Semester: Spring 2023
Course: MAE4120 - Item Response Theory
Participating students: 18
Answering frequency: 7/18 (39%)
Date: 2023-05-22

Comments from the responsible teacher on the implementation and outcome of the course

The course is an introduction to item response theory (IRT) with an emphasis on various applications of IRT in the social sciences. Students are introduced to IRT models for binary and ordinal data and how tools in IRT can be used to evaluate individuals in terms of a latent construct, infer group differences, construct scales and tests and to investigate hypotheses regarding item properties or the relationship between performance and covariates. The course mainly focuses on unidimensional IRT but gives a brief introduction to multidimensional models. The course involves teaching in the form of lectures which introduces the topics and areas of applications, computer labs which allow for using the methods in practical situations and seminars which allow the students to present IRT studies and results of IRT analyses. The exam consists of an analysis using IRT that students propose and receive feedback on before handing in the final report.

The evaluation was conducted anonymously via Nettskjema after the exam but before the exam results were given. The evaluation form consisted of the parts *General*, *Course topic emphasis*, *Work load*, *Assignments*, *Exam,* and *Overall*. The evaluation without free response comments is attached to this report and a summary of the free response comments is given as an appendix.

The spring 2023 course was conducted with a similar structure as the previous year. The largest change involved the format of the last seminar, where the topics for the final exam were presented. This year, we implemented a system of peer review where the students had to read through the proposed exam topic of another student and present the exam topic, along with suggestions for improvements on the proposal. Written feedback for each exam topic was also provided by the responsible teacher afterwards. As per the suggestions from the previous year, the material regarding polytomous IRT were particularly revised and the labs were partly re-worked be more practically-oriented. The lectures, labs, assignments and exam had the same structure and focus as previous years. The course was taught by an associate professor and a PhD candidate.

The course was mainly evaluated positively by the respondents. All respondents strongly agreed or agreed with the statement “I would recommend this course”. One free-response comment concerned the usage of simulated data, where it was suggested to use real data to a higher extent. The course primarily uses real data in the examples of the lectures and during several of the labs. Simulated data are used in both assignments largely because it allows for each student to have a similar but not identical data set to work on in solving the same tasks. For this reason, no changes are planned regarding the assignments in this aspect. Additional real data can instead be utilized in the labs to prepare for the real data tasks in the final exam. The free response comments on emphasis suggested to

devote more time to multidimensional IRT, emphasize the connection between classical test theory and IRT, and additional material on the distinction between the Rasch and two-parameter logistic models. For next year's course, these subjects in particular will be reviewed. The assignments were well-regarded but two free-response comments indicated that there could be more than two assignments. Previously there were three assignments in the course but this proved difficult to manage in terms of grading ahead of the exam registration deadline and hence the current system of two assignments will be kept.

The exam was generally well-regarded but some students indicated that was not enough time complete the exam. Two students indicated that the exam questions were surprising. To note here is that the exam questions and the grading guide were posted far in advance of the exam hand-in, and it is not meaningful to post these even earlier. In addition, to avoid conflicts with other courses, it is not appropriate to have the exam deadline later than it is. There were conflicting opinions regarding the exam in the free-response comments, where one student indicated that the exam format was not ideal but where another student suggested that the exam format was good.

Proposed changes/comments/measures

- Keep the format of lectures, labs and seminars
- Maintain the assignments but add some further smaller tasks in addition to the assignments
- Review the course material pertaining to multidimensional IRT, the relationship between classical test theory and IRT, and the distinction between the Rasch and 2-PL/3-PL models
- Keep the exam format and peer-review system for the topic proposals but add a more detailed description of what is expected regarding the structure of the final exam

Björn Andersson,
Responsible teacher

Appendix: Summary of student viewpoints and suggestions

In the following, the main viewpoints reflecting the free-response student comments are given.

General

- Work more with real data examples which prepares better for challenges in realistic settings.

Course topic emphasis

- Not enough time devoted to multidimensional IRT.
- Not enough emphasis on introducing how IRT differs to classical test theory and factor analysis.
- More on the distinction between Rasch and 2-PL/3-PL models.

Work load

- Ranged from 12 hours to 40 hours.

Assignments

Strengths

- The assignments were fine, with the second being challenging (which was good).
- The depth of the assignments and the coherence of the questions.
- Enabled good practice which was useful to understand basic concepts of the course.

Weaknesses

- Two assignments not enough.
- Perhaps too few assignments, could also have smaller assignments.

Exam

- The final exam was too unstructured which made the work too difficult if the chosen topic was complicated.
- Getting to choose the topic was good in order to prepare for the thinking of the master's thesis.

Overall

Strengths

- Freedom to choose the exam topic a very good idea.
- The showcase of all the different aspects of IRT.
- Labs were structured well (especially the later labs where some preparation was instructed ahead of time).

Suggestions for improvement

- More time for feedback on the exam topic choice and some guiding questions to focus the topic.
- Have the same exam topic for everyone, to ensure the same conditions for the exam.
- Provide an example structure for an IRT article to better understand what to include in the final exam.

- Give a more detailed introduction to how IRT developed and why it is used.

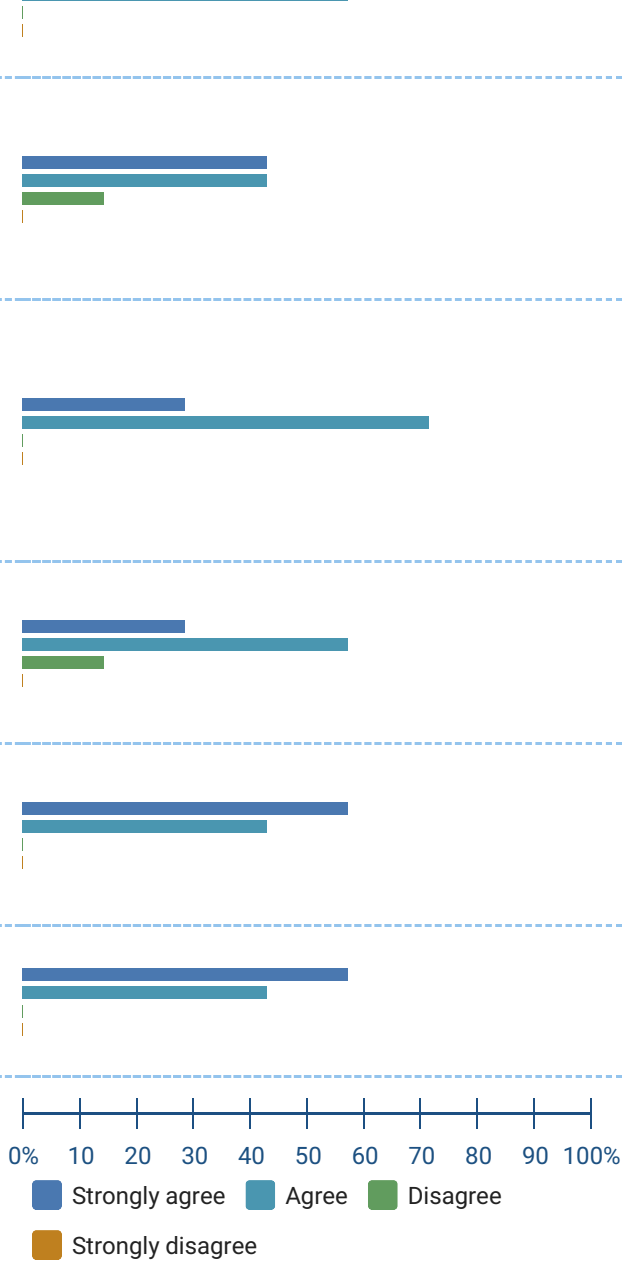
Nettskjema

MAE4120 Item Response Theory: Course Evaluation V23

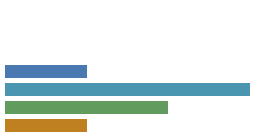
Updated: 9 May 2023 at 9:15

Question without text

Submissions	Strongly agree	Agree	Disagree	Strongly disagree	Chart
The information provided on the UiO course page was sufficiently clear	5	2	0	0	
The information provided on the CANVAS course page was sufficiently clear	5	2	0	0	
The learning outcomes of the course were met	4	3	0	0	
The instructor (s) explained the topics clearly	3	4	0	0	
The instructor (s) demonstrated concern about whether I was learning	3	3	1	0	
The instructor (s) inspired and motivated me and encouraged my interest in the course content	2	5	0	0	
The speed at which the course proceeded was exactly right for me	2	4	1	0	
The course improved my critical thinking skills	4	3	0	0	
I would recommend this course	4	3	0	0	



Question without text

Submissions	Strongly agree	Agree	Disagree	Strongly disagree	Chart
There was sufficient time to prepare before the scheduled exam	3	3	1	0	
There was sufficient a priori information given on the nature of the exam	3	2	2	0	
The time to complete the exam was sufficient	3	2	2	0	
The exam questions did not come as a surprise to me	3	2	2	0	
The exam adequately covered the whole span of the course contents	1	3	2	1	
The exam questions were clearly formulated	1	4	1	0	
I feel I have a pretty good idea about how I will score on the exam	1	5	1	0	

Nettskjema

MAE4120/UV9293: Item Response Theory Interim Evaluation V23

Updated: 15 May 2023 at 14:05

I think the pace of the course so far has been:

Number of submissions: **7**

Submissions	Count	% of submissions	
Much too slow	0	0%	0%
Too slow	0	0%	0%
Just right	3	42.9%	42.9%
Too fast	3	42.9%	42.9%
Much too fast	1	14.3%	14.3%

I think the difficulty level of the course so far has been:

Number of submissions: **7**

Submissions	Count	% of submissions	
Much too low	0	0%	0%
Too low	0	0%	0%
Just right	4	57.1%	57.1%
Too high	2	28.6%	28.6%
Much too high	1	14.3%	14.3%

To what degree do you think the amount of teaching has been sufficient?

Number of submissions: **7**

Submissions	Count	% of submissions	
Not at all	0	0%	0%
A little bit	0	0%	0%
A fair amount	6	85.7%	85.7%
A high amount	1	14.3%	14.3%