

令和 6 年 6 月 5 日現在

機関番号：12701

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K11856

研究課題名（和文）リスク回避型データサイエンス技術を開発するための統計的因果推論の研究

研究課題名（英文）Statistical causal inference for developing the risk-averse data science technique.

研究代表者

黒木 学（Kuroki, Manabu）

横浜国立大学・大学院工学研究院・教授

研究者番号：60334512

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では、統計的因果推論の観点から、リスク回避行動の有効性などを定量的に評価するためのデータサイエンス技術を開発するために、(1) 総合効果の推定精度を向上させる統合型推定量の開発、(2) 潜在反応タイプの存在割合の統計的推測法の開発、(3) 平均および分散に対する因果効果の統計的推測法の開発、(4) 総合効果の分解性を利用した公平性配慮型尺度の開発、を行った。

研究成果の学術的意義や社会的意義

本研究の成果をとおり、リスク回避行動の有効性などを定量的に評価するためのデータサイエンス技術を開発するうえで重要な役割を果たす統計的因果推論技術のいくつかを構築したことに学術的意義があると考えます。

研究成果の概要（英文）：In order to develop data science techniques to evaluate the effect of risk-averse behavior based on statistical causal inference, we conducted the following research:(1) the development of integrated estimators to improve the estimation accuracy of total effects, (2) the development of statistical estimation methods for estimating the proportion of potential outcome types, (3) the development of statistical estimation method for causal effects on the mean and variance, and (4) the development of fairness-aware measures based on the effect decomposition.

研究分野：統計科学

キーワード：構造的因果モデル 潜在反応タイプ

## 様式 C - 19, F - 19 - 1 (共通)

### 1. 研究開始当初の背景

膨大な個人データが集積され、高度なデータマイニング技術が簡単に利用できるようになった現在、リスク回避行動を適切に選択する情報技術を開発するために、統計的因果推論に対する期待が高まりつつある。それにもかかわらず、統計的因果推論という考え方がデータマイニング技術に積極的に導入された形跡はなかった。そこで、本研究では、構造的因果モデルのフレームワークに基づいて、現行のリスク回避行動の有効性などを定量的に評価するためのデータサイエンス技術を開発することが重要であると考えた。

### 2. 研究の目的

本研究では、以下の項目を中心に研究を行った。

- (1) 総合効果の推定精度を向上させる統合型推定量の開発
- (2) 潜在反応タイプの存在割合の統計的推測法の開発
- (3) 平均および分散に対する因果効果の統計的推測法の開発
- (4) 総合効果の分解性を利用した公平性配慮型尺度の開発

### 3. 研究の方法

本研究は、国内では、横浜国立大学内で組織した統計的因果推論研究会および日本品質管理学会の計画研究会であるテクノメトリクス研究会を母体として実施した。特に、日本品質管理学会の計画研究会であるテクノメトリクス研究会には積極的に参加・議論することで、幅広い学術的視野を取り入れるよう心がけてきた。また、統計的因果推論研究会では、博士後期課程の学生との共同研究を中心に、統計的因果推論の理論構築を行った。加えて、産業界からのニーズに応えるべく、いくつかの共同研究を実施した。一方、海外では、Eli Lilly and CompanyのYongming Qu博士、Iowa State UniversityのJin Tian教授との共同研究を推し進めた。

### 4. 研究成果

2019年度は、(1) 統合型条件付き操作変数推定量の定式化とその数理的性質の解明、(2) 平均に対する因果効果の予測区間の定式化、を行った。(1)について、複数の総合効果推定量を統合することでより総合効果をより精度よく推定することが可能となることを明らかにした。(2)について、平均に対する因果効果について予測区間を与えるとともに単回帰モデルの予測区間との比較を行った。

2020年度は、(1) 神田・黒木 (2020) によって提案された重み付き条件付き操作変数推定量の拡張、(2) 中間変数に基づく平均に対する因果効果の予測区間の定式化、を行った。(1)について、IV-pairと呼ばれる概念を用いることで、神田・黒木(2020)を包括した総合効果の一致推定量が得られることを明らかにした。(2)について、フロントドア基準に基づいて推定された平均に対する因果効果について厳密な予測区間を与えるとともに単回帰モデルの予測区間との比較を行った。

2021年度は、新たな原因の確率の識別可能条件の導出と統計的推測法の開発を行った。原因の確率を定量的に評価することは実質科学における主要課題の一つであるが、事象間の因果関係に密接に関係する問題であるため、一般には観測情報だけでは評価できないことが指摘されてきた。この問題に対して、ある条件の下では、共変量情報を用いることで、原因の確率が推定可能となることを示したうえで、拡張ラグランジュ法を用いた推定法を開発した。

2022年度は、(1) 総合効果の推定精度を向上させる統合型推定量の開発、(2) 実験データにおける潜在反応タイプの存在割合の識別可能条件の導出と統計的推測法の開発、を行った。(1)について、線形構造方程式モデルのフレームワークの下で、条件付き操作変数推定量と効果復元推定量を統合し

た新たな総合効果の一致推定量を開発し、その統計的性質を明らかにした。(2)について、2021年度に引き続き、実験研究のフレームワークの下で潜在反応タイプの存在割合を定量的に評価する問題に取り組み、共変量情報を用いることで原因の確率が推定可能となることを示した。

2023年度は、(1) 総合効果の推定精度を向上させる統合型推定量の開発、(2) 実験データと観察データを融合させた潜在反応タイプの存在割合の識別可能条件の導出と統計的推測法の開発、(3) ノンパラメトリック構造的因果モデルに基づく効果の分解性を利用した公平性配慮型尺度の開発、を行った。(1)について、線形構造方程式モデルのフレームワークの下で、条件付き操作変数推定量とフロントドア基準を統合した新たな総合効果の一致推定量を開発し、その統計的性質を明らかにした。(2)について、2022年度に引き続き、実験データと観察データを融合させることにより、単一の代替共変量を観測するだけで潜在反応タイプの存在割合が識別可能となる条件を提案した。(3)について、公正配慮型機械学習における直接的差別と間接的差別を構造的因果モデルにおける直接効果と間接効果に対応させた公平性配慮型尺度を開発した。

5. 主な発表論文等

〔雑誌論文〕 計20件（うち査読付論文 16件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 吉田 悠夏、黒木 学	4. 巻 53
2. 論文標題 偏相関係数行列の逆行列を利用した主変数選択規準とグラフィカル・モデルに基づく特徴づけ	5. 発行年 2023年
3. 雑誌名 品質	6. 最初と最後の頁 239 ~ 251
掲載論文のDOI (デジタルオブジェクト識別子) 10.20684/quality.53.4_239	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 -

1. 著者名 吉田 悠夏、黒木 学	4. 巻 50
2. 論文標題 情報論的主変数選択規準とその性質について	5. 発行年 2023年
3. 雑誌名 行動計量学	6. 最初と最後の頁 49-59
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Tezuka Taiki、Kuroki Manabu	4. 巻 197
2. 論文標題 An unbiased estimator of the causal effect on the variance based on the back-door criterion in Gaussian linear structural equation models	5. 発行年 2023年
3. 雑誌名 Journal of Multivariate Analysis	6. 最初と最後の頁 105201 ~ 105201
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.jmva.2023.105201	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 -

1. 著者名 Kuroki Manabu、Tezuka Taiki	4. 巻 -
2. 論文標題 The estimated causal effect on the variance based on the front-door criterion in Gaussian linear structural equation models: an unbiased estimator with the exact variance	5. 発行年 2023年
3. 雑誌名 Statistical Papers	6. 最初と最後の頁 -
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s00362-023-01401-8	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 -

1. 著者名 Y. Kawakami, M. Kuroki, and J. Tian	4. 巻 40
2. 論文標題 Instrumental variable estimation of average partial causal effects	5. 発行年 2023年
3. 雑誌名 The 40th International Conference on Machine Learning	6. 最初と最後の頁 16097-16130
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Shingaki Ryusei, Kuroki Manabu	4. 巻 37
2. 論文標題 Probabilities of Potential Outcome Types in Experimental Studies: Identification and Estimation Based on Proxy Covariate Information	5. 発行年 2023年
3. 雑誌名 Proceedings of the AAAI Conference on Artificial Intelligence	6. 最初と最後の頁 12287 ~ 12294
掲載論文のDOI (デジタルオブジェクト識別子) 10.1609/aaai.v37i10.26448	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 -

1. 著者名 Kawakami Yuta, Shingaki Ryusei, Kuroki Manabu	4. 巻 37
2. 論文標題 Identification and Estimation of the Probabilities of Potential Outcome Types Using Covariate Information in Studies with Non-compliance	5. 発行年 2023年
3. 雑誌名 Proceedings of the AAAI Conference on Artificial Intelligence	6. 最初と最後の頁 12234 ~ 12242
掲載論文のDOI (デジタルオブジェクト識別子) 10.1609/aaai.v37i10.26442	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 -

1. 著者名 東川 知樹、田口 千恵、黒木 学、宮川 雅巳	4. 巻 52
2. 論文標題 効果復元法と条件付き操作変数法を用いた総合効果の統合型推定量	5. 発行年 2022年
3. 雑誌名 品質	6. 最初と最後の頁 165 ~ 177
掲載論文のDOI (デジタルオブジェクト識別子) 10.20684/quality.52.3_165	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 -

1. 著者名 Nanmo, H. and Kuroki, M.	4. 巻 180
2. 論文標題 Partially adaptive regularized multiple regression analysis for estimating linear causal effects	5. 発行年 2022年
3. 雑誌名 Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence	6. 最初と最後の頁 1456-1465
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kawakami, Y., Shingaki, R. and Kuroki, M.	4. 巻 37
2. 論文標題 Identification and Estimation of the Probabilities of Potential Outcome Types using Covariate Information in Studies with Non-Compliance	5. 発行年 2022年
3. 雑誌名 Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence	6. 最初と最後の頁 -
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Shingaki, R. and Kuroki, M.	4. 巻 37
2. 論文標題 Probabilities of Potential Outcome Types in Experimental Studies: Identification and Estimation based on Proxy Covariate Information	5. 発行年 2022年
3. 雑誌名 Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence	6. 最初と最後の頁 -
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 川上祐大・黒木学	4. 巻 50
2. 論文標題 クラスターサイズ多変量カーネルリッジ回帰分析法とその応用	5. 発行年 2021年
3. 雑誌名 応用統計学	6. 最初と最後の頁 1-20
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Nanmo, H. and Kuroki, M.	4. 巻 185
2. 論文標題 Exact variance formula for the estimated mean outcome with external intervention based on the front-door criterion in Gaussian linear structural equation models	5. 発行年 2021年
3. 雑誌名 Journal of Multivariate Analysis	6. 最初と最後の頁 104766
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 黒木学, 田口千恵, 川上裕大, 小林輝樹	4. 巻 51
2. 論文標題 外的操作による線形回帰モデルの変化について	5. 発行年 2021年
3. 雑誌名 品質	6. 最初と最後の頁 239-246
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 無
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 -

1. 著者名 Shingaki, R. and Kuroki, M.	4. 巻 34
2. 論文標題 Identification and Estimation of Joint Probabilities of Potential Outcomes in Observational Studies with Covariate Information	5. 発行年 2021年
3. 雑誌名 Advances in Neural Information Processing System	6. 最初と最後の頁 26475--26486
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Kuroki Manabu, Nanmo Hisayoshi	4. 巻 104
2. 論文標題 Variance formulas for estimated mean response and predicted response with external intervention based on the back-door criterion in linear structural equation models	5. 発行年 2020年
3. 雑誌名 ASTA Advances in Statistical Analysis	6. 最初と最後の頁 667 ~ 685
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/s10182-020-00372-7	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 -

1. 著者名 Kuroki Manabu, Shingaki Ryusei, Qu Yongming	4. 巻 63
2. 論文標題 Proportion of treatment effect mediated by surrogate endpoints	5. 発行年 2020年
3. 雑誌名 Biometrical Journal	6. 最初と最後の頁 105 ~ 121
掲載論文のDOI (デジタルオブジェクト識別子) 10.1002/bimj.202000119	査読の有無 無
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 -

1. 著者名 神田博・黒木学	4. 巻 50
2. 論文標題 重み付き条件付き操作変数推定量の性質と工程解析への応用	5. 発行年 2020年
3. 雑誌名 品質	6. 最初と最後の頁 -
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 オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 黒木学・松浦峻	4. 巻 49
2. 論文標題 予測型主変数選択規準と統計的品質管理への応用-基本的なアイデア-	5. 発行年 2019年
3. 雑誌名 品質	6. 最初と最後の頁 11-16
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 無
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 -

1. 著者名 黒木学・松浦峻	4. 巻 50
2. 論文標題 予測型主変数選択規準と統計的品質管理への応用-ケーススタディ-	5. 発行年 2020年
3. 雑誌名 品質	6. 最初と最後の頁 4-11
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 無
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 -



〔学会発表〕 計29件（うち招待講演 6件 / うち国際学会 5件）

1. 発表者名 T. Tezuka and M. Kuroki
2. 発表標題 An Unbiased Estimator of the Causal Effect on the Mean/Variance under the Adaptive Control and Its Application to Quality Control
3. 学会等名 The 22th ANQ Quality Congress (ANQ 2022) (国際学会)
4. 発表年 2022年

1. 発表者名 新垣隆生・黒木学
2. 発表標題 共変量情報に基づいた潜在結果変数の同時確率の識別と推定
3. 学会等名 第17回日本統計学会春季集会
4. 発表年 2023年

1. 発表者名 佐藤光一・黒木学
2. 発表標題 多重共線性を考慮したクラスターワイズ回帰分析法の変数選択規準
3. 学会等名 日本計算機統計学会第36回シンポジウム
4. 発表年 2022年

1. 発表者名 吉田悠夏・黒木学
2. 発表標題 グラフィカルモデルに基づく新たな主変数選択規準とその性質について
3. 学会等名 日本計算機統計学会第36回シンポジウム
4. 発表年 2022年

1. 発表者名 手塚大樹・黒木学
2. 発表標題 フロントドア基準に基づく分散に対する因果効果の不偏推定量の定式化
3. 学会等名 日本計算機統計学会第36回シンポジウム
4. 発表年 2022年

1. 発表者名 田口千恵・東川知樹・宮川雅巳・黒木学
2. 発表標題 総合効果の識別可能条件の選択基準- 総合効果の推定精度の観点から -
3. 学会等名 日本品質管理学会第125回年次大会研究発表会
4. 発表年 2021年

1. 発表者名 東川知樹・田口千恵・宮川雅巳・黒木学
2. 発表標題 効果復元法と条件付き操作変数法を用いた総合効果の統合型推定量
3. 学会等名 日本品質管理学会第125回年次大会研究発表会
4. 発表年 2021年

1. 発表者名 田口千恵・宮川雅巳・黒木学
2. 発表標題 Qualitative Identification of Linear Causal Effects based on Bad Control
3. 学会等名 2021年度統計関連学会連合大会
4. 発表年 2021年

1. 発表者名 黒木学・川上裕大・新垣隆生
2. 発表標題 共変量情報を利用した「原因の確率」の定量的評価-識別可能性と存在範囲-
3. 学会等名 2021年度 統計関連学会連合大会（招待講演）
4. 発表年 2021年

1. 発表者名 黒木学
2. 発表標題 原因の確率-必要性，十分性，必要十分性とその定量的評価-
3. 学会等名 第24回情報論的学習理論ワークショップ（招待講演）
4. 発表年 2021年

1. 発表者名 Y.Kawakami and M.Kuroki
2. 発表標題 Clusterwise Multivariate Kernel Ridge Regression Analysis and the Visualization of the Analytical Results
3. 学会等名 The 20th ANQ Quality Congress (ANQ 2020) (国際学会)
4. 発表年 2020年

1. 発表者名 R.Hasegawa and M.Kuroki
2. 発表標題 Collapsibility Condition for Ridge Regression Coefficients with an Application to Estimating Total Effects
3. 学会等名 The 20th ANQ Quality Congress (ANQ 2020) (国際学会)
4. 発表年 2020年

1. 発表者名 Y.Kawakami and M.Kuroki
2. 発表標題 Identification of Joint Probabilities of Potential Outcomes using Instrumental Variable Method with a Proxy Variable
3. 学会等名 Joint Quantitative Political Science Conference for Asia and Australasia (国際学会)
4. 発表年 2021年

1. 発表者名 川上祐大・黒木学
2. 発表標題 クラスタワイズカーネルMT法と解析結果の可視化について
3. 学会等名 日本品質管理学会第50回年次大会研究発表会
4. 発表年 2020年

1. 発表者名 長谷川亮・黒木学
2. 発表標題 リッジ回帰モデルにおける併合可能条件とその応用
3. 学会等名 日本品質管理学会第50回年次大会研究発表会
4. 発表年 2020年

1. 発表者名 川上祐大・黒木学
2. 発表標題 MTシステムのクラスタワイズ・カーネル化とその応用
3. 学会等名 日本信頼性学会第33回秋季信頼性シンポジウム
4. 発表年 2020年

1. 発表者名 田口千恵・黒木学・宮川雅巳
2. 発表標題 測定誤差をとまなう中間特性を利用した総合効果の定量的評価について
3. 学会等名 日本信頼性学会第33回秋季信頼性シンポジウム
4. 発表年 2020年

1. 発表者名 川上 裕大・黒木 学
2. 発表標題 Instrumental Variable-based Identification for Causal effects using Covariate Information
3. 学会等名 第15回日本統計学会春季集会
4. 発表年 2021年

1. 発表者名 田口千恵・東川知樹・宮川雅巳・黒木学
2. 発表標題 Variance based Difference between Graphical Identification Conditions of Causal Effects in Linear Structural Equation Models
3. 学会等名 第15回日本統計学会春季集会
4. 発表年 2021年

1. 発表者名 K.Kawakami, Sakurai, S., M.Kuroki and H. Gotoh
2. 発表標題 Clusterwise Kernel MT method with Multiple Normal Groups
3. 学会等名 The 17th ANQ Quality Congress (ANQ 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 神田博・黒木学
2. 発表標題 統合型条件付き操作変数推定量の性質に関する考察
3. 学会等名 日本品質管理学会第49回年次大会研究発表会
4. 発表年 2019年

1. 発表者名 南茂尚義・黒木学
2. 発表標題 線形構造方程式モデルに基づいた平均に対する因果効果の統計的推測
3. 学会等名 日本品質管理学会第49回年次大会研究発表会
4. 発表年 2019年

1. 発表者名 永峰拓哉・黒木学
2. 発表標題 在反応モデルに基づく必要治療数の定式化とその統計的品質管理分野への応用
3. 学会等名 日本品質管理学会第49回年次大会研究発表会
4. 発表年 2019年

1. 発表者名 黒木学・黒木景心
2. 発表標題 特定地域店舗における野菜価格の初等的データ解析
3. 学会等名 日本品質管理学会第49回年次大会研究発表会
4. 発表年 2019年

1. 発表者名 松浦 峻・黒木学・山下遥
2. 発表標題 数の目的変数を解析対象とした予測型主変数選択規準について
3. 学会等名 日本品質管理学会第49回年次大会研究発表会
4. 発表年 2019年

1. 発表者名 黒木学
2. 発表標題 Counterfactual based NNT
3. 学会等名 計量経済学ワークショップ（招待講演）
4. 発表年 2019年

1. 発表者名 黒木学
2. 発表標題 原因を評価することの難しさ-原因の必要性・十分性・必要十分性とその識別問題-
3. 学会等名 日本学術振興会プロセスシステム工学第143委員会第220回委員会・平成31年度第1回研究会（招待講演）
4. 発表年 2019年

1. 発表者名 黒木学
2. 発表標題 相関とは違うのだよ！相関とは！ by 因果関係
3. 学会等名 日本品質管理学会第112回クオリティトーク（招待講演）
4. 発表年 2019年

1. 発表者名 黒木学
2. 発表標題 原因を評価することの難しさ -原因の必要性・十分性・必要十分性とその周辺-
3. 学会等名 順天堂大学医学部附属順天堂医院臨床研究・治験センター2019年度セミナー（招待講演）
4. 発表年 2019年

〔図書〕 計1件

1. 著者名 黒木学	4. 発行年 2020年
2. 出版社 共立出版	5. 総ページ数 256
3. 書名 数理統計学	

〔産業財産権〕

〔その他〕

-

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		
米国	Iowa State University		