

Performance of Robust Estimators: Sampling, Variables and Dimensions

Elizabeth Reis*

BRU-ISCTE-IUL, Lisboa, Portugal ear@iscte.pt

Maria do Carmo Botelho

BRU-ISCTE-IUL, Lisboa, Portugal maria.botelho@iscte.pt

The use of data collected from market research and opinion surveys is common in social and business areas. Probability samples are usually the first option for data collection though they are quite often excluded due to the non existence of a suitable sampling frame. In addition to the lack of randomness of sample data, further problems are caused by inadequate sample representativeness for some population characteristics, limitations of the measuring instrument and occurrence of measurement errors. In these situations, robust statistical techniques can be a valid option for estimation purposes as they are not sensitive to sample biases. The main aim of this study is to evaluate the performance of robust estimators, particularly Huber M-estimator, Tukey's biweight and Least Trimmed Squares (LTS) estimators, when compared to the sample mean and median, and applied to different types of variables, diverse sampling methods and dimensions. Quantitative and qualitative ordinal Likert type variables with 4, 6 and 10 point were used. Samples were generated by stratified and quota methods, both with dimensions 50, 100 and 300. Results show the best behavior of the Huber and the Tukey's biweight estimators in most situations, particularly for quantitative variables, for both sample methods. The LTS estimator performed worse than any other estimator, being better solely in the case of ordinal variables with a 4 point scale, sharp skewness and high kurtosis.

Key Words: estimation, M-estimators, robustness, sampling methods.