Kursanalys SF2930 (Regression Analysis) VT2022

Staff

Course coordinator and lecturer: Malin Palö Forsström
Exercise assistant: Isaac Ren
Examiner: Mykola Shykula

Guest lecturers: Mattias Villani, John Pålsson, and Oscar Allerbo.

Course representatives

Bogdana Vashchuk, Anna Melnichenko, and Alexander Westman were course representatives. Two meetings with the course representatives were held. The first meeting was held a few weeks into the course, and the second meeting was held after the course was finished. At the time of the second meeting, we also had the results from the course evaluation.

Focus and level of the course

There seemed to be two distinct groups of students taking the course; students with a background in computer science, and students with a background in engineering mathematics/physics. Some of the engineering mathematics students felt that the course was far too applied and not advanced enough for being an advanced-level math course, while the computer science students generally seemed very satisfied.

One suggestion for next year is to discuss the level of the course (applied vs. non-applied) at the beginning of the course next year so that the students know what to expect.

Lectures

Regular lectures

The lectures and exercise sessions this year were held exclusively on Zoom. Several students asked for these to be recorded, but due to a math department policy, and also because I did not have time to further look into how and if to do this in a good way, this was not done. Roughly half of the registered students attended the lectures, and this level was relatively constant throughout the course. For each lecture, R code and slides were provided to the students.

The slides were intentionally kept relatively simple, with incomplete information, to encourage the students to attend and listen to the lectures, and also for me to be able to fill in information during the lectures. Some students asked for the slides to be more detailed, but I am skeptical as to whether this would be beneficial to the students. Some students also asked that more of the information on the slides would be added at the lectures (handwritten), and this was implemented for the second half of the lectures.

Guest lectures

The course had three guest lecturers, lecturing on Bayesian regression (MV), general linear models (JP), and kernel regression (OA) respectively. These lectures seemed to be appreciated by the students in general. One thing that could perhaps be improved for next year is to make sure there are references to reading material provided for these lectures, which was not the case this year.

In a meeting with JP after the course was finished, he suggested that next year the lecturer of the course will take over the theory about general linear models and that they in their lectures concentrate on their particular application.

Projects

In the course, there were two projects (referred to as Project 1 and Project 2 below), and in general, the students seemed to believe that they were good.

For Project 1, the students could choose between two different projects, which were quite different. The theory needed for the second of these projects was not presented in the lectures until quite late in the course, which pushed more of the course work until the end of the course. One suggested solution was to split this project up into two separate projects. Then all students would be required to do both projects, and thus learn about both of the two very different topics.

Another suggestion, in particular for Project 1, was to include references in the project instruction about which lecture corresponds to which part of the project so that it would be easier for the students to plan when to work on which parts of the project.

Looking at the submitted reports for Project 1, I would recommend that the following is stressed more in the lectures next year.

- The model assumptions should always be verified for a proposed model *before* one considers statistics such as confidence intervals.
- The model assumptions needs to be rechecked every time one looks at a new model, such as the model corresponding to a subset of the variables.

For Project 2, there was some criticism that it seemed to be more about data management than general linear models. In particular, since code for the GLM part was provided in the project.

For next year all the project instructions need to be updated with the correct year and deadline. However, since I did not have access to the tex-files, I did not do this for this year.

Feedback

A general comment about the projects in the course evaluation was that the students would have liked to get more feedback. In previous years, the deadline for the projects was on the same day as the exam. For this reason, I assume that the purpose of these has never been to give the students feedback ahead of the exam. This year, even though many students handed in the project well ahead of the deadline, most students only got comments on the mistakes they made. One student suggested that this could be solved by replacing the project reports with lab presentations. The idea would be that by doing this the students' work could be discussed in more detail, without using more teaching hours. Also, it would save time for the students as they then would not have to write reports. However, the course representatives felt that they actually learned a lot from writing the report itself, and hence was not convinced that this would be an actual improvement of the course.

A difference compared to previous years was that If changed the format on their last lecture, and replaced the previous discussion of the projects with an "Insurance game". As a consequence, the students did not have to hand in any factors to if. Since the students would in general want more feedback, a discussion was had with JP about whether or not this should be reinstated somehow. One midway solution would be to have some automated feedback indicating how good the obtained model was.

Deadlines

The deadline for both projects was on the day of the exam. Several students said that this encouraged doing very little work at the beginning of the course, and then having a lot of work to do at the end. Also, several students said that it was stressful to have the deadlines on the same day as the exam. A suggested solution, which might fix this, would be to have the deadline for, in particular, Project 1 earlier since all the material needed for this project is presented in the first weeks of the course anyway.

Examination

As has been the case in previous years, the students were given a list of 36 questions a few weeks ahead of the exam, and six of these questions were then chosen for the examination.

The students seemed to be satisfied with this format, both before and after the exam. Also, as can be seen from the distribution of grades, it was certainly not the case that all students got all points even if they had all seen the questions before the exam.

As a consequence of the format, many of the students collaborated on writing up solutions to the questions on the list of questions for the exams. Unfortunately, it became evident when grading the exams that several of these solutions were either wrong or incomplete. Also, many students thought that since they had memorized such "solutions", they would all get full points, which was not the case, leaving some students disappointed or confused. It is not clear what the best solution to this problem is, but perhaps one should at least stress that memorizing a solution that one does not fully understand is probably not a very good method if one wants a good grade on the exam.