A Multivariate Statistical Analysis of Reporting Error in Age Data of India

Barun Kumar Mukhopadhyay* and Prasanta Kumar Majumdar

Population Studies Unit, Indian Statistical Institute, 203 B.T.Road, Kolkata 700 108, West Bengal, India
*E-mail: barun_mukhopadhyay@yahoo.com

ABSTRACT It is now a general concept that education improves the quality of age reporting. However, in the paper an attempt has been made to analysis the age reporting error vis-à-vis different socio economic cultural and developmental factors in addition to literacy factor in order to find out the different unique factors and their commonalities in influencing the variations in the age reporting error in India. The data for this type of analysis have been obtained from the reports of National Family Health Survey (NFHS) for two periods 1992-93 and 1998-99. In those reports a measure of age reporting error separately for male and female populations has been given by Myers' indices. Since handling with large number of variables which may not be significant or may not be worth at all, at the initial stage of analysis all the variables were undergone through backward regression process where a number of variables were deleted as they were found insignificant at a certain level of significance. The remaining variables then were analyzed through commonality analysis which gave some interesting results. Still literacy of male played a significant role in improving the quality of age data. In female part, scheduled tribe population has some influence on the reporting error. Mean household size also played some role in influencing the age reporting error. Apart from these some commonalities have been found between urban households and mean household size or between household response rate and many other. Apart from these there was some difference between the two periods of 1992-93 and 1998-99.