

これまで行ってきたベイジアンモデルへの t 分布への適用をもちいた heavy tail modeling の応用を行っていききたい。主な課題は以下の2点である。

(1) Robust Bayesian CART

統計的推論において、外れ値の影響を取り除くことは大変重要である。外れ値に対する頑健性についての研究は数多くなされている。しかし、ベイズ統計学に基づく理論研究は極めて少ない。統計的推論の分野において、近年、解釈の容易さから決定木が予測モデルとして、マーケティング等の応用分野で頻繁に用いられるようになってきた。この中で、予測の対象となる変数が量的変数であるものを回帰木と呼ぶ。本研究では、ベイジアン CART に対するロバストなモデルの構築を考えている。回帰木分析においては、データに偏りがある場合には結果がうまく得られないことが知られている。この解決策としてランダムフォレストやバギングといったブートストラップ法に基づく手法が用いられる。外れ値を含むデータは、これらの手法によって外れ値の影響が弱められると考える。一方で、これらの手法は計算量が重くなる。

回帰木を生成する代表的なアルゴリズムは、CART (Classification and Regression Trees) と呼ばれ、Breiman (1984 Chapman and Hall)によって提案された。この CART の、ベイジアンモデルに基づく理論的分析として Chipman et al. (2000 Stat. Comput.)による階層ベイジアンモデルを用いた Bayesian CART がある。Rockva and Van der pa (2020, Ann. Stat.) では、外れ値がない場合に、Bayesian CART の Tree が成長しすぎずに一定の大きさでとどまることを理論的に示している。しかし、回帰木における外れ値が含まれた場合についてはまだ研究がなされていない。

本研究では、ベイジアン CART に対して t 分布をもちいた heavy tail modeling を適用することでこの問題を解消するための十分条件を導くことを目的とする。

(2) Empirical Local Bayes Correction for Robust Bayesian modeling

James-Stein 推定量は、最尤推定量よりもより良い推定値をもたらす縮小推定量として多くの関心を集めてきた。しかしこれに対して、大規模データで取り扱われる問題においては、十分に近いとは考えられない平均を持つ混合分布となる。それゆえ、James-Stein 推定量を適用することは適切ではない。こうした大規模データを対象とした問題に対して James-Stein 推定量と同様により良い推定値をもたらす、Efron によって提案された局所的な経験ベイズ修正がある。この研究では、局所的経験ベイズ修正をベイジアンモデルに適応し、heavy tail modeling に基づく階層ベイジアンモデルを用いて外れ値に対して頑健である局所的経験ベイズ修正の提案と理論的解釈を行っていく。