

今後の研究計画

林 由子

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3点である。

(1) Robust Bayesian CART (慶応義塾大学菅澤翔之助先生との共同研究)

ベイジアン CART への適用を考えている。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。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) Objective Robust FDR

この研究は、heavy tail modeling に基づく大規模研究のための頑強ベイズ分析による分析方法を開発することを目的とする。遺伝子発現データにベイズ解析を適用する場合、FDR (false discovery rate) の閾値は、頻度論の有意水準と同様に任意にその水準が決定される。本研究では、大規模データセットにおける外れ値問題を解決するために、個々の遺伝子の差異を求める検定量を計測する際に t 分布をもちいていく。これらの統計量を用いてベイジアン FDR の閾値についての定義を行ったのち、Storey's q -value method によって null ratio の推定を行っていく。そして、この推定された値を用いて、FDR のカットポイントを客観的に設定することを提案するものである。

(3) Empirical Local Bayes Correction for 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用し、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強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。

