

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 24 年 5 月 14 日現在

機関番号：15401

研究種目：基盤研究（C）

研究期間：2009～2011

課題番号：21500278

研究課題名（和文）

非正規母集団に対する2段階標本抽出法の開発

研究課題名（英文）

Development of two-stage methods for non-normal populations

研究代表者

若木 宏文（WAKAKI HIROFUMI）

広島大学・大学院理学研究科・教授

研究者番号：90210856

研究成果の概要（和文）：非正規母集団に対する母平均ベクトルの領域推定のための2段階標本抽出法として、漸近展開近似を用いる方法、ブートストラップ法を用いる方法を構築し、その特徴を明らかにした。改良すべき点として、漸近展開近似を用いる場合の、高次キュムラントの推定量の改良、ブートストラップ法によるパーセント点の過小評価の問題がある。球形信頼領域を楕円領域で置き換える従来の過大評価の問題点を補うブートストラップ手法を考案した。

研究成果の概要（英文）：Two new methods of two-stage procedure of constructing a confidence region of the mean vector for non-normal populations are suggested. The one is a method using asymptotic expansion of the distribution, and the other is a method using a bootstrap method. We clarified character of each methods including their drawback. We found another bootstrap method to overcome a problem of over estimating the sample size needed in the original Healy's two-stage method.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,100,000	330,000	1,430,000
2010年度	1,000,000	300,000	1,300,000
2011年度	700,000	210,000	910,000
年度			
年度			
総計	2,800,000	840,000	3,640,000

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：2段階標本抽出、漸近展開、単調変換、ブートストラップ

## 1. 研究開始当初の背景

2段階標本抽出法は、Stein (1945) によって提案された推測法である。正規母集団の母平均の区間推定問題において、指定した信頼区間幅でかつ、信頼係数が指定した値以上となるために必要な標本数は母分散に依存しているため、母分散が未知の場合には標本抽出前に、必要な標本数を決めることができな

い。Stein (1945) は、標本抽出を2段階に分けて、1段階目に抽出されたデータから得られる情報を利用することで、区間幅と信頼係数に関する条件を満たすのに必要な追加標本数を決定し、信頼区間を構成する方法を提案した。

2段階法は、初期標本を測定した段階で、必要な総標本数を決定することができるため、実験の費用・時間を見積もることができ

る。新薬の効能を検定するための臨床試験などで有効な手段であるが、実際の現場で2段階抽出法が利用されていることはほとんどないようである。大きな理由の一つは、これまでに開発された2段階手法のほとんどは、母集団の正規性に強く依存したものであり、実際の現場で母集団の厳密な正規性が保証されることはほとんどないからである。そこで、本研究では非正規母集団に対しても有効な2段階抽出法を開発を目指した。

## 2. 研究の目的

非正規母集団の母平均ベクトルの領域推定問題において、信頼領域の幅と信頼係数の条件を満たすような球形の信頼領域を2段階標本抽出法によって構築することを目的とした。

正規母集団に対しては、厳密にしん来領域と幅の条件を満たすような信頼領域を構築することができるが、非正規母集団に対しては、その母集団分布を正確に表すパラメトリックモデルを指定することは現実的ではないため、指定するとしても楕円型などの分布形を想定するのみとなる。したがって信頼係数については、近似的に条件が満たされるようなものしか構築できない。本研究では、漸近的に2次のオーダーまで信頼係数の条件が満たされるような信頼領域の構築を目的とした。具体的には、統計量の分布の漸近展開近似を利用して必要な追加標本数を推定する2段階法と、ブートストラップ法を用いて必要な追加標本数を推定する方法を提案し、実用可能なものとするとともに、その性能や特徴を明らかにすることを目的とした。

## 3. 研究の方法

Stein (1945)の2段階法は、Healy (1956)によって多変量正規母集団の母平均の領域推定問題に拡張された。初期標本によって推定される標本分散共分散行列に関する、全標本の標本平均ベクトルの偏差の2次形式(以下  $Tm^2$  と表す)の分布が未知母数に依存せず、初期標本数  $m$  と変数の次元  $p$  のみで決まることを利用したものである。非正規母集団の場合には、 $Tm^2$  の分布の上側パーセント点(以下  $Am$  と表す)は未知母数に依存するため、 $Am$  を初期標本により推定する必要がある。本研究では、 $Am$  の推定法として、(1)  $Tm^2$  の分布の漸近展開近似を利用する方法と、(2) ブートストラップを用いて推定する方法、および、(3) 漸近展開を利用してブートストラップ近似を改良する方法を考えた。

### (1) $Tm^2$ の分布の漸近展開の導出法

初期標本を与えた  $Tm^2$  の分布の条件付分布は追加標本の標本平均の関数の分布であるから、母平均の周りで確率展開して項別に期待値を評価することで条件付分布の特性関数の展開式を得ることができる。得られた結果は、初期標本の標本共分散行列の関数であるから、母分散共分散行列の周りで確率展開して期待値を評価すれば良い。ここで得られる展開式は、数百の項を含むため、手計算で期待値を評価してまとめることは不可能である。そこで、数式処理ソフトである MATHEMATICA を利用して、統計量を展開し、各項の期待値を評価するアルゴリズムを構築した。漸近展開導出の理論的な側面では、追加標本数は、ある統計量の整数部分であることが問題となる。小数部分が漸近的に一樣分布に従うことを証明することで、その問題を解決した。

### (2) ブートストラップによるパーセント点と追加標本数の決定

非正規母集団の場合、 $Tm^2$  の分布は追加標本数によって決まり、一方、追加標本数は  $Tm^2$  の分布のパーセント点によって決まる。追加標本数 ( $n$  と表す) を指定して、 $Tm^2$  のパーセント点を推定し、それを基に推定される必要な追加標本数 ( $f(n)$  と表す) が、最初に指定した追加標本数と等しい ( $n = f(n)$ ) ような  $n$  が求めたいものである。一般に  $f(n)$  は  $n$  の単調減少関数であるので、 $n_1 > f(n_1)$ ,  $n_2 < f(n_2)$  なる  $n_1, n_2$  の区間を2分割してゆく分割アルゴリズムによって  $n$  を求める方法を用いた。

### (3) 漸近展開を利用したブートストラップ近似の改良

一般に、統計量の分布がピボタル(極限分布が母集団分布に依存しない)ならば、標本分布のブートストラップ近似は、1次までの漸近展開近似と同等の精度を持つことが知られている。もし、統計量の分布の1次までの漸近展開近似が母集団分布に依存しないならばブートストラップ近似は2次までの漸近展開と同等の精度を持つことが期待される。(1)で導出した漸近展開公式を利用して、 $Tm^2$  を単調変換した統計量の分布の漸近展開式が1次まで、母集団に依存しないようにできるならば、単調変換後にブートストラップ法によりパーセント点を近似し、逆変換を施すことで  $Tm^2$  のパーセント点のより精度の高いブートストラップ推定が得られることが予想される。

#### 4. 研究成果

非正規母集団に対する、 $T_m^2$  の分布の漸近展開式を利用した2段階抽出法、およびブートストラップ法を用いた2段階抽出法を実現した。その有効性については、数値実験の行っているところであり、次のような特徴が明らかになった。

(1) 漸近展開法では、高次キュムラントを初期標本から推定する必要があるため変数の次元が高く、初期標本数がそれほど大きくない場合には、精度が良くない。

(2) ブートストラップ法によるパーセント点の推定は、真のパーセント点より若干小さめに推定される傾向がある。これは、必要な追加標本数を過小評価することを意味し、期待した信頼係数よりも小さな信頼係数となることが懸念される。しかし、Healy (1956) による  $T_m^2$  を用いた標本数の決定方法は、球形の信頼領域の信頼係数を、それに含まれる楕円領域の信頼係数で小さめに評価するため、必要な追加標本数を過大評価している。ブートストラップ法による過小評価より、楕円領域を用いることによる過大評価の影響の方がずっとお大きいと、ブートストラップ法による過小評価は、むしろ楕円領域を用いることのデメリットを補う方向に働いている。ただし、1次元の場合には、過小評価の問題点は改善すべき課題として残る。

(3) 楕円領域を用いることによる必要な追加標本数の過大評価は、Healy (1956) の方法の大きな問題点である。いくつかの数値実験では、90パーセントの信頼領域を構築するはずが、98パーセントの信頼領域になってしまっている。Healyの方法において、球形の信頼領域を内側の楕円領域に置き換える理由は、統計量の分布が母集団分布に依存しないようにするためであるが、ブートストラップ法は母集団分布を経験分布で置き換える手法であるから、球形の信頼領域のままパーセント点を推定することも考えられる。扱う統計量がピボタルでないため、その推定精度が心配であったが、ある程度初期標本数を大きくすると、信頼係数は、望む値に結構近い値となった。初期標本数が大きくない場合には、やはり必要な追加標本数を過小評価してしまう問題点があるが、楕円領域を用いることによる過大評価の解決法として魅力的な手法であり、今後継続して研究する価値がある手法である。

得られた成果は3編の論文として現在執筆中で、プレプリントは下記記載のホームページにテクニカルレポートとして掲載予定で

ある。

#### 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 2 件)

① Shimizu, H. and Wakaki, H. (2011) Asymptotic expansions for a class of tests for a general covariance structure under a local alternative. *Journal of Multivariate Analysis*. 102, issue 6, 1080-1089. (査読あり)

② Akita, T., Jin, J., and Wakaki, H. (2010) High{dimensional Edgeworth expansion of a test statistic on independence and its error bound. *Journal of Multivariate Analysis*. 101, issue 8, 1806-1813. (査読あり)

[学会発表] (計 5 件)

統計学会

① 若木宏文 「Box 型大標本漸近展開近似の誤差限界」  
統計関連学会連合大会、2011年9月5日、アクロス福岡

② 若木宏文、藤越康祝 「正準相関分析の追加情報の冗長性に関する検定統計量の分布の漸近展開近似の誤差評価」  
統計関連学会連合大会、2010年9月6日、早稲田大学早稲田キャンパス

③ 清水暁裕、秋田智之、若木宏文 「Asymptotic expansions for a class of tests for a general covariance structure under a local alternative」  
統計関連学会連合大会、2009年9月8日、同志社大学京田辺キャンパス

④ 若木宏文、青嶋 誠 「非正規母集団に対する二段階標本抽出法」  
広島統計グループ金曜セミナー、2009年7月24日、広島大学理学部

⑤ A. Shimizu and H. Wakaki, "An asymptotic expansion of the local power of a general test for testing a general covariance structure." *Hirshoshima Research Group Seminar*, 2009.10.30, Department of Statistics of Radiation Effects Research Foundation, Hiroshima

〔その他〕  
ホームページ等

広島大学理学部数学科数理統計グループの  
ホームページ（セミナー、テクニカルレポート等の情報）

<http://www.math.sci.hiroshima-u.ac.jp/stat/>

## 6. 研究組織

### (1) 研究代表者

若木 宏文 (WAKAKI HIROFUMI)  
広島大学・大学院理学研究科・教授  
研究者番号：90210856

### (2) 研究分担者

( )

研究者番号：

### (3) 連携研究者

( )

研究者番号：