

Recent Developments

Nooreen Dabbish, Daqian Huang, and Shinjini Nandi

April 13, 2015

Abstract

1 Instructions

Evaluation: You will be evaluated on the basis of BOTH the quality and clarity of in-class presentation and the final report. All the students of any group will get the SAME grade (depending on their contribution on the project). Before presentation you MUST TURN IN the printed final report to me.

Format of the Report: Clear Title, Names of Teammates (and their individual contributions) and Abstract on the first page. Summary of the main idea in the remaining 3-4 pages. The LAST section of your report should discuss what are the possible ways you think this work could be further extended. Use Latex to prepare the manuscript. Clear reference list should be provided. While writing the manuscript your emphasize will be to summarize the main findings in your own words (do not copy paste). Presentation should be simple yet comprehensive. You are highly recommended to use numerical examples (when applicable) to better illustrate the ideas.

Format of presentation: Board work. Use your manuscript while presenting. Time: 40 mins + 5-7 mins Q&A session. . Illustration using data set and case studies (while presenting your idea) is highly valuable.

Start as soon as possible with your group members.

This is a GREAT opportunity. Successful projects in the past have gone on to become full-fledged research papers !

2 Recent Developments in Bootstrap Methodology

2.1 Introduction

This article sets out to “give a bird’s eye overview of the current state of bootstrap research.” The authors cover basic ideas with references to bootstrap literature as well as giving in-depth explanation and examples of extensions. Topics include parametric inference using bootstrap stimulations, non-uniform nonparametric sampling, bootstrap failure, hypothesis testing, bagging, dependent data and other topics. Our goal is to summarize and highlight the examples and ideas in the paper, and we conclude with a section on future directions suggesting what next steps can take this work further.

2.2 Basic Ideas: Bootstrap approaches to confidence intervals and hypothesis testing

- Nonparametric confidence intervals
 - either Studentized pivots or direct use of quantiles
 - Studentized bootstrap uses estimated var V^* are second-order accurate $1-\alpha+O(n^{-1})$
 - improvement over $O(n^{-1/2})$
 - BC_a also second-order but additionally transformation-invariant.
- Using a pivot
 - avoids the need to modify the sampling plan in hypothesis testing
- Model-based bootstrapping
 - when data are not identically distributed.
 - time-series/autoregressive moving average
 - Notes on Notation:
 - * \hat{F} empirical distribution
 - * $F(y;\psi)$ parametric model with parameter ψ
 - *
 - Topics to explore/look-up
 - * conditions under which bootstrap is consistent (Bickel and Freedman 1981)
 - * Edgeworth correction

- * Edgeworth expansion
- * permutation tests

2.3 Bootstraps for Parametric Likelihood Inference

- profile log-likelihood $l_p(\gamma)$
- ratio statistic $w_p(\gamma) = 2(l_p(\hat{\gamma}) - l_p(\gamma)) \sim \chi_1^2 + O(n^{-1})$
- signed root likelihood ratio statistic

$$r_p = \text{sgn}(\hat{\gamma} - \gamma)w_p(\gamma)^{1/2} \sim N(0, 1) + O(n^{-1/2})$$

- r_a adjusted r_p

2.3.1 Example 1 Exponential Regression

Lawless(1982)

Survival in weeks	log wbc
156	2.88
108	4.02
143	3.85
56	3.97
1	5.0

2.4 Weighted Non-parametric Bootstrapping

2.5 Subsampling and the m out of n bootstrap

2.6 Bootstrapping Superefficient estimators

2.7 More on Significance Tests

2.8 Bagging and Classification

2.9 Bootstrapping Dependent Data

2.10 Other topics

2.11 Final Remarks

3 Future Directions