EVALUATING THE QUALITY OF THE MASTER'S PROGRAM IN STATISTICS: A STUDY OF THE CURRICULUM EFFECT ON THE QUALITY OF STUDENTS' MASTER THESES

Tatjana von Rosen
Department of Statistics
Stockholm University, SE-106 91Stockholm, Sweden
tatjana.vonrosen@stat.su.se

In 2012, the Swedish Higher Education Authority evaluated the quality of the Master's Program in Statistics with a main focus on the students' master theses. In this work, the course outcomes of all master students (N=55) registered at the Stockholm University, between 2012 and 2017, were used to investigate whether a particular practice at the department that can be characterized by program syllabus, supervision, teachers' expertise, is associated with the quality of master theses, or the quality has been mostly affected by student related characteristics. The obtained results showed significant effect of the course results in Inference, Statistical Computations and Multivariate Analysis. No significant effects of student's gender and age have been found however the supervisor's effect turned to be significant.

BACKGROUND

In the last decade, learning outcomes have played a significant role in the reforms of education policies in many European countries (e.g. Cedefop 2008, 2009). To some extent, this was caused by need to compare educational programmes and ensure both the quality and homogeneity of qualifications across universities/countries (Cedefop 2008; Adam 2008). In Sweden, Swedish Higher education institutions (HEIs) share responsibility for quality assurance in higher education with the Swedish Higher Education Authority (UKÄ). In 2011-2014, as a part of work with quality improvements in higher education, UKÄ conducted the evaluation of education programmes at Swedish HEIs. In this evaluation, targeting the outcomes of programmes one wanted to assesses to what extent the learning outcomes achieved by the students were consistent with the expected/intended learning outcomes (LOs). The outcome-based approach to evaluation is based on the idea that LOs capture the essential issues of learning and knowledge, and ensures that students have achieved the required qualifications upon completing the degree/education programme. Assessment material consisted of the HEI's self-evaluations, interviews with students and representatives of the reviewed programme. In their self-evaluations HEIs had to illustrate using concrete examples how does a specific degree programme ensure that students were given good opportunities to achieve the intended LO of the degree programme. Since 2011, student's degree project (DP) become ultimate criterion for the quality evaluation system being a final "product" of the educational programmes. Hence, evaluating DPs one examined the extent to which the students' DPs (actual study results) corresponded to the expected LOs for the respective education.

DEGREE PROJECTS AS QUALITY INDICATORS IN HIGHER EDUCATION

Nowadays, student's degree projects (undergraduate and master theses) have become of utmost importance in Swedish higher education, partly because of the Bologna-related reforms and the quality evaluation/audit system of 2011–2014. In this evaluation system, the students' DPs were considered as a quality indicator function for the entire education/degree programme, and hence provided a solid basis for the overall assessment of e.g. master's programme in Statistics. These DPs were randomly collected from statistical departments and the external experts assessed how well the students have achieved certain learning outcomes (predetermined in the qualification descriptors). Those programmes that had showed lack of quality, i.e. they failed to attain the required standards for higher education, were evaluated as having inadequate quality. Further, they were followed-up after one year in order check what actions they had taken to improve the quality of the programme. Among others, the emphasis was put on safeguarding a lowest acceptable level of DP.

Hence, in order to enhance the programme to adequately meet required standards, e.g. in terms of learning outcomes, we aim to investigate whether a particular practice at the department that can be characterized by e.g. programme syllabus, examination forms, supervision, teachers' expertise, course administration, is strongly associated with a quality of student's DPs reflected in its grade, or the quality/grade of DPs has been mostly affected by student related characteristics, e.g. entrance score and study achievements during the whole education.

DEGREE PROJECTS, QUALITY, CURRICULUM

In Sweden, a number of structural changes took place in higher education partly related to the Bologna process which could have influenced the status and conditions for the students' DP which is compulsory within most of the master's programmes. In recent years, due to an increased focus on DPs as a product characterizing student's actual knowledge and scientific literacy, there has been seen a revived interest toward research about student's DPs. A substantial part of Swedish research is concerned with professional education, e.g. teacher education (Svärd, 2013). Though, the master's DP is the final and often crucial part of the education programme, in general, there is little research focusing on it (Anderson, Day & McLaughlin, 2008; Drennan & Clarke, 2009; Eklund, 2009; Feldt, Höst & Lüders, 2009). Some general perspectives can be found in research about thesis supervision (Dysthe, 2002; Holmberg, 2006) and assessment (e.g.Hand & Clewes, 2000; Pathirage et al., 2007).

The quality evaluation of the university degree programmes including Master's programme in Statistics, stimulated discussion about what a statistics curriculum should contain. For the most part, attention has been focused on the compulsory courses. In Sweden, there are commonly several compulsory courses: probability, inference, statistical computing, statistical theory of Science, degree project. Additionally, multivariate analysis, mathematics and some more courses on statistical methods are requited. In contrast, there has been much less attention paid to updating the course DP. This may be partly due to the vision of DP as a natural and final part of the master's education where students are supposed to use the knowledge acquired during the two year master's programme.

In spite of dynamic changes in many scientific branches and technology and their impact on the statistical theory and practice, the statistics curricula have not been modified to follow new trends, at least not at all Departments of Statistics in Sweden. Traditional statistics courses commonly give focus to computing and scarcely discuss issues related to "statistical thinking" and statistical/practical experience. Come courses, like Statistical Methods, demonstrate using few examples how certain statistical methods are used in practise and how do they perform. This is often because of course budget limitations. The DPs however, give the students possibility to complement missing academic knowledge from the curricula and integrate it in practical applications.

In this work, we present results from a study of all 55 DPs in Statistics produced by master students between 2012 and 2017 at the Department of Statistics, Stockholm University. We can argue that Master's programme in Statistics at Stockholm University is quite representative in relation to other master's programmes in Statistics in Sweden with respect to the curriculum and students' general/average performance and teachers' competence. It is worth mentioning that the course work is quite regulated by the course plans which are commonly similar among the Statistics Departments due to the collaboration and pedagogical dialog between teachers. Students within the Master's programme in Statistics at one university can read statistical courses at other universities without hindrance. Later, some of them apply to PhD programs in Statistics and their course outcomes will be treated/evaluated in the same way. Furthermore, the courses in Sweden are mostly connected to a certain text book (can be complimented with other materials), and the course literature is quite homogeneous between Statistics Departments. For example, in Multivariate Analysis course one is using Johnson & Wichern book "Applied Multivariate Statistical Analysis", in Probability and Inference courses the book of Casella & Berger "Statistical inference" is used.

The present study is partly inspired by the pilot study of all DPs collected from the five largest History departments in Sweden, where the relationship between upper secondary school grades and undergraduate DPs quality was studied (Ekecrantz et al., 2015). A significant correlation between thesis quality and upper secondary grades could be shown only among theses

that passed with distinction, suggesting a mock research phenomenon discussed in Ekecrantz et al. (2015) which is caused by the traditionally used professional research model where DP is the main intended outcome. This differs from core principles in research-based learning, where the primary outcome would be student learning. Here, we matched the quality of master DPs with individual student grades from other subjects included in Master's programme in Statistics, and intended to discuss several related issues: how student's knowledge accumulated from master programme curriculum relates to DPs's quality in different groups, is DP's quality directly influenced by specific course outcomes, can the DP's quality be influenced by factors independent of student learning, e.g. supervisor-related factors.

It is important to recognize that in real-life there is a lot of factors influencing DP quality, including student's motivation, personality traits, institutional and cultural aspects, among others (Diseth, 2011; Giota, 2010; Poropat, 2009). In this study we shall use a simple model which allows to explore the master students' DPs using the data available at the department without extra work on acquiring information related to personal or outside-department study data. The latter would demand additional resources to cover personal costs. The data available at the department can be efficiently used to discover possible patterns in students' performance and the factors influencing their study outcomes, which could be used to improve the teaching and learning at the department, and hence the educational programmes in general.

DATA AND METHOD

The data for this study comprises the course outcomes of all master students (N=55) registered at the Department of Statistics, Stockholm University, between 2012 and 2017. The Degree Projects were evaluated according to the assessment criteria available at the beginning of the course and graded according to a seven-point scale (A to F). Then the original grades were converted into the quantitative values on a 100-points scale using routines to generate uniformly distributed random numbers over a pre-specified interval. For example, grade A corresponds to a generated value from the interval [90,100], grade C to a value generated from [70,80), etc. Additionally, the gender and age of master students, the field of the master thesis (MT/DP) and the scientific competence of the supervisor (PhD, ass professor, professor) were recorded. From the analysis eight observations have been excluded: when the corresponding supervisor had had only one or two DPs under supervision during 2012-2017. Among master students 45% were females (average age 30, STD=5.08) and 55% (average age 30, STD=5.67) males, respectively. The correlation analysis was conducted to explore correlations among course outcomes. A linear model was used to identify factors significantly affecting the grade of DP.

RESULTS

As expected, some of the course outcomes were significantly correlated, for example, the grade in Statistical Methods was significantly correlated with grade in Probability (r=0.37, p-value=0.019) and Inference (r=0.31,p-value=0.045).The grade in Probability was correlated with grade in Multivariate Analysis, MVA, (r=0.49,p-value=0.006), Mathematics (r=0.54,p-value=0.0002) and Stat. Computations (r=0.43,p-value=0.005), correspondingly. The grade of DP was significantly associated with the grade in MVA (r=0.35,p-value=0.049). Removing gender effect, significant correlations are observed only between DP and MVA (r=0.49, p-value=0.017), between grade in Stat Methods and Probability (r=0.55, p-value=0.007), Stat Methods and MVA (r=0.65, p-value=0.01), Stat. Computations and Inference (r=0.63, p-value=0.001).

In Figure 1 one can explore the spread of grades within subjects which exhibit relatively large variation among subjects.
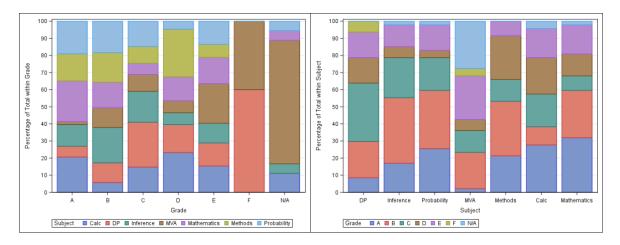
Figure1. The distribution of subjects for the corresponding grade (left); The distribution of grades (A-F) within subjects (right). Here, DP stands for Degree Project, MVA for Multivariate Analysis, Calc for Statistical Computations, Methods for Statistical Methods, respectively. N/A corresponds to a missing value.

Plotting the course outcomes, i.e. individual trajectories which can be regarded as study profiles, discovers a lot of variation among students in each group corresponding to the grade of DP (see Figure 2).
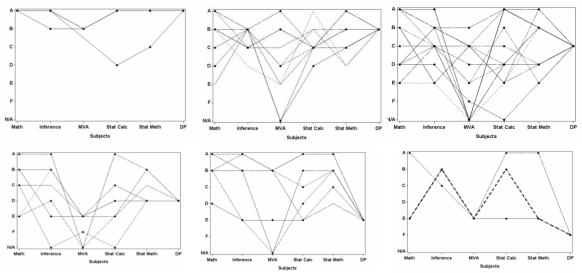

Figure 2. Student's course outcomes (Grades) in Mathematics (Math), Inference, Multivariate Analysis (MVA), Statistical Computations (Stat Calc), Statistical Methods (Stat Meth) according to their performance in Degree Project (DP), from A to F.

To determine factors which could potentially influence the grade of DP, and in a certain way its quality, an Anova-model has been used. The results of the statistical modelling show that the course outcomes in Inference (p-value=0.032), Statistical Computations (p-value=0.033), Multivariate Analysis (p-value=0.01) significantly influence the grade of DP. The grade in Mathematics is significant at 10% significant level. Furthermore, there is a significant effects of DP's supervisor (p-value=0.032). To better understand how the response variable (score of DP), varies across factors in the model, least squares means have been calculated and compared. As result, the significant difference in the average DP score has been observed between students having in Inference grade A, B, C, or D comparing to those having E grade. The average DP scores differed significantly for students having in MVA B or D grades compared to those having E grade. The average results in DP for students with A grade in Statistical Computations were significantly different from students with grade C. The analysis did not discover any significant

effect of student's gender or age on the DP score, similar result were obtained for the supervisor's scientific competence.

CONCLUSION

The master DP in nowadays the most work intensive module in Swedish Master's programme in Statistics, it is often used when evaluating PhD candidates, and is therefore associated with high assessment standards. The DP is usually a mixture of an original empirical research and a solid theoretical background, written by a single author. Prior to the DP the students have finished reading twelve master courses in statistics, some can be in related subjects, e.g. computing, mathematical statistics, economy.

Traditionally, the relationship between a student's actual knowledge and skills and the finished DP are assumed to match perfectly. This is why DPs have been used in the quality evaluation of the educational programmes. In practice, the judgement about the quality of the whole program based on students' DP should be done with care (Hamilton, Johnson & Poudrier, 2010). The causality is very difficult to establish due for example the complexity of the educational environment, learning and teaching processes.

It is important to highlight the significant effect of supervisor on the DP outcome which could be found in our study. This is in line with the discussion concerning the role of supervision (Cook, 1980; Todd, Smith & Bannister, 2006; Heinze & Heinze, 2009), and the potential supervision effect, both positive and negative, on student's learning and DP. The supervision issue is often neglected due to teaching planning at the department, and possibly due to its complexity, since it can be considered as a collaborative work given the authority of a supervisor. It is well understood that a finished DP may in fact be better than the student's ability to produce it. This may be explained by supervisor's help and various support structures assisting to improve the quality of a DP without influencing the student's independent learning abilities to the same degree (Ekecrantz et al., 2015). It should be noted here, that there was the same examiner of DPs during 2012-2017 who worked out clear assessment criteria for DPs which were communicated to the master students as well as supervisors.

The obtained results showed significant effect of the course results in Inference, Statistical Computations and Multivariate Analysis on the grade of DP. These results are not surprising if one considers working on the DP as a profound experience based on learning-by-doing principle when students within certain time limits integrate statistical knowledge and skills for solving multifaceted real-life problems. These three courses provide students with necessary theoretical basis of univariate and multivariate statistics and computational skills. Possibly, students having good course results in these subjects can successfully handle complicated real-life problems treated in DP.

No significant effect of student's characteristics as gender and age has been found. This is also not surprising considering university/department systematic work on equal rights and opportunities for students.

As mentioned above, the Swedish Higher Education Authority has a responsibility to assess the quality of Swedish higher education. This places new demands on HEIs to evaluate and improve the quality of educational programmes locally. A dominant issue is whether an assessment of the quality of a DP can be used for evaluating the quality of the entire education/degree programme and whether to consider the thesis work as a process or a product, or as a combination of both (Penny and Grover, 1996). The practicalities around this issue, e.g. the course design and organization, can vary a lot among departments. Hence, there is a need at the department level to create a relatively simple but reliable tool to determine what is needed to maintain and improve the quality of education programmes. Consequently, more research is needed to support monitoring students' learning outcomes and the quality of DPs in relation to a specific subject.

Our further goal is to cooperate with other Departments of Statistics in Sweden and conduct the first thorough study of DPs in Statistics at advanced level, exploring the issues related to the quality of DPs in particular, and statistics education in general. This would help to meet challenges of modern statistics education for both teachers and students, improve teaching and learning by integrating in an efficient way the academic/subject knowledge and practice.

REFERENCES

Adam, S. (2008). *Learning Outcomes Current Developments in Europe: Update on the Issues and Applications of Learning Outcomes Associated with the Bologna Process.* http://www.academia.edu/32806775/Learning_Outcomes_Current_Developments_in_Europe_ Update_on_the_Issues_and_Applications_of_Learning_Outcomes_Associated_with_the_Bolo gna_Process

Anderson, C., Day, K. & McLaughlin, P. (2008). Student perspectives on the dissertation process in a masters degree concerned with professional practice. *Studies in Continuing Education*, *30*(1), 33–49.

Cedefop. (2008). *The shift to learning outcomes: Conceptual, political and practical developments in Europe.* Luxembourg: Office for Official Publications of the European Communities.

Cedefop. (2009). *The development of national qualifications frameworks in Europe; main tendencies and challenges*. Luxembourg: Publications Office of the European Union.

Cook M. C. F. (1980). The Role of the Academic Supervisor for Undergraduate Dissertations Science and Science-Related Subjects. *Studies in Higher Education*, *5*(2), 173–185.

Diseth, Å. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences*, *21*(2), 191-195.

Drennan, J. & Clarke, M. (2009). Coursework master's programmes: the student's experience of research and research supervision. *Studies in Higher Education. 34*(5), 483–500.

Dysthe, O. (2002). Professors as Mediators of Academic Text Cultures: An Interview Study with Advisors and Master's Degree Students in Three Disciplines in a Norwegian University. *Written Communication, 19*(4), 493–544.

Ekecrantz, S., Parliden, J., & Olsson, U. (2015). Teaching-research nexus or mock research? – Student factors, supervision and the undergraduate thesis in History. In D. Ludvigsson & A. Booth (Eds.), *Enriching history teaching and learning: Challenges, possibilities, practice* (pp. 15-32). Linköping: ISAK, Linköpings universitet.

Eklund, G. (2009). *Student teachers' views of the scientific thesis within teacher education in Finland.* Åbo: Faculty of Education, Abo Akademi University, Finland.

Feldt, R., Höst, M. & Lüders, F. (2009). Generic Skills in Software Engineering Master Thesis Projects: Towards Rubric-Based Evaluation. *Konferensbidrag i 22nd IEEE-CS Conference on Software Engineering Education and Training.*

Giota, J. (2010). Multidimensional and hierarchical assessment of adolescents' motivation in School. *Scandinavian Journal of Educational Research*, *54*(1), 83-97.

Hamilton, P., Johnson R. & Poudrier C. (2010). Measuring educational quality by appraising theses and dissertations: pitfalls and remedies. *Teachingin Higher Education,* 15(5), 567–577.

Hand, L. & Clewes, D. (2000). Marking the Difference: an investigation of the criteria used for assessing undergraduate dissertations in a business school. *Assessment & Evaluation in Higher Education*, *25*(1), 5–21.

Heinze, A. & Heinze, B. (2009). Blended e-learning skeleton of conversation: Improving formative assessment in undergraduate dissertation supervision. *British Journal of Educational Technology*, *40*(2), 294–305.

Holmberg, L. (2006). Coach, Consultant or Mother: Supervisors' views on quality in the supervision of bachelor theses. *Quality in Higher Education, 12*(2), 207–216.

Pathirage, C., Haigh, R., Amaratunga, D. & Baldry, D. (2007). Enhancing the quality and consistency of undergraduate dissertation assessment. A case study. *Quality Assurance in Education*, *15*, (3), 271–286.

Penny, A. J. & Grover, C. (1996). An analysis of student grade expectations and marker consistency. *Assessment & Evaluation in Higher Education, 21*(2), 173–183.

Poropat, A.E. (2009). A meta-analysis of the five-factor model of personality and academic performance', *Psychological bulletin*, 135(2), 322-338.

Svärd, O. (2013). *Examensarbetet i högre utbildning - En litteraturöversikt*. Uppsala: Uppsala universitet. (in Swedish)

Todd, M., Smith, K., & Bannister, P. (2006). Supervising a social science undergraduate dissertation: staff experiences and perceptions. *Teaching in Higher Education*, *11*,161-173.