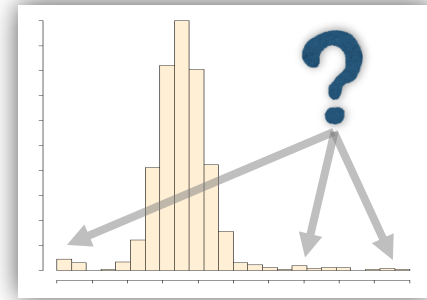


背景と動機・現状 あらゆる分野のデータには様々な由来を持った「**外れ値**」が付き物

外れ値：他のデータから離れた値や傾向を持ったデータ

飛び抜けた外れ値がデータの中に混ざっていると、分析結果に悪影響を与える
定義は曖昧で、どこからどこまでが「外れている」のかの線引きは極めて困難



➡ 0か1かで除外するのではなく、**影響を適切に(連続的に)抑制し、頑健な分析を行いたい**
特に「**外れ値に頑健なモデル選択**」についての研究は未整備な部分・課題が多い

Q. 外れ値の影響をどのように軽減する？ ➡ **A. 確率(密度)関数に応じた重み付けを行う**

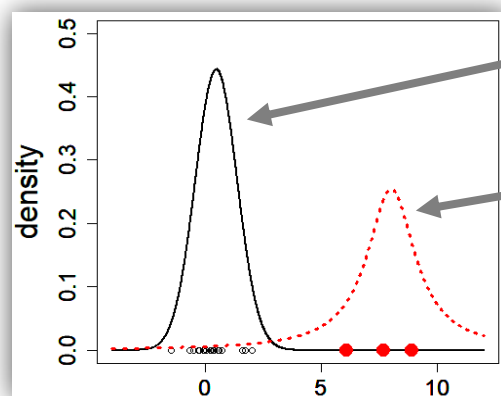
母数 θ に対する
推定方程式の例

$$\sum_{i=1}^n \left\{ \frac{\partial \log f(Y_i | \theta)}{\partial \theta} \underbrace{f(Y_i | \theta)^\alpha}_{\text{weight}} - \int \frac{\partial \log f(y | \theta)}{\partial \theta} f(y | \theta)^{\alpha+1} dy \right\} = 0$$

g : 真の分布の確率(密度)関数
 f : 統計モデルの確率(密度)関数
 α : 正値チューニングパラメータ

貢献① 母数推定で用いられている重み付けのアプローチを応用し、外れ値に影響され難いモデル選択規準を導出！

Q. 選択の頑健さをどうやって定義する？ ➡ **A. 分布の混合を想定して選択法の「ズレ」を定義**



大部分のデータが従う確率分布 **G**

外れ値を生む確率分布 **H**

混合分布 $\Omega = (1-r)G + rH$ ($r > 0$) を想定し、
モデル選択規準が外れ値に対して頑健であるか否かを、
 $|\text{Criterion}(\Omega) - \text{Criterion}(G)|$ が有界であるかで議論する

貢献② 有界性の条件を精査し、選択の頑健性は推定の頑健性と一致しないことを解明！

貢献③ 様々な規準の、意思決定分析・因果推論・官能評価分析 等への応用を実施！

展開

- ✓ 影響を軽減するアプローチのとり方や、モデル選択規準の形の改良
- ✓ より広範な問題設定・モデルに拡張した手法の確立を目指す (例：高次元データ、関数データ)