

Symbolic Interval-Valued Regression Model

Liang-Ching Lin

Department of Statistics and Institute of Data Science, National Cheng Kung University, Tainan, Taiwan

Abstract

From micro data to macro data, symbolic data analysis can provide statistical inferences in the macro data while preserving as much information as possible from the micro data. In this study, we focus on the symbolic interval-valued regression model. The microdata are reorganized into intervals by using the largest and smallest order statistics. Then, we develop innovative symbolic interval-valued regression models to construct the relationship between two or more intervals. By defining the negative sign to the intervals, we can always keep the natural order that the higher value of the dependent variable will be larger than the lower value of the dependent variable, even when the parameters are negative. Firstly, we develop the simple linear symbolic interval-valued regression model and derive the corresponding maximum likelihood estimators (MLEs). In addition, the Fisher information matrix of the MLEs is shown, and the asymptotic normality is given. Next, we extend the aforementioned model to a multiple linear symbolic interval-valued regression model, and the corresponding maximum likelihood estimators are also provided. Monte Carlo simulations and real data analysis confirm the validity of the proposed method.

Keywords: asynchronous observations, order statistics, regression, symbolic data analysis.