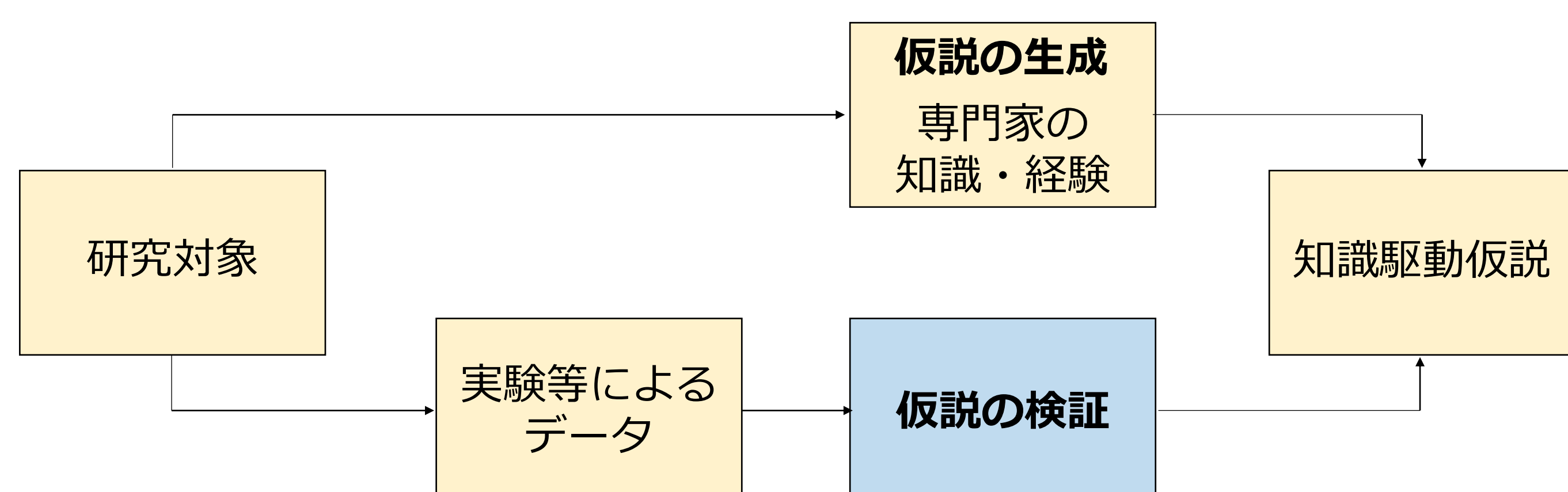


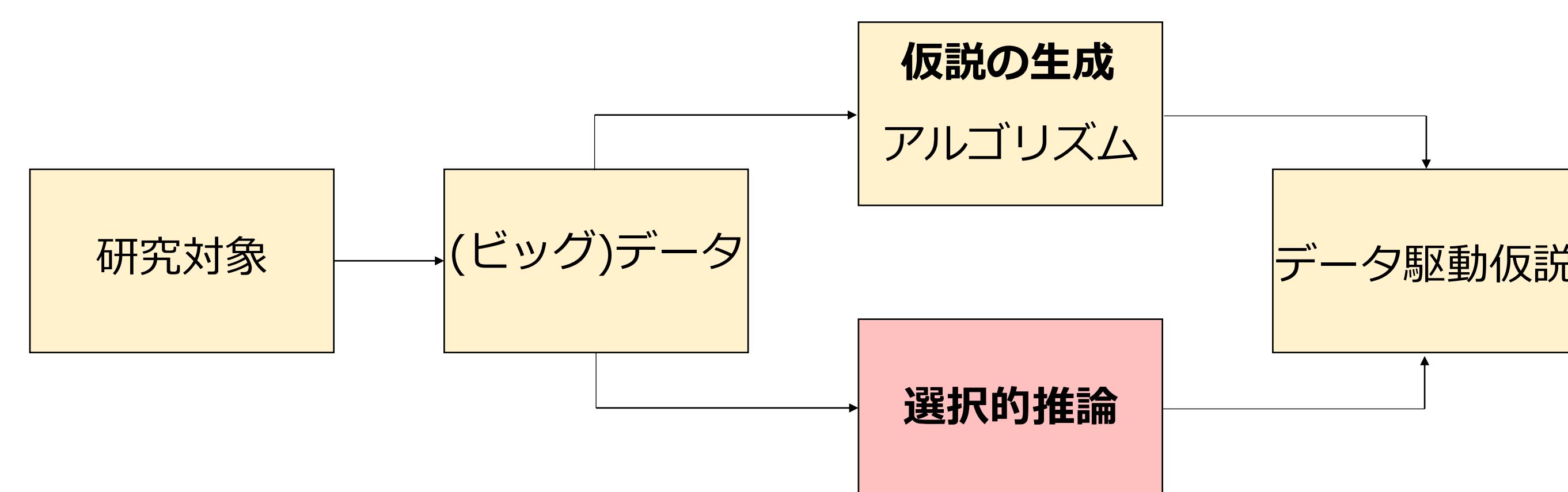
データ駆動仮説の統計的信頼性評価に関する研究

概要：様々な分野でデータを活用して科学技術を発展させるデータ駆動型アプローチが有望視されている。機械学習などのデータ分析によって発見される仮説（データ駆動型仮説）は、専門家の知識に基づく仮説（知識駆動型仮説）とは異なり、新たな科学的発見につながる可能性がある。一方、前者はデータから選択された仮説であるため、伝統的な統計的推測法を使って信頼性評価を行うことができない。本チームでは、データ駆動型仮説の統計的信頼性を評価するための新たな枠組である**選択的推論（Selective Inference）**に関する理論研究・実践研究を行っている。

（伝統的な）科学研究における仮説検証の枠組



データ駆動科学における仮説検証の枠組

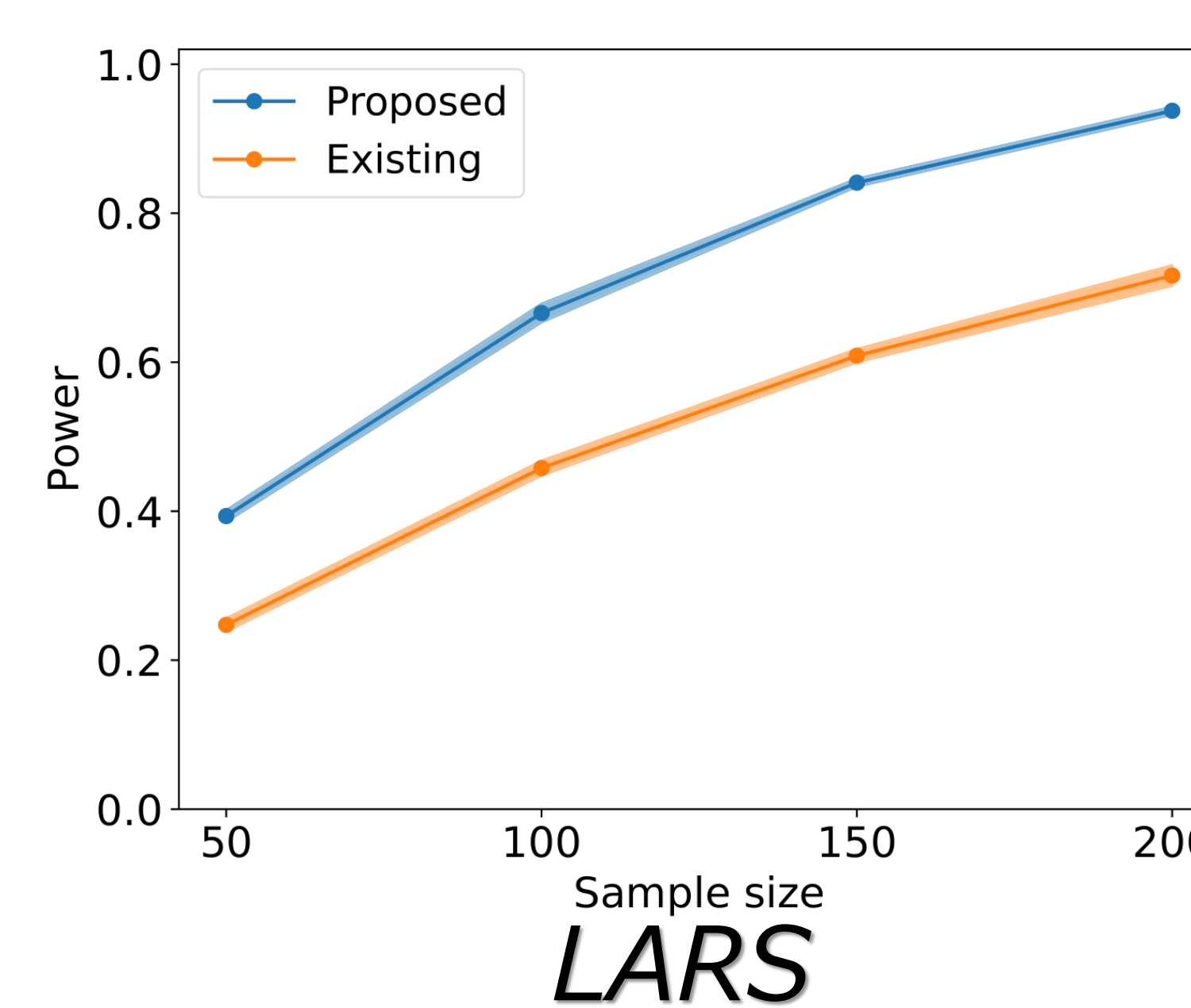
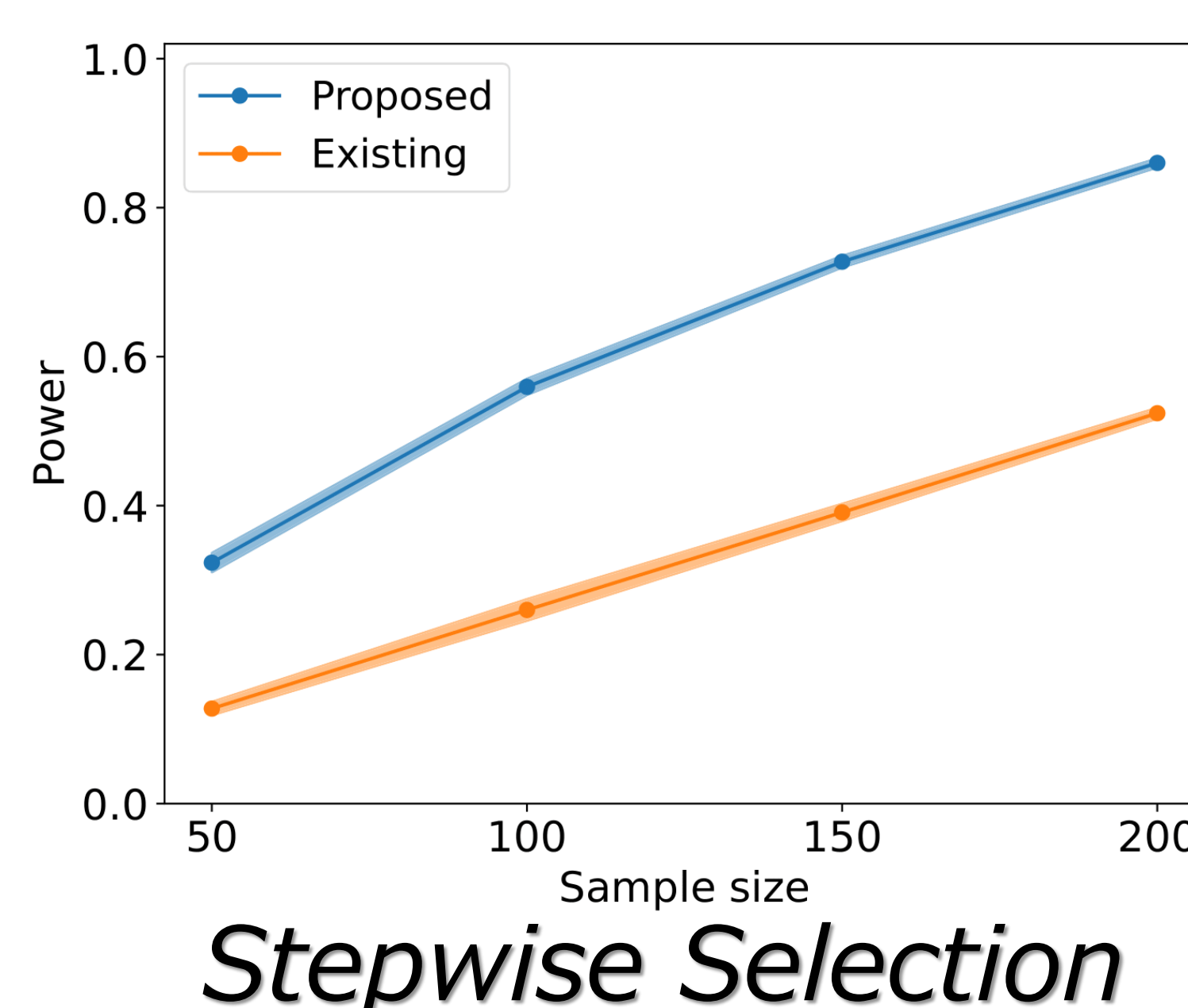
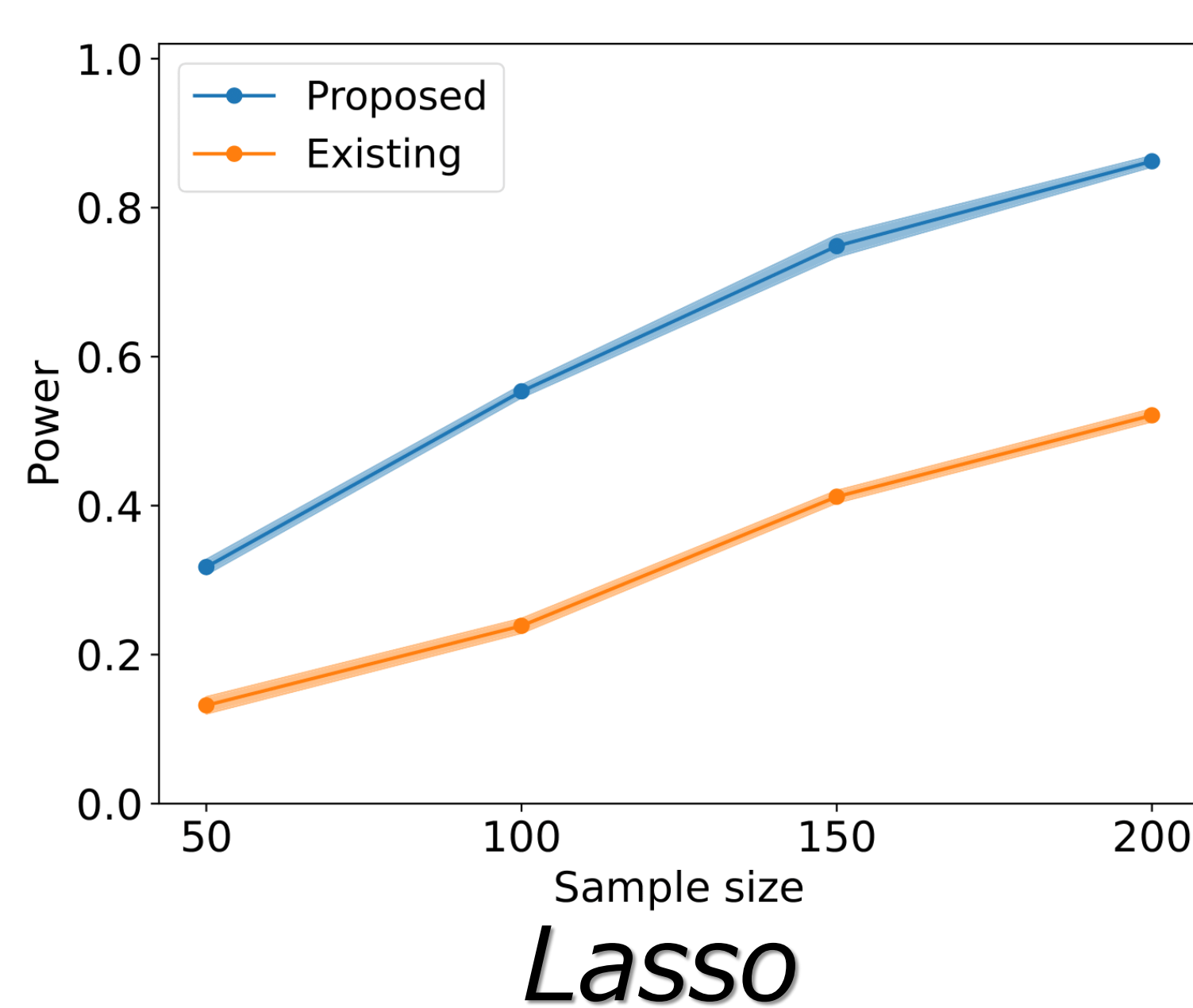


1. 選択的推論の理論研究

概要：従来の選択的推論は計算を容易にするため、過剰な条件付けを行っていたが、パラメトリック計画法に基づくアルゴリズムの開発により、過剰な条件付けを回避し検出力が向上した。

既存の選択的推論法を大きく上回る検出力を有する理論と方法を構築

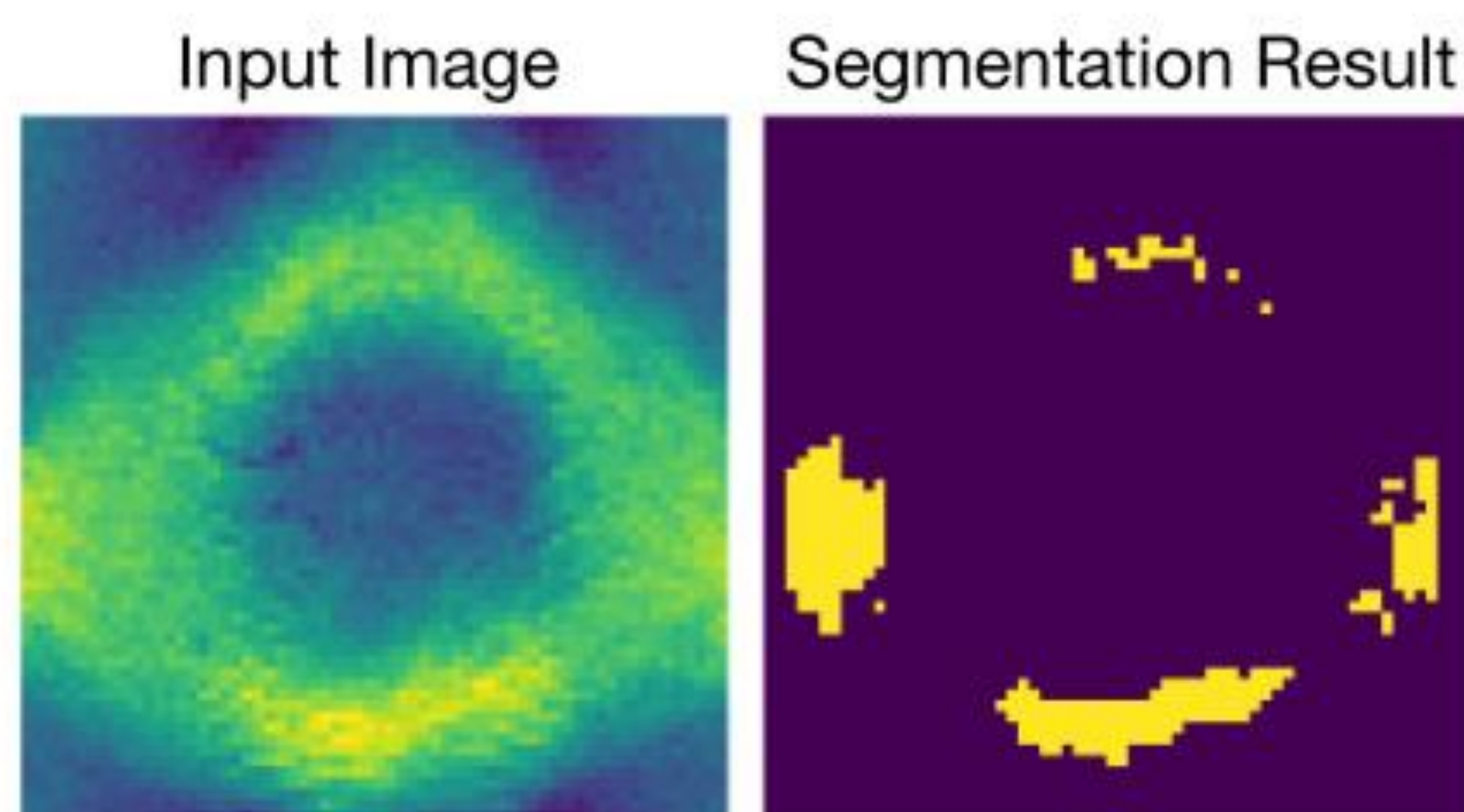
Duy et al. (AISTATS2021), Sugiyama et al. (ICML2021) etc.



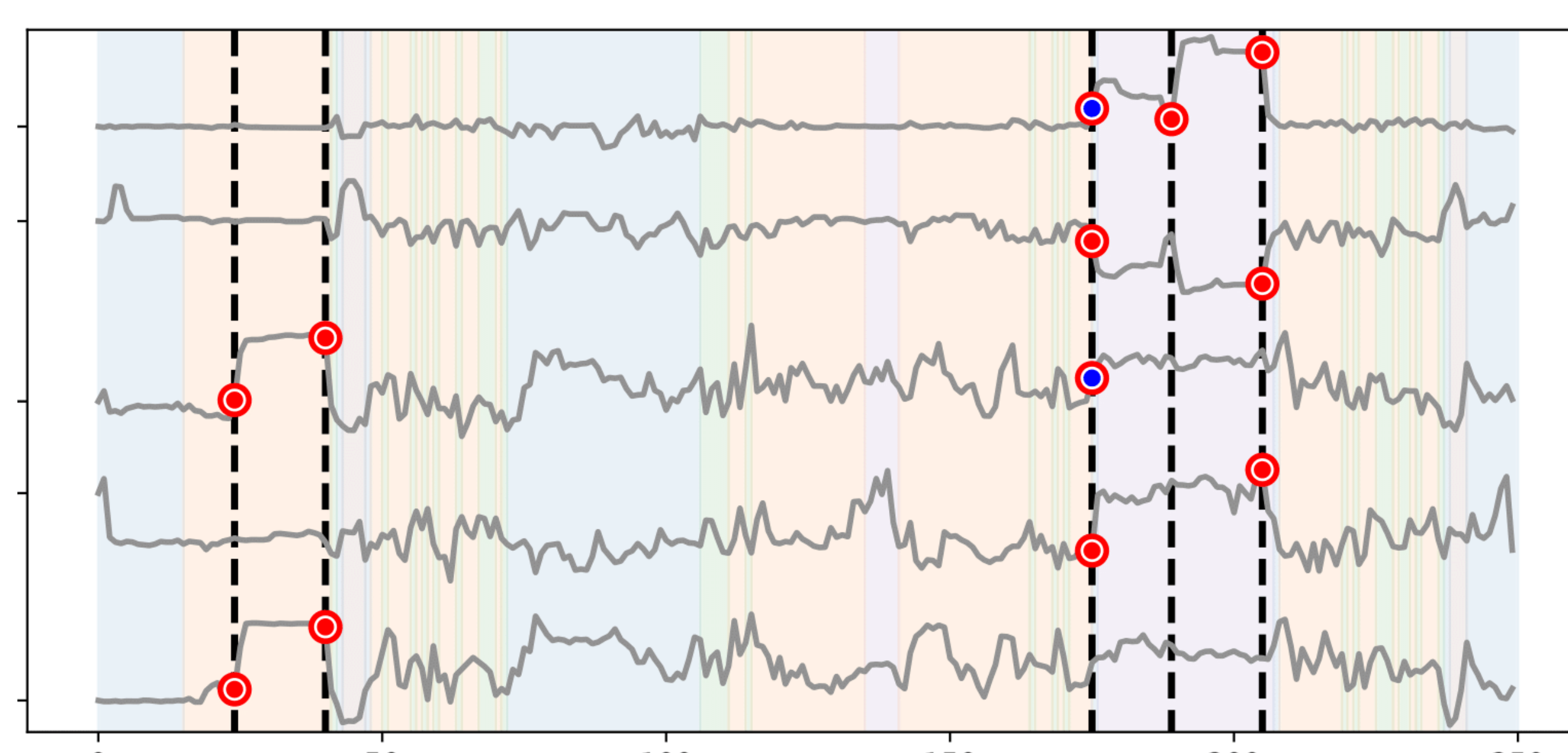
Python Toolbox: <https://github.com/takeuchi-lab/parametric-si>

2. 選択的推論の応用研究

概要：従来の選択的推論は仮説選択イベントが線形不等式の集合などで記述できる場合に限定されていたが、パラメトリック計画法により複雑な選択イベントの選択的推論が可能となり、応用範囲が拡大した。



深層学習モデルによるセグメンテーション結果のSI



多次元時系列データ（例: IoT）の変化点検知のSI