

## 科学研究費助成事業 研究成果報告書

平成 27 年 5 月 29 日現在

機関番号：17102

研究種目：挑戦的萌芽研究

研究期間：2012～2014

課題番号：24650151

研究課題名(和文) 高次有効性をもつスムーズなノンパラメトリック推測法の開発

研究課題名(英文) Improvement of nonparametric inference which has smoothness and higher order efficiency

研究代表者

前園 宜彦 (Maesono, Yoshihiko)

九州大学・数理(科)学研究科(研究院)・教授

研究者番号：30173701

交付決定額(研究期間全体)：(直接経費) 2,400,000円

研究成果の概要(和文)：本研究ではノンパラメトリックな検定統計量である順位検定に対して、有意確率の離散性を解消する連続化を提案し、その理論的な性質を明らかにした。手法としては、スムーズな推測結果を与えるカーネル法に基づく連続化を提案し、提案した統計量が順位検定と同等であることを示した。その上で、エッジワース展開を利用した有意確率の近似精度の改良を求めた。得られた近似はカーネルをうまく構成すると、新たな推定量が不要で、元の母集団分布に依存しないものになっており、有意確率も連続化されている。この成果は世界に先駆けてのユニークなものである。

研究成果の概要(英文)：In this project, we propose smoothed rank tests based on the kernel method which gives us smooth statistical inference. The proposed tests conquer the problem of the discreteness of the distribution for rank tests. We also obtain theoretical properties of the smoothed rank tests, and show the proposed tests are asymptotically equivalent to the ordinary rank tests. Further we obtain Edgeworth expansions of these tests, which are improvements of the normal approximations. If we choose proper kernels, we can get the Edgeworth expansions which do not depend on the population distribution. These results are unique and forefront of this area.

研究分野：統計科学

キーワード：ノンパラメトリック 順位検定 カーネル型推定量 符号検定 正規近似 高次漸近理論 ウィルコクソン検定 エッジワース展開

## 1. 研究開始当初の背景

(1) 符号検定および順位検定の漸近相対効率とその理論的性質は 70 年代までにかなり研究され、帰無仮説の下での分布は元の母集団分布に依存しないことが知られている (distribution-free). 一方、統計量が離散的な分布を持つために、有意水準を設定する方法では検定できず、有意確率を評価する方法で利用されている。しかし Lehmann(1975, Wiley) も指摘しているように、標本数が小さい時は離散分布の刻みの細かいほうが有意確率は小さくなる傾向があり、恣意的に検定統計量を選ぶ危険性が指摘されている。他方、カーネル法は密度関数の推定量に関して知られているように、バンド幅の影響で収束のオーダーが遅いという問題がある。これは当然スムーズ化ブートストラップ推測においても問題になる。研究代表者はこれまでノンパラメトリック統計量の漸近理論を中心に研究し、カーネル法についてもバンド幅の影響を込みにした高次の漸近理論を構築している (Maesono and Penev, 2011, Ann. Inst. Math. Stat. pp.617-644). またブートストラップ法についても Hall and Maesono (2000, J.R.S.S., pp.137-144) の業績を上げており、「高次有効性をもつスムーズなノンパラメトリック推測法の開発」の研究を遂行していくことが十分可能な状況であった。

(2) 1980 年代以降、リサンプリング法の研究がノンパラメトリック法の主流になったために、順位型統計量に基づく推測法の改良は停滞気味である。またカーネル法による推測の 1 次漸近理論は構築され、推測のオーダーの意味ではパラメトリックな手法に劣ることが知られている。しかし推測のスムーズさは重要な要素で、オーダーが劣っていても良く使われる手法になってきている。したがってスムーズな性質を持ちながら、推測の精度を保つ手法の開発は非常に重要な課題になっている。本研究は推測のスムーズさを失うことなく高次の推測精度を保つノンパラメトリック手法の開発を目指す極めて独創的と言えるものである。

## 2. 研究の目的

(1) ノンパラメトリックな統計的推測において生じる離散的な結果をスムーズにする研究を行うと同時に、精度の改良を目指した。ノンパラメトリック推測における符号検定や順位検定に使われる統計量の分布は離散的なものになるために、なめらかな推測結果が得られない。また汎用性があるぶん、効率の面では落ちることが知られている。このような離散的なものに対してはカーネル法を使ったスムーズ化が利用される。ブートストラップ推測においてもスムーズ化した経験分布関数を利用した手法が提案されている。しかしカーネル法による推測法はバンド幅の影響を受けて、通常の統計量もつ収束のオーダーより遅くなることが知られてい

る。本研究では収束のオーダーを改善する手法を提案し、その理論的な性質を明らかにすることを目指した。

(2) 具体的には、スムーズ化ノンパラメトリック統計量の高次の漸近理論を構築し、推測の精度を上げることを図った。研究代表者・前園は統計的漸近理論を研究し、多くの有益な結果を得ている。特に条件付期待値で射影する Hoeffding 分解を使って、ノンパラメトリックな設定の下での漸近理論を発展させてきた。本研究では、この手法を用いてスムーズ化されたノンパラメトリック法に基づく統計量の漸近的な表現を求め、その表現を利用して近似精度の改良を目指す。具体的には順位型統計量の分布をカーネル法によってスムーズ化して、その分布を元に高次漸近理論を導出し、その上で精度の高い推測法の構築を図る。またカーネル型推定量の高次漸近理論を構築し、それを利用してスムーズ化ブートストラップ推測の精度の改良に結び付けていくことを目的とした。

## 3. 研究の方法

統計的推測の結果をなめらかにする手法について、今までに提案された手法とその研究状況を調査した。その上で分布に依存しない順位型統計量をスムーズ化することにより、分布の刻みを細かくする方法を検討した。その後高次漸近理論を構築し、推測の精度を上げることを目指した。またカーネル法を用いたスムーズ化ブートストラップ法の高次漸近理論を求め、新しいノンパラメトリック法の提案を行った。特にカーネル法におけるバンド幅の影響を取り除くための手法を開発し、その妥当性を理論的に証明した。このときの基本的な道具となるのが Hoeffding 分解と呼ばれる一般の統計量についての直交分解である。これを利用することによって、高次の漸近理論の構築が可能になり、スムーズな統計手法の精度の改良法を見つけることができた。

高次漸近理論を実際に利用するとき必要となる未知な部分の推定量の構成法を求め、その理論的性質を研究した。具体的には利用するスムーズ化統計量の分散及び高次のキュムラントの漸近表現を求め、未知の母数に個々に推定量を代入する方法の有効性を検証する。それと同時にジャックナイフ法やブートストラップ法などのリサンプリング法とカーネル法を融合した新しい推測法の提案とその理論的性質を明らかにすることを目指した。本研究では大偏差確率の評価とエシンのスムーズ化・レンマを利用して、確率的に残差項を評価してスムーズ化統計量の漸近表現を求め、高次漸近理論を求めた。その後高い精度を持つ推測法を構築し、シミュレーションや実データへの適用で有効性を確認した。

具体的には順位型統計量の分布のスムーズ化を求め、その分布を持つように順位型統

計量のスムーズ化を行った。これらの成果を元にして、高次漸近理論を構築し、ブートストラップ法への適用も研究した。またロジスティック回帰におけるスムーズ化の高次漸近理論の構築、コックスモデルにおけるスムーズ化による影響の研究、およびこれらノンパラメトリック回帰における最適なバンド幅の選択について理論的な研究を行った。また個々の統計手法に応じたスムーズ化の修正を行い、目的に沿って精度を上げる方法を開発した。その中で新しく生じた問題点を統一的に解決し、スムーズな新しい統計手法確立の可能性を探った。

上記の研究遂行のために、国内の他の研究機関を訪問し、他の研究者との情報交換及び資料収集を行った。同時に海外の研究者からレビューを受けるために、外国出張した。また高性能のパソコンを使って、計算機ソフトによる研究成果の実用性のチェックを行うと同時に、計算機上でのシミュレーションを行った。これらの研究で得られた成果を論文にまとめジャーナルに投稿し、他の研究者の客観的な評価をもとに本研究の位置づけと適用範囲の拡大を目指して修正を行った。

#### 4. 研究成果

ジャックナイフ分散推定量を利用したチューデント化カーネル型確率点推定量に関するエッジワース展開を求め、展開の有効性を理論的に示した。またその展開に基づいて新しい確率点のノンパラメトリック信頼区間の構成を提案し、その優位性をシミュレーションで示すことができた。さらにカーネル型確率密度関数推定量の高次エッジワース展開を具体的な形で求めることに成功した。またその展開の有効性を保証する条件を求めた。これらは世界に先駆けた成果で、論文として出版することができた。カーネル型分布関数推定量について正規近似の精密化を具体的に求め、その有効性をシミュレーションで示した。正規近似を精密化したものは、カーネル関数と母集団分布のモーメントに依存するものになっており、それらを具体的に求めたのは世界に先駆けての成果である。経験分布関数の連続化であるカーネル型分布関数推定量を利用して、ノンパラメトリック検定として良く利用される符号検定の連続化に成功した。連続化した検定統計量が Pitman の漸近相対効率の意味で、符号検定と同等であることを示し、また符号検定では不可能であった正規近似の精密化の有効性を理論的に証明することに成功した。この連続化検定は世界に先駆けてのもので、離散型検定統計量における有意確率の離散性という問題を解消できるものになっている。これまでに得られたカーネル型分布関数推定量正規近似の精密化をさらに発展させて高度の精密化を具体的に求め、その有効性をシミュレーションで示した。また経験分布関数の連続化であるカーネル型分布関数推定

量を利用した符号検定の連続化の漸近展開について細かく分類することに成功し、その展開を具体的な形で求めた。求めた成果は 3 次の漸近理論に相当するもので、初めて得られたものである。さらに符号検定の連続化において、推定に利用するカーネルをうまく選ぶことによって、エッジワース展開の各項が分布に依存しないようにすることができるということを見出した。符号検定の連続化のアイデアを利用し、ウィルコクソンの符号付き順位検定の連続化を提案した。この連続化により、有意確率の離散性の問題を解消し、連続化したウィルコクソン検定統計量が Pitman の漸近相対効率の意味で、元のウィルコクソン検定と同等であることを示した。新たに提案した連続化ウィルコクソン統計量のエッジワース展開を求めることにも成功し、正規近似の精密化を得ることができた。求めた展開はカーネルをうまく選択することにより、データの従う母集団分布に依存せず求められることを示した。これにより符号検定の連続化と同様に離散型検定統計量における有意確率の離散性という問題を解消できることが示された。これらの成果は現在論文として投稿準備中である。

#### 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 4 件)

Atsufumi Nishimoto and Yoshihiko Maesono, On jackknife variance estimator for kernel density estimator and its application, Bulletin of Informatics and Cybernetics, 査読有, 掲載確定

Zhong Huang and Yoshihiko Maesono, Edgeworth Expansion for Kernel Estimators of a distribution function, Bulletin of Informatics and Cybernetics, 査読有, Vol.46, 2014, pp.1-10

Shota Umeno and Yoshihiko Maesono, Improvement of Normal Approximation for Kernel Density Estimator, Bulletin of Informatics and Cybernetics, 査読有, Vol.45, 2013, pp.11-24

Yoshihiko Maesono and Spiridon I. Penev, Improved confidence intervals for quantiles, Annals of the Institute of Statistical Mathematics, 査読有, Vol. 65, 2013, pp.167-189

〔学会発表〕(計 15 件)

森山卓, 前園宜彦, 符号付き順位検定の連続化と有意確率の近似, 研究会集「ノンパラメトリック統計解析とベイズ統計」, 2015年3月25-26日, 慶応大学

前園宜彦, カーネル型推定量に対する正規近似の改良, 第9回日本統計学会春季集会, 2015年3月8日(招待講演), 明治大学

前園宜彦, ハザード関数のカーネル型推定量の漸近表現とその応用, 研究集会「多様な分野における統計科学の教育・理論・応用の新展開」2014年10月24-26日, 新潟大学

前園宜彦, ハザード関数のカーネル型推定量の高次漸近表現, 日本数学会総合分科会, 2014年9月25-28日, 広島大学

前園宜彦, 森山卓, ノンパラメトリック検定統計量の有意確率と連続化統計量について, 統計関連学会連合大会, 2014年9月13-16日, 東京大学

Yoshihiko Maesono, Smoothing of sign test and approximation of its p-value, The 3rd Institute of Mathematical Statistics Asia Pacific Rim Meeting, 2014年6月29日-7月4日, Howard International House

前園宜彦, 分布関数のカーネル型推定量の高次漸近分布とその応用, 研究集会「ノンパラメトリック統計解析とベイズ統計」, 2014年3月19-20日, 慶応大学

前園宜彦, 符号検定の連続化と有意確率の近似について, 日本数学会総合分科会, 2013年9月24-27日, 愛媛大学

前園宜彦, 分布関数のカーネル型推定量の高次漸近分布, 統計関連学会連合大会, 2013年9月8-11日, 大阪大学

Mengxing Lu and Yoshihiko Maesono, Smoothing of sign test and approximation of its p-value, Bernoulli Society Satellite Meeting to the ISI World Statistics Congress 2013, "Asymptotic Statistics and Related Topics: Theories and Methodologies", 2013年9月2-4日, 東京大学

前園宜彦, 魯夢欣, 符号検定の平滑化と有意確率の近似について, 日本数学会年会, 2013年3月20-23日, 京都大学

前園宜彦, 魯夢欣, 符号検定の連続化とその有意確率の近似について, 研究集会「ノンパラメトリック統計解析とベイズ統計」, 2013年3月14-16日, 慶応大学

前園宜彦, Spiridon Penev, Improved confidence intervals for quantiles, 研究集会「統計科学における深化と横断的展開」, 2012年10月24-26日, 松江テルサ

前園宜彦, カーネル型密度関数推定量の高次漸近理論, 統計関連学会連合大会, 2012年9月9-12日, 北海道大学

前園宜彦, Spiridon Penev, Asymptotic distributions of kernel type estimators, IMS-APRM2012 サテライトシンポジウム, 2012年7月2-4日(招待講演), 早稲田大学

前園宜彦, Spiridon Penev, Nonparametric confidence intervals for quantiles for moderate sample sizes, 第2回 IMS-APRM2012, 2012年7月2-4日, エポカルつくば

〔図書〕(計 0 件)

〔産業財産権〕  
出願状況(計 0 件)

名称:  
発明者:  
権利者:  
種類:  
番号:  
出願年月日:  
国内外の別:

取得状況(計 0 件)

名称:  
発明者:  
権利者:  
種類:  
番号:  
出願年月日:  
取得年月日:  
国内外の別:

〔その他〕  
ホームページ等

## 6. 研究組織

### (1) 研究代表者

前園 宜彦 (MAESONO, Yoshihiko)  
九州大学・大学院数理学研究院・教授  
研究者番号: 30173701