究課題名(和文) 欠測値を含む高次元データに対するいくつかの多変量統計的推測法の開発とその応用

研究課題名(英文) Development of Some Multivariate Statistical Inference Procedures for Missing and High Dimensional Data and Its Application

研究代表者

瀬尾 隆 (Seo, Takashi)

東京理科大学・理学部・教授

研究者番号：00266909

交付決定額(研究期間全体)：(直接経費) 2,200,000円、(間接経費) 660,000円

研究成果の概要(和文)：データが欠測値を持つ場合の平均ベクトルに関する検定について、いくつかの成果を得た。特に、2ステップ単調欠測データの下で1標本問題と2標本問題に対する平均ベクトルに関する検定問題について、その尤度比検定法とホテリング型検定統計量による検定法を与え、同時信頼区間の構成に成功した。近似精度についてはモンテカルロ・シミュレーションを行い数値的に評価している。平均ベクトルと分散共分散行列の同時検定についても、2ステップ単調欠測データの下での尤度比検定について、いくつかの研究成果をあげている。高次元データについては、判別分析をはじめ、多変量正規性検定等の問題に対して、いくつかの結果を得た。

研究成果の概要(英文)：We obtained some results for the tests of mean vectors when the data have missing observations. In particular, we gave the test procedures for mean vectors using the likelihood ratio test statistic and Hotelling type test statistic for the one sample and two sample problems when the data set has as two-step monotone missing data. We also gave the simultaneous confidence intervals for any and all linear compounds of the mean. Further, the accuracy of the approximation is investigated by Monte Carlo simulation. Testing specified values for the mean vector and the covariance matrix with a two-step monotone missing data were considered and derived the likelihood ratio test statistic and some results. As for the study of high dimensional data, we obtained some results for the problem of discriminant analysis and testing for multivariate normality and so on.

研究分野：総合領域

科研費の分科・細目：情報学・統計科学

キーワード：統計数学 統計理論 多変量解析 漸近展開

1. 研究開始当初の背景

従来の統計的手法は、得られたデータに対して、以下の仮定をおくことが多い。

- ・データが正規分布に従っている。
- ・データに欠損がなく、すべてある。

さらに、多変量解析法(特に漸近理論)を議論する場合には、上記の仮定に加えて以下の仮定が設けられることが多い。

- ・変数の数(次元数)を固定したもとの、次元数に比べ標本数が十分大きい。

しかしながら、現実のデータにおいては、必ずしも、これらの仮定を満たすとは限らず、これらの仮定が成り立たない場合の統計的推測および統計的漸近理論の構築が重要な研究課題となっている。このような動機の下で、データが欠測値をもつ場合や高次元データである場合の統計的推測理論の構築と統計的検定法の開発を行う。

2. 研究の目的

上記の背景の下、平均ベクトル・分散共分散行列に関する検定および多重比較法と判別分析・多変量正規性検定における統計理論に焦点をしばり、以下の内容を研究目的として行う。

- (1) 1標本問題における2ステップ単調欠測データの下での平均ベクトルのホテリング型検定統計量と尤度比検定統計量の帰無分布の漸近的性質を明らかにする。
- (2) (1)の研究結果を2標本問題に拡張する具体的には、2つの正規母集団の平均ベクトルの同等性検定に対する検定統計量を導出し、特に、2つの母集団の分散共分散行列が異なる場合への拡張、さらには、プロフィール分析への拡張を行う。また、単調欠測データのもとで、線形判別分析における誤判別確率の推定についても同様の議論を行い、さらには、データが正規分布であるかどうかの検定問題について議論し、検定統計量の提案および数値的評価を行う。

3. 研究の方法

- (1) 先行研究である学術論文やこれまでに得た研究成果や導出法などを参考に、現在考えている問題へ適用する。
- (2) 数式処理ソフトを用いて、統計量の計算や計算チェックを行う。
- (3) 乱数を用いた大規模なモンテカルロ・シミュレーションを行うことによって理論的結果の検証と数値的評価を行う。

4. 研究成果

- (1) 平成23年度
 - (a) 観測ベクトルの成分が一部分欠損している2ステップ単調欠測データの下で、1標本問題における平均ベクトルの尤度比検定と T^2 型検定統計量によるものを与え、それらの帰無分布の近似について良い近似法を開発した。さらに、この問題を2標本問題へと拡張することに成功した。
 - (b) (a)の議論をプロフィール分析問題に適用することを考え、2ステップ単調欠測データの下でプロフィール分析に対する尤度比検定統計量と T^2 型検定統計量を導出し、それらの帰無分布の数値評価を行った。
 - (c) 高次元データの下での判別分析に関する検定問題について、その検定法の開発と性質についていくつかの研究成果を与えた。さらに、母集団間に相関のある繰り返し測定データなどの平均ベクトルの検定統計量について、その帰無分布に対する漸近展開近似を楕円母集団のもとで導出し、非正規性の影響について調べた。また、関連する研究として、多次元データの正規性検定についても、いくつかの研究成果を得ることができた。
- (2) 平成24年度
 - (a) 平成23年度に得た2ステップ単調欠測データの下での1標本問題および2標本問題における平均ベクトルの検定についての研究成果を一般の k 標本問題に拡張することを考え、平均ベクトル間の対比較に対する近似同時信頼区間を構成することに成功した。
 - (b) 標本数よりも次元数の方が大きい場合の高次元大標本理論の下でのプロフィール分析について、平行性仮説検定問題に対する検定統計量の提案とその帰無分布に対する漸近展開の導出に成功した。ただし、この成果はデータに欠測値はなく、完全データの下での議論となっている。
 - (c) その他の研究成果として、多変量正規性検定問題について、多変量歪度と多変量尖度を用いた検定統計量をいくつか提案し、その統計量の帰無分布を与え、さらに、高次元データの下での多群判別分析におけるモデル変数選択基準の提案や新たな修正線形判別分析などについていくつかの研究成果を得た。結果は学術雑誌に掲載されている。
- (3) 平成25年度
 - (a) 1標本問題における平均ベクトルに関する検定について、3ステップ単調欠測データの下で、検定統計量の導出と帰無

分布に対する近似上側パーセント点について議論し、標本数が少ない場合にも良い近似パーセント点を得ることに成功した。また大規模なモンテカルロ・シミュレーションを行い、近似精度について数値評価を行った。

- (b) 平成24年度の(a)の応用として、2ステップ単調欠測データの場合を含む平均成分の同等性検定について、平均成分間の近似同時信頼区間の導出に成功した。さらに、2ステップ単調欠測データの下で、平均ベクトルと分散共分散行列の同時検定についても議論し、尤度比検定統計量の導出と帰無分布に対する近似上側パーセントを提案し、近似精度について数値評価を与えている。
- (c) 高次元データに対する問題については、歪度と尖度を用いた多変量正規性検定について新たな検定法とその性質を与え、また、プロフィール分析における並行性仮説検定などの検定問題に対して、検定統計量とその分布に対する漸近展開式の導出に成功した。

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16件)

Onozawa, M., Takahashi, S. and Seo, T. (2014). Tests for profile analysis based on two-step monotone missing data. *Discussiones Mathematicae Probability and Statistics*, **33**, 171–190. 査読有

Okamoto, N. and Seo, T. (2014). Multivariate multiple comparisons with a control in elliptical populations. *Discussiones Mathematicae Probability and Statistics*, **33**, 191–205. 査読有

Yamada, T., Hyodo, M. and Seo, T. (2013). The asymptotic approximation of EPMC for linear discriminant rules using a Moore-Penrose inverse matrix in high dimension. *Communications in Statistics - Theory and Methods*, **42**, 2629–2638. 査読有

Takahashi, S., Nishiyama, T. and Seo, T. (2013). Pairwise comparisons among components of mean vector in elliptical distributions. *American Journal of Mathematical and Management Sciences*, **32**, 53–74. 査読有

Nishiyama, T., Hyodo, M. Seo, T. and Pavlenko, T. (2013) Testing linear hypotheses of mean vector for high-dimensional data with unequal covariance matrices. *Journal of*

Statistical Planning and Inference, **143**, 1898–1911. 査読有

Enomoto, R., Okamoto, N. and Seo, T. (2013). On the asymptotic normality of test statistics using Song's kurtosis. *Journal of Statistical Theory and Practice*, **7**, 107–119. 査読有

Takahashi, S., Nishiyama, T. and Seo, T. (2012). Simultaneous tests for independence among components of random vector by step-down multiple comparison procedure. *Communications in Statistics - Simulation and Computation*, **41**, 1998–2005. 査読有

Shutoh, N., Hyodo, M., Seo, T. and Pavlenko, T. (2012). Constrained linear discriminant rule via the Studentized classification statistic based on monotone missing data. *SUT Journal of Mathematics*, **48**, 55–69. 査読有

Seko, N., Yamazaki, A. and Seo, T. (2012). Tests for mean vector with two-step monotone missing data. *SUT Journal of Mathematics*, **48**, 13–36. 査読有

Hyodo, M., Yamada, T., Himeno, T. and Seo, T. (2012). A modified linear discriminant analysis for high-dimensional data. *Hiroshima Mathematical Journal*, **42**, 209–231. 査読有

Hyodo, M. and Seo, T. (2012) A model selection criterion for discriminant analysis of several groups when the dimension is larger than the total sample size. *Communications in Statistics - Theory and Methods*, **41**, 2419–2436. 査読有

Enomoto, R., Okamoto, N. and Seo, T. (2012). Multivariate normality test using Srivastava's skewness and kurtosis. *SUT Journal of Mathematics*, **48**, 103–115. 査読有

Shutoh, N., Hyodo, M. and Seo, T. (2011). An asymptotic approximation for EPMC in linear discriminant analysis based on two-step monotone missing samples. *Journal of Multivariate Analysis*, **102**, 252–263. 査読有

Seko, N., Kawasaki, T. and Seo, T. (2011). Testing equality of two mean vectors with two-step monotone missing data. *American Journal of Mathematical and Management Sciences*, **31**, 117–135. 査読有

Miyagawa, C., Koizumi, K. and Seo, T. (2011). A new multivariate kurtosis

and its asymptotic distribution. *SUT Journal of Mathematics*, **47**, 55–71. 査読有

Seo, T. and Ariga, M. (2011). On the distribution of sample measure of multivariate kurtosis. *Journal of Combinatorics, Information & System Sciences*, **36**, 179–200. 査読有

[学会発表](計 56 件)

八木 文香, 瀬尾 隆, A Test for Equality of Two Mean Vectors with Three-step Monotone Missing Data, 日本数学会, 2014 年 3 月 17 日, 学習院大学.

細谷 美貴, 八木 文香, 瀬尾 隆, 2-step 単調欠測データにおける分散共分散行列と平均ベクトルの同時検定, 第 13 回西東京統計研究会, 2014 年 3 月 11 日, 中央大学.

篠崎 絢, 岡本 直也, 瀬尾 隆, 関連のある楕円母集団の下でのプロフィール分析, 第 13 回西東京統計研究会, 2014 年 3 月 11 日, 中央大学.

川崎 玉恵, 瀬尾 隆, 分散共分散行列が異なる平均ベクトルの検定に対する検定統計量の近似分布とその評価, 第 8 回日本統計学会春季集会, 2014 年 3 月 8 日, 同志社大学.

細谷 美貴, 瀬尾 隆, Simultaneous Tests of Covariance Matrix and Mean Vector for Two-step Monotone Missing Data, 第 8 回日本統計学会春季集会, 2014 年 3 月 8 日, 同志社大学.

大竹 杏奈, 瀬尾 隆, 2-step 単調欠測データの下での平均成分の一様性検定について

第 8 回日本統計学会春季集会, 2014 年 3 月 8 日, 同志社大学.

脇村 真隆, 三ツ井 誠, 小泉 和之, 瀬尾 隆, ブロック対角化構造法による多変量歪度及び尖度について, 科研費シンポジウム「統計的推測の新展開とその応用」, 2013 年 12 月 4 日, 筑波大学.

渡邊 弘己, 兵頭 昌, 瀬尾 隆, ある高次元線形判別関数に伴う誤判別確率の漸近挙動について, 科研費シンポジウム「高次元データ解析の理論と方法論, 及び, 関連分野への応用」, 2013 年 11 月 26 日, 筑波大学.

脇村 真隆, 三ツ井 誠, 小泉 和之, 瀬尾 隆, 高次元データに対する多変量歪度及び尖度について, 日本計算機統計学会, 2013 年 11 月 16 日, 市民会館崇城大学ホール(熊本).

小川 裕一郎, 山田 隆行, 瀬尾 隆, 対角線形判別法における判別点の提案とその数値的評価, 日本計算機統計学会, 2013 年 11 月 15 日, 市民会館崇城大学ホール(熊本).

Ayaka Yagi, Takashi Seo, On the Test

for Mean Vector with Three-step Monotone Missing Data, Ishigaki International Conference on Modern Statistics Theories, Practices, and Education in the 21st Century(石垣統計国際会議), 2013 年 11 月 8 日 ~ 11 月 10 日, ANA インターコンチネンタル石垣リゾートホテル.

渡邊 弘己, 兵頭 昌, 瀬尾 隆, 高次元大標本データのための判別分析における判別分岐点の決定方法, 日本数学会, 2013 年 9 月 26 日, 愛媛大学.

兵頭 昌, 西山 貴弘, 瀬尾 隆, T^2 統計量と Dempster 統計量の線形結合による検定統計量について, 日本数学会, 2013 年 9 月 26 日, 愛媛大学.

八木 文香, 瀬尾 隆, 3-step 単調欠測データにおける平均ベクトルに関する検定と同時信頼区間, 日本数学会, 2013 年 9 月 26 日, 愛媛大学.

小野沢 瑞季, 西山 貴弘, 瀬尾 隆, 高次元データにおけるプロフィール分析とその検出力について, 統計関連学会連合大会, 2013 年 9 月 11 日, 大阪大学.

八木 文香, 瀬尾 隆, 3-step 単調欠測データの下での平均ベクトルの検定統計量について, 統計関連学会連合大会, 2013 年 9 月 10 日, 大阪大学.

川崎 玉恵, 瀬尾 隆, 多変量 Behrens-Fisher 問題に対する検定統計量の近似分布の提案, 統計関連学会連合大会, 2013 年 9 月 9 日, 大阪大学.

Junichiro Ogawa, Takayuki Yamada, Takashi Seo, A cut-off point for diagonal discriminant analysis in high dimension, 22nd International Workshop on Matrices and Statistics, 2013 年 8 月 14 日, Bahen Centre, トロント, カナダ.

川崎 玉恵, 瀬尾 隆, 多変量 Behrens-Fisher 問題の近似解について, 応用統計学会, 2013 年 5 月 25 日, パルセいいざか(福島).

小野沢 瑞季, 西山 貴弘, 瀬尾 隆, Two Sample Profile Analysis with Unequal Covariance Matrices in High Dimensional Data, 応用統計学会, 2013 年 5 月 25 日, パルセいいざか(福島).

① 小野沢 瑞季, 西山 貴弘, 瀬尾 隆, 高次元データにおける分散共分散行列が異なるもとでの平行性仮説検定について, 第 12 回西東京統計研究会, 2013 年 3 月 5 日, 東海大学高輪キャンパス.

② 兵頭 昌, 西山 貴弘, 瀬尾 隆, 母共分散行列の逆行列の推定と線形判別分析への応用, 第 12 回西東京統計研究会, 2013 年 3 月 5 日, 東海大学高輪キャンパス.

③ 篠崎 絢, 瀬尾 隆, 楕円母集団の下での

- paired T^2 統計量の分布と多重比較法について, 第 12 回西東京統計研究会, 2013 年 3 月 5 日, 東海大学高輪キャンパス.
- ②4 兵頭 昌, 奥山 愛美, 西山 貴弘, 瀬尾 隆, 高次元大標本データにおける平均ベクトルに関する検定統計量の検出力比較について, 科研費シンポジウム「統計科学の基礎的理論とその応用」, 2012 年 11 月 22 日, 国際奈良学セミナーハウス.
- ②5 小野沢 瑞季, 瀬尾 隆, 2-step 単調欠測データのもとでのプロフィール分析について, 科研費シンポジウム「統計科学の基礎的理論とその応用」, 2012 年 11 月 22 日, 国際奈良学セミナーハウス.
- ②6 澄川 琢磨, 小泉 和之, 瀬尾 隆, Cornish-Fisher 展開を用いた多変量 JB 型統計量の改良, 日本計算機統計学会, 2012 年 11 月 2 日, 東京大学.
- ②7 世古 規子, 瀬尾 隆, Tests for mean vector with two-step monotone missing data, 科研費シンポジウム「統計推測理論の展開と諸モデルへの応用」, 2012 年 10 月 3 日, 釧路市生涯学習センター.
- ②8 川崎 玉恵, 瀬尾 隆, 異なる分散共分散行列をもつ 2 つの平均ベクトルの検定に対する検定統計量の近似分布について, 科研費シンポジウム「統計推測理論の展開と諸モデルへの応用」, 2012 年 10 月 3 日, 釧路市生涯学習センター.
- ②9 川崎 玉恵, 瀬尾 隆, 分散共分散行列が異なる下での 2 標本問題における平均ベクトルの検定, 統計関連学会連合大会, 2012 年 9 月 12 日, 北海道大学.
- ③0 世古 規子, 川崎 玉恵, 瀬尾 隆, Tests for mean vector with two-step monotone missing data for k-sample problem, 統計関連学会連合大会, 2012 年 9 月 12 日, 北海道大学.
- ③1 中水流 嘉臣, 瀬尾 隆, k 標本問題における部分分散共分散行列の均一性の検定について, 統計関連学会連合大会, 2012 年 9 月 12 日, 北海道大学.
- ③2 奥山 愛美, 兵頭 昌, 瀬尾 隆, 判別分析におけるリッジ型モデル選択基準について, 統計関連学会連合大会, 2012 年 9 月 11 日, 北海道大学.
- ③3 榎本 理恵, 岡本 直也, 瀬尾 隆, Song 型多変量標本尖度による正規性検定統計量について, 統計関連学会連合大会, 2012 年 9 月 11 日, 北海道大学.
- ③4 澄川 琢磨, 小泉 和之, 瀬尾 隆, 新たな多変量尖度を用いた正規性検定統計量の提案, 統計関連学会連合大会, 2012 年 9 月 11 日, 北海道大学.
- ③5 渡邊 弘己, 兵頭 昌, 瀬尾 隆, 高次元大標本データにおける Dempster 型判別基準とその誤判別確率の制御法, 統計関連学会連合大会, 2012 年 9 月 10 日, 北海道大学.
- ③6 三谷 知海, 兵頭 昌, 瀬尾 隆, 高次元線型判別分析における CPMC 及び EPMC の区間推定, 統計関連学会連合大会, 2012 年 9 月 10 日, 北海道大学.
- ③7 小野沢 瑞季, 瀬尾 隆, 2 標本問題における 2-step 単調欠測データのもとでのプロフィール分析, 応用統計学会, 2012 年 5 月 24 日, 統計数理研究所.
- ③8 川崎 玉恵, 瀬尾 隆, 分散共分散行列が異なる場合における平均ベクトル間の同等性検定の比較, 応用統計学会, 2012 年 5 月 24 日, 統計数理研究所.
- ③9 西山 貴弘, 兵頭 昌, 瀬尾 隆, Testing equality of mean vectors for high-dimensional data with unequal covariance matrices, 日本数学会, 2012 年 3 月 28 日, 東京理科大学.
- ④0 篠崎 絢, 瀬尾 隆, The Effects of Nonnormality on Test for Equality of Pairs of Mean Vectors in Elliptical Distributions, 第 11 回西東京統計研究会, 2012 年 3 月 2 日, 統計数理研究所.
- ④1 小野沢 瑞季, 瀬尾 隆, 2-step 単調欠測データにおける平行性仮説検定統計量の漸近分布について, 第 11 回西東京統計研究会, 2012 年 3 月 2 日, 統計数理研究所.
- ④2 西山 貴弘, 兵頭 昌, 瀬尾 隆, Testing equality of mean vectors for the case of high-dimension and unequal covariance matrices, 科研費シンポジウム「高次元データの推測理論の開発と応用」, 2012 年 1 月 29 日, 駿河大記念館.
- ④3 瀬尾 隆, Asymptotic Expansions for the Distributions of Some Statistics in Multivariate Analysis, 科研費シンポジウム「多変量解析の新展開」, 2012 年 1 月 20 日, てんぶす那覇.
- ④4 世古 規子, 川崎 玉恵, 瀬尾 隆, 2-step 単調欠測データにおける平均ベクトルの T^2 型検定と尤度比検定, 科研費シンポジウム「生物統計における統計的推測とその応用」, 2011 年 12 月 11 日, 宮崎県婦人会館.
- ④5 齋藤 めい, 澄川 琢磨, 小泉 和之, 瀬尾 隆, On some tests for assessing multivariate normality based on sample moments, 科研費シンポジウム「経験尤度法と判別・分類解析の理論と応用」, 2011 年 12 月 2 日, 和歌山ビッグ愛.
- ④6 榎本 理恵, 岡本 直也, 瀬尾 隆, On the Estimation of Kurtosis Parameter with Missing Data in Elliptical Population, Joint Meeting of the Korea-Japan Conference of Computational Statistics, and the 25th Symposium of Japanese Society of Computational Statistics, 2011 年 11 月 12 日, 海雲台グランドホテル, 釜山, 大

- 韓民国 .
- ④7 小野沢 瑞季, 高橋 翔, 瀬尾 隆, Test for Parallelism Hypothesis of Several Groups with Two-step Monotone Missing Data, Joint Meeting of the Korea-Japan Conference of Computational Statistics, and the 25th Symposium of Japanese Society of Computational Statistics, 2011年11月12日, 海雲台グランドホテル, 釜山, 大韓民国 .
- ④8 福本 裕子, 首藤 信通, 瀬尾 隆, 2-step 単調欠測データのもとでの分散共分散行列の同等性検定統計量の漸近分布について, 統計関連学会連合大会, 2011年9月6日, 九州大学 .
- ④9 榎本 理恵, 岡本 直也, 瀬尾 隆, 2-step 単調欠測データにおける多変量標本尖度について, 統計関連学会連合大会, 2011年9月5日, 九州大学 .
- ⑤0 兵頭 昌, 首藤 信通, 瀬尾 隆, Tatjana Pavlenko, Comparison of two high-dimensional linear discrimination methods, 統計関連学会連合大会, 2011年9月5日, 九州大学 .
- 51 渡邊 弘己, 兵頭 昌, 瀬尾 隆, 高次元データにおける分散共分散行列の構造に関する検定について, 統計関連学会連合大会, 2011年9月5日, 九州大学 .
- 52 川崎 玉恵, 世古 規子, 瀬尾 隆, 2 標本問題における 2-step 単調欠測データの下での平均ベクトルの同等性検定について, 統計関連学会連合大会, 2011年9月5日, 九州大学 .
- 53 小野沢 瑞季, 瀬尾 隆, On the Distributions of Some Test Statistics for Profile Analysis with Two-step Monotone Missing Data, The 7th International Conference on Multiple Comparison Procedures, 2011年8月31日, ヒルトンホテル, ワシントン DC, アメリカ .
- 54 篠崎 絢, 瀬尾 隆, Testing the Equality of Pairs of Mean Vectors and Simultaneous Confidence Intervals in Elliptical Distributions, The 7th International Conference on Multiple Comparison Procedures, 2011年8月31日, ヒルトンホテル, ワシントン DC, アメリカ .
- 55 高橋 翔, 西山 貴弘, 瀬尾 隆, Multiple Comparisons among Components of Mean Vector under an Elliptical Population, The 7th International Conference on Multiple Comparison Procedures, 2011年8月31日, ヒルトンホテル, ワシントン DC, アメリカ .
- 56 川崎 玉恵, 瀬尾 隆, Tests for Two Mean Vectors and Simultaneous Confidence Intervals with Unequal Covariance

Matrices in Two-step Monotone Missing Data, The 7th International Conference on Multiple Comparison Procedures, 2011年8月31日, ヒルトンホテル, ワシントン DC, アメリカ .

〔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〕なし

6. 研究組織

(1) 研究代表者

瀬尾 隆 (SEO TAKASHI)
東京理科大学・理学部・教授
研究者番号: 00266909

(2) 研究分担者

なし

(3) 連携研究者

首藤 信通 (SHUTOH NOBUMICHI)
神戸大学大学院・海事科学研究科・講師
研究者番号: 00634099