Topics in Probability (STAT 547C)
Fall 2020
https://ben-br.github.io/stat-547c-fall-2020/

Zhipeng Zhu
zhipeng.zhu@stat.ubc.ca

Concentration inequalities in Statistical learning

Project Outline

1 Background

Inequalities in statistics provide a means of bounding measures and quantities. They are usually used to specify bounds on quantities when these bounds are particularly difficult or intractable to compute. Inequalities play an important role in the algorithm of statistical learning and machine learning. They are involved in underpinning methods or approaches used in actual cases.

Here are several famous concentration inequalities involved in statistical learning.

- Markov's Inequality
- Chebyshev's Inequality
- Bounded Differences Inequality

2 Technical aspects

First of all, basic statistical knowledge, including distributions and calculus, are required.

Depending on different inequalities, specific techniques will be used including moment generation function, sums of independent random variables and martingale methods. The other basic mathematical techniques will also be used.

3 Literature

The key references for this project are:

1. Bartlett, P. (2020, November 10). CS281B/Stat241B. Statistical Learning Theory. Lecture 4. Lecture. Retrieved from `https://bcourses.berkeley.edu/courses/1409209/files/65720942/download?wrap=1`
2. O. Bousquet, S. Boucheron, and G. Lugosi. Introduction to statistical learning theory". In: Summer School on Machine Learning. 2003, pp. 169207. url: `http://www.econ.upf.edu/~lugosi/mlss_slt.pdf`

4 Plan

I will carry out this project with the following sequence of steps:

1. Introduce the concentration inequalities. What it is and how it functions.

2. Introduce the main techniques in proving or constructing the ineuqalities.

3. As there are several versions of concentration inequalities, each single inequality will have its own section including proof and extensions. Through each section (for each inquality), there will be several propositions, lemmas and sub-proofs.

4. After all the inequalities are introduced and varified, there will be some related exercises focusing on inequalities' applications. Along with these exercises, more practical applications of these inequalities in statistical learning will be introduced and discussed.

5 Why I'm interested in this topic

Personally, I have been interested in inequalities from primary school. The techniques used in solving and proving inequalities attract me. In past years, I encountered many questions involving inequalities. The participation of inequalities reduced a lot of work in determining or specifying some range or bound of quantities.

I am also interested in machine learning. Besides actual practice in computers, the background theory is also what I want to explore. Therefore, I choose the topic regarding the inequalities involved in statistical learning.