

平成 30 年 6 月 12 日現在

機関番号：12501

研究種目：若手研究(B)

研究期間：2015～2017

課題番号：15K17036

研究課題名(和文)多様な形態のデータに対する分位点回帰モデルと内生性についてのベイズ解析

研究課題名(英文)Bayesian quantile regression with endogeneity for various type of data

研究代表者

小林 弦矢(Kobayashi, Genya)

千葉大学・大学院社会科学研究院・准教授

研究者番号：00725103

交付決定額(研究期間全体):(直接経費) 3,000,000円

研究成果の概要(和文):本研究では、内生変数を含むベイズ分位点回帰モデルとその推定方法の開発を行った。提案するモデルは通常の変数法と同様に2本の式からなり、一段目の回帰残差を二段目にコントロール変数として導入することにより、内生性によるバイアスの除去を可能とした。また、内生変数の非線形の効果の推定も試みた。非線形分位点回帰は各分位点を柔軟に推定することができるが、サンプル数が小さい場合や分布の裾にあたる分位点を推定する場合には、推定結果が不安定になる。よって本研究では、データの分布としては形状に大きな制約がある非対称ラプラス分布の代わりに一般化非対称ラプラス分布を導入し、新たなモデルの定式化を行った。

研究成果の概要(英文): In this project, we developed the Bayesian quantile regression models with endogenous covariates and estimation methods for the proposed models. The proposed model consists of two equations as in the conventional instrumental regression where the second stage regression includes the residuals from the first stage regression in order to remove the bias from the endogeneity. We also considered introducing nonlinear effects of endogenous covariates on the quantiles of the response variable. Although the nonlinear quantile regression is a flexible approach to quantile estimation, the estimates are known to be unstable when the sample size is small or when the quantiles in the tails are estimated. We replace the asymmetric Laplace distribution, whose shape is severely restrictive for a data distribution, with the generalised asymmetric Laplace distribution for improved estimation performance.

研究分野：ベイズ統計学

キーワード：分位点回帰 ベイズ統計学 非線形回帰

1. 研究開始当初の背景

分位点回帰モデル (quantile regression model) は被説明変数の条件付確率分布の分位点と説明変数との関係を分析するための計量モデルである。分位点回帰モデルはデータの分布について詳しく調べることができるため、分位点回帰モデルは理論と応用の両方の観点から多くの注目を集めてきている。特に近年では、打切りデータや、質的反応データ、空間データ、時系列データ、個人の異質性を考慮したパネルデータなどを分析する分位点回帰モデルの開発が盛んに行われてきている。

従来の最小二乗推定量と同様に、回帰分析に内生変数が含まれる場合には分位点回帰モデルに対する通常の推定量はバイアスを持つことが知られている。内生性のある分位点回帰モデルに対しては、二段階推定法や操作変数法などが提案されているが、これらは現在のところ非説明変数が連続な横断面データにのみ適用可能である。またコントロール変数 (CV) 法においてはモデル特定化や推定方法において多くの問題が未解決である。例えば、誤差分布の分位点に対する条件 (α 分位点がゼロとなるような α の値) を分析者が決定する必要がある。誤差分布のそのような値を先験的に特定するのは極めて困難であり、未だに明示的な指針が存在しない。

2. 研究の目的

本研究ではミクロ計量分析において近年多くの関心を集めてきている打切りデータ、空間データ、パネルデータなどを取り上げ、これらに対する計量モデルの作成を行う。内生性に対処するために本研究では CV 法を参考にしつつ、従来の CV 法における計量経済学上の問題点を克服するために、階層ベイズモデルによる定式化を行う。本研究で提案するモデルは潜在変数を多く含む階層モデルであるため、解析的に解けないことが予想される。そのため、シミュレーションによる推定方法 (主にマルコフ連鎖モンテカルロ法) の開発を目指す。

3. 研究の方法

(1) CV を用いた計量モデルの開発を行った。内生性に対処するために、CV を導入し、打ち切りのある変数に対して潜在変数を導入した階層モデルを考えた。まずは一段階目の回帰モデルの誤差分布の α 分位点がゼロであるパラメトリックな分布 (非対称なラプラス、正規、指数べき乗分布) を導入し、簡便なパラメトリックモデルを作成した。非対称ラプラスや非対称正規と行った分布は裾の形状に柔軟性を欠くため、柔軟性を向上するためにディリクレ過程混合モデルを用いたセミパラメトリックモデルもこれらの分布を基に考えた。 α はモデル特定化の一部であり事前にこの値についての情報はないため、本研究ではデータから推定するパラメ

ータとして扱った。上記の分布を用いたモデルでは α 分位点と分布のモードが一致するため、本研究での一段階目の回帰モデルはモード回帰モデルとしても見ることができる。二段階目の分位点回帰モデルでは、一段階目の回帰モデルの残差を説明変数 (コントロール変数) として導入し、その回帰係数によって内生性の度合いを推測することができるようにし、誤差分布は通常のベイズ分位点回帰モデルに従い非対称ラプラス分布を用いた。

(2) 内生的な共変量がモデルに含まれる際の、被説明変数の条件付分布を柔軟に推定するために、操作変数法を用いたノンパラメトリック回帰モデルを考案した。具体的には位置尺度モデルを考え、一段階目と二段階目の回帰モデルに操作変数と内生変数の非線形な関数を導入し、さらに二段階目の回帰モデルの誤差の分散に対しても共変量の非線形な関数を導入した。そして誤差分布を柔軟にモデル化するために、誤差分散にコレスキー分解を適用し、ディリクレ過程によって二変量正規分布のディリクレ過程混合モデルを考えた。

提案するモデルには複雑であるが、ディリクレ過程に関する変数をサンプリングするスライスサンプラーと、分散関数に関する変数をサンプリングする混合サンプラーからなるマルコフ連鎖モンテカルロ法を適用することによって、効率的にモデルを推定することが出来る。提案するモデルは戦前の日本の各都道府県の死亡率と医師数に関するデータに対して適用した。ここで、医師数は各都道府県の観測することができないウエルスとの相関を持っていることが考えられるので内生変数として扱う。操作変数は各都道府県の面積を用いることにした。医師が増えるに従って死亡率が減少するという結果が得られた。また DIC や WAIC などの情報量基準に基づき、医師の死亡率に対する効果が非線形的であるモデルが支持されるという結果も得られた。

(3) 被説明変数の分位点に非線形な効果をもつ分位点回帰モデルを取り上げた。非線形なコンポネントを持つ分位点回帰モデルは、通常の線形分位点回帰のように非対称ラプラス分布を誤差項に仮定することで定式化することが可能であるが、サンプル数が少ない場合や、分布の裾にあたる分位点を推定する場合には、過適合などにより推定結果が不安手になってしまい、分位点の単調性を満たさなかつたり解釈可能性に欠けたりしてしまうことが起きる。本研究課題で取り上げるような内生変数を含むモデルにおいては、これは大きな問題である。よって、データの分布としては形状に大きな制約がある非対称ラプラス分布の代わりに、誤差項をより柔軟な一般化非対称ラプラス分布を分位点回帰モデルによって、また非線形コンポネントはガウス過程によつ

て新たな非線形分位点 回帰モデルの定式化を行った。またガウス過程を用いた非線形関数の形状制約の導入も分位点回帰モデルの枠組みで試みた。

4. 研究成果

(1) 様々な設定を用いた数値実験により、本研究で提案するモデルは多くの状況で関心のあるパラメータをうまく推定することができ、セミパラメトリックモデルのパフォーマンスが比較的良好なことがわかった。また、推定結果は事前分布の設定に対して頑健であり、提案する推定方法も効率的であることがわかった。既婚女性の労働参加に関する実証分析も行い、労働時間の打ち切りのある周辺の分位点において、妻以外による家計収入が妻の労働時間に対して内生的な影響があることがわかった。ここまでの研究結果は Bayesian Analysis に掲載された。

(3) 提案するモデルは戦前の日本の各都道府県の死亡率と医師数に関するデータに対して適用した。ここで、医師数は各都道府県の観測することができないウェルスとの相関を持っていることが考えられるので内生変数として扱う。操作変数は各都道府県の面積を用いることにした。医師が増えるに従って死亡率が減少するという結果が得られた。また DIC や WAIC などの情報量基準に基づき、医師の死亡率に対する効果が非線形的であるモデルが支持されるという結果も得られた。ここまでの研究結果は Korean International Statistical Society などで発表された。

(3) 数値実験の結果より、提案するモデルのほうが非対称ラプラス分布を用いたモデルよりも安定的に分位点を推定できることがわかった。また労働・環境・所得などの実データへの適用から、提案するモデルからはより安定的で解釈可能性が高い推定結果が得られることがわかった。ここまでの研究結果は 国際学会 EcoSta2017 での招待講演で報告を行った。

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 3 件)

1. Ogasawara, K., Shirota, S. and Kobayashi, G. (2018). Public health improvements and mortality in Tokyo: a Bayesian disease mapping approach, *Cliometrica*, 12, 1-31.
2. Kobayashi, G. (2017). Bayesian endogenous Tobit quantile regression, *Bayesian Analysis*, 12, 161-191.
3. Kobayashi, G. (2016). Skew exponential power stochastic volatility model for analysis of skewness, non-normal tails, quantiles and expectiles,

Computational Statistics, 31, 49-88.

〔学会発表〕(計 7 件)

1. Kawakubo, Y. and Kobayashi, G.. Small Area Estimation for grouped data, Small Area Estimation 2017, Paris, France.
2. Kobayashi, G., Choi, T. and Roh, T.. Bayesian spectral analysis quantile regression models with shape restrictions, The 1st International Conference on Econometrics and Statistics, 2017, HKUST, HongKong.
3. Kobayashi, G.. Bayesian nonparametric instrumental variable regression approach to quantile inference, The First Eastern Asian Meeting on Bayesian Statistics, 2016, Jiao Tong University, China.
4. Kobayashi, G.. Bayesian nonparametric instrumental variable regression approach to quantile inference, Korean International Statistical Society, 2016, Daejeon, South Korea.
5. Kobayashi, G.. Bayesian endogenous Tobit quantile regression, The 9th International Conference on Computational and Financial Econometrics, 2015, University of London, UK.
6. Kobayashi, G. and Kakamu, K. Approximate Bayesian computation for Lorenz curves from grouped data, Joint Statistical Meeting 2015, Seattle, USA.
7. Ogasawara, K., Shirota, S. and Kobayashi, G.. The impact of piped water on the mortality in inter-war Tokyo: Bayesian disease mapping approach, Joint Statistical Meeting 2015, Seattle, USA.

〔図書〕(計 件)

〔産業財産権〕

出願状況(計 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 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

千葉大学・大学院社会科学研究院・准教授

小林 弦矢 (GENYA KOBAYASHI)

研究者番号：00725103

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：

(4) 研究協力者

()