

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 6 日現在

機関番号：34504

研究種目：若手研究(B)

研究期間：2012～2015

課題番号：24730195

研究課題名(和文) ベイジアンアプローチを用いた離散連続選択モデルのノンパラメトリック推定

研究課題名(英文) Nonparametric Bayesian estimation of discrete/continuous choice models

研究代表者

宮脇 幸治 (MIYAWAKI, Koji)

関西学院大学・経済学部・准教授

研究者番号：40550249

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)： ベイジアンアプローチを用いた離散連続選択モデルのノンパラメトリック推定に関する研究を行った。まず、単純な場合の統計的モデル及びその推定手法を開発することができた。しかし、より現実的なデータを分析していく中で、選択肢の数を減らすための部分が、データの特徴を生かすためには不十分であることが分かり、モデルの修正が必要であると分かった。そのためモデルを修正し、選択肢の数をある程度減らした統計的モデルを開発することに成功した。

研究成果の概要(英文)： I conducted the research on the nonparametric Bayesian approach to the discrete/continuous choice model. I have developed a statistical model and its estimation method for a simple case. However, in analyzing the practical dataset, I found the component that reduces the number of choices is not appropriate for the dataset to be used. Thus, I modified the model, so that the resulting statistical model is appropriate for the dataset as well as reduces the number of choices.

研究分野：計量経済学

キーワード：ベイズ統計学

1. 研究開始当初の背景

離散連続選択モデルとは、離散的な選択と連続的な選択を同時に考慮した統計的モデルを指す。例えば Dubin and McFadden (1984) では暖房器具の選択と電力消費量の選択を同時に考慮した分析を離散連続選択モデルを用いて行っている。その他にも労働供給(Burtless and Hausman (1978)), 通信プラン選択(Iyengar (2004)), 電力需要(Reiss and White (2005)), 消費者のブランド選択(Song and Chintagunta (2007)), 水道需要(Miyawaki, Omori, and Hibiki (2016)), ガス需要(Miyawaki, Omori, and Hibiki (2015))などで離散連続選択モデルが応用されている。そこでは、離散選択の性質にもよるが、離散選択の部分にプロビットモデルと呼ばれる統計的モデルを用いることが推奨される。しかし選択肢の数が多い場合、プロビットモデルのパラメータの推定が困難となる。その理由としては、プロビットモデルはその背後に効用最大化問題を考えているが、選択肢の数が多くなれば効用最大化問題における比較すべき効用の数が増えてしまう。そして比較すべき効用の数が増えれば、パラメータが満たすべき不等式制約が多くなり、推定が困難となる。

しかし現実のデータを見てみると、消費者が自動車を購入するような場合や企業が工場を立地する場合の意思決定などでは、通常のプロビットモデルが想定している選択肢の数より多い選択肢を含むような離散選択はしばしば見られる。そのようなケースにおける分析では、ロジットモデルやネストしたモデルなどを用いることが主であるが、それらには別の問題が存在する。つまり、ロジットモデルには無関係な代替案からの独立と呼ばれる問題があり、またネストしたモデルにはどのようなネストを考えるかという問題が存在する。前者は、意思決定に無関係な選択肢に関する問題である。より具体的には、そのような選択肢が含まれている場合とそうでない場合において、他の選択肢が選ばれる確率を考える。合理的な意思決定をしている場合、そのような選択確率は同じであると考えられるが、ロジットモデルにおいては確率はそのような選択肢に影響を受けてしまう。なおプロビットモデルではこのような問題はない。また後者は、自動車の選択の場合、まず国産車とそれ以外に分けて、さらに国産車をいくつかのカテゴリに分けてといったように選択肢の階層構造を考えることである。しかし消費者がこのようなネストを考慮して意思決定しているとは限らない。以上の議論から、既存の統計的手法で選択肢の数が多い場合の離散選択を扱うのは困難であるといえる。

一方ノンパラメトリックな推定を用いることができれば、データから選択肢のグループを判別できると期待される。そして、もし

選択肢をうまくグループ化できれば、離散選択における選択肢の数を減らすことができるため、例えばネスト構造に関する仮定が不要になると考えられる。またプロビットモデルを基にすれば、無関係な代替案からの独立という仮定は不要となる。以上のような問題意識の下、離散連続選択モデルにおけるノンパラメトリックなベイズ推定に関する研究を行った。

参考文献

Burtless, G. and J. A. Hausman (1978). The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment. *Journal of Political Economy* 86(6), 1103-1130.

Dubin, J. A. and D. L. McFadden (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica* 52(2), 345 - 362.

Iyengar, R. (2004). A structural demand analysis for wireless services under nonlinear pricing schemes. unpublished working paper, Marketing Department, Wharton School, University of Pennsylvania.

Miyawaki, K., Y. Omori, and A. Hibiki (2015). A Discrete/Continuous Choice Model on a Nonconvex Budget Set. Forthcoming at *Econometric Reviews*.

Miyawaki, K., Y. Omori, and A. Hibiki (2016). Bayesian estimation of demand functions under block rate pricing. *Econometric Reviews* 35 (3), 311-343.

Reiss, P. C. and M. W. White (2005). Household electricity demand, revisited. *Review of Economic Studies* 72, 853-883.

Song, I. and P. K. Chintagunta (2007). A discrete-continuous model for multicategory purchase behavior of households. *Journal of Marketing Research* 44(4).

2. 研究の目的

本研究の目的は、離散連続選択モデルにおけるノンパラメトリックなベイズ推定に関する統計的モデル及び推定手法の開発である。より具体的には、離散連続選択モデルにおけるノンパラメトリックなベイズ推定の研究とそれを用いた離散選択パートにおける選択肢の数を減らすような統計的モデルの開発及び推定手法の開発である。

3. 研究の方法

離散連続選択モデルにおけるノンパラメ

トリックなベイズ推定に関する研究を行う。ノンパラメトリックなベイズ推定に関しては、様々な方法が提案されているが、主にディリクレ過程事前分布を用いたものに関する研究を行う。ディリクレ過程事前分布とは、Ferguson (1973, 1974) によって提案されたベイズ推定の一つで、その特徴はパラメータをいくつかのグループに（ある程度）ノンパラメトリックに分けることが可能であるということである。もしこれらのグループを用いて選択肢を潜在的にグループ化できれば、離散選択における選択肢の数をノンパラメトリックに推定することが可能となり、その結果意思決定に用いる選択の数を減らすことができる。実際、ディリクレ過程事前分布を用いた統計的モデルのベイズ推定は、近年の計算機の発達とマルコフ連鎖モンテカルロ法の研究に伴い、Escobar (1994) を始めとして、盛んに行われている (Ghosh and Ramamoorthi (2003) も参照)。ここでは推定手法をどのように効率化するかといったものから、より柔軟な統計的モデルの提案といったものまで幅広く行われている。またディリクレ過程事前分布を用いた統計的モデルがうまく機能しなかった場合を考慮して、ディリクレ過程事前分布を用いた基本的な手法以外の潜在的なグループ化に関する他の統計手法についてサーベイを行う。これにより、ディリクレ過程事前分布の特質をより深く理解でき、また代替的な手法に関する検討を併せて行うことができる。

参考文献

Escobar, M. D. (1994). Estimating normal means with a Dirichlet Process Prior. *Journal of the American Statistical Association* 89(425), 268-277.

Ferguson, T. (1973). Bayesian analysis of some nonparametric problems. *Annals of Statistics* 1(2), 209-230.

Ferguson, T. (1974). Prior distributions on spaces of probability measures. *Annals of Statistics* 2(4), 615-629.

Ghosh, J. K. and R. V. Ramamoorthi (2003). *Bayesian Nonparametrics*. New York: Springer.

4. 研究成果

まずディリクレ過程事前分布を用いたサンプルセレクションモデルの研究を行った。サンプルセレクションモデルとは、離散的な選択と連続的な選択を同時に考慮している統計的モデルであり、最も単純な形の離散連続選択モデルの一つと見ることができ。そのモデルにおいてディリクレ過程事前分布を用いることで、どのようにノンパラメトリックなベイズ推定が行われているか、またそ

れがサンプルセレクションモデルにおいてどのように機能するかなどに関して研究を行った。そして、特にディリクレ過程事前分布の構成方法として、Sethuraman (1994) によるものが近年しばしば用いられており、応用範囲が広いことが分かった。続いてディリクレ過程事前分布を組み込んだサンプルセレクションモデルを開発した。その結果、通常の推定手法ではパラメータの推定が難しくなることが分かったため、推定手法についてのサーベイを行った。そして、離散連続選択モデルのような複雑なモデルの場合、近似的な推定手法が有用であることが分かった。近似的な手法とは Sethuraman (1994) によるディリクレ過程事前分布の構成方法を基にして Ishwaran and James (2001) によって提案されているものである。近似的な手法を用いる利点としては、パラメータをベイズ推定するための手法であるマルコフ連鎖モンテカルロ法がギブスサンプラーと呼ばれる比較的簡単なものになることが挙げられる。得られた成果は論文としてまとめ、紀要への投稿を行った。

次により現実的な分析を行うため、消費者の自動車の購買に関するデータを分析し、消費者の意思決定をどのように推定できるかについて研究を行った。データにはどのような消費者がどのような自動車を購入したかという履歴が含まれており、離散的な意思決定を行っていると見ることができ。そしてそのような意思決定は、通常プロビットモデルと呼ばれる統計的モデルで扱われる。そこにおいて、まずノンパラメトリックなベイズ推定をどのように組み込むかに関して研究を行った。また潜在的なグループ化に関する代替的な統計的な手法についてのサーベイも同時に行った。その結果、ディリクレ過程事前分布の応用的手法からディリクレ過程事前分布とは異なるアプローチによる手法まで様々な手法があることが分かった。応用的手法としては、例えば Quintana and Iglesias (2003) が挙げられ、本手法はマーケティング分野での応用が見られることが分かった。また異なるアプローチとしては Fraley and Raftery (2002) などが挙げられる。

しかし自動車の購買データの分析を行った結果、主にデータの特徴を生かすためには統計的モデルの大幅な修正が必要なが分かった。そのため、ディリクレ過程事前分布によるノンパラメトリックな推定ではなく、データの特徴を生かし、かつディリクレ過程事前分布における選択肢のグループ化というアイデアを生かせるような、新しい離散選択モデルの開発を行った。ここでは、ディリクレ過程事前分布に関するこれまでの理解と潜在的なグループ化が統計的にどのように考えられているかについてのサーベイを生かすことができた。新しいモデルはいくつかの特徴を有している。まずプロビッ

トモデルを基にしたモデルであるため、意思決定の際に効用最大化を行うと仮定するミクロ経済学の考え方と整合的である。また、データの特徴を生かしている点に加えて、選択枝のグループ化を考慮しているため、意思決定に用いられる選択の数を減らしている点も特徴として挙げられる。特に選択枝のグループ化に関して、他研究者との議論から、より柔軟な統計的モデルにするために階層的な構造を導入した。

また新しいモデルに関する推定手法の開発を行った。手法はマルコフ連鎖モンテカルロ法を用いた通常のものとなった。ただし一部に他研究者からの意見を基にしたより効率的な手法を実装した。その結果、比較的許容できる計算時間での推定が可能となった。

手法の開発においては、パラメータの識別問題についての議論も行った。新しいモデルでは、選択枝のグループ化を導入した結果、それらのグループのラベリングが任意であるという特徴がある。この特徴は、一般には識別問題と呼ばれる統計学上の研究トピックに含まれており、例えば Fruhwirth-Schnatter (2001) ではどのような推定手法がより効率的に推定を行えるかについて議論されている。本研究において分析を行った結果、ラベリングの問題は推定手法の効率に関する問題を引き起こしていることが分かった。そのため、どのような制約条件を課せばパラメータを効率的に推定できるかについて検討を行った。検討の結果、全てのパラメータを識別するような制約条件の数はパラメータ数が増えるに従って爆発的に増えることが分かったため、そのような制約条件を課すことは難しいことが分かった。しかし、直観的かつある程度弱い条件を課すことは可能であり、弱い形で識別が可能であることが分かった。現在、実データを用いた分析を行っており、分析が終わり次第得られた成果を論文にまとめ、査読誌に投稿する予定である。

参考文献

Fraley, C. and A. E. Raftery (2002). Model-Based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association* 97(458), 611-631.

Fruhwirth-Schnatter, S. (2001). Markov chain Monte Carlo estimation of classical and dynamic switching and mixture models. *Journal of the American Statistical Association* 96(453), 194-209.

Ishwaran, H. and L. F. James (2001). Gibbs sampling methods for stick-breaking priors. *Journal of the American Statistical Association* 96(453), 161-173.

Quintana, F. A. and P. L. Iglesias (2003). Bayesian clustering and product partition models. *Journal of the Royal*

Statistical Society 65(2), 557-574.

Sethuraman, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica* 4, 639-650.

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 1 件)

(1) 宮脇幸治, ディリクレ過程事前分布を用いたベイズ推定, 商学論究, 査読無, 61巻, 2014, pp. 173 - 186.

[学会発表](計 0 件)

[図書](計 0 件)

6. 研究組織

(1) 研究代表者

宮脇 幸治 (MIYAWAKI, Koji)

関西学院大学・経済学部・准教授

研究者番号: 40550249