

科学研究費助成事業 研究成果報告書

令和 6 年 6 月 24 日現在

機関番号：32660

研究種目：基盤研究(C) (一般)

研究期間：2021～2023

課題番号：21K11795

研究課題名(和文) 欠測値を含む高次元および非正規データに対する漸近理論の開発とその応用

研究課題名(英文) Development of asymptotic theory for high-dimensional and non-normal data with missing value and its application

研究代表者

瀬尾 隆 (Seo, Takashi)

東京理科大学・理学部第一部応用数学科・教授

研究者番号：00266909

交付決定額(研究期間全体)：(直接経費) 2,000,000円

研究成果の概要(和文)：多次元データが単調型欠測データの下で、次の3つの課題：(1)多変量正規性検定問題、(2)平均ベクトルや分散共分散行列に関する検定問題、(3)成長曲線モデルの下での適合性検定問題について、それぞれ研究成果を得た。ここで単調型欠測データは、欠測構造として完全ランダム欠測を仮定している。

(1)については、主に多変量標本尖度を用いた検定統計量の導出とその正規近似の精度評価、(2)については、分散共分散行列に関する尤度比検定統計量および修正尤度比検定統計量の導出と部分平均ベクトルの検定統計量の提案とその帰無分布の導出、(3)については、検定統計量とその帰無分布に対するF近似などである。

研究成果の学術的意義や社会的意義

多変量統計解析の理論研究、特に検定問題において、検定統計量の分布の導出は容易ではなく、近似によるものがほとんどである。そのような背景の下、より良い近似分布、すなわち、近似上側パーセント点を与えることが重要な問題となる。本研究は、欠測データや多変量正規性が成り立たないデータ、高次元データに対して、特に、単調型欠測データの下でのいくつかの検定問題に対する検定統計量の帰無分布の漸近展開近似や近似精度のよい変換統計量、正規近似、F近似などを導出しており、この結果は学術的に大きな意義があり、実データへの応用につながるもので社会的意義のあるものと確信している。

研究成果の概要(英文)：When the type of multi-dimensional data has a monotone pattern of missing observations, the research achievements for the following three problems were obtained: (1) test for multivariate normality, (2) test for mean vectors and test for covariance matrix, and (3) test for adequacy in growth curve model. We assume that the data are missing completely at random (MCAR). The main topics of (1) are the test statistic using multivariate sample kurtosis and the accuracy of normal approximation, (2) the likelihood ratio test statistic and the modified likelihood ratio test statistic for the covariance matrix, the test statistic for the sub-mean vector and the derivation of their null distributions, and (3) the test statistic for adequacy in growth curve model and F-approximation for its null distribution.

研究分野：数理統計学

キーワード：多変量解析 欠測データ 正規近似 漸近展開

様式 C - 19、F - 19 - 1 (共通)

1. 研究開始当初の背景

本研究は、統計科学における数理データサイエンス分野の基盤となる新たな統計理論を展開する研究である。その中で特に多次元データの母集団が多変量正規分布に従っているかどうかを調べる多変量正規性検定問題、多次元データの特徴を表す平均ベクトルや分散共分散行列に関する検定問題、そしてより現実的なデータ構造である成長曲線モデルというモデルの下での適合性検定問題について、高次元データや欠測値が生じた場合などのより現実的な場面を想定した統計的手法の開発を行う研究である。

2. 研究の目的

上記の背景の下、(1)多変量正規性検定問題、(2)平均ベクトル及び分散共分散行列の構造に関する検定問題、(3)成長曲線モデルの下での適合性検定及び平均パラメータなどの検定問題について、以下のような状況(仮定)の下で、新たな統計理論及び統計的手法の開発を行う。与えられている多次元データが欠測を含む場合や、母集団分布が多変量正規分布でない非正規分布である場合、さらには高次元データである場合、そしてこれらを融合した場合について議論し、検定統計量の分布やその性質などについて研究を行う。

3. 研究の方法

多変量正規性検定問題については、これまでの研究成果として、特に多変量標本尖度を用いた検定統計量の正規近似に関して、正規化変換統計量を提案しているが、これを改良する検定統計量の提案、およびその発展として、データが単調型欠測構造をもつ場合に適用できる検定統計量を導出し、その近似分布について研究を行う。また、欠測データの下での分散共分散行列の検定に関する研究についても並行して実施していく。計画としては随時、結果が得られ次第、学会発表、論文投稿を行う予定である。成長曲線モデルについては、これまでの最尤推定量などの研究成果を基盤として検定問題への適用が実施可能であると考えており、まずはこれまでの成果を再確認し議論を進めていく。これらの研究においては膨大な統計計算を行う必要があり、研究を円滑に実施するために、大学院生の補助・協力を得る予定である。またこの分野に詳しい国内外の研究者と連絡を取り合い、レビューなどを受け進めていく。また、理論結果の検証が必要不可欠であり、そのために大規模なモンテカルロ・シミュレーションを行う。

4. 研究成果

(1) 令和3年度

大きく2つの課題である、多変量正規性検定問題、平均ベクトルや分散共分散行列に関する検定問題についていくつかの成果を得た。

多次元データの母集団が多変量正規分布であるかどうかを検定する多変量正規性問題において、先行研究にある完全データの下での多変量標本尖度を用いた検定統計量に着目し、データ構造として単調型欠測データ、特に2ステップ単調欠測データに対する検定統計量の提案を行った。さらに、その期待値と分散に対する漸近展開近似を用いて評価することによって、標準化した検定統計量を導出し、モンテカルロ・シミュレーションにより高次元の場合を含めて正規近似に対する近似精度の数値的評価を与えた。

平均ベクトルの検定問題については、2標本問題に対する平均ベクトルの同等性検定問題を一般の単調欠測データの下で議論し、検定統計量の帰無分布に対する漸近展開及び近似上側パーセント点を与えることができた。さらにカイ二乗近似を改良する変換統計量も導出した。

分散共分散行列の検定問題については、2ステップ単調欠測データの下で、分散成分がすべて等しく、共分散成分がすべてゼロであるスフェリシティ構造を帰無仮説とする検定問題、そして、2標本問題における分散共分散行列が異なる場合の部分平均ベクトルの検定問題について、それぞれ検定統計量の導出とその近似上側パーセント点を与えた。最後にモンテカルロ・シミュレーションを行い、近似精度の評価を行った。

(2) 令和4年度

令和3年度に引き続き、大きく分けて以下の3つの研究成果を得た。

多変量標本尖度を用いた多変量正規性検定統計量については、令和3年度に得た2ステップ単調型欠測の結果を3ステップ単調型欠測に拡張した結果を得ることができた。また、完全データの下で、多変量尖度による検定統計量として正規化変換統計量が知られているがそのさらなる改良となる検定統計量を提案し、モンテカルロ・シミュレーションを通して、帰無分布の正規近似評価、そして、いくつかの非正規分布による検出力の数値的評価を与えた。さらに多変量歪度を用いた検定統計量の帰無分布について、自由度調整による新たなカイ二乗近似を提案した。

平均ベクトルと分散共分散行列の同時検定について、2ステップ単調欠測データの場合について議論し、尤度比検定統計量のバートレット補正を導出した。また、部分平均ベクトルの検定問題については、2ステップ単調欠測データの下での Rao の U 型検定統計量とその帰無分布に対する漸近展開近似を導出し、モンテカルロ・シミュレーションにより、近似精度の評価を行った。

成長曲線モデルの適合性検定問題について、先行研究にある完全データの場合の検定統計量を基に2ステップ単調型欠測の場合に拡張することを考え、検定統計量の提案と、その近似帰無分布として検定統計量の期待値に対する漸近展開近似を与えることによって、自由度調整による F 近似を与えた。

(3) 令和5年度

令和3年度と令和4年度に引き続き、以下の3つの研究成果を得た。

多変量正規性検定問題については、令和4年度の で述べた多変量標本歪度を用いた検定統計量のカイ二乗近似を改良し、統計量の性質を与えた。また、多変量尖度と多変量歪度を組み合わせた MJB 統計量の修正統計量を与え、モンテカルロ・シミュレーションによる数値評価を行った。

令和3年度の で述べた分散共分散行列のスフェリシティ構造に関する検定問題について、2ステップ単調型欠測から一般の単調欠測の場合に拡張することに成功した。関連して令和4年度の で述べた、平均ベクトルと共分散行列の同時検定について、局所対立仮説に対する検定統計量の分布の漸近展開を与え、検出力の評価を行った。さらに部分平均ベクトルの検定については、仮説のパターンが異なるいくつかの検定問題を考え、それらに対する Rao の U 型検定統計量の分布の漸近展開及び変換統計量を導出した。

令和4年度の の続きとして、検定統計量の期待値ばかりでなく分散に対する漸近展開(漸近分散)を与えることによって、新たな自由度調整による F 近似を与え、モンテカルロ・シミュレーションによって近似精度の数値評価を行った。

今後の展望として、本研究において議論が不十分であった非正規母集団、特に楕円分布などの下でのこれらの議論、また高次元データの議論を行う必要がある。これらの研究は、国内外を通して多くの研究者が取り組んでおり、本研究は欠測データ及び成長曲線モデルに関するものの発展・拡張となるものである。特に、欠測構造にランダム性と単調性の仮定をおくことにより解析的な解を与える理論研究という位置づけであり、今後さらに発展させていく予定である。

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 5件 / うち国際共著 2件 / うちオープンアクセス 1件）

1. 著者名 Ayaka Yagi, Takashi Seo, Yasunori Fujikoshi	4. 巻 41
2. 論文標題 AIC for Growth Curve Model with Monotone Missing Data	5. 発行年 2022年
3. 雑誌名 American Journal of Mathematical and Management Sciences	6. 最初と最後の頁 185 ~ 199
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/01966324.2021.1946667	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Ayaka Yagi, Takashi Seo, Zofia Hanusz	4. 巻 52
2. 論文標題 Testing equality of two mean vectors with monotone incomplete data	5. 発行年 2023年
3. 雑誌名 COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION	6. 最初と最後の頁 506 ~ 522
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/03610918.2020.1861291	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 Eri Kurita, Takashi Seo, Zofia Hanusz	4. 巻 1
2. 論文標題 A Modified Normalizing Transformation Statistic Based on Kurtosis for Testing Multivariate Normality -Empirical Power-	5. 発行年 2023年
3. 雑誌名 Technical Report, Statistical Research Group, Hiroshima University	6. 最初と最後の頁 1 ~ 25
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 Eri Kurita, Takashi Seo	4. 巻 188
2. 論文標題 Multivariate normality test based on kurtosis with two-step monotone missing data	5. 発行年 2022年
3. 雑誌名 JOURNAL OF MULTIVARIATE ANALYSIS	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.jmva.2021.104824	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Remi Sakai, Ayaka Yagi, Takashi Seo	4. 巻 18
2. 論文標題 Simultaneous Tests for Mean Vectors and Covariance Matrices with Three-Step Monotone Missing Data	5. 発行年 2024年
3. 雑誌名 Journal of Statistical Theory and Practice	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s42519-023-00355-2	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Yuichiro Saeki, Takashi Seo, Hiroto Hyakutake	4. 巻 52
2. 論文標題 Testing parallelism and confidence intervals of level difference in an intraclass correlation model with monotone missing data	5. 発行年 2023年
3. 雑誌名 COMMUNICATIONS IN STATISTICS-THEORY AND METHODS	6. 最初と最後の頁 6147-6160
掲載論文のDOI (デジタルオブジェクト識別子) 10.1080/03610926.2022.2026961	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計29件 (うち招待講演 0件 / うち国際学会 3件)

1. 発表者名 栗田 絵梨, 瀬尾 隆, Zofia Hanusz
2. 発表標題 A modified multivariate kurtosis test statistic for multivariate normality
3. 学会等名 応用統計学会
4. 発表年 2022年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 A test for multivariate normality based on kurtosis with three-step monotone missing data
3. 学会等名 日本計算機統計学会
4. 発表年 2022年

1. 発表者名 八木 文香, 瀬尾 隆
2. 発表標題 Multivariate pairwise comparisons among mean vectors with monotone missing data
3. 学会等名 日本数学会
4. 発表年 2022年

1. 発表者名 橋田 航平, 八木 文香, 瀬尾 隆
2. 発表標題 2-step単調欠測データの下での平均ベクトルと分散共分散行列の同時検定に対する尤度比検定統計量のパートレット補正
3. 学会等名 日本計算機統計学会
4. 発表年 2022年

1. 発表者名 細沼 璃玖, 川崎 玉恵, 瀬尾 隆
2. 発表標題 2-step 単調型欠測をもつ部分平均ベクトルの検定に対する検定統計量について
3. 学会等名 日本計算機統計学会
4. 発表年 2022年

1. 発表者名 栗田絵梨, 瀬尾 隆
2. 発表標題 Approximate Null Distribution of the Test Statistic for Multivariate Skewness
3. 学会等名 日本計算機統計学会
4. 発表年 2022年

1. 発表者名 尾崎 冬弥, 八木 文香, 瀬尾 隆
2. 発表標題 A Test for Adequacy in Growth Curve Model with Two-step Monotone Missing Data
3. 学会等名 日本計算機統計学会
4. 発表年 2022年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 3-step 単調欠測データの下での多変量尖度の検定
3. 学会等名 統計関連学会連合大会
4. 発表年 2022年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 多変量歪度を用いた多変量正規性検定統計量の帰無分布について
3. 学会等名 日本数学会
4. 発表年 2023年

1. 発表者名 八木 文香, 瀬尾 隆
2. 発表標題 単調型欠測をもつ成長曲線モデルにおける推定問題とその周辺
3. 学会等名 研究集会『多変量統計学・統計的モデル選択の新展開』
4. 発表年 2023年

1. 発表者名 八木 文香, 尾崎 冬弥, 瀬尾 隆
2. 発表標題 On the test for adequacy in growth curve model with two-step monotone missing data
3. 学会等名 日本数学会
4. 発表年 2023年

1. 発表者名 菊地 武生, 八木 文香, 瀬尾 隆
2. 発表標題 2-step 単調欠測データの下での分散共分散行列におけるスフェリシティ検定
3. 学会等名 応用統計学会
4. 発表年 2021年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 2-step単調欠測データの下での多変量尖度を用いた多変量正規性検定
3. 学会等名 統計関連学会連合大会
4. 発表年 2021年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 2-step単調型欠測をもつ多変量尖度による多変量正規性検定統計量について
3. 学会等名 日本数学会
4. 発表年 2021年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 2-step単調欠測データにおけるMardia型多変量標本尖度
3. 学会等名 日本計算機統計学会
4. 発表年 2021年

1. 発表者名 八木 文香, 瀬尾 隆
2. 発表標題 単調型欠測をもつ2つの平均ベクトルに対する同等性検定問題
3. 学会等名 日本計算機統計学会
4. 発表年 2021年

1. 発表者名 Tamae Kawasaki, Takashi Seo
2. 発表標題 T^2 type test statistic and simultaneous confidence intervals for sub-mean vectors in two-sample problem
3. 学会等名 14th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics 2021) (国際学会)
4. 発表年 2021年

1. 発表者名 Eri Kurita, Takashi Seo and Zofia Hanusz
2. 発表標題 A test statistic based on normalized transformation for multivariate sample kurtosis
3. 学会等名 IASC-ARS2022 (The 11th Conference of the IASC-ARS) (国際学会)
4. 発表年 2022年

1. 発表者名 佐藤 哲也, 八木 文香, 瀬尾 隆
2. 発表標題 単調欠測データの下でのスフェリシティ検定における検定統計量の帰無分布に対する漸近展開
3. 学会等名 RIMS 共同研究による研究会「確率モデルと統計的推測」
4. 発表年 2024年

1. 発表者名 森本 直紀, 八木 文香, 瀬尾 隆
2. 発表標題 3-step 単調欠測データにおける T^2 型検定統計量を用いた平均ベクトルの検定
3. 学会等名 科学研究費シンポジウム「統計科学・機械学習・情報数学の最前線」
4. 発表年 2024年

1. 発表者名 佐藤 哲也, 八木 文香, 瀬尾 隆
2. 発表標題 An Asymptotic Expansion of the Null Distribution for Sphericity Test Statistic with Monotone Missing Data
3. 学会等名 東京理科大学 研究推進機構 総合研究院 (RIST) データサイエンス医療研究部門・デジタルトランスフォーメーション研究部門合同シンポジウム
4. 発表年 2023年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 多次元データにおける多変量歪度検定について
3. 学会等名 東京理科大学 研究推進機構 総合研究院 (RIST) データサイエンス医療研究部門・デジタルトランスフォーメーション研究部門合同シンポジウム
4. 発表年 2023年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 Multivariate kurtosis test for multivariate normality with monotone missing data
3. 学会等名 日本数学会
4. 発表年 2023年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 Modified multivariate Jarque-Bera test statistics for multivariate normality test
3. 学会等名 統計関連学会連合大会
4. 発表年 2023年

1. 発表者名 Tamae Kawasaki, Takashi Seo
2. 発表標題 On the asymptotic distribution of mean slippage test statistic in multivariate normal data
3. 学会等名 The 2023 Joint Statistical Meetings (JSM2023) (国際学会)
4. 発表年 2023年

1. 発表者名 栗田 絵梨, 瀬尾 隆
2. 発表標題 多変量歪度に対する検定統計量の分布のカイ二乗近似
3. 学会等名 日本計算機統計学会
4. 発表年 2023年

1. 発表者名 松尾 直紀, 八木 文香, 瀬尾 隆
2. 発表標題 成長曲線モデルにおける平均構造の適合性検定
3. 学会等名 日本計算機統計学会
4. 発表年 2023年

1. 発表者名 佐藤 哲也, 八木 文香, 瀬尾 隆
2. 発表標題 An Asymptotic Expansion of the Distribution for Sphericity Test Statistic with Monotone Missing Data
3. 学会等名 日本計算機統計学会
4. 発表年 2023年

1. 発表者名 細沼 璃玖, 川崎 玉恵, 瀬尾 隆
2. 発表標題 2-step 単調型欠測をもつ2つのタイプの部分平均ベクトルの検定について
3. 学会等名 日本計算機統計学会
4. 発表年 2023年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------