At-distance college-level training in statistical methods: Prospects and considerations
Maria Meletiou-Mavrotheris, Efi Paparistodemou, Ana Serrado Bayes

To cite this version:
Maria Meletiou-Mavrotheris, Efi Paparistodemou, Ana Serrado Bayes. At-distance college-level training in statistical methods: Prospects and considerations. CERME 10, Feb 2017, Dublin, Ireland. hal-01927786

HAL Id: hal-01927786
https://hal.science/hal-01927786
Submitted on 20 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
At-distance college-level training in statistical methods: Prospects and considerations

Maria Meletiou-Mavrotheris¹, Efi Paparistodemou² and Ana Serrado Bayes³

¹European University Cyprus; m.mavrotheris@euc.ac.cy
²Cyprus Pedagogical Institute, e.paparistodemou@cytanet.com.cy
³La Salle-Buen Consejo, Spain; ana.serrado@gm.uca.es

This article focuses on how information and communication tools made available online could be effectively exploited to help improve the quality and efficiency of college-level, at-distance statistics training. The paper first provides an overview of the content, structure, and pedagogical and didactical approach underlying Quantitative Educational Research Methods, an online course targeting post-graduate education majors that has been built based on contemporary visions of web-based statistics instruction and computer-mediated communication. It then presents some of the insights gained from a case study of a group of students (n=49) participating in a recent offering of the course. The article concludes with some instructional and research implications.

Keywords: Statistics education, distance education, statistical inference, model-eliciting activities.

Introduction

The affordances offered by modern Internet technologies provide new opportunities for statistics instruction, making it possible to overcome restrictions of shrinking resources and geographical locations, and to offer, in a cost-effective and non-disruptive way, high-quality learning experiences to geographically dispersed students. In recent years, we have witnessed an exponential growth of distance education worldwide. Online course delivery has become common in a wide variety of disciplines, including statistics. This expansion is likely to continue, given the expanding access to the Internet and the greater emphasis to lifelong learning. Several advantages associated with distance education have been identified in the literature. In addition to the flexibility and convenience it offers, the distance option may also allow students the opportunity to take courses from established experts in their field of study that might not be available locally. From the viewpoint of statistics, it creates some unique opportunities for enhancing instruction, including the provision of a vast array of technological tools and resources for better understanding of statistical methods and concepts (e.g. interactive applets, virtual laboratory experiments, etc.). Several successful examples of successful programs of teaching statistics via distance have been documented in the literature (e.g., Evans et al., 2007; Everson & Garfield, 2008; Meletiou-Mavrotheris, Mavrotheris, & Paparistodemou, 2011).

Despite its proliferation and the unique opportunities for enhancing statistics teaching and learning that it offers, online statistics course delivery also presents several unique challenges. There are several pedagogical and technical issues that need to be incorporated into the design of an online statistics course for it to provide an effective learning environment. Review of the existing research literature alerts us to the fact that the quality and effectiveness of online statistics training currently offered is variable and inconsistent (Evans et al., 2007). While most of the conducted studies indicate that students taking courses with an online component have similar achievement and satisfaction levels compared to students in traditional, face-to-face classrooms (Mathieson, 2010), there is...
growing evidence of many web-based distance learning courses failing to meet the expectations raised. For example, while it is well-documented in statistics education research that the incorporation of discussion and active learning in the classroom can help learners to think and reason about statistical concepts, bringing these important learning approaches to an online course has proved very challenging (Gould, 2005; Meletiou-Mavrotheris & Serrado, 2012).

Early attempts at web-based instruction assumed that setting up an attractive website with interesting multimedia applications was adequate for learning to occur. However, it has now been recognized that the level of success of distance education is determined by multiple factors, including underlying theory, technologies, teaching strategies, and learner support. Elements in the design of an online course such as its content and structure, the tools and cognitive technologies employed, and the amount of interaction between instructors and learners as well among learners, are important factors affecting students’ learning and attitudes (Tudor, 2006). A particularly important criterion for the level of success of at-distance statistical training is also the extent to which instruction allows learners to experience the practice of statistics and to apply statistical tools in order to tackle real-life problems.

This article provides an overview of a post-graduate quantitative research methods course that has adopted a non-conventional approach, which promotes online participation and collaboration of students using contemporary technological and educational tools and resources. After describing the course pedagogical approach, and content and structure, the paper presents some of the insights gained from a case study of a group of students participating in a recent offering of the course.

**Nature of the Quantitative Educational Research Methods course**

“Quantitative Educational Research Methods” is a graduate-level course targeting students enrolled in the M.A. in Educational Studies program offered at European University Cyprus. Although originally developed for a face-to-face setting, the course was in 2013 redesigned as an online course to make it accessible to students enrolled in this program through distance education. In designing the e-learning course, efforts were made to preserve the pedagogical approach, and content and structure of the classroom-based course.

Contemporary visions of web-based instruction and computer-mediated communication underpin the course design. Concurring with Roseth, Garfield, and Ben-Zvi (2008), the online learning environment has been built upon the premise that instruction of statistical methods ought to resemble statistical practice, an inherently cooperative problem-solving enterprise. Students enrolled in the course are provided with ample opportunities for interactive and collaborative learning. They are actively involved in constructing their own knowledge, through participation in authentic educational activities encouraging enculturation such as projects, experiments, computer explorations with real and simulated data, group work and discussions. Statistical thinking is presented as a synthesis of statistical knowledge, context knowledge, and the information in the data in order to produce implications and insights, and to test and refine conjectures. There is a focus on modeling and simulation—along with inference—which is being facilitated by having students use the dynamical statistical software package TinkerPlots2 (Konold & Miller, 2011) for all modeling and analysis. This software was selected because it is designed explicitly to support integration of exploratory data analysis approaches and probabilistic models, and to allow for generation of data (e.g., drawing samples from a model) and experimentation (e.g., improving models, conducting simulations).
The course lasts 15 weeks. It is made up of 7 modules, that are concept driven and focused on enriching students’ knowledge of quantitative research methods (mainly inferential) by exposing them to innovative learning situations, technologies, and curricula. Each module involves a range of activities, readings, contributions to discussion, and the completion of group and/or individual assignments. The activities and assignments mirror those completed in the classroom-based course.

Throughout the course, students use TinkerPlots2 to work on a set of carefully designed open-ended Model-Eliciting Activities (MEAs) (Lesh et al., 2000) in which they create and test statistical models in order to solve real world problems of statistics (Garfield, delMas & Zieffler, 2012). The activities are carefully designed to support but, at the same time, also explore students’ evolving understandings of fundamental ideas related to statistical inference. Some of the MEAs are completed individually, and others collaboratively in groups of 3-4 students. The MEA “How many tickets to sell?” (adapted from http://new.censusatschool.org.nz/resource/using-tinkerplots-for-probability-modelling/) is a typical example of these activities. It is based on the following fictitious scenario: “Air Zland has found that on average 2.9% of the passengers that have booked tickets on its main domestic routes fail to show up for departure. It responds by overbooking flights. The Airbus A230, used on these routes, has 171 seats. How many extra tickets can Air Zland sell without upsetting passengers who do show up at the terminal too often?” In this MEA, students use Tinkerplots2 to model the Air Zland flight (e.g. model the scenario in which AirZland sells five extra tickets, i.e. books 176 tickets). They repeat the experiment a large number of trials using the “Collect Statistic” feature of Tinkerplots2 to keep track of the number of passengers not showing up, and draw the resulting distribution of collected sample statistics. Students then decide whether their model should be adjusted or not and, based on that, make recommendations to the airline as to how many extra tickets it should issue. Finally, they use the properties of the binomial distribution to determine theoretical probabilities when booking a certain number of seats (e.g. 176 seats) and compare the results with those they get through the Tinkerplots simulation. (see Meletiou-Mavrotheris et al. 2015 for more details).

A progressive formalization approach is being employed in the course. The first part focuses on building a teaching pathway towards formal inference by helping students experience and develop the ‘big ideas’ of informal inference. Through their engagement with the open-ended MEA activities, students learn where these ideas apply and how. Later in the course, students are introduced to confirmatory or formal inference methods, and begin comparing empirical probabilities with the theoretical ones. They learn the formal procedures for building sampling distributions, constructing confidence intervals, and conducting hypothesis testing using different statistical tests. The similarities and differences between ideal, mathematical models of reality, and statistical models based on experimental data are being emphasized throughout the course. From informal uses of models early in the course to formal uses as part of significance tests at the later part, instruction encourages explicit discussion of how every model is essentially an oversimplification of reality which involves loss of information, and of how the success of probability models depends on their practically and potential to give useful answers to our research questions.

The course is delivered completely online using the instructional content and services of the project platform (on the LMS Blackboard system). In addition to the course content (video lectures, PowerPoint presentations, video tutorials, links to statistics resources available on the internet, etc.) the site offers access to various tools for professional dialogue and support (email, videoconferencing, etc.)
chat rooms, discussion forums, wikis, etc.). The course instructor acts as a facilitator of a deeper learning experience through guiding discussions, encouraging full, thoughtful involvement of all participants, and providing feedback, in both asynchronous and synchronous activities.

**Methodology**

The case studied was a group of students taking the online version of the Quantitative Research Methods course during the Fall 2014 semester. The first author was the course instructor. There were forty-nine (n=49) students enrolled in the course, residing in Greece (n=38) or Cyprus (n=11). Course participants were characterized by a wide diversity in a number of parameters including age, and professional and academic background. Their age ranged from 23 to 55. Some had an academic background in primary education (n=18), while the rest were secondary school teachers in different domains (languages, humanities, natural sciences, physical sciences etc.). While the majority were experienced educators with several years of teaching experience, a sizeable proportion were either unemployed or employed in non-education related occupations. Students also had a varied background in statistics. Most of the older participants had very limited prior exposure to statistics, while the younger ones had typically completed a statistics course while at college. Even students who had formally studied statistics had attended traditional lecture-based courses that made minimal use of technology. Thus, upon entering the course, almost all students had very weak statistical reasoning and/or a tendency to focus on the procedural aspects of statistics.

Documenting online student activity and collaborative knowledge construction is a multifaceted phenomenon that requires complementary methods of data collection and analysis in order to understand how learning is accomplished through interaction, how learners engage in knowledge building, and how designed media support this accomplishment (Hmelo-Silver, 2003). Consequently, to increase understanding of the research setting, the current study employed a variety of both qualitative and quantitative data collection techniques, including: (i) The contents of the online discussion boards, chats, and wikis, in which students had been participating during the course; (ii) Bi-weekly collaborative assignments, in the form of Model Eliciting Activities (MEAs); (iii) videotaped synchronous sessions taking place weekly throughout the semester using Blackboard Collaborate as a communication tool; (iv) final course examination administered to both students enrolled in the course under study, and students enrolled in a face-to-face version of the same course again taught by the same instructor; (v) an open-ended survey administered at the course completion, aimed at determining students’ perceptions, opinions, and feelings regarding the course; (vi) Quantitative statistics automatically collected by the system (e.g. number of students participating in a discussion forum or successfully completing group assignments, etc.).

The text-based and video-based data collected during the course (MEAs, discussion forums, videotaped synchronous sessions, open-ended survey at course completion) were eventually analysed in order to examine how students’ engagement with Tinkerplots2, with MEAs and with each other impacted their motivation and participation levels, and how it scaffolded and extended their understanding of the big statistical ideas encountered during the course. We did not use an analytical framework with predetermined categories. What we instead did was a content analysis aimed at identifying, though careful reviewing of the transcripts, the recurring themes or patterns in the data. Quantitative data (system statistics, performance on final examination) were analysed using descriptive and inferential statistics. Linking the depth of qualitative data with quantitative breadth
provided a more holistic picture of the course impact on students’ attitudes and learning of statistics.

Results

Analysis of the data obtained during the case study, indicates that the online Quantitative Educational Research Methods course provided students with experiences parallel to those provided in its face-to-face version. The course was characterized by high levels of student engagement in online discussions and participation in videoconferencing sessions, and by successful collaborations for the completion of group assignments. Findings also suggest that the adoption of a pedagogical approach focused on modeling, using a dynamic statistics software like Tinkerplots2 for the conduct of statistical investigations, and of technological tools for facilitation of communication and collaboration among learners, is a viable option for online statistics instruction. The informal approach to statistical inference espoused by the course, using TinkerPlots2 as a tool for investigating authentic, open-ended model-eliciting activities (MEAs), fostered students’ ability to reason about the stochastic, while also developing their appreciation for the practical value of statistics. Through their engagement in MEAs in which they collaboratively built models and used them to evaluate research claims and hypotheses, the graduate students in our study developed relatively coherent understandings of fundamental concepts related to statistical inference.

The affordances offered by Tinkerplots2 for building and experimenting with data models to make sense of the situation at hand, proved instrumental in supporting student understanding of both informal and formal inferential statistical ideas. Of course, similarly to other researchers we also witnessed a number of challenges associated with the adoption of a modelling approach (Konold, Harradine, & Kazak, 2007), and different levels of student reasoning and understanding of the role of models and modelling, and of the key assumptions underlying the models simulated by the computer (for more details, interested readers could refer to Meletiou-Mavrotheris, Paparistodemou & Serrado, 2015). Nonetheless, use of Tinkerplots2 enabled students to build and modify their own models of real world phenomena, and to use them to informally test hypotheses and draw inferences. Their engagement with data-driven inferences helped them to develop sound informal understanding of the logic of hypothesis testing and its related statistical ideas (significance level, p-value, null and alternative hypothesis etc.), and served as a foundation for the formal study of inferential statistics.

Student performance on assignments and assessments was comparable to what was observed in the face-to-face version of the course concurrently taught by the first author. When the end of the semester, both groups of students were administered an identical assessment instrument (as a final exam) with several open-ended tasks aimed at investigating their understanding of the main ideas and concepts related to statistical inference covered in the course, both groups obtained very similar results (Mean Score: At distance=76.1, Face-to-face=77.06). A two-sample unequal variance t-test (conducted after checking all assumptions) indicated that there was no significant difference in mean scores ($p=0.73$) between the two groups of students (see Table 1).

<table>
<thead>
<tr>
<th>Course</th>
<th>No. of students</th>
<th>Mean Score</th>
<th>Standard deviation</th>
<th>t test for equality of means</th>
</tr>
</thead>
<tbody>
<tr>
<td>At-distance</td>
<td>49</td>
<td>76.10</td>
<td>14.68</td>
<td>-0.3463</td>
</tr>
<tr>
<td>Face-to-face</td>
<td>34</td>
<td>77.06</td>
<td>10.49</td>
<td>Two-tail Sig.</td>
</tr>
</tbody>
</table>

Table 1: Two-tailed t-test Comparison of Mean Differences in final exam scores

In the survey administered at the end of the course, students were asked to indicate what they liked
most about the course. The overall feedback regarding the course content, services, and didactical approaches was generally very positive. The flexibility and convenience associated with distance education was an aspect appreciated by all course participants, since it made it possible for them to determine their own place, pace and time of study. Another aspect also much appreciated by the majority of the participants was the fact that both the discussions and the assignments were carefully designed to be learner-centered, and to make explicit ties between theory and practice by utilizing students’ own experiences as learning resources. The promotion of communication and collaboration was also considered to be an important strength of the course by most learners. Students, in general, enjoyed the interaction and the sharing of experiences and ideas, although there were a few who expressed a preference for individual assignments, arguing that “group assignments are less flexible since you have to regularly meet online with your group”, or that “in group tasks, some members do minimal work while the rest work very hard, but the end they all get the same score… this is unfair.”

In a previous study conducted by the authors in the context of a transnational online teacher professional development course in statistics education, the biggest difficulty experienced was the limited success in establishing a functional online community of practice (Meletiou-Mavrotheris and Serrado, 2012). Similarly to other statistics education researchers (e.g. Gould and Peck, 2005), there was a much lower than anticipated level of learner-to-learner interaction in the course. Although community building was a main objective, and while at the course outset there was big enthusiasm and very high participation in discussion forums, interaction dropped off over time. The vast majority of messages (around 80%) had been sent during the first half of the course, while in the second half only a handful of learners actively participated in the discussion forums, while the rest had made minimal or no contributions. In the current study, by contrast, students’ level of engagement in the discussion forum was consistently high throughout the semester. All discussion forums created during the course were characterized by vibrant interaction and rich dialog.

We consider the active participation of students in the discussion forums witnessed throughout the semester to be an important success of this course since, as the literature indicates, leading a discussion of substance on a “discussion board” is much more challenging and difficult to achieve than in a real classroom (Gould & Peck, 2005). We believe that, in the current study, the adoption of the following strategies led to more successful community building compared to our prior research: (i) Making participation in group activities a compulsory element of the course that counts towards learners’ grade; (ii) Establishing a clear set of criteria in the course syllabus to help learners better understand the academic expectations and increase the intellectual depth of their contributions; (iii) Providing sufficient time for group members to make meaningful interpersonal connections before the assignment of the first cognitive task; (iv) Increasing the duration of each discussion forum to allow adequate time for learners to formulate and articulate their contributions; (v) Providing more prompt and effective moderation of online interactions.

Despite the overall success of the course, analysis of the collected data has allowed us to identify a number of issues and student concerns that adversely affected the online participation of course participants. The biggest shortcoming identified was the course overload. When requested, in the end-of-course survey, to indicate what they liked the least about the course, most participants mentioned the course workload that made it extremely difficult for them to keep up with the course requirements due to their overburdened schedules. Also, participation in videoconferencing and other activities that
required synchronous communication (e.g. chat sessions) proved very difficult to schedule, as it was almost impossible for all of the students to be available at the same time.

The Quantitative Educational Research Methods course team has adopted a continuous improvement iterative model. Insights from the current study informed the revision of the course, so as to further improve its quality and effectiveness. The heavy workload was corrected in subsequent offerings, and more realistic work expectations were set so as not to overburden students. There has also been more careful scheduling of course activities to offer students more flexibility.

Discussion

Teaching online courses is a new, unexplored territory for most statistics instructors. Distance education is similar yet different from classroom-based instruction, and requires new teaching skills and strategies. Several pedagogical and technical issues should be taken into account in the course design to provide an effective online learning environment. Using the case study of a distance-based approach to a quantitative educational research methods course as an example, the paper has provided some suggestions on how to best exploit the affordances offered by modern e-learning technologies to improve the quality and attractiveness of the online learning experience through the promotion of hands-on and collaborative knowledge construction. In accord with contemporary visions of web-based instruction that support collaborative and participatory models of online learning, the article has offered some insights on how to build an online learning environment in ways that resemble statistical practice, an inherently cooperative, problem-solving enterprise involving participation in projects, modelling and experimentation with real and simulated data, group work, and discussions.

Statistics education research in distance education settings is still at a developmental stage. More research is needed to advance our understanding of how to best take advantage of computer-mediated communication tools to support the development of effective virtual learning environments. By exploring the forms of collaboration and shared knowledge building undertaken by the group of students participating in our online course, the current case study has contributed some useful insights into the factors that may facilitate or impede the successful implementation of distance education. These insights have helped to further improve the quality and effectiveness of the course, and sketch a road map for our future research work, and for other similar endeavours.

References


