EXPERIENCES TEACHING AN INTRODUCTORY STATISTICS MOOC

Alison L. Gibbs
Department of Statistical Sciences, University of Toronto, Canada
alison.gibbs@utoronto.ca

Can a Massive Open Online Course (MOOC) meet the needs of learners interested in introductory statistics? In the spring of 2013, Statistics: Making Sense of Data, an 8-week, non-calculus, introductory statistics course was taught on the Coursera platform, with over 60,000 learners enrolled while the course was taking place. The course was developed in response to a call for MOOCs focused on high-enrolment, low-success, general education courses. We describe the course, and summarize what we have learned so far from the data we collected about the learners, including what we know about who enrolled, their initial intentions, and how these are related to indicators of success in the course.

INTRODUCTION

Massive Open Online Courses (MOOCs) have attracted a great deal of attention, particularly because of the massive number of learners enrolling. The website for Coursera, currently the largest platform for MOOCs, proclaims as its mission: "We believe in connecting people to a great education so that anyone around the world can learn without limits." Might an introductory statistics course serve a useful purpose in this massive, open environment?

In the spring of 2013, we taught an 8-week non-calculus introductory statistics course on Coursera. *Statistics: Making Sense of Data* attracted almost 40,000 learners when it began, with total enrolment over 60,000 by the end of the course. The course was developed with financial support from the Bill & Melinda Gates Foundation. The Foundation was interested in investigating whether the MOOC format, which had been dominated by advanced courses, could work for introductory courses, possibly creating learning opportunities for learners who had not yet experienced success in post-secondary education.

In this paper, we share some of our experiences with our MOOC, including a description of the structure of the course, the learners, and the factors that were associated with learner retention in the course. We also briefly discuss some of the open questions about learning through a MOOC.

A BRIEF HISTORY OF MOOCS

The continuing rapid evolution of the MOOC can be followed on Wikipedia (2014), and we give a brief summary here. The term Massive Open Online Course was introduced in 2008 to describe a course in connectivist knowledge taught for credit to 25 fee-paying students at the University of Manitoba and opened to learners worldwide, resulting in over 2,000 learners who participated without cost and without the ability to earn credit. MOOCs gained greater attention in the fall of 2011 with the launch of three Stanford courses in computer science with a total enrolment over 300,000. The Stanford MOOCs are examples of xMOOCs, to differentiate them from courses emphasizing connected learning (cMOOCs). xMOOCs are instructor-driven, with the focus on transmitting rather than constructing knowledge, and organized around instructor-led resources, reflecting the didactic methods often used in on-campus credit offerings. Key features of MOOCs include open access and scalability to any number of learners (Yuan & Powell, 2013).

Dennen & Chauhan (2013) present a Strengths, Weaknesses, Opportunities, and Threats analysis of the decision to offer a MOOC. They identified institutional-level opportunities, which they summarized as "reputation, recruitment and research." These reflect the motivation of the University of Toronto to become involved with MOOCs. In the summer of 2012, the University of Toronto announced a partnership with Coursera ("Online courses for anyone, anywhere", 2012). The university's interest was sharing its expertise through open educational resources, showcasing itself as a leading institution for education and research, leveraging open educational resources into its on-campus offerings, and leading research in the evaluation of online platforms and pedagogical approaches (Harrison, 2013).

In a short time, the development of MOOCs has included their own literature (see, for example, Liyanagunawardena et al. (2013) for a review) and a dedicated journal (see *MOOCs*

Forum). The hype around MOOCs grew quickly, with the *New York Times* declaring 2012 the "Year of the MOOC" (Pappano, 2012) and debates of lofty claims about their potential influence, such as "MOOCs will democratize education around the globe" (Sonwalker et al., 2013).

Whether or not MOOCs will be a truly disruptive force in education remains to be seen. Many instructors are experimenting with blended learning involving wrapping a MOOC within on-campus courses (Griffiths, 2013, Flipped Classroom Field Guide, 2014), some using their own MOOCs, others using MOOCS created by other instructors. MOOCs provide learners from around the globe with exciting opportunities, such as the chance to engage with instructors from leading universities, to develop collaborative study networks across borders, to access materials for self-study, or to choose a combination of instructor-led courses, peer learning, or self-study that best suits them. With the advent of MOOCs, we can begin to envision a future post-secondary education that is flexible according to individual learners desires and needs (Irvine et al., 2013). However, MOOCs have many detractors. Criticisms of MOOCs include "a loss of personal interaction and educational quality to concerns that MOOCs may signal an undesirable turn toward corporatized learning" (Dennen & Chauhan, 2013).

THE DEVELOPMENT OF *STATISTICS: MAKING SENSE OF DATA*

The development of *Statistics: Making Sense of Data* was funded in part by the Bill & Melinda Gates Foundation. The request for proposals (RFP) from the Gates Foundation called for MOOCs in general and developmental education courses (Bill & Melinda Gates Foundation, 2012). The Foundation was interested in courses "likely to provide value to a broader range of learners than is currently served by existing MOOCs," in particular "high-enrollment, low-success introductory level course[s] that [are] a barrier to success for many students." The Foundation was also interested in learning for which disciplines and types of learners the MOOC format may be more successful, and what data are necessary to best support improvements in learning. And the RFP expressed interest in how institutions might blend the MOOC content into formal on-campus courses.

At the time of the Gates Foundation RFP, the University of Toronto Department of Statistical Sciences was planning an alternative method of teaching our large introductory, non-calculus statistics course. These plans included the creation of reusable, modular, online content that could be used by different instructors in flexible ways, mainly in an inverted classroom environment. The Gates Foundation funding provided us with the opportunity to explore the MOOC learning environment, while preparing material for our on-campus course.

Like most MOOCs, our MOOC was an xMOOC. Learners could not gain credit at any university by completing it, but they could earn a "Statement of Accomplishment."

Our goals when creating the course material included:
- Make use of interesting data with which we had personal connections,
- Keep the lecture videos dynamic,
- Create content that can be repurposed,
- Have no barriers to earning a Statement of Accomplishment for learners who engage in the course material for the entire period of the course and complete all evaluations, and
- In the spirit of open resources, use R for computation. We provided instruction in R, but did not require demonstration of proficiency to earn a Statement of Accomplishment.

Because a partnership between the University of Toronto and Coursera was already in place, the course was delivered on the Coursera platform. See Yuan & Powell (2013) for a description of many of the MOOC platforms. Courses on the Coursera platform follow a specific pedagogical framework, based on Coursera research into online learning ("About Coursera", 2014). The format is designed to allow learners to learn at their own pace, with emphasis on mastery learning, and to provide frequent evaluation and feedback so that learners can monitor their own learning. Coursera provides tools for mass grading, including machine grading and peer-assessment, allowing the course to scale to any number of learners. The following features of the Coursera platform were used in the design of *Statistics: Making Sense of Data:*
- *Short videos.* Although some of our videos were longer, Coursera recommends videos of length 8-12 minutes. As a result, it was necessary to re-imagine our traditional lectures into small

chunks, focused on one main concept per chunk. Coursera provided subtitles and transcripts for each video. Learners could interact with the videos at a pace appropriate to them, adjusting the speed and rewinding as necessary.

- *In-video quizzes*. At several points, a video would stop and learners would be presented with a question about the material just covered. These quiz questions did not count as part of the evaluation, but learners were provided with immediate feedback and discussion of the correct response and common misconceptions. These questions were designed to keep learners engaged and focused, and they are thought to improve retention by forcing students to retrieve and think critically about what they have just heard.

- *Weekly machine-graded quizzes, emphasizing mastery learning*. For these quizzes, learners were allowed unlimited attempts and for each attempt the questions were randomly selected from a large pool. On completing a quiz, learners were given immediate feedback with their score and where they want wrong. They could then take remedial action on the concepts they did not understand, seeking help from the videos, discussion forums, or any other external source.

- *Peer-assessed assignments*. The heart of the course evaluation resided in two lengthy assignments. For each assignment, learners were given a scenario and computer output, including summary statistics and plots, and were asked to describe and interpret the findings, including the strength and generalizability of the conclusions they could make. Each learner then provided feedback on the responses of five of their peers. These assignments required a significant time commitment from the learners, and about 75% of learners who completed the quiz for the week corresponding to each assignment also completed the assignment.

- *Community building and opportunities for peer engagement*. Participation in a discussion group was not a required component of the course. Students were encouraged to discuss the course on the forums, share resources on the course wiki, and organize local meet-ups with learners in their geographic area. As an example of the online community that developed, learners who desired slides of lectures (even though we did not teach from prepared slides) crowd-sourced the creation of summary documents consisting of screen captures of the videos. This collaboration was organized and shared through the course wiki.

A PORTRAIT OF THE LEARNERS: ENROLMENT, COMPLETION, AND PERSISTENCE

Enrollment

The hype around MOOCs is typically associated with the massive number of learners who enroll in the courses. As a MOOC instructor, I am regularly asked "How many students were in the course?"

At the time that *Statistics: Making Sense of Data* went live, almost 40,000 learners had enrolled. Of these, approximately 17,500 learners completed a pre-course survey, intended to capture demographics and intent. Of those who responded:

- 62% were male.
- The average age was about 35 years old.
- 88% had already completed a university or college degree with 46% holding a post-secondary degree. Although it is well-known that MOOC learners are extremely well-educated, the proportion of people with post-secondary degrees who enrolled in *Statistics: Making Sense of Data* was more than 10% larger than in other University of Toronto MOOCs (Demographic Report, 2013).
- Although the MOOC covered material from an introductory statistics course, assuming no background in statistics, 74% had learned statistics previously, more than half of these with experience equivalent to more than one course.
- 36% were from North America, 28% were from Europe, and 21% were from Asia.
- English was the first language of 37%.
- 77% of respondents planned to complete the requirements for a Statement of Accomplishment.

Completion and Persistence

MOOC learners are self-selected to be motivated learners, passionate about the opportunity for open access to knowledge and to experts from leading post-secondary institutions. Yet retention and persistence with the material are typically low. (See, for example, Jordan (2013) for examples of course completion rates.) After asking about the number of learners enrolled, the second question people typically ask about my MOOC experience is "And how many finished?" In the period from when the course became available for enrolment to the course completion, approximately 60,000 learners enrolled, with just under 3,000 earning a Statement of Accomplishment, a "completion rate" of approximately 5%. This is typical of xMOOCs, particularly those requiring a peer-assessed assignment which tend to have lower completion rates.

However, 5% does not reflect the true attrition in the course. Learners could enroll in the course at any time, including while the course was in session, and many of the 60,000 enrolled arrived to the course too late to meet the necessary deadlines for the requirements to earn a Statement of Accomplishment. Enrolment in the course had no cost or risk, and less than half of the learners who enrolled actually participated in the course. Moreover, learners enrolled with various intents, and earning a Statement of Accomplishment, which has no recognized value as a credential, may not be a meaningful measure of the success of the course in meeting learners' needs. If, instead, we consider persistence with the course materials as a measure of success, the attrition, while still large, is less alarming. Some statistics about persistence:

- Of the 16,891 people who completed the week 1 quiz, 4,437 (26%) completed the final quiz.
- Of the 19,757 people who watched or downloaded at least one video during the first week of the course, 7,794 (39%) watched or downloaded a video during the final week.

These statistics are typical of MOOCs on Coursera (see Koller et al., 2013) and, although these completion rates would not be acceptable in our for-credit courses, they are not necessarily reflective of an inherent problem with MOOCs. Koller et al. (2013) compare a MOOC to a book in a public library. Would we claim a book was a success if and only if a large proportion of borrowers read it cover to cover by the due date?

In order to better understand the characteristics of learners who completed the course, we examined factors associated with earning a Statement of Accomplishment and with persisting with quizzes or videos throughout the course. The learner characteristics were collected from a pre-course survey and a pre-course background knowledge quiz. The analysis we report here only includes learners who completed the survey and quiz.

Logistic regression was used to examine the factors associated with earning a Statement of Accomplishment. For persistence with quizzes and videos, we considered only the learners who watched a video or attempted the quiz in the first week of the course, and considered the time until they were no longer active in that aspect of the course, that is, the time until they stopped watching videos or attempting quizzes. Cox-proportional hazards regression was used to examine the factors associated with dropping out. Because this analysis is exploratory, and we are have a very large sample size (as is inherent with MOOC enrollments) providing sufficient power to detect effects that are not practically important, we only report here on results for which the P-value indicating a association is < 0.001. No significant interactions were found. Table 1 gives the results of this analysis.

Most of the results shown in Table 1 are not surprising. Learners with better background knowledge, higher education level, and those who intended to work towards a Statement of Accomplishment and spend more time on the course were more likely to earn a Statement of Accomplishment and to persist with the course. Older learners, learners from Australia and Europe, and learners whose first language is not English, were also more likely to persist with the course. Sex was not a significant factor associated with completing the evaluations, but females were more likely than males to stop watching the videos.

Table 1: Factors associated with the odds of earning a Statement of Accomplishment (SoA) (logistic regression), or persisting with quizzes and videos over the weeks of the course (Cox proportional hazards regression). Results are reported for associations with P-value < 0.001.

Predictor	Statement of Accomplishment (SoA)	Persistence with quizzes	Persistence with videos
Ability: Score on background knowledge quiz	Learners with better scores were more likely to earn SoA (OR=1.15)	Hazard decreases by 7% with increase of 10% in score	Hazard decreases by 5% with increase of 10% in score
Demographic: Age in years	Older learns were more likely to earn SoA (OR=1.02)	Hazard decreases by 0.5% with increase of one year in age	Hazard decreases by 1% with increase of one year in age
Region	19% of Australians and Europeans earn SoA; 14% of learners from other regions earn SoA	Australians and Europeans have lower hazard functions than other regions; Asians have the highest	Australians and Europeans have lower hazard functions than other regions; Asians and North Americans have the highest
Education level	College graduates were more likely to earn SoA (OR=1.61)	Hazard decreases by 19% for college graduates	Hazard decreases by 22% for college graduates
English	Learners whose first language was not English were more likely to earn SoA (OR=1.23)	Hazard increases by 13% for learners whose first language is English	Hazard increases by 11% for learners whose first language is English
Sex			Hazard decreases by 7% for males
Intent: Planned to work towards SoA	18% of learners who planned to earn an SoA did; 9% of learners who did not plan to earn an SoA, did earn one	Hazard decreases by 31% for those planning to work towards an SoA	Hazard decreases by 19% for those planning to work towards an SoA
Intended time to spend on course in hours / week	Learners intending to spend more time were more likely to earn SoA (OR=1.08)	Hazard decreases by 4% with increase of one hour	Hazard decreases by 3% with increase of one hour

DISCUSSION

Our experience teaching a MOOC has given us a glimpse into the global demand for knowledge of statistics, including some indications of the characteristics of motivated learners and a profile of those who are more likely to follow a course to completion. However, truly understanding the learners is challenging.

While there is no scarcity of data (every click for ever learner is available), accessing the data requires a great deal of processing. Learners have a variety of motivations for enrolling in the course and we continue to work on defining and understanding their patterns of activity. Moreover, accessing all learners so that their progress through the course can be monitored is challenging. Coursera collects no demographic information on its learners. It is particularly challenging to understand the reasons that learners disengage with the course. There is no risk to staying enrolled in a course after it has met or failed to meet a learner's needs, so capturing the timing and the events that might trigger attrition is difficult. As Breslow et al. (2013) noted, "If educational researchers studying conventional brick and mortar classrooms struggle to operationalize variables like attrition and achievement, it is doubly difficult to do so for MOOCs." With the ease of enrolment, no need to withdraw, and learners who are extremely diverse in their motivations and intents, both participation and success bear little resemblance to what we might consider them to be in our on-campus classrooms. What is clear from our experience is that there is a huge demand,

from a well-educated group of learners, for learning introductory statistics in an open environment. And this demand exists even without the opportunity to earn a meaningful credential.

There is still a great deal of interesting work to be done in developing best practices for teaching in this environment. The Coursera platform offers many excellent pedagogical innovations that encourage mastery learning and constant opportunities for formative evaluation, and the platform's development continues at a rapid pace. There is not yet the capacity for building a course with fully adaptive learning. And while the intent of MOOCs includes the open sharing of educational resources, xMOOCs typically involve one-way transmission of knowledge, and there may be benefits to be gained through increased use of social networking for peer-to-peer learning and evolution of learning resources.

The discussion about MOOCs is often focused on the massive number of learners who enroll, and the less massive, though still impressive, number of learners who earn Statements of Accomplishment. Perhaps it is time to give more attention to the question: "But did they learn?"

REFERENCES

About Coursera® (2014). https://www.coursera.org/about

Bill & Melinda Gates Foundation (2012). Request for Proposals. https://www.qualitymatters.org/files/MOOC%20RFP%20and%20Application.pdf

Breslow, L., Pritchard, D., DeBoer, J., Stump, G., Ho, A. & Seaton, D. (2013). Studying Learning in the Worldwide Classroom: Research into edX's First MOOC. *Research & Practice in Assessment 8*, 1325.

Demographic Report on University of Toronto Coursera MOOCs (2013, June 16). http://www.ocw.utoronto.ca/demographic-reports

Dennen, V.P., & Chauhan, A. (2013). Shall we MOOC? A SWOT Analysis at the Program Level. *MOOCs Forum, 1*(P), 17-21.

Flipped Classroom Field Guide (2014). https://class.coursera.org/mooc/wiki/FlippedLectures

Griffiths, R. (2013). *MOOCs in the Classroom*. Report from Ithaka S+R. http://www.sr.ithaka.org/blog-individual/moocs-classroom

Harrison, L. (2013, August). Open UToronto MOOC Initiative: Report on First Year of Activity. http://www.ocw.utoronto.ca/open-utoronto-mooc-initiative

Irvine, V., Code, J., & Richards, L. (2013). Realigning Higher Education for the 21st-Century Learner through Multi-Access Learning. *MERLOT Journal of Online Learning and Teaching 9*(2), 172-186.

Jordan, K. (2013). MOOC Completion Rates: The Data. http://www.katyjordan.com/MOOCproject.html

Koller, D., Ng, A., Do, C., & Chen, Z. (2013, June 3). Retention and Intention in Massive Open Online Courses: In Depth. *EDUCAUSE Review Online*. http://www.educause.edu/ero/article/retention-and-intention-massive-open-online-courses-depth-0

Liyanagunawardena, T., Adams, A., & Williams, S. (2013). MOOCs: A Systematic Study of the Published Literature 2008-2012. *The International Review of Research in Open and Distance Learning, 14*(3). http://www.irrodl.org/index.php/irrodl/article/view/1455/2531

MOOCs Forum (2013). http://www.liebertpub.com/overview/moocs-forum/619

Online courses for anyone, anywhere (2012, July 17). *U of T News*. http://www.news.utoronto.ca/online-courses-anyone-anywhere

Pappano, L. (2012, November 2). The Year of the MOOC. *The New York Times*. http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html

Sonwalkar, N., Wilson, J., Ng, A., & Sloep, P. (2013). State-of-the-Field Discussion. *MOOCs Forum 1*(P), 609. http://online.liebertpub.com/doi/full/10.1089/mooc.2013.0006

Wikipedia (2014). Massive Open Online Course. http://en.wikipedia.org/wiki/Massive_open_online_course

Yuan, L., & Powell, S. (2013). Open Education: Implications for Higher Education, A White Paper. *Centre for Education Technology, Interoperability and Standards*. http://publications.cetis.ac.uk/2013/667