



大阪医学統計学セミナー 第8回

Osaka Biostatistics Seminar

4月25日 (水)
18:00~19:30



ブートストラップアンサンブルからの代表的な
樹木構造の同定手法

講演者: 中谷 英仁 (医療データ科学寄附講座)

概要:

医学研究におけるアウトカムの予測問題にて、回帰・分類樹木分析法により得られた二値分割された共変量の組み合わせによる分類法すなわち樹木構造は、使用者にとって解釈が容易であり高い需要がある。しかしわずかなデータ摂動により劇的にその樹木構造が変化してしまうため、予測モデルとして不安定（弱い学習器）である。機械学習の分野では、ブートストラップ抽出データを基にしたアンサンブル学習にて、弱い学習器である樹木構造を多数生成し、各症例に対して生成される多数の予測値候補を統合することで、予測性能を強化させることに成功している（Random Forest、Bagging、Boosting）。しかしこれら方法では、多数の樹木構造を生成しているため、単一の樹木構造としての視覚化が不可能である。一方で多数の樹木の中から予測性能の良い樹木を選出するBumpingという方法があるが、これは樹木構造の複雑性（共変量の数等）を固定した下で選出するため、複雑性に依存した樹木構造を選んでしまう。

このような背景から本発表では、アンサンブル学習のよさである安定性を保ちつつ、複雑性を制御した下で、単純な樹木構造を同定する統計手法を提案する。二値をアウトカムとした場合の9つのシナリオにてシミュレーションを行い、真の樹木構造を含むかどうかを確認する。

場所：
医学系研究科基礎研究棟L階
医学統計学研究室

お問い合わせ：大阪大学大学院医学系研究科・医学部 医学統計学講座

E-Mail : office@biostat.med.osaka-u.ac.jp TEL : 06-6879-3301