UNCERTAINTY ESTIMATION BY BOOTSTRAP APPROACH

ALBEANU, Grigore
Spiru Haret University, Bucharest, Romania
g.albeanu.mi@spiruharet.ro

BURTSCHY, Bernard
ENST, INFRES, France
burtschy@enst.fr

POPENŢIU-VLĂDICESCU, Florin
University of Oradea, UNESCO IT Chair
popentiu@imm.dtu.dk

Abstract

This paper provides a methodology to investigate uncertainty evaluation by bootstrap and a procedure to obtain confidence bands for linear and nonlinear models used in data analysis and design measurements.

Keywords: *Bootstrap method, Resampling methods, Data modelling, Sampling variance, Median, Z-scores, Confidence Intervals, Confidence bands.*

AMS classification: 62P30

1. Introduction

All modelling applications in social and engineering sciences give rise to an estimation process (parameters or functional relations). The numerical estimation process always depends on the previously data collection. Even data filtering and outlier's elimination procedures were applied, the ill-conditioned effect could appear. To increase the confidence on the estimations or to improve the estimates, some computer intensive methods were recommended.

The methods are suitable for any level of modelling being useful for fully parametric, semiparametric, and completely nonparametric analysis. These approaches are not only in use by statisticians, but also are applied anywhere statistics can be used: life sciences, social sciences, econometrics, reliability etc. For the aim of this paper we outline the application of bootstrap sampling for accuracy estimation and the method of simultaneous confidence bands for uncertainty management.

2. **Uncertainty estimation and bootstrap**

For a set (sample) of uncorrelated N random variables, xi, i = 1, 2, …, N, let us denote by Mean[X] the sample mean given by:

Mean[X] = $\frac{1}{N}\sum_{i=1}^{N} x_i$.

The median is the middle value of the group for a particular sample, i.e. half of the results for the sample are higher than it and half are lower (Q2). It is calculated from the sorted values (from lowest to highest). If N is even, the median is the average of the two central values. Let us denote this value by Median[X]. Interquartile range (IQR[X]) is the difference between the lower and upper quartiles = Q3 - Q1. The lower quartile is the value below a quarter of the results lie. Similarly, the upper quartile is the value above a quarter of the results lie. Normalised IQR (NIQR[X]) is a measure of the variability of the results which basically is a robust standard deviation. It is equal to the IQR[X] multiplied by the factor 0.7413.

Robust CV (coefficient of variation) is equal to the NIQR[X] divided by the median, expressed as a percentage (i.e. multiplied by 100), it allows for the variability in different tests to be compared.

An accepted statistical method for analysis test results in proficiency testing is to calculate a Z-score for each laboratory's result. The standard form for the calculation of Z-scores is

$$Z_i = \frac{X_i - A[X]}{B[X]},$$

where A[X] is the assigned value (sample mean), and B[X] is an estimate of the spread of all results (standard deviation). The classical approach based on mean and standard deviation is significantly influenced by the presence of extreme values (outliers). Therefore, a robust approach based on median and interquantile range is better to be used.

Robust Z-scores are calculated by replacing A[X] and B[X] in the "classical" Z-score by the median and NIQR, respectively. For measurements and proficiency testing both between-laboratory and within laboratory Z-scores can be used.

The standardised sum (S) and the standardised difference (D) for the pair of results are:

S = (A+B)/$\sqrt{2}$

and

D = (A-B)/$\sqrt{2}$

(median of A is less than median of B) or

D = (B-A)/$\sqrt{2}$,

otherwise.

The between-laboratory Z-score (ZB) is the robust Z-score of S and the within-laboratory Z-score (ZW) is the robust Z-score of D:

ZB = (S-Median[S])/NIQR(S),

and

ZW = (D - Median(D))/NIQR(D).

A methodology based on bootstrap approach can be used to study the robust Z-scores ZB and ZW.

A Z-score can be computed repeatedly by simultaneous resampling in order to obtain a *resample mean and standard deviance.* The process is repeated by a number of steps and the final results can be obtained considering the best performance. An 80% approach can be used when a strong acceptance is required. Generally speaking, the best test could be based on a 51% approach.

Also, a robust Z-score can be computed repeatedly by simultaneous resampling in order to obtain a resample *median and normalised interquantile range*. The process is repeated by a number of steps and the final results can be obtained considering the best performance. Also, an 80% approach can be used when a strong acceptance is required. Generally speaking, the best test could be based on a 51% approach.

3. **Summary**

The paper provides a methodology to investigate uncertainty evaluation by bootstrap. The approach can be used both for theoretical applications, and it may addresses practical applications for accuracy assessment, and uncertainty analysis for all kind of models.

References

1. Albeanu G., *The Statistical and Numerical Analysis of Some Nonlinear Regression Models*, PhD Disertation (in Romanian), Bucharest University, 1996.
2. Albeanu G., *Resampling Simultaneous Confidence Bands for Nonlinear Explicit Regression Models*, Mathematical Reports, 50(5-6), 289-295, 1998.
3. Albeanu G., Popentiu F., *On the Bootstrap Method: Software Reliability Assessment and Simultaneous Confidence Bands*, Annals of Oradea University, Energetics Series, 7(1), 109-113, 2001.
4. Cowling A.; Hall P., Phillips M.J., *Bootstrap Confidence Region for the Intensity of a Poisson Point Process*, JASA, 91, 1516-1524, 1996.
5. Efron B., *Bootstrap Methods: Another Look at the Jacknife*, Ann. Stat., 7, 1-27, 1979.
6. Efron B., Tibshirani R.J., *An Introduction to the Bootstrap*, Chapman and Hall, London, 1997.
7. Efron B., *Computer-Intensive Methods in Statistical Regression*, Siam Review, 30, 3, 421-449, 1988.
8. Frey H. C., Burmaster D. E., *Methods for Characterizing Variability and Uncertainty: Comparison of Bootstrap Simulation and Likelihood-Based Approaches*, Risk Analysis, 19, 1, 109-130, 1999.
9. Gonçalves S., White H., *Bootstrap Standard Error Estimates for Linear Regression*, JASA, 100, 471, 970-979, 2005.

10. Jolliffe I., *Estimation of Uncertainty in Verification Measures*, International Verification Methods Workshop, The McGill Faculty Club, Montreal, Quebec, Canada, http://www.met.rdg.ac.uk/~swx03itj/ITJ InvitedTalk.ppt, 2004.
11. Siersma V., Johansen C., *The Use of the Bootstrap in the Analysis of Case-Control Studies Missing Data*, Research Report - Institute of Public Health, University of Copenhagen, http://pubhealth.ku.dk/bs/publikationer/rr-04-1.doc/, 2004.
12. Văduva I., *Simulation Models by Computer*, Technical Publishing House (in Romanian), Bucharest, 1977.
13. Verdonck F., Jaworska J., Thas O., Vanrolleghem P.A., *Uncertainty Techniques in Environmental Risk Assessment*, Med. Fac. Landbouw. Univ. Gent, 65/4, 247-252, 2000.
14. ***, *Testing Interlaboratory Comparisons*, Asia Pacific Laboratory Acreditation Cooperation, http://www.aplac.org/documents/web_docs/pt_002.pdf