

令和 6 年 6 月 21 日現在

機関番号：32660

研究種目：基盤研究(C)（一般）

研究期間：2018～2023

課題番号：18K11198

研究課題名（和文）多変量楕円母集団の下での新たな検定手法の開発とその応用

研究課題名（英文）New test procedures under multivariate elliptical population and its application

研究代表者

岩下 登志也（Iwashita, Toshiya）

東京理科大学・教養教育研究院野田キャンパス教養部・教授

研究者番号：20266919

交付決定額（研究期間全体）：（直接経費） 800,000円

研究成果の概要（和文）：楕円対称性 - 等確率面が楕円面 - の検定方法は、漸近正規性を利用して近似的な検定方法が提案されていた。本研究では、楕円対称性を検定するための検定統計量で、正確な分布をもつような統計量を考案し、その正確な分布がStiefel多様体上の一様分布であることを明らかにした。そしてその検定統計量に基づく検定手順の構築し、国内外の学会で報告、学術雑誌に2編の論文を発表することができた。

研究成果の学術的意義や社会的意義

正規性を仮定しない多変量解析で母集団モデルとして頻繁に用いられる楕円分布の本質にたどり着けることができた。従来では、検定統計量の分布を、漸近展開を用いて近似式を求めることが主流であったが、本研究では、正確な分布を導き出し、かつ、その本質がStiefel多様体上の一様分布であることを突き止めた。そして、多くの研究者が試みていた「データが楕円母集団から採取されたか否か」の問題に対する一つの解答を与えることができた。これは、複雑化するデータを解析する一助になるものと考えられる。

研究成果の概要（英文）：Conventional methods for testing elliptical symmetry have been based on asymptotic normality. In this research, we consider a test statistic for elliptical symmetry that has an exact distribution, and derive the exact distribution of which distribution is a uniform distribution on the Stiefel manifold. We also try to construct a procedure based on the test statistics we obtain, and we present them at domestic and international conferences, in addition, the result is published as a paper in an academic journal.

研究分野：多変量解析

キーワード：楕円分布 楕円対称性 Stiefel多様体

1. 研究開始当初の背景

採られた標本が楕円母集団から採られたが否か、という問題について、サンプル $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$ の Student 化した残差 $\hat{\mathbf{X}}_i = \Sigma^{-1/2}(\mathbf{X}_i - \boldsymbol{\mu})$ ($\boldsymbol{\mu}, \Sigma$ はそれぞれ母平均ベクトル, 母分散共分散行列) を用いて母集団分布の分布特定を行う。母平均ベクトル・母分散共分散行列が未知である場合, 未知母数をその (不偏) 推定量で置き換えた Student 化した残差である $\mathbf{Y}_i = S^{-1/2}(\mathbf{X}_i - \bar{\mathbf{X}})$ ($\bar{\mathbf{X}}, S$ は標本平均ベクトル, 標本分散共分散行列) を大標本理論のもと広く利用されてきた。しかしながら, Fang and Zhang (1990, p. 186) が指摘するように $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_N$ は独立ではない, \mathbf{Y}_i の分布は, $\hat{\mathbf{X}}_i$ の分布と一致しないという大きな問題があった。そのため $\hat{\mathbf{Y}}_i$ を用いた検定法は, 大標本の場合でも検定の精度について問題があった。

2. 研究の目的

Iwashita and Klar (2014) で証明するに至らなかった確率行列 Y または Y' の分布の特定とそれを利用した楕円対称性の検定問題の解決が本研究の主たる目的であった。 \mathbf{Y}_i が独立ではないことから, \mathbf{Y}_i の同時分布, すなわち $N \times p$ 確率行列 $Y' = [\mathbf{Y}_1 \mathbf{Y}_2 \dots \mathbf{Y}_N]' = QX'S^{-1/2}$, $X = [\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_N]$, $Q = I_N - (1/N)\mathbf{1}_N\mathbf{1}_N'$ の分布を (i) **大標本近似を利用せず求め**, (ii) サンプルが楕円母集団から採取されたか否かを検定するための**検定統計量の開発**とそれに基づく**検定手順を考案**すること, その検定方法の (iii) **性能評価**を行い, 提案する検定手法が利用可能なものであるかを考察することが本研究の目的である。

3. 研究の方法

Student 化した残差の正確な同時分布, すなわち Y の分布を導出するために必要な文献収集を行った。Fang and Zhang (1990) で提供されている参考文献は, 非常に有益であった。 Y の正確な分布が導出されたのち, 本補助金で購入した高速演算可能なワークステーションにより大規模な数値実験を行った。

4. 研究の成果

(1) $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N \stackrel{iid}{\sim} EC_p(\boldsymbol{\mu}, \Lambda)$ であるとき, 楕円母集団より得られた標本のスケール化された残差統計量 $\mathbf{Y}_i = S^{-1/2}(\mathbf{X}_i - \bar{\mathbf{X}})$ ($i = 1, 2, \dots, N$) ここに, $\bar{\mathbf{X}} = \sum_{i=1}^n \mathbf{X}_i$, $S = \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})'$ ($n = N - 1$) により構成される $N \times p$ 確率行列 $Y' = [\mathbf{Y}_1 \mathbf{Y}_2 \dots \mathbf{Y}_N]' = QX'S^{-1/2}$ ここに, $X = [\mathbf{X}_1 \mathbf{X}_2 \dots \mathbf{X}_N]$, $Q = I_N - (1/N)\mathbf{1}_N\mathbf{1}_N'$ が Stiefel 多様体上の一様分布 $\text{Unif}(\mathcal{U}_{n,p})$ に従うことを示すことを試みた。定数行列 Q の直交射影行列による分解を利用するアイデア $Q = KK'$, $K'K = I_n$, $K'\mathbf{1}_N = \mathbf{0}$ により $XQX' = nS$, Student 化した残差の一部である $Z = K'X'S^{-1/2}$ が左右直交変換分布不変である *Matrix Spherical distribution* に従うこと, このことから $U = Z(Z'Z)^{-1/2} = (1/\sqrt{n})Z$ が, 正確に Stiefel 多様体上の一様分布 $\text{Unif}(\mathcal{U}_{n,p})$ に従うことの証明が完了した。Stiefel 多様体上の一様性検定について, いくつかの先行研究があるが, いずれも我々の研究には応用できないことが既に判明していた。そこで, \mathbb{R}^p の球面 \mathbb{S}^{d-1} 上の一様性検定を非母数統計解析で用いられる V 統計量を応用することを考えた。 $\mathbf{U}_1, \mathbf{U}_2, \dots, \mathbf{U}_m$ を球面 \mathbb{S}^{d-1} 上の一様分布に従う確率ベクトルとする。この仮定の下で $V = (1/m^2) \sum_{i=1}^m \sum_{j=1}^m h(\mathbf{U}_i' \mathbf{U}_j) = (1/m^2) \sum_{i=1}^m \sum_{j=1}^m h(\Theta_{ij})$ (Θ_{ij} は, \mathbf{U}_i と \mathbf{U}_j のなす角) を検定統計量として, 母集団の楕円対称性について, 数値実験により提案する検定手法の有意性について確認をした。また, アヤメの実データについても我々の検定方法の有効性を検証した。検証数値実験結果の一部は以下の Table 1-3 である。

(2) 本研究の課題とは直接的な関連はないが, 楕円分布族には裾の重い分布が含まれている。この裾の重みに関する漸近特性について, 海外研究協力者から研究の補助依頼があったため, 本研究課題の一部として協力をした。非負値をとる確率変数 X, X_1, X_2 に対して関数 $g_X(d) = E[|X_1 - X_2| / (X_1 + X_2) | X_1 + X_2 > d]$ を定義する。 X_1 と X_2 が同等の値をとるような場合, この関数 $g_X(d)$ は 0 に近い値をとるようになり, X_1 また X_2 のいずれかが $X_1 + X_2$ のとる値に近いとき, 大きな d に対して, 1 に近い値をとるようになる。この性質を利用して裾の長い分布, 例えば, 対数正規分布 (球形分布の一種) やワイブル分布に対して $d \rightarrow \infty$ のとき, $g_X(d) \rightarrow 1$ となることが, 先行研究で示されている。母集団分布が *log-convex* であるか *log-concave* であるかを $\limsup_{d \rightarrow \infty} g_X(d)$ が 1/2 を閾値として判定できることを, 海外研究協力者である PD Dr Klar 氏より情報が与えられ, これをさらに発展させて, 雨量に関する実データの解析することを提案された。岩下の本研究の貢献は, $X_i (i = 1, 2)$ がそれぞれ独立にパラメータが (α_i, β) のガンマ分布に従うとき, $E[|X_1 - X_2| / (X_1 + X_2)]$ および

$E[(|X_1 - X_2|/(X_1 + X_2))^2]$ の数式処理言語による計算, ならびに数値実験の補佐であった。

(3) 行列型楕円分布の下での Wishart type 確率行列の最大固有値の分布に関する研究について, 研究協力の依頼があった。行列変数型楕円分布の固有値に関する分布論は, 本研究課題と密接な関連があり, 本研究課題の成果をさらに発展させるために必要との判断から協力をした。正規母集団の下での Wishart (type) 行列の最大固有値や最小固有値の分布論についての先行研究は, 相当数ある。近年の行列変数型楕円分布の理論が発展するに伴い, Wishart type 確率行列が楕円分布に従うという仮定の下で, 最大固有値の分布に関する研究が国外で進められている。本研究では, 最大固有値の分布について楕円分布族の下で考察し, 行列変数 t 分布を具体例として, 分布の挙動を数値実験により精度の確認を行った。

<引用文献>

- (1) Fang, K. T. and Zhang, Y. T. (1990). *Generalized Multivariate Analysis*, Science Press Beijing and Springer-Verlag, Berlin. *Statistics and Probability Letters*. **79**, 1935–1944. *Statistics and Probability Letters*. **128**, 89–96.
- (2) Iwashita, T. and Klar, B. (1994). The joint distribution of Studentized residuals under elliptical distributions. *Journal of Multivariate Analysis*. **128**, 203–209.

Table 1. Upper tail percentage points for $p = 3, N = 5$

α	$m \setminus$ Statistic	Rayleigh	$V_{\sqrt{2/3}}$	\tilde{V}_2	\tilde{V}_3	\tilde{V}_4
0.10	5	18.36	12.58	4.66	4.25	4.27
	10	18.49	12.74	4.77	4.33	4.38
0.05	5	20.66	13.15	5.33	4.87	4.89
	10	20.95	13.21	5.48	5.02	5.05
0.01	5	25.38	14.60	6.84	6.01	6.04
	10	26.10	14.24	7.02	6.36	6.37

Table 2. Monte Carlo type I error rates and empirical powers ($p = 3, N = 5, m = 10$)

α	Statistic	Normal	t	Kotz	Exp. 1	Exp. 2	χ^2	Skew Normal
0.10	Rayleigh	0.100	0.101	0.100	0.100	0.100	0.098	0.100
	$V_{\sqrt{2/3}}$	0.101	0.101	0.101	0.161	0.130	0.228	0.101
	\tilde{V}_2	0.101	0.100	0.099	0.101	0.102	0.102	0.100
	\tilde{V}_3	0.100	0.100	0.099	0.226	0.159	0.351	0.099
	\tilde{V}_4	0.101	0.101	0.100	0.108	0.106	0.137	0.098
0.05	Rayleigh	0.050	0.050	0.049	0.050	0.049	0.048	0.049
	$V_{\sqrt{2/3}}$	0.050	0.050	0.050	0.089	0.069	0.138	0.050
	\tilde{V}_2	0.051	0.049	0.050	0.050	0.052	0.050	0.050
	\tilde{V}_3	0.049	0.050	0.049	0.137	0.088	0.241	0.050
	\tilde{V}_4	0.050	0.050	0.050	0.056	0.054	0.074	0.049
0.01	Rayleigh	0.009	0.009	0.010	0.009	0.010	0.009	0.009
	$V_{\sqrt{2/3}}$	0.010	0.010	0.010	0.023	0.016	0.044	0.010
	\tilde{V}_2	0.010	0.010	0.009	0.010	0.010	0.011	0.010
	\tilde{V}_3	0.010	0.009	0.010	0.042	0.022	0.094	0.009
	\tilde{V}_4	0.009	0.010	0.010	0.011	0.011	0.018	0.009

Table 3. Iris (Setosa) data for $p = 4, m = 5, N = 10$

Statistic	Rayleigh	r	$V_{\sqrt{2/3}}$	\tilde{V}_2	\tilde{V}_3	\tilde{V}_4
$\alpha \setminus$ Values	29.09	1	7.08	-4.11	3.51	4.17
		2	7.00	-3.40	3.61	2.83
		3	7.04	-4.03	3.26	4.17
		4	7.22	-4.14	0.16	4.05
0.10	46.98		7.36	-3.33	4.07	6.72
0.05	50.61		7.47	-2.99	4.51	7.16
0.01	57.86		7.72	-2.23	5.26	7.94

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 2件/うちオープンアクセス 2件）

1. 著者名 Iwashita Toshiya, Klar Bernhard	4. 巻 78
2. 論文標題 A gamma tail statistic and its asymptotics	5. 発行年 2023年
3. 雑誌名 Statistica Neerlandica	6. 最初と最後の頁 264 ~ 280
掲載論文のDOI (デジタルオブジェクト識別子) 10.1111/stan.12316	査読の有無 有
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 Toshiya Iwashita and Bernhard Klar	4. 巻 56
2. 論文標題 A necessary test for elliptical symmetry based on the uniform distribution over the Stiefel manifold	5. 発行年 2020年
3. 雑誌名 SUT Journal of Mathematics	6. 最初と最後の頁 129, 145
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 有
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 Aya Shinozaki, Hiroki Hashiguchi, Toshiya Iwashita	4. 巻 30
2. 論文標題 Distribution of the largest eigenvalue of an elliptical Wishert matrix and its simulation	5. 発行年 2018年
3. 雑誌名 Journal of the Japanese Society of Computational Statistics	6. 最初と最後の頁 1, 12
掲載論文のDOI (デジタルオブジェクト識別子) 10.5183//jjscs.1708001_244	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2件（うち招待講演 0件/うち国際学会 1件）

1. 発表者名 岩下登志也
2. 発表標題 楕円対称性の必要性検定
3. 学会等名 科研費シンポジウム「多変量データ解析法における理論と応用」
4. 発表年 2018年

1. 発表者名 Toshiya Iwashita
2. 発表標題 A Necessary test for Elliptical Symmetry
3. 学会等名 DAGStat (国際学会)
4. 発表年 2018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			
ドイツ	Karlsruhe Institute of Technology			
ドイツ	Karlsruhe Institute of Technology			
ドイツ	Karlsruhe Institute of Technology			
ドイツ	Karlsruhe Institute of Technology			