

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 13 日現在

機関番号：12601

研究種目：基盤研究(B)

研究期間：2011～2013

課題番号：23300104

研究課題名(和文) ベイズ統計学の予測理論としての再構築とその応用

研究課題名(英文) A reformulation of Bayesian statistics as prediction theory and its applications

研究代表者

駒木 文保 (Komaki, Fumiyasu)

東京大学・情報理工学(系)研究科・教授

研究者番号：70242039

交付決定額(研究期間全体)：(直接経費) 5,900,000円、(間接経費) 1,770,000円

研究成果の概要(和文)：ベイズ統計学において必要となる事前分布を、条件付相互情報量と呼ばれる量を最大化することにより構成する研究を行った。特に、統計的検定の問題を、帰無仮説を0、対立仮説を1として、0であるのか1であるのかを予測する問題として定式化し、ベイズ予測の理論の枠組みから研究を行った。従来のベイズ検定で自然とされ想定されていた対立仮説に関する事前分布の枠に当てはまらない事前分布が自然に現れることを示した。さらに、計算統計学的手法として、条件付き相互情報量を最大化することにより事前分布を構成するアルゴリズムの開発研究を行った。

研究成果の概要(英文)：Constructing methods for prior distributions in Bayesian statistics based on maximization of the conditional mutual information are investigated. In particular, statistical testing problems are formulated as a prediction problem by setting the null hypothesis as 0 and the alternative hypothesis as 1. It is shown that prior distributions that have not been considered in previous studies naturally appear. Numerical algorithms for maximizing the mutual information are developed.

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：予測理論 ベイズ統計 情報量

1. 研究開始当初の背景

研究開始当初、情報幾何と条件付相互情報量を用いた方法により事前分布の構成について統一的な扱いに関する見通しが得られつつあった。このアプローチを拡張することにより、従来の理論では種々のパラドックスがつきまとい適切に扱うことが困難であった、複数のモデルが想定される場合のベイズ検定やモデル選択についての理論を構築することが可能になると予想された。さまざまな設定において、ベイズ統計学の最も基礎的で未解決であった問題について扱うことが可能になる。条件付相互情報量と情報幾何を用いて、予測の観点からベイズ統計学の基礎に対する統一的な議論をする試みは従来には無く独創性の高いものになることが期待された。

2. 研究の目的

研究代表者は、パラメトリックモデルに基づくベイズ推測の問題を、予測分布の視点から定式化することにより、多くの場合に性能の良い事前分布を構成できる統一的なアプローチとなることを目的として研究する。特に、複数のモデルが想定される場合のベイズ推測のための事前分布の選択には、条件付相互情報量を最大化する事前分布を用いることによりミニマックスな予測分布を構成することを目指す。この条件付相互情報量に基づく方法は、観測される確率変数と予測したい確率変数の分布が異なる場合や独立でない場合にも適用できると考えられる。また、ベイズ検定における問題に対して一定の解決を与え、複数のモデルがある場合の予測・推測問題において有効な手法を構築する。ベイズ検定の問題は、2つの仮説に関する予測問題として定式化することができる。このことにより、予測分布理論の適用が可能になり、モデル選択等の関連した問題を自然に扱うことができる。

本研究課題で扱う統計手法を、統計モデルを利用した解析に応用する際に高次元積分の評価が必要になる。疑似乱数・準乱数とそのオペレーションズ・リサーチへの応用の分野で実績のある研究分担者（諸星穂積教授・政策研究大学院大学）と協力し、準モンテカルロ法を用いた数値積分を用いた解析手法を開発する。

本研究のアプローチは、観測される量と予測される量を自由に設定できるため、1期先予測の性能評価を基に構成される赤池情報量規準と、1期先予測の評価とは異なる基準により提案された Bayes factor, prequential analysis, MDL 等の手法とを橋渡しする理論に成長することが期待できる。

3. 研究の方法

研究代表者が研究を進めてきた、情報幾何に基づく予測分布の理論と条件付相互情報量を最大化に基づく方法とを結び付けることにより研究を行った。従来の手法では扱えない、一般の予測問題における事前分布の構成について、漸近理論・計算統計学的方法の両面から研究を進めた。予測と条件付き相互情報量との関係に基づく理論を発展させることにより、複数のモデルが想定される場合のモデルの予測・推測問題への適用に関する研究を進めることをめざした。

研究分担者の諸星穂積教授と協力して、本研究により開発するベイズ統計学の手法を準乱数を用いて複雑な統計モデルに対しても応用可能にする研究を進めた。金融工学におけるオプションの評価等で利用されている準モンテカルロ法による高次元積分は、ベイズ統計学においても有用である場面が多いと考えられるが、その応用研究はまだ十分にこなされていない。特に準乱数を用いた数値積分により、準モンテカルロ法を用いた、粒子フィルタ、マルコフ連鎖モンテカルロ法の拡張の可能性についても研究を進めた。

さらに、平成 23 年度に日本で開催された統計・確率分野の国際会議 “The 2nd Institute of Mathematical Statistics Asia Pacific Rim Meeting, Tokyo, July 4-6, 2011” において invited discussant としてペンシルベニア大学の E. I. George 教授を招待して議論する機会をもった。

マルチコア計算機を用いて、現実のデータに対する開発したベイズ統計の手法を用いた本格的なモデリングと解析に応用できる手法の開発・研究を行った。

4. 研究成果

ベイズ予測においては、Fisher 情報行列に基づく計量 (Fisher-Rao 計量) とそれに基づく体積要素である Jeffreys 事前分布が基本的な役割を果たすことが従来の研究により知られていた。本研究で、観測される確率変数の確率分布と予測する確率変数の確率分布が、同じパラメータをもつ異なる統計モデルに属するときのベイズ予測では、従来知られていた Fisher 情報行列に基づく計量とは異なる新しい計量構造が自然な役割を果たすことを発見した。この設定は統計学においては、回帰モデルなどでしばしば現れる自然なものである。新しい計量をもつリーマン多様体に基づく考察により、従来知られていなかった性能の良いベイズ予測・推測を構成することが可能となった。また、この結果を多変量ポアソン分布の予測問題に適用し、その有限サンプルにおける性質を調べた。観測されるデータと予測する量が、同じパラメータを持つ異なる多変量ポアソン分布にしたがう問題において、新しい計量に基づ

く優調和関数に基づく事前分布を構成した。さらに、統計的検定の問題を、帰無仮説を0, 対立仮説を1として、0であるのか1であるのかを予測する問題として定式化し、条件付き相互情報量に基づくベイズ予測の理論の枠組みから考察を行った。従来ベイズ検定で自然とされ想定されていた対立仮説に関する事前分布の枠に当てはまらない事前分布がミニマックス予測の観点から自然に現れることを示した。事前分布は、従来想定されていた、単峰型や一様分布型とは異なり、特異な分布となる。また、この方法により求められる仮説の事後確率と、従来のベイズ検定による事後確率と、さらに頻度法によるp値とを比較検討した。

計算統計学的手法として、条件付き相互情報量を最大化することにより事前分布を構成するアルゴリズムの開発研究を行った。

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)

Komaki, F. (2013). Bayesian testing of a point null hypothesis based on the latent information prior, *Entropy*, vol. 15, pp. 4416-4431.

Hirose, Y. and Komaki, F. (2013). Edge selection based on the geometry of dually flat spaces for Gaussian graphical models, vol. 23, pp. 793-800.

Komaki, F. (2012) Asymptotically minimax Bayesian predictive densities for multinomial models, *Electronic Journal of Statistics*, vol. 6, pp. 934-957.

Komaki, F. (2011) Bayesian predictive densities based on latent information priors, *Journal of Statistical Planning and Inference*, vol. 141, pp. 3705-3715.

[学会発表](計10件)

Komaki, F. Bayesian predictive densities when the distributions of data and target variables are different, WITMSE 2013: The Sixth Workshop on Information Theoretic Methods in Science and Engineering (招待講演), 2013年08月28日, The University of Tokyo.

Morohosi, M. and Furuta, T. Optimization model and simulation for improving ambulance service system, The 11th International Symposium on Operations

Research & Its Applications, 2013年08月23日, Huangshan, China.

Morohosi, M. A computational study of likelihood estimation by randomized quasi-Monte Carlo method for filtering problems, Ninth IMAC seminar on Monte Carlo Methods, 2013年07月15日, Annecy-le-Vieux, France.

Furuta, T. and Morohosi, M. Minimizing the spatial mismatch of ambulance service in Tokyo, INFORMS Healthcare 2013, 2013年06月25日, Chicago, USA.

Komaki, F. On the Use of Latent Information Priors, ISBA 2012 World Meeting (招待講演), 2012年6月25日~6月29日, Kyoto Terrasa Conference Center, Kyoto, Japan.

Morohosi, H. and Furuta, T. Hypercube simulation analysis for a large-scale ambulance service system, The 2012 Winter Simulation Conference, 2012年12月09日~12月12日, Berlin, Germany.

Tanaka, T. and Komaki, F. Second-order Matching Priors Depending on Sample Sizes, The 2nd Institute of Mathematical Statistics Asia Pacific Rim Meeting (IMS-APRM 2012), 2012年7月02日~7月04日, Tsukuba International Congress Center, Tsukuba, Japan.

Komaki, F., Bayesian Inference Based on Latent Information Priors, The 7th International Congress on Industrial and Applied Mathematics - ICIAM 2011, 2011年7月18日~7月22日, Vancouver, Canada.

Komaki, F., Bayesian Inference from a Predictive Perspective, The 2nd Institute of Mathematical Statistics Asia Pacific Rim Meeting (IMS APRM 2012)(招待講演), 2012年7月02日~7月04日, Tsukuba International Congress Center, Tsukuba, Japan.

Morohosi, H., Hybrid Quasi-Monte Carlo Method for the Simulation of State Space Models, The Tenth International Symposium on Operations Research and Its Applications (ISORA 2011), 2011年8月28日~8月31日, Dunhuang, China.

[図書](計0件)

〔産業財産権〕

出願状況（計0件）

取得状況（計0件）

6. 研究組織

(1) 研究代表者

駒木 文保 (KOMAKI, Fumiyasu)
東京大学・情報理工学研究科・教授
研究者番号： 7 0 2 4 2 0 3 9

(2) 研究分担者

諸星 穂積 (MOROHOSHI, Hozumi)
政策研究大学院大学・政策研究科・教授
研究者番号： 1 0 2 7 2 3 8 7

(3) 連携研究者

田中 冬彦 (TANAKA, Fuyuhiko)
大阪大学・基礎工学研究科・准教授
研究者番号： 9 0 4 5 6 1 6 1