

Computer-intensive Statistics and Data Mining

1MS009

Silvelyn Zwanzig

Uppsala University, Period 3 , 2010

1 Goal

The goal of this lecture course is to present the different ideas behind the new methods in statistics due to increasing computer capacity. The course will give an introduction in both in theory and practice. Real data sets are studied using the GNU software R. You can download R from the net. See www.r-project.org.

2 Organization

Lectures. Software demonstrations are given during the lectures.

Students presentations

Last lecture repetition - before the written examination, 16/12

3 Examination

4 Obligatory home works (Inlupp 1-4), team work is strongly recommended

Obligatory presentation (Inlupp5) of the statistical analysis of an own chosen data set: 5-10 minutes (team work!)

written examination 16/12

4 Literature

There is no textbook as course literature. The script can be downloaded from my home page <http://www.math.uu.se/zwanzig>

Strongly recommended as additional literature:

Hastie, T., Tibshirani, R. and Friedman, J. (2001), *The Elements of Statistical Learning*. Springer

The course touch only Chapter 2,3,4,5,6,7,8. It is useful to remind the regression course and to read 3.1-3.3 by your own.

5 Contact

Do not hesitate to contact me: Silvelyn Zwanzig, zwanzig@math.uu.se, rum 74106