

令和 4 年 5 月 7 日現在

機関番号：14501

研究種目：若手研究

研究期間：2019～2021

課題番号：19K13670

研究課題名（和文）因果推論、欠損データ分析、コピュラモデルを結びつける革新的アプローチ

研究課題名（英文）An innovative approach for unifying causal inference, missing data analysis, and copula models

研究代表者

茂木 快治（MOTEGI, KAIJI）

神戸大学・経済学研究科・准教授

研究者番号：60742848

交付決定額（研究期間全体）：（直接経費） 3,200,000円

研究成果の概要（和文）：本研究課題では、カリブレーションという革新的アプローチを用いて、因果推論、欠損データ分析、コピュラモデルという3つの研究分野を融合させた。因果推論、欠損データ分析、コピュラモデルは、それぞれ統計学や計量経済学において盛んに研究されているが、これら3分野を統一的なアプローチの下で結びつけたのは本研究が初めてである。本研究により、欠損データに対するコピュラ回帰が実行可能となり、また様々な処置や処置効果を許容する因果推論が実行可能となった。

研究成果の学術的意義や社会的意義

社会科学系の統計分析では、様々な理由により頻繁にデータの欠損が発生する。したがって、欠損データに対する新たな統計分析の手法を確立した本研究は、学術面・実務面で重要な貢献を果たしたと言える。実際、ドイツ製造業の個別企業データやアメリカ大統領選挙の広告・献金データに対して提案の分析手法を応用した結果、従来の分析手法では得られない新たな知見を得ることに成功した。これらは企業の経営戦略や政党の選挙戦略の進化に資する重要な知見である。

研究成果の概要（英文）：In this research project, I have successfully unified causal inference, missing data analysis, and copula models by adopting a novel approach called calibration. Each of causal inference, missing data analysis, and copula models has been well studied in statistics and econometrics, but this project is the first one that unifies these three topics. I have established copula-based regression with data missing at random and causal inference for general treatment types and treatment effects.

研究分野：計量経済学

キーワード：因果推論 欠損データ コピュラ カリブレーション 回帰分析

## 1. 研究開始当初の背景

本研究の開始以前より、因果推論、欠損データ分析、コンピュータモデルという3つの研究分野は、統計学や計量経済学などの学問領域においてそれぞれ盛んに研究されていた。しかし、因果推論、欠損データ分析、コンピュータモデルは別々に研究されることが多く、これら3分野を統一的なアプローチの下で結びつけようとする分野横断的研究は存在していなかった。そのような背景もあり、これらの研究分野には次の3つの課題が存在していた。

### 【研究課題1: 欠損データに対するコンピュータモデルの実行方法の開発】

コンピュータモデルとは、多変数の複雑な相互依存関係を比較的少ないパラメータで記述することのできる有力なモデルである。従来のコンピュータモデルは、すべての個体について完備データ（欠損のないデータ）が手に入るという前提の下で構築されていた。一部の個体が不完備データを有するときのコンピュータモデルの実行方法は、未知の研究課題として残されていたのである。社会科学系の統計分析では、様々な理由によりデータの欠損が頻繁に生じる。したがって、欠損データに対するコンピュータモデルの当てはめ方を確立することは重要な課題であった。

### 【研究課題2: 欠損データに対するコンピュータ回帰の実行方法の開発】

欠損データに対するコンピュータ回帰の実行方法も、未解決の研究課題であった。コンピュータ回帰は、一般性の高い回帰関数とパラメータの少なさを両立する回帰分析として近年注目を集めている。コンピュータ回帰は中間ステップにおいてコンピュータモデルの当てはめを必要とする。そのため、本研究開始当初のコンピュータ回帰の対象は完備データに限定されており、不完備データへの拡張が学術的・実務的に望まれている状況であった。

### 【研究課題3: 因果推論における処置と処置効果の一元化】

因果推論とは、ある処置  $T$  が結果  $Y$  に対して与える効果（処置効果）を推定することである。因果推論の実行時、重要なポイントが2つある。第一のポイントは、処置  $T$  のとり得る値の範囲である。最も単純なケースでは、処置  $T$  は二項的な値をとる（処置を行うか否かの二者択一）。より一般的なケースでは、処置  $T$  は多項的もしくは連続的な値をとる。第二のポイントは、どのような種類の処置効果を推定のターゲットとするかである。最も単純なケースでは、平均的な処置効果がターゲットとなる。より複雑なケースでは、中央値や四分位点などの分位点効果がターゲットとなる。連続的処置や分位点効果などの発展的ケースは、それぞれ別個に研究されていたものの、それらすべてを網羅した包括的研究は存在していなかった。つまり、本研究開始以前の因果推論では、処置と処置効果の種類に関する統一的な視点が欠けていたのである。

## 2. 研究の目的

本研究の目的は、上述の研究課題1~3を解決することである。研究課題1では、欠損データに対するコンピュータモデルの実行方法を提案する。そして、大標本および小標本において、提案の手法が完備データを前提とする既存の手法よりも高い統計的精度を有することを示す。

研究課題2も同様にして、欠損データに対するコンピュータ回帰の実行方法を提案し、それが統計的に望ましい性質を有することを理論的・数值的・実証的に示す。

研究課題3では、二項的・多項的・連続的な処置を含み、かつ平均処置効果と分位点処置効果の両方を含む極めて一般性の高い枠組みの下で、因果推論の手法を構築する。さらに、提案の手法が統計的に望ましい性質を有することを理論的・数值的・実証的に示す。

## 3. 研究の方法

本研究は、カリブレーションという革新的なアプローチを駆使して、研究課題1~3を統一的に解決する。カリブレーションはもともと因果推論の分野で提案された手法であり、処置を施した個体と施さなかった個体との間のバランスを考慮して、各個体に適切な重み（ウェイト）を付与するというアプローチである。欠損データ分析においては、データが観察されるという事象をひとつの処置と見なし、データの観察される個体と観察されない個体との間のバランスが保たれるように各個体を重み付けすべきである。カリブレーションを欠損データ分析に応用したのは、本研究が初めてである。

研究課題1, 2に関しては、カリブレーションにより各個体に適切な重みを付与した上で、コンピュータモデルおよびコンピュータ回帰を当てはめる。適切な重み付けという工夫により、データの欠損メカニズムが **Missing at Random (MAR)** のとき、統計的に望ましい分析結果が得られるようになる。研究課題3に関しては、一般性の高い処置および処置効果を想定した上でカリブレーションを実行し、統計的に望ましい因果推論が可能となることを示す。

#### 4. 研究成果

研究課題 1~3 に関して、カリブレーションを駆使した提案手法が統計的に望ましい性質を備えていることを示した。第一に、サンプルサイズが十分に大きいとき、提案手法が既存手法よりも高い統計的精度を有することを数学的に証明した。第二に、数値実験の結果、現実的なサンプルサイズの下で提案手法が既存手法よりも高い統計的精度を有することが明らかとなった。カリブレーションという統一的アプローチに基づいて研究課題 1~3 を解決したという点で、本研究は因果推論、欠損データ分析、コンピュータモデルの融合に成功したと言える。

研究課題 2 の実証分析では、ドイツの製造業 500 社の売上と利益にコンピュータ回帰を応用した。売上は全社について観察されるが、利益は一部の企業についてのみ観察されるため、提案の手法が有効となる。実証分析の結果、データの欠損を明示的に考慮した場合と考慮しない場合では得られる回帰曲線が有意に異なることが明らかとなり、提案手法の有用性が示唆された。これは企業の経営戦略の進化に資する新たな知見である。

研究課題 3 の実証分析では、米国大統領選挙における政治広告の政治献金に対する処置効果を分析した。政治広告を二項的に定義した場合は統計的に有意な処置効果が観察された一方、連続的に定義した場合は有意な処置効果が観察されなかった。これは政党の選挙戦略の進化に資する興味深い知見である。

#### 研究成果物（査読付き英文学術雑誌掲載論文）

- [1] Shigeyuki Hamori, [Kaiji Motegi](#), and Zheng Zhang (2019). Calibration estimation of semiparametric copula models with data missing at random. *Journal of Multivariate Analysis*, vol. 173, pp. 85-109. (※ 研究課題 1 の成果物. JMVA は SCIE 該当の統計学系ジャーナル.)
- [2] Shigeyuki Hamori, [Kaiji Motegi](#), and Zheng Zhang (2020). Copula-based regression models with data missing at random. *Journal of Multivariate Analysis*, vol. 180, #104654. (※ 研究課題 2 の成果物. JMVA は SCIE 該当の統計学系ジャーナル.)
- [3] Chunrong Ai, Oliver Linton, [Kaiji Motegi](#), and Zheng Zhang (2021). A unified framework for efficient estimation of general treatment models. *Quantitative Economics*, vol. 12, pp. 779-816. (※ 研究課題 3 の成果物. QE は SSCI 該当の経済学系トップジャーナル.)

## 5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件/うち国際共著 3件/うちオープンアクセス 1件）

1. 著者名 Shigeyuki Hamori, Kaiji Motegi, and Zheng Zhang	4. 巻 180
2. 論文標題 Copula-based regression models with data missing at random	5. 発行年 2020年
3. 雑誌名 Journal of Multivariate Analysis	6. 最初と最後の頁 #104654
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jmva.2020.104654	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Shigeyuki Hamori, Kaiji Motegi, and Zheng Zhang	4. 巻 173
2. 論文標題 Calibration estimation of semiparametric copula models with data missing at random	5. 発行年 2019年
3. 雑誌名 Journal of Multivariate Analysis	6. 最初と最後の頁 85-109
掲載論文のDOI（デジタルオブジェクト識別子） 10.1016/j.jmva.2019.02.003	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Chunrong Ai, Oliver Linton, Kaiji Motegi, and Zheng Zhang	4. 巻 12
2. 論文標題 A unified framework for efficient estimation of general treatment models	5. 発行年 2021年
3. 雑誌名 Quantitative Economics	6. 最初と最後の頁 779-816
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 該当する

〔学会発表〕 計3件（うち招待講演 2件/うち国際学会 2件）

1. 発表者名 Kaiji Motegi
2. 発表標題 Copula-based regression models with responses missing at random: A unified approach
3. 学会等名 2019 Japanese Joint Statistical Meeting（国際学会）
4. 発表年 2019年

1. 発表者名 Kaiji Motegi
2. 発表標題 Copula-based regression models with data missing at random: A unified approach
3. 学会等名 Departmental seminar, Department of Economics, University of Essex (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 茂木快治
2. 発表標題 Copula-based regression models with data missing at random: A unified approach
3. 学会等名 慶應義塾大学経済研究所「計量経済学ワークショップ」(招待講演)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

研究代表者(茂木快治)の個人ウェブサイト <a href="http://www2.kobe-u.ac.jp/~motegi/">http://www2.kobe-u.ac.jp/~motegi/</a> 研究代表者の所属機関の教員紹介ページ <a href="http://www.econ.kobe-u.ac.jp/faculty/fields/econometrics/motegi.html">http://www.econ.kobe-u.ac.jp/faculty/fields/econometrics/motegi.html</a>
--

6. 研究組織		
氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関			
中国	中国人民大學			
中国	CUHK, Shenzhen			
英国	ケンブリッジ大学			