

# 統計科学セミナー・早慶戦 2024

2024年12月21日(土)

@早稲田大学 西早稲田キャンパス (52号館 102)

プログラム ※講演は一人15分(発表)+5分(質疑)

13:25 ~ Opening 清水泰隆 (早稲田)

13:30 - 14:30 座長: 清水泰隆 (早稲田)

[1] (慶応・白石研 D3) 林宣安 (Xuanan Lin)

「Estimating the Effective Reproduction Number Using a State-Space Model with the Kalman Filter」

[2] (早稲田・劉研 M2) 一箭凜太郎

「Detection of Chaotic Behavior in Financial Data」

[3] (慶応・林研 M2) 板谷勇輝

「項目反応理論におけるロバスト推定」

15:00 - 16:00 座長: 白石 博 (慶応)

[4] (慶応・小林研 M2) 奥土康太

「 $\alpha$ サブ指数分布における二値分類問題の良性過適合バウンドと学習率の条件」

[5] (早稲田・清水研 M2) 山野邊敦史

「Win-Lose Prediction in Team Sports and Player's Potential Beyond STATS」

[6] (慶応・林研 M2) 一色修平

「Quantitative Method for Weight Selection in SGDDP with Multiple Co-primary Endpoints」

16:30 – 17:30 座長：小林 景（慶応）

[7]（早稲田・清水研 M2）新井康太

「エルゴード性をもつ定常ガウス過程に対する共分散関数の推定」

[8]（慶応・白石研 M1）飯田 優希

「Local Fréchet regression の理論と方向データへの応用」

[9]（慶応・林研 M2）槻 智也

「共変量シフト下における密度比推定を用いた転移学習と M 推定としての定式化」

17:30 ~ Closing 白石 博（慶応）

18:00 ~ 総合討論@ ちから 高田馬場店

■お店所在地：東京都新宿区高田馬場3-5-5 ACN 高田馬場ビル 1F

<https://www.hotpepper.jp/strJ003807018/?vos=cphpporgmail120621004>

■ご来店日時：2024年12月21日（土）18時00分

■コース：【大山鶏のちゃんこ鍋食べ放題コース】3時間飲み放題付 全8品 3980円

## アブストラクト集

[1] (慶応・白石研 D3) 林宣安 (Xuanan Lin)

「Estimating the Effective Reproduction Number Using a State-Space Model with the Kalman Filter」

Accurate estimation of the effective reproduction number ( $R_t$ ) is crucial for understanding the transmission dynamics of infectious diseases. However, a significant challenge in such estimations arises from the reporting delay problem, where delays in case reporting introduce bias and uncertainty into real-time analyses. To address this, this study develops a state-space model leveraging a modified Kalman filter framework to estimate  $R_t$  in real time, explicitly accounting for reporting delays in the epidemiological context. Building on the methodology proposed by Atherino et al. (2010), the reported values of positive cases are organized in a run-off triangle structure, commonly used in actuarial sciences for claims reserving. This structure effectively captures reports over successive time intervals. By stacking the rows of the triangle, the data are transformed into a one-dimensional “time-series-like” format, suitable for dynamic modeling. This transformation enables the effective reproduction number ( $R_t$ ) to be treated as a hidden state in the state-space framework. Through the application of the Kalman filter,  $R_t$  can be estimated iteratively, allowing for real-time monitoring and prediction of disease transmission dynamics.

[2] (早稲田・劉研 M2) 一箭凜太郎

「Detection of Chaotic Behavior in Financial Data」

Chaos is a complex phenomenon in deterministic processes, marked by sensitive dependence on initial conditions. The Lyapunov exponent is widely used to measure this sensitivity, but its application to noisy data presents challenges. To address this, Wolff (1992) introduced the Local Lyapunov Exponent (LLE), which examines the divergence of local trajectories. Building on this, we propose the Local Block Lyapunov Exponent (LBLE), an improved version of the LLE, and introduce the Diagonal Lyapunov Dispersion Ratio

(DLDR). We use these statistics to construct confidence intervals and provide a framework for testing and verifying chaos. Finally, we demonstrate the effectiveness of our approach in detecting chaotic behaviors in financial data.

[3] (慶応・林研 M2) 板谷勇輝

「項目反応理論におけるロバスト推定」

項目反応理論 (Item Response Theory, IRT) は、心理測定や教育測定において被験者と項目の特性をモデル化するための統計的手法である。IRT は古典的テスト理論と比較して、被験者の能力や項目の特性 (難易度, 識別力など) を個別に推定できる点で優れており、異なるテスト形式でも一貫した尺度で能力を評価できるという利点がある。しかし、応用上重要な理論にもかかわらず、IRT におけるロバスト推定に関する研究は十分に進んでいないのが現状である。

本発表では、まず IRT の基本的な理論について簡潔に説明し、その後、ダイバージェンスに基づくロバストな推定方法を新たに提案する。提案手法の有効性についてはシミュレーションを通じて既存の手法と比較し、その性能を評価する。

[4] (慶応・小林研 M2) 奥土康太

「 $\alpha$  サブ指数分布における二値分類問題の良性過適合バウンドと学習率の条件」

良性過適合とは、過剰パラメータモデルがノイズを含むデータに適合しながら、未知データでも良好な汎化性能を示す現象である。本研究では、従来のサブガウス分布を重い裾を持つ  $\alpha$  サブ指数分布に置き換え、線形分類器の汎化誤差の上界を導出した。さらに、noisy rare-weak model を解析し、誤分類率が次元  $p$  の増加とともにノイズレベルへ収束することを示した。また、学習率  $\beta$  に関しては、 $n, p \rightarrow \infty$  の漸近設定に基づき、良性過適合の発現が分布の裾の重さに依存することを理論的に明らかにし、数値シミュレーションで検証した。

[5] (早稲田・清水研 M2) 山野邊敦史

「Win-Lose Prediction in Team Sports and Player's Potential Beyond STATS」

本発表では、確率過程モデルを基に、各種チームスポーツにおいてリアルタイムに試合の優勢度を判定するための一般的方法を提案し、これを用いて既存の STATS には現れないが潜在的には活躍している選手の抽出法を目指す。特にバスケットボールのデータを例に、試合の流れ、リアルタイムな勝利確率の算出、選手の潜在能力評価を行い、既存の評価方法との比較からその有用性を示す。

[6] (慶応・林研 M2) 一色修平

「Quantitative Method for Weight Selection in SGDDP with Multiple Co-primary Endpoints」

Simultaneous global drug development program (SGDDP; Huang et al, 2012) は、2段階のプロセスに分けられる。第1段階では、複数の国からなる治験により、国際的な新薬の有効性が検証される。第2段階では、特定地域のみでの治験を行ない、特定地域での有効性を検証する。この検証のために、特定地域以外の患者集団の情報を圧縮し、統計的仮説検定に用いる。Quantitative Method for Weight Selection (QMWS; Huang et al, 2015)では、SGDDPにおいて、情報を借りることに妥当性を与えるように被験者数を定める。本研究では、QMWSをco-primary endpoint(複数の評価指標)を設定する試験に拡張した。

[7] (早稲田・清水研 M2) 新井康太

「エルゴード性をもつ定常ガウス過程に対する共分散関数の推定」

本研究の目的は、平均関数が0、共分散関数が $K(t,s)$ で与えられるガウス過程を離散観測で得た際に、共分散関数の推定量を構成することである。先行研究では、ガウス過程の任意の $n$ 時点での値が $n$ 次元正規分布に従うという性質を利用して尤度を構成し、それから定まる推定量を用いているが、 $n$ 次元正規分布の尤度では分散共分散行列の逆行列を計算することが要求されており、計算量・時間上のコストが多である。本研究では観測されるガウス過程が定常性、エルゴード性を有することを仮定し、Kessler (1997)の手法を一部模倣しながら推定量を提案する。当日の発表では提案した推定量の性質の他、従来の方法との比較について言及する。

[8] (慶応・白石研 M1) 飯田 優希

「Local Fréchet regression の理論と方向データへの応用」

本発表では、距離空間上の反応変量に対するノンパラメトリック回帰手法である Local Fréchet Regression(LFR) [Petersen and Müller, 2019]に焦点を当て、ヒルベルト空間に限定した場合の漸近理論について紹介する。さらに、ヒルベルト空間でない距離空間上のデータを、正定値カーネルを用いてRKHSに写し、RKHS上でLFRを行うことで、計算コストを下げつつ距離空間上のデータを解析できることを確認し、漸近分布を用いた近似的な信頼領域の構成を考える。最後に、その周期性により通常の実数値データに対する統計的手法が意味をなさない場合がある方向データへの応用を考える。

[9] (慶応・林研 M2) 梶 智也

「共変量シフト下における密度比推定を用いた転移学習と M 推定としての定式化」

転移学習とは、過去の知識を新しい学習に応用する手法であり、トレーニングデータとテストデータの標本空間や従う分布が異なる場合に学習性能を向上させることを目的としている。一方で、転移学習を行うことで、かえって学習性能が落ちてしまう「負の転移」という問題が存在する。本研究では、共変量シフト下における密度比推定を用いたトランスダクティブ転移学習を M 推定として定式化し、パラメータの漸近分散による負の転移の定義を提案する。