Multiple Imputation Using Weight Adjustment Method

Jin-Young Kim*
Hankuk University of Foreign Studies, Seoul, Republic of KOREA,
suwonfis@naver.com

Key-Il Shin
Hankuk University of Foreign Studies, Seoul, Republic of KOREA,
keyshin@hufs.ac.kr

It is well known that survey results are affected by errors arising from several sources. Among them, one crucial effect could be introduced by unit and item non-responses with the consequence that these last could produce bias and distortions of distributions.
Multiple imputation is a widely used method for handling missing data in survey. The performance of imputation is influenced by various factors, especially outlier. Hence detecting outliers and reducing the effect of them are crucial to improve performance of imputed values. Therefore in this study we investigate imputation methods which reduce the effect of outlier using various weight adjustment methods including removing outliers.
For detecting outliers, we consider two methods, studentized deleted residual and Hidiroglou-Berthelot methods. The student deleted residual method is popularly used in regression analysis and Hidiroglou-Berthelot method is useful to detect outliers for panel type data.
In order to reduce the effect of outliers, removing the outliers in data set is a simple and effective approach. However several methods have been developed to reduce the outlier effect. Among them, in this study, we consider three weight adjustment methods. The first one is removing outliers which is the same as assigning zero weight to outlier. The second one is to assign final weight of 1 to outlier. The last one is the method suggested by Hidiroglou and Srinath (1981).
With the final weight obtained by adjusting outliers, we conduct multiple imputation by appling the final weight to the variables in data set. In this study we use the regression method in SAS/PROC MI for the multiple imputation. Generally in this approach the weight assigned to each values does not considered. However since the final weight to each value is different and should be counted on to impute, we transform the each value in data set according to the assigned final weight before imputing. The obtained imputed values are transformed back properly to get final imputed values.
Small simulation studies are conducted to compare performances obtained by weight adjustment methods and detecting outlier methods. To establish population data of simulation, we use the same steps as in Lee et. al. (1995). Also for real data analysis, we use total wages and total number of employers in Korea monthly labor statistics.

Key Words : Business survey, non-response, outlier, regression method