

MÉTODOS DE ESTIMAÇÃO ROBUSTOS EM MODELOS DE REGRESSÃO LINEAR

IV Encontro de Programas de Educação Tutorial

Joao Victor Bastos de Freitas, Juvencio Santos Nobre

Em modelagem estatística/estocástica, um outlier ou valor atípico/discrepante, é uma observação que é numericamente distante das demais, muitas vezes justificado pela variabilidade dos dados ou por um erro de algum procedimento do planejamento amostral. Já valores influentes, são denominados valores nos quais tornam as inferências sensíveis a sua presença, fazendo inclusive que a sua presença modifique a tomada de decisão. Tipicamente tais observações tornam o modelo sensível a sua presença, implicando em prejuízos na interpretação dos dados. O presente trabalho tem por objetivo apresentar à comunidade acadêmica métodos alternativos de estimação para modelos de regressão linear que visem reduzir a influência dessas observações (outliers, influentes e de alavanca) nas estimativas dos parâmetros do modelo e/ou na análise de variância, tais como, Regressão Quantílica, Regressão L1 e Regressão por mínima mediana dos quadrados, apresentando suas vantagens em relação aos métodos mais comuns, como o método dos mínimos quadrados ordinários. A metodologia utilizada para consolidação desse estudo ocorreu através de revisão bibliográfica e aplicações usando o software R com conjuntos de dados reais. A partir de então foi possível descobrir e aprender a importância dos valores atípicos e como tratar corretamente esse tipo de observação, produzindo estimativas que se adequam mais a realidade da massa de dados.

Palavras-chave: Modelos de Regressão. Estimação L1. Outlier. Curva de Influência.