

2023年度北海道大学情報基盤センター萌芽型共同研究成果報告書

1. 研究類型 B) 研究集会開催支援型

2. 研究課題名 データサイエンスにおける統計的データ分析手法と諸問題

3. 研究期間 令和5年 5月 23日 ～ 令和6年 3月 31日

4. 研究代表者

氏 名	所属機関・部局名	職 名	備 考
飯塚 誠也	岡山大学 教育推進機構	教授	

5. 研究分担者

氏 名	所属機関・部局名	職 名	備 考
栗原 考次	京都女子大学・データサイエンス研究所	教授	
宿久 洋	同志社大学・文化情報学部	教授	
森 裕一	岡山理科大学・経営学部	教授	
黒田 正博	岡山理科大学・経営学部	教授	
石橋 雄一	株式会社スタットラボ	代表取締役	
田中 豊	岡山大学	名誉教授	
富田 誠	横浜市立大学・データサイエンス学部	教授	
水田 正弘	統計数理研究所・大学統計教員育成センター	特任教授	
石岡 文生	岡山大学・学術研究院環境生命自然科学学域	教授	
南 弘征	北海道大学・情報基盤センター	教授	

6. 共同研究の成果

データサイエンスが社会に浸透しつつある現在,多様なデータに対応する統計的データ分析手法の開発等が必要となってきた。そこで, データサイエンスにおける統計的データ分析手法と諸問題の課題に対し, 複数の分野の専門家が集い議論することにより, 新たな手法の開発や諸問題の解決を目指した。

因果推論, 地理空間統計, 質的データ, 非構造化データ, 医学統計における専門家が研究成果を発表し, 大規模データ・複雑データに対する新たなデータ分析手法の開発や, 具体的な解析方法, データサイエンスに関する課題等について議論することを目的として, 次の研究会を企画し実施した。

研究会名:

データサイエンスにおける統計的データ分析手法と諸問題

開催日時: 2024年3月6日(水) 10:00-17:30

開催場所：北海道大学情報基盤センター（北館4階会議室）

10:00-10:05 開会

座長 富田誠（横浜市立大）

10:05-10:40 %エシェロン解析を用いた空間クラスターの信頼性の評価について%

竹村祐亮（京都女子大）、石岡文生（岡山大）、栗原考次（京都女子大）%

10:40-11:15 データ同化手法を組み込んだ変分推論による非線形時系列モデルの学習アルゴリズム
春日井海，鎌倉稔成(中央大)

11:15-11:50 ファジィクラスタリングにおける高次元質的データの扱い

赤木辰伎，森 裕一，黒田正博（岡山理科大），飯塚誠也（岡山大）

昼休み

座長 山田隆行（京都女子大）

13:30-14:05 二値アウトカムに対する非対称なリンク関数を用いた因果媒介分析

坪田有司，山本倫生（大阪大）

14:05-14:40 外れ値にロバストなダイバージェンスを用いた Flow Contrastive Estimation について

木元篤郎，宿久洋（同志社大）

14:40-15:25 隣接情報を考慮した非負値複合テンソル分解について

久保幸平，宿久洋（同志社大）

休憩

座長 石橋雄一（中央大）

15:30-16:05 時間依存性処置の効果推定における傾向スコアモデルの共変量選択について

山下璃久，坪谷圭恵，宿久洋（同志社大）

16:05-16:40 半教師あり次元削減を用いた多重代入法の性能評価について

後藤寛治，宿久洋（同志社大）

16:40-17:15 Multivariate two-mode cluster elastic net について

大井海渡，柚木慎太郎，谷岡健資，宿久洋（同志社大）

17:15- 17:30 総合討論 座長 飯塚誠也（岡山大）

<研究打ち合わせ>

17:30-18:00 データサイエンスにおける統計的データ分析手法と諸問題に関する検討会

この研究会では、10件の発表と検討会、総合討論、データサイエンスにおける統計的データ分析手法と諸問題に関する検討会を実施した。理論面からは、近年のデータサイエンスの手法として重要となっている、因果推論、クラスタリングに関する分析手法の開発、隣接情報を伴うデータ分析手法の開発、多重代入法、密度推定、非線形時系列モデルに関する報告があった。それぞれの発表では現実の課題を解決するための分析方法を検討しており、データサイエンスにおいて取り扱う各種のデータタイプ等に対して開発した分析方法を適応した結果などについて、統計的な観点を持ったデータ分析に関する最新の研究報告がなされた。それぞれの報告について活発な議論がなされ、課題の共有と今後の研究の発展に寄与することができた。