

Bayesian Tree Models for Data from a Complex Design

Daniell Toth

Bureau of Labor Statistics

We demonstrate an algorithm to producing sample-design consistent tree models that allow for incorporating more complex data models in the node and allow for easy variance estimation. By use of a prior distribution assumptions we obtain Bayesian tree model estimates using data collected under an informative sample design. We demonstrate this proposed method using consumer expenditure data. Using a repeated sample simulation we show that this proposed method produces models that have strong predictive power with variance estimates that provide credible intervals with good coverage for the estimates. We show that ignoring the design can lead to the wrong posterior distribution of the model and produce biased estimates with inflated variance.