

ベイジアンモデルに基づく t 分布を用いた外れ値問題の理論的研究を主に行ってきた。これは、一般に **heavy tail modeling** と呼ばれ、尤度に t 分布を用いることで外れ値を事前に識別することなく、自動的にモデルが外れ値を過小評価、あるいは棄却するモデルである。

この分野の理論研究においては、主に極値理論に基づいた分析方法と **log-regularly varying** を誤差項に仮定する二つの方法が存在する。

(1) Bayesian Analysis on Limiting the Student- t Linear Regression Model (現在投稿中)

この研究では、極値理論に基づく手法を用いて、**heavy tail modeling** を線形回帰分析に適用する場合の頑強モデルとなる十分条件の導出と頑強モデルが機能しない区間の提示を行っている。線形回帰分析への **heavy tail modeling** は、外れ値の定義が **crucial** であり、モデル自体が観測値を外れ値として認知できるか否かが重要である。そこで、この研究では、**Hat matrix** を用いて、モデルが頑強モデルとして機能しうる区間を示した。さらにこうした区間において、**Jeffreys prior** の下での t 分布に基づく回帰モデルが頑強モデルとなる十分条件を示している。

(2) Theoretical properties of Bayesian Student- t linear regression (Statistics and Probability Letters)

この研究では、**log-regularly varying** を誤差項に仮定する手法を用いて **heavy tail modeling** を線形回帰分析に適用する場合に外れ値が複数個存在する場合に頑強モデルとなる十分条件の導出とモデルの効率性の考察を行った。その結果、極値理論に基づく手法と一致する結果が示された。

(3) Bayesian analysis for mixtures of discrete distribution with a non-parametric component (Journal of Applied Statistics)

この研究では、外れ値問題において多く用いられる外れ値の分布と外れ値ではない分布の混合分布を仮定した分析を行っている。外れ値と外れ値以外の混合分布は、同じ分布が仮定されることが多い。この研究では、外れ値に対する分布にノンパラメトリックな分布を仮定することで、分布への制約を緩和している。その結果、**FDR(false discovery rate)** および **FNDR(false non-discovery rate)** で評価した場合に、より良い結果が得られることを示した。