Programsensorrapport for Bachelorprogram i statistikk, Integrert master i aktuarfag og Masterprogram i statistikk ved Matematisk institutt, UiB.

Programsensor: Tore Selland Kleppe (UiS)

Rapporten leveres for fristen 1 juli 2016.

Background

This report considers the master and bachelor programs in statistics as well as the integrated master program, as the three share many elements and may be regarded as tightly interwoven.

In the run up to writing this report, in order to gauge some of the sentiment, I wrote an email to the members of the statistics group asking them whether they would like me to consider 1) how to relate to the Big Data / Data Science "revolution", 2) try to get some feedback to former students who currently work in industry or 3) anything else that may be of interest to the employees (at that point in time I was thinking about discussing one for the former two). This email sparked a number of replies and some email-based debate, and much of this report details and discusses ideas that came up during these exchanges. In addition, the issue of increasing the number of applicants to the bachelor in statistics was mentioned, but I do not have much to add in that respect.

Some general comment

Even though the separation into different sub-discipline (stats, applied, pure) bachelor studies and the integrated master are relatively recent constructs, the overall statistics education at the UiB appears mature and very well worked through. I therefore comment more on specific details rather than the overall programs.

What to do with respect to Data Science / Big Data

Using a definition taken from the Wikipedia page for "Data science", "Data science is an interdisciplinary field about processes and systems to extract knowledge or insights from data in various forms, either structured or unstructured, which is a continuation of some of the data analysis fields such as statistics, data mining, and predictive analytics…". Big Data generally refers to large amounts of data, where what is complicating statistical analysis is generally about handling/processing large amounts of information, rather than the making sure we have sufficient information so that the asymptotic theory holds.

Though the definitions above can be discussed, the phenomenon is relatively tangible and is very much discussed these days[1]. Moreover, it seems that candidates with such skills are highly sought after these days. In my opinion, both the stats bachelor/master and the integrated program should contain some form of exposure to Data Science in order not to loose some relevance / edge against corresponding educations elsewhere. E.g. UiO has the Big Insight center and the Informatics and Mathematics departments are jointly hiring academics in Data Science.

---

[1] Interestingly, I saw an article about the scarcity of experts in Big Data analysis in Norway in Aftenposten when preparing this report: http://www.aftenposten.no/okonomi/Selskaper-kappes-om-Ellie-og-Alexandras-kompetanse-8422984.html

A good starting point would be to *include one course in Data Science in the programs* (probably late bachelor or early master). This should at least give the students some exposure to the Data Science and Big Data, and how to handle such problems in practice using e.g. R and Hadoop. It is possible that such a course also would be of interest to the Informatics Department, *and it seems reasonable to discuss this with them*. (I had a quick look the Informatics course portfolio, but it seems that they primarily have a database course and a machine-learning course). Another possible avenue, as was pointed out is to consider *using a MOOC*, e.g. from Coursera to implement such a course, most likely in conjunction with a locally organized exam.

Of course, it should not be taken lightly to change the program in this manner, as it is likely that something may/has to be taken out to compensate for the added material. I do not have an obvious candidate for this, and this is something the statistics group must discuss.

Reliability
Another potential addition to the course portfolio was a course in reliability analysis, which is sought after skills for the local industries. I agree fully that such a course would be good to have. However, it is certainly an investment to make and run a new course, and in my mind I would have preferred a Data Science course over a reliability course if the budget is only one new course (see also comments regarding cooperation across institutions below).

Internationalization
Another issue that was discussed was to make the master program an all-English program (does the University have a distinction between programs that require/does not require proficiency with Norwegian?) I think this would be a relatively modest investment since, from my recollection; most of the teaching/exam is in English anyway. If foreign students would feel certain that language would not be a problem (given that such a label exists), this could probably lead to a somewhat increased number of applicants. Thus at least *figuring out if the University has an "English only label"* would possibly be worthwhile.

It is of course possible that such a step would possibly alter the profile of who applies for the master program, and this must be taken into account if such steps are taken.

Exams
Both in conjunction with Data Science and teaching in English only, the exam was discussed as a possible obstacle. In the case of including a Data Science course, and I think also in some of the existing courses that uses R / programming a lot, there should be a way of testing the skills of the students in solving practical tasks using a computer. In theory, this should be rather straight forward, i.e. let the student solve some tasks using R and include the code in the answers to be handed in. However, it requires a lot of work and a willing IT-department in practice (we do such a thing at UiS, but arguably it is a lot of work).

Rumors around here are that every university exam will soon be performed on a computer, and as far as I know, several solutions for this are in the pipeline. At the UiS, we use (in production for many "text only" subjects, math and physics are somewhat skeptical given the current rather crude ways of typing formulas) the solution http://www.inspera.no, and we have been told that future versions will allow selected software (e.g. R or a compiler for the programming courses) to be available in the exam environments.

My take on this is in short to *wait for the production digital exam solutions to be rolled out*, and hope/require/ask for a possibility to use R in these.

The exam was also mentioned in the internationalization context, and I guess it would be reasonable to *make only an English exam if the master program had an English-only label* to save unnecessary work. This naturally has to be OKed by the UiB, but I do not think there are any other formal obstacles to giving the exam in English only (we do this routinely at UiS at the master level. Sometimes we are encouraged to make a Norwegian translation; so far I have refused to do so).

Further discussion
Most of the changes discussed above a relatively minor, and the overall picture is that the statistics educations at the UiB are in good shape, and that small changes could improve them further.

Another aspect that is worth mentioning is that the academic statistics community in Norway is very small, and that it may be a good idea to share the courses across the universities, possibly by using modern technology. In particular, I wish to mention that at the UiS we have many courses related to reliability/industrial risk management[2], and also a very vibrant group that does Data Science[3], including courses. Thus making it easier to follow courses across the universities would be a good thing. I.e. remove barriers against taking courses across institutions, (I recently gave up trying to send one of our master students to take an advanced course in Bergen) as it seems this may be the way forward.

Tore Selland Kleppe
Stavanger, April 2016

---

[2] See e.g. http://student.uis.no/subject?code=RIS510_1
[3] See e.g. http://student.uis.no/subject?code=DAT500_1 and the webpage of the lecturer: http://www.wiktorski.net/