Effects of a single outlier on the coefficient of determination: an empirical study

ABSTRACT

This article investigates the effects of outliers on the coefficient of determination, $R^2$ which is computed by Ordinary Least Squares (OLS) estimator. It is now evident that the OLS is greatly affected by outliers and hence the $R^2$ is also affected. This problem can be solved by using the robust estimators such as Least Trimmed Squares (LTS) estimator. In this article, we compare the value of $R^2$ which is computed by OLS and LTS estimators. We modify a regression data set to effectively generate outliers in both X and Y directions. Then the coefficient of determination (OLS and LTS) is investigated from the modified data sets (data with outliers). The numerical results show the merit of using the LTS based $R^2$ estimator compared to the OLS estimator.

Keyword: Outlier; Coefficient of determination; LTS estimator