

## 科学研究費助成事業 研究成果報告書

平成 30 年 6 月 4 日現在

機関番号：17102

研究種目：挑戦的萌芽研究

研究期間：2015～2017

課題番号：15K11995

研究課題名(和文)ノンパラメトリック統計量の平滑化と高精度推測法の構築

研究課題名(英文)Improvement of statistical inference based on nonparametric smoothing statistics

研究代表者

前園 宜彦(Maesono, Yoshihiko)

九州大学・数理学研究院・教授

研究者番号：30173701

交付決定額(研究期間全体)：(直接経費) 2,200,000円

研究成果の概要(和文)：本研究では平滑化統計量の提案とそれを応用したノンパラメトリック統計量の平滑化による高精度の推測法の開発に成功した。具体的にはカーネル法を利用した比の形の関数の平滑化推定の改良に成功した。提案した手法は推定量の変動を小さくする直接型推定量で、分母の推定量を積分の変換を通して得るもので、バイアスは増えるときもあるが、分散は一様に縮小するものになっている。また順位統計量の離散性に起因する問題点をカーネル法の活用によって解決することに成功し、有意確率の高次近似の公式の導出にも成功した。さらに有界なサポートを持つ密度関数推定に対する境界問題を解消する推定量を求めた。

研究成果の概要(英文)：In this project, we succeed to construct direct kernel estimators of ratio functions such as a density ratio, a hazard ratio and a conditional density. These estimators are sort of transformation estimators that reduce variances. We also succeed to solve problems of discreteness of rank test statistics. Furthermore, we propose new kernel type estimators which eliminate boundary biases and apply them to a mean residual life function.

研究分野：統計科学

キーワード：ノンパラメトリック ハザード関数 順位検定 カーネル型統計量 平均二乗誤差 漸近理論 境界問題

## 1. 研究開始当初の背景

(1) ノンパラメトリックな検定統計量については、漸近相対効率など標本数を無限大にしたときの比較が主流で、分布が離散型であるという固有の問題点についてはあまり議論されてこなかった。順位検定統計量の一番の強みは帰無仮説の下での分布が、標本の従う母集団分布に依存することなく決まることである。したがって特定の分布を仮定しなくても妥当性を失わない検定や信頼区間を構成することができる。しかし統計量の分布は離散型であるために、有意確率を使った検定となり、信頼区間も保守的なものになる。ここで注意しないといけないのは、標本数が少ない時は母集団分布がどのような分布であっても、検定統計量の分布の刻みの細かい方が有意確率は小さくなる傾向があるということである。一般に有意確率が小さい方が検定としては良い検定となる。これは帰無仮説の下での議論であり、従来の漸近相対効率による比較や、対立仮説の下での漸近検出力による比較などのように、検定の良さが元の母集団分布に依存して決まるという主張と整合していない。このような齟齬が起こるのは、離散分布を連続分布で近似して議論することに起因する。この問題解決へのアプローチとしてカーネル法を利用した統計量の平滑化の利用が妥当と思われた。この研究プロジェクトでは順位検定の連続化を利用して有意確率の離散性の問題点の解決を目指した。

(2) Maesono and Penev (2011, Ann. Inst. Math. Stat. pp.617-644)は、カーネル型確率点推定量についての高次漸近理論を求め、近似分布の収束のオーダーを改善している。また確率密度関数のカーネル推定についても高次の漸近展開を求めている。この一連の成果は世界に先駆けてのもので、これらの成果を使えば、カーネル法などで平滑化した統計量に基づく推測精度の向上が期待できた。本研究ではカーネル法などに基づく平滑化を利用して離散統計量を連続化し、離散に起因する問題の解決を図るものであった。しかしカーネル法はそれ自身バンド幅の影響で収束のオーダーが遅くなることが知られており、その改善のために高次漸近理論の構築が必要とされている。さらに関数比についての推測においては、分母に自然な推定量を代入すると、変動が大きくなる傾向にあり、密度比にはその変動を抑制するために積分を利用した回避法が提案されている。この方法は他の関数比にも拡張が可能と思われた。さらに密度のサポートが有界な場合には、その平均二乗誤差の改良が必要であった。

## 2. 研究の目的

(1) ノンパラメトリックな統計量を利用するときに生じる離散的な推測結果をスムーズにする研究を行うと同時に、推測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り、得られる有意確率も離散的になる。また比率の推定においても、推定値は離散的で、ロジット変換を行ってもこの問題は解決しない。これらの解決法としてはカーネル法を使った平滑化が有効と思われた。ここでは1つの目的として平滑化順位検定統計量の構成と有意確率の近似の改良を目指した。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た。本研究ではノンパラメトリック推測の精度を上げるとともに、滑らかさと効率のバランスがとれる規準を導出することにより、より精度の高い平滑化の手法を構築することを目指した。

(2) 関数の比に対する平滑化推定量の安定化を目指し、密度関数に対して提案されている直接型推定量の拡張を目指した。特に生存時間解析におけるハザード比及び条件付き密度関数の直接型推定量を導出し、その有効性を理論及びシミュレーションによって検証した。また統計量のエッジワース展開をバンド幅が組み込まれた形で求め、得られた結果に基づいて高精度信頼区間の構成を目指した。さらに推測精度の改良において、バンド幅と効率のバランスをとる新たな規準を導出して、得られた成果の有用性について理論的に検証し、密度関数が有界なサポートを持つときの境界バイアスを縮小するカーネル型推定量の改良を図った。

## 3. 研究の方法

統計的推測の結果を滑らかにする手法について、パラメトリック、ノンパラメトリックを問わず、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などを平滑化することにより、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平滑化ブートストラップ法の高次漸近理論を求め、新しい平滑化ノンパラメトリック推測法の提案を行った。このときの基本的な道具となるのが直交分解とマルチンゲールのモーメント評価である。これら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提案できた。

その上で、高次の漸近理論を利用するときに必要な未知母数の推定量の構成法を求め、その有効性を研究した。具体的には利用する平滑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、表現に含まれる未知母数に推定量を代入する手法の有効性を検証した。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平滑化推測法の提案とその理論的性質を明らかにした。また大偏差確率の評価とエシンのスムージング・レンマを利用して、確率的に残差

項を評価して平滑化統計量の漸近表現を求め、高次漸近理論を確立した。その後高い精度を持つ推測法を構築し、シミュレーションで有効性を確認した。

具体的には順位型統計量の平滑化をまず構成し、元の分布との差を考察した。その上で、高次のエッジワース展開を求めて、展開が未知母数を含まない分布を持つように順位型統計量の平滑化を行った。また個々の統計手法に応じた平滑化の修正を行い、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する可能性を検証した。同時に開発された手法をシミュレーションで理論的成果との整合性を検証した。

これらの研究で得られた成果を論文にまとめて国際的なジャーナルに投稿し、客観的な評価を受けた。その上で本研究の位置づけと適用範囲の拡大を目指して、研究の総括を行った。

#### 4. 研究成果

(1) カーネル法を利用したウィルコクソン統計量の平滑化を提案し、その漸近的な性質が元の統計量と同等であることを明らかにした。また有意確率の高次近似を導出し、シミュレーションで検証して理論との整合性を確認した。この成果は世界で初めて得られたものであり、現在国際誌に掲載が確定している。またカーネル型密度関数推定量の分散に対するジャックナイフ分散推定量の一致性を示し、それを利用したスチューデント化カーネル型密度推定量のエッジワース展開を求めることに成功し、国際誌に掲載された。(2) 判別分析などで利用される密度関数の比のカーネル型推定量の漸近的性質を明らかにして、統計量の分布の近似の精密化であるエッジワース展開の導出に成功した。カーネル法に関連した高次漸近理論の議論は殆どなく、この成果は高く評価され国際誌に掲載が確定している。(3) 関数比のカーネル型推定量の分散を小さくする直接型推定量の漸近的性質を明らかにすることにも成功した。またハザード関数の直接型推定量を新たに提案して、その漸近的性質を明らかにした。同時に、条件付き密度関数の新たなカーネル型推定量を提案し、その性質を明らかにした。またノンパラメトリック回帰に応用して、ナダラヤ・ワトソン推定量に対抗する直接型推定量を提案しその漸近平均二乗誤差を求めた。(4) またカーネル型分布関数推定量についての積分で表現される推定量を提案し、漸近平均二乗誤差を求めるとともにその有効性をシミュレーションで検証した。有界なサポートを持つ確率密度関数について、従来提案されている推定量を漸近平均二乗誤差の意味で改善する統計量を考案した。(5) 統計的リサンプリング法のカーネル型推定量への適用を研究し、ジャックナイフ法の密度関数推定に直接適用するのは問題があること

を理論的に明らかにした。同時に分布関数のカーネル型推定量への適用は有効に働くことを示した。(6) 非負性を保つ密度関数の推定法についてこれまで提案されている方法との組み合わせを考察し、その性質を求めた。これらの成果は2本の論文として掲載確定で、他の成果も国際誌に投稿準備中である。

#### 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 5 件)

Taku Moriyama and Yoshihiko Maesono, Smoothed alternatives of the two-sample median and Wilcoxon's rank sum tests, *Statistics*, 査読有, 掲載確定

Yoshihiko Maesono, Taku Moriyama and Mengxin Lu, Smoothed nonparametric tests and approximations of p-values, *Annals of the Institute of Statistical Mathematics*, 査読有, 掲載確定

Rizky Reza Fauzi and Yoshihiko Maesono, Error reduction for kernel distribution function estimators, *Bulletin of Informatics and Cybernetics*, 査読有, Vol.49, pp.53-66, 2017

Taku Moriyama and Yoshihiko Maesono, Asymptotic properties of a kernel type estimator of a density ratio, *Bulletin of Informatics and Cybernetics*, 査読有, Vol.48, pp.37-46, 2016

Atsufumi Nishimoto and Yoshihiko Maesono, On jackknife variance estimator for kernel density estimator and its application, *Bulletin of Informatics and Cybernetics*, 査読有, Vol.47, pp.1-9, 2015

〔学会発表〕(計 20 件)

Rizky Reza Fauzi, 前園宜彦, Boundary-free Estimators of Distribution Function with Transformation, 研究集会「ノンパラメトリック統計解析とベイズ統計」, 2018年3月28日-3月29日, 慶応大学  
森山卓, 前園宜彦, 関数の比のノンパラメトリック推定について, 研究集会「ノンパラメトリック統計解析とベイズ統計」, 2018年3月28日-3月29日, 慶応大学

森山卓, 前園宜彦, Smoothed two-sample nonparametric tests and their asymptotic properties, 日本数学会年会, 2018年3月18日-3月21日, 東京大学

森山卓, 前園宜彦, Smoothed two-sample nonparametric tests and their asymptotic properties, 研究集会「Current topics on algebraic statistics and related fields」, 2018年3月7日-3月8日, 神戸大学

Rizky Reza Fauzi, 前園宜彦, Boundary free estimators of distribution function with transformation, 研究集会「生命・自然

科学における複雑現象解明のための統計的アプローチ」, 2018年2月16日-2月17日, 滋賀大学彦根キャンパス

本山真誠, 前園宜彦, On direct kernel estimator of density ratio, 日本数学会総合分科会, 2017年9月11日-9月14日, 山形大学

Rizky Reza Fauzi, 前園宜彦, Error Reduction for Kernel Distribution Function Estimators, 統計関連学会連合大会, 2017年9月4日-9月6日, 南山大学

Rizky Reza Fauzi, 前園宜彦, Modified gamma kernel density estimator by variance reduction, 研究集会「ノンパラメトリック統計解析とベイズ統計」, 2017年3月28日-3月29日, 慶応大学

前園宜彦, Rizky Reza Fauzi, Modified gamma kernel density estimator, 日本数学会年会, 2017年3月24日-27日, 首都大学東京

森山卓, 前園宜彦, 二標本ノンパラメトリック検定の連続化と局所漸近検出力, 研究集会「統計科学の現代的課題」, 2017年1月27日-1月29日, 金沢大学サテライトプラザ

Rizky Reza Fauzi, 前園宜彦, Modifying Gamma Kernel Density Estimator by Reducing Variance, 研究集会「統計科学における数理と理論の最先端について」, 2016年10月27日-10月29日, 北海道大学

森山卓, 前園宜彦, 条件付き密度関数のカーネル型直接推定と回帰分析への応用, 日本数学会総合分科会, 2016年9月15日-9月18日, 関西大学

Rizky Reza Fauzi, 前園宜彦, Modified Gamma Kernel Density Estimator, 統計関連学会連合大会, 2016年9月5日-9月7日, 金沢大学

森山卓, 前園宜彦, Asymptotic properties of kernel type estimators of ratios, 「The 4th Institute of Mathematical Statistics Asia Pacific Rim Meeting」, 2016年6月27日-6月30日(招待講演), The Chinese University of Hong Kong

森山卓, 前園宜彦, カーネル型直接推定量の漸近的性質について, 研究集会「ノンパラメトリック統計解析とベイズ統計」, 2016年3月29日, 慶応大学

前園宜彦, 森山卓, ハザード関数のカーネル型直接推定量の漸近的性質について, 日本数学会年会, 2016年3月17日, 筑波大学

森山卓, 前園宜彦, 条件付き密度関数のカーネル型推定と回帰分析への応用, 「熊本国際シンポジウム」, 2016年3月5日, 熊本大学

森山卓, 前園宜彦, 比のカーネル型推定量の漸近的性質について, 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」, 2015年10月26日, 富山大学

前園宜彦, 森山卓, カーネル法による密度比推定量の漸近表現について, 日本数学会総

合分科会, 2015年9月15日, 京都産業大学  
森山卓, 前園宜彦, 二標本ノンパラメトリック検定の連続化と有意確率の近似について, 統計関連学会連合大会, 2015年9月8日, 岡山大学

〔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:  
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・数理学研究院・教授  
研究者番号: 30173701