Refining bootstrap methods in dependent data structures: is there room for improvement?

Giancarlo Manzi
University of Milan, Italy

1. Objectives

According to Efron and Tibshirani (1993), the traditional road to statistical knowledge is blocked, for most, by a formidable wall of mathematics. The usefulness of the bootstrap method is particularly appreciable when dealing with complex estimators or non-smooth statistics. Our work is intended to find out if bootstrap methods other than the standard bootstrap can be satisfactory used to break a "thicker wall": the dependence in data structure. In this paper we present some variants of existing bootstrap methods for complex surveys and nested multilevel models.

2. Methods/Models

Some modifications (ModB) of existing bootstrap algorithms for complex surveys and for multilevel models are presented. These existing bootstrap algorithms are the Booth, Butler and Hall (1994) method (BWO), the Rao and Wu (1988) method (BWR) and some bootstrap methods for multilevel models proposed by Goldstein (1998) (MLB). The proposed modifications are compared with their own original algorithm via some articulated simulation studies.

3. Results and Conclusions

ModB1 algorithm performs better than the BWO method whereas ModB2 is not significantly different from the Booth, Butler and Hall method, but is less computer-demanding than the original. ModB3 seems to perform better than the ML algorithm for some estimators but not for others. In general, most of the proposed bootstrap methods are good for certain properties but not for other ones. In this study we also show that the flexibility of bootstrap methods for dependent data is confirmed when compared with other resampling methods, like the jackknife: the bootstrap seems to perform better in median and percentile estimation. In multilevel models the finding of a "good" bootstrap seems more difficult. Some methods retain the structure of data, but do not retain the intrinsic correlation structure. A further investigation on the bootstrap for ML models is therefore needed.