13th International Congress on Mathematical Education
Hamburg, 24-31 July 2016

DESIGN OF A COURSE FOR LEARNING PROBABILITY VIA SIMULATIONS WITH TINKERPLOTS

Susanne Podworny

University of Paderborn, Germany

This paper describes the development of a one semester course for preservice teachers at university level, which is embedded in a larger study. The course is on deepening the understanding of probability and ranges from basic simulations to the accuracy of simulations to independence and dependence and ends with informal hypothesis testing via p-values. TinkerPlots is used as simulation software during the whole course. In this paper some design ideas for the course are presented. At the end there is some insight via short excerpts of students' reflection of the course.

Introduction

Simulations can enrich the teaching and learning of probability, as was determined over 25 years ago (e.g. Gnanadesikan, Schaeffer, & Swift, 1987; Biehler, 1991). Recent research confirms the benefit of simulation in teaching probability in school and at university level (Batanero, Biehler, Maxara, Engel & Vogel, 2005; Konold, 2007). In Germany, simulations are included in the national educational standards for mathematics, and the German special working group of stochastic strongly recommends the use of computers for data analysis and probability (Arbeitskreis Stochastik, 2003), so it is necessary to prepare future teachers to teach these topics. For this reason we developed a one semester course on simulation as a follow-up of an introductory course on statistics. In the introductory course, the preservice teachers learned the basics of statistics and probability. The introductory course is compulsory and the newly developed course titled "Applied stochastics – Understanding and solving complex problems via simulations with TinkerPlots" is an elective course. Learning to simulate chance experiments, deepening the understanding of probability and inferential reasoning (Cobb, 2007) are the main goals of this course.

Course design

The course design was based on some ideas of two existing courses that were developed in the working group of Rolf Biehler; both used Fathom for simulations (Meyfarth, 2008; Prömmel, 2013) and with ideas from the CATALST curriculum (Zieffler, 2013). The designed course was for preservice teachers of primary and lower secondary school. These students are obliged to attend an introductory course "Elementary statistics and probability" at the University of Paderborn. After this introductory course, they are free to choose an elective course to deepen some of the topics of the introductory course, though they are not obligated to deepen their statistics knowledge. At this point our course "Applied stochastics" comes in with the focus on simulating chance experiments and culminating in inferential reasoning, which is not a topic of the introductory course. Passing the exam of the introductory course is the precondition for attending the course "Applied stochastics". This means that students entering the new course are experienced with Fathom for data analysis and with basic ideas on probability.

The design of the course follows the six principles of the "Statistical Reasoning Learning Environment" (SRLE) by Garfield & Ben-Zvi (2008) and includes some aspects of the model-eliciting activity (MEA) approach by Lesh, Hoover, Hole, Kelly, & Post (2000) and Lesh & Doerr (2003). The combination of materials, class activities, discussions, technology, teaching approach and assessment is taken from the SRLE approach. The course is student-centered and not teacher-centered. By working in pairs of two or three on tasks and activities, rather than listening to a teachers' presentation, the development of statistical ideas and concepts is the focus. One major goal is to enable participants to view simulations as one solution to solving probability problems and to be aware of the limitations that come along with simulations. The final aim is to introduce students to inferential reasoning. Monitoring of the learning process happens with weekly homework and a cumulative portfolio with regular reflections. TinkerPlots is used as simulation software, because it is adequate software for preservice teachers for primary and secondary teaching, and it is quite easy to use. As a type of "expressive tool" (Hoyles & Noss, 2003) for "expressive modeling" (Doerr & Pratt, 2008) the simulation process in TinkerPlots is visual and not based on commands, which simplifies the procedure. New learning goals related to the software are presented briefly at the beginning of a lesson and afterwards students work in groups of two or three on selected statistical topics. Each topic deals with one major statistical concept; within the activities the focus is on understanding rather than the tool use. "Real and motivating data" (Garfield & Ben-Zvi, 2008, p. 50) and a variety of contexts are used for the activities. This is combined with the model-eliciting activity approach (MEA) as proposed by Lesh, et al. (2000) for mathematical team problem-solving. MEAs offer a variety of learning, because "model-eliciting activities usually involve mathematizing – by quantifying, dimensionalizing, coordinatizing, categorizing, algebratizing, and systematizing relevant objects, relationships, actions, patterns, and regularities" (Lesh & Doerr, 2003, p. 5). Some of the problems are designed to be worked on over one or two class periods. Due to their social aspects, MEAs are trend-setting, because "Model-eliciting activities have embedded in the social dimensions that reflect the teamwork and communication required in designing and using new technologies" (Zawojewski, Lesh, & English, 2003, p. 338). According to the principles, one main aspect is to construct models appropriate to the problem with the help of TinkerPlots sampler. These models should reproduce reality and enable users to make predictions or identify trends or patterns. In MEA the context of the activity plays an important role, as this is what makes problems meaningful and relevant to the students. Some (fictitious) audience as addressee of the solution helps to state the purpose and the need of a solution. Self-assessment should be conducted during the working process so that students are able to select, refine, and elaborate their models meaningfully. A "client" who is making requests may motivate the self-assessment and refinement in MEA. Another important aspect is the documentation. The solution should integrate the thinking process, beginning with assumptions, goals and a solution path. Cooperative learning can help facilitate this and is highly recommended during the course. A general way of thinking should be represented by the model, and this ensures that students' models can be used by others too. A model should be as simple as possible, so that it represents kind of a prototype for other problems with the same structure. A problem-solving iteration of MEA is like a cycle with several runs. This cycle starts with a *description and modeling* of a real-world problem, a *manipulation* of the model, *translation* of relevant results back to the real world, and finally *verification* of the usefulness of solutions completes the cycle (Lesh & Doerr, 2003, p. 17). Revising a model is an important part of MEAs, and TinkerPlots offers very good possibilities for changing models. For the design of the course, discussions in teams and whole-class discussions are as important as weekly homework, and they are used as an assessment "to monitor the development of their statistical learning as well as to evaluate instructional plans and progress" (Garfield & Ben-Zvi, 2008, p. 48). As supporting material we developed a simulation scheme for simulations with TinkerPlots (Podworny & Biehler, 2014) and used it throughout the course. The simulations scheme is a combination of TinkerPlots screenshots and text fields to explain several steps in the simulation process. It may be used for planning, structuring or documenting a simulation or as a worked example. All tasks and additional reflections on each topic are collected in portfolios by every participant. For research purposes, all files of students (tp-files, reports, simulation schemes and portfolios) are collected and many activities of students working with TinkerPlots are recorded with a screen-capture software. The portfolios are used to monitor students' understanding and to evaluate material and tasks.

Topics of the course

The course covers five topics of different lengths for a total of 15 sessions (90 minutes each). The first and the last sessions are reserved for organizational details and a pre- and posttest. The topics are: (1) Data analysis with TinkerPlots; (2) Basic simulations with TinkerPlots; (3) Accuracy of simulations; (4) Independence and dependence; (5) Hypothesis testing with *p*-values and randomization tests. Please find in the following a short description of the topics.

(1) The first topic on data analysis covers one session. The software is introduced and connected with the participant's previous knowledge from the introductory course.

(2) The second topic on basic simulations, which is a major topic of the course, covers four sessions. There are two goals for this topic. First, participants learn technical aspects of simulations with TinkerPlots and, more importantly, they are introduced to inferential reasoning to estimate probabilities of an event via simulation. Then the six steps of a simulation process according to Maxara (2009) in combination with the MEA-cycle mentioned above are discussed and implemented. The steps are: Step M – Modeling a real situation, step 1 – Choosing the model of the chance experiment, step 2 – Defining the interesting events or random variables, step 3 – Realization and repetition of the simulation, step 4 – Data analysis, step I – Interpretation of the data. Step 1-4 are software related and represented in the TinkerPlots simulation scheme (Podworny & Biehler, 2014), which is used since the third session. To face the question of how many repetitions of a simulation are required, we establish a "rule of thumb" concerning the accuracy of simulations in the second session. This is the distance between the probability and the relative frequency for several numbers of repetitions (for n=100, 1000, 5000, 10000). Several tasks from different contexts are included in this topic, each of them with the focus on a special feature of TinkerPlots and the different modeling possibilities offered by TinkerPlots. At the end of this topic participants should know different options for setting up a model in the sampler, defining random variables or chance events by a formula in the table or by plotting the results, analyzing the data and interpreting the outcomes. The first task for the autonomous work of the participants is the following: *Some pupils want to build a wheel of fortune booth on Christmas market. They want to earn 50 cents per game on average. The wheel allows the following outcomes: in 1 % you get 50 Euro; in 45 % you get 1 Euro; in 54 % you get nothing. Rule: The visitor may turn the wheel three times in one game and receives the sum as*

benefit. Question 1: What should be the price for one game in order for the pupils to make 50 cents on average? Question 2: What is the probability for a random visitor to win nothing? With some homework and some more subtasks, this problem is used to introduce many aspects of simulations with TinkerPlots.

(3) A hands-on activity on throwing coins starts the third topic on accuracy of simulations. It lasts for two sessions and deals with the empirical law of large numbers and ends with the $1/\sqrt{n}$ – law. Important aspects of this law are discussed and are used from this point on in every task. The whole topic is a more formal one and focusses mainly on throwing coins. Only the last task of this topic is situated in a special context again. It is about overbooked planes and the probability of overbooking depending on the number of seats in different sizes of planes. The knowledge gained in these two sessions facilitates more accurate answers on problems solved via simulations.

(4) The fourth topic is about independence and dependence and their modeling in TinkerPlots. During these sessions, TinkerPlots is introduced as a data factory (Konold, et al., 2007). Two paths are discussed in two sessions: from model to data and from data to model. In the approach from model to data, TinkerPlots is used to make predictions in a conditional probability taskdesign. The second direction is led by the question 'which model fits best to given data?' There is room for plenty discussions in the small groups and in the plenum during this topic.

(5) The last topic of the course is on hypothesis testing with *p*-values, which introduces first ideas of inferential reasoning in two sessions and results in randomization tests in another two sessions. Together with the second topic this is the main focus of the course. The starting point is a model-eliciting activity about a fictitious person who identified the quality of music (mp3 or CD quality) and scored 10 out of 12 qualities. The question is whether this person has the ability of hearing differences in quality of music or if she is just guessing. This task takes a whole session with several subtasks and discussions and results in deepening questions according to the ideas of error type I and II in a second session. For supporting the inferential reasoning the following steps are suggested: *note an observation, choose a null hypothesis, simulate under this null hypothesis, identify the test statistics, estimate the p-value, interpret the results* with regard to Cobb's three 'Rs': randomize, repeat, reject (Cobb, 2007). For setting up the null hypothesis we focus on a "just guessing" or "randomly chosen" vocabulary. For the last topic of randomization tests, TinkerPlots shows its great capacity by visualizing the random assignment process between two groups.

Students' reflections in their portfolios

There were twenty-five students participating in the course, twenty-one of them are female and four are male. Twenty students are at university to become primary school teachers and five want to become teachers at lower secondary schools, which in Germany range from grade five to grade ten. This is a typical audience for a course like this, because there are many preservice primary teachers at the University of Paderborn and a very large percentage of them are female. All of the preservice primary teachers get the teacher license up to grade ten and therefore should be introduced into probability too, especially since every teacher should know more mathematical background than just what he will be teaching.

Students have to choose one course during their studies at university that deepens any mathematical topic. Because there is limited space in these advanced courses they may not get into the course they want, but into another course with free space. So not everybody is in the course because of interest, but because of free space. All of the participants passed the compulsory introductory course, but this occurred one or two years ago, if not earlier. The prior knowledge of the participants is rather mixed, if the written test at the end of the intro course is taken into account (see Table 1).

	Exam passed in 2010 or 2011
Less than 50 % of points	11
50 % of points and more	15

Table 1: Participants' exam results of the intro course as evidence for prior knowledge

Four of them who had failed in 2010 succeeded in 2011. Only one participant scored more than 70%. Three of the participants attended an earlier course than 2010 and I don't have results for them.

The course is divided into five topics (see above) and for each topic the participants write a reflection on their individual learning success, their difficulties and suggestions for improvement. The reflections will be used later for evaluation and redesign of the course. To gain insight into students' thinking I present some excerpts of different reflections. Every reflection on each topic was at least one page long, but rather mixed in content. Some students reported single tasks, some summarized results and some reflected their own learning process. Very few difficulties were written down at all. The selection of the excerpts is arbitrary, due to the amount of space the leading principle is the idea of "interesting or representative" episodes (Voigt, 1984).

Excerpts related to topic I, data analysis with TinkerPlots (one session):

> "[…] The software could be discovered through the questions and many functions could be used within the plot. […] The homework related to this topic trained the use of the software additionally. Starting difficulties like 'what happens, if I push this button' or 'how can I change the size of a dot in the plot' could be overcome by trial and error. With the tasks and the homework there were no new teaching contents [they were known from the intro course], but I got to know the software well." (Melissa, in her portfolio)

> "I learned in the first topic how to plot data. If there are data of people with their sex, size, weight and so on, I am able to compare the distribution of the variable size distinguished by the variable sex or to prove if there is a relationship between size and weight." (Anita, in her portfolio)

Like both excerpts related to topic I show, participants could refresh their knowledge on data analysis in the first session and got to know the software TinkerPlots for data analysis. Some did not like the homework, where they had to create twenty graphs to get used to the capacity of the plot, some others reported that they liked this homework, but would have liked to discuss their results in small groups, so this is one minor aspect for redesign on the first topic.

Excerpts related to topic II, basic simulations with TinkerPlots (four sessions):

> "Altogether, I liked this topic. Especially task "Tegut" [German supermarket] and "cereal bar" [both tasks are on trading stickers as kind of a complete series problem], because they had an immediate reference to everyday life. I rate it as very valuable, that we were engaged with problems of our everyday life and which are very interesting." (Rebecca, in her portfolio)

> "I like it very much, that it is possible to simulate every step in TinkerPlots as it is in my mind. This was very helpful for task 'sea voyage'." (Sara, in her portfolio)

Topic II was a major topic of the course, because all basics were taught in these four sessions for using TinkerPlots as an instrument (Rabardel, 2002). Many students reported that they liked the orchestration of "work and walk-by" of the teacher (Drijvers, 2012), so this is one aspect that will be kept in the next iteration. The selection of problems seemed to be interesting for students, at some point nearly everybody reported a task that related to his everyday life, like Rebecca did in her reflection. The simulation capacity, that was analyzed by the author in detail before the course, was represented well in these four sessions but not as detailed as it could be, and this will be part of a redesign, although there were students like Sara, who noticed TinkerPlots as a good expressive medium. But this aspect should be more emphasized in this topic in the next iteration.

Excerpts related to topic III, accuracy on simulations and $1/\sqrt{n}$ – law (two sessions):

> "For me the benefit of our computations [in topic III] was not clear enough. The conclusion of the formula [$1/\sqrt{n}$] was not emphasized enough. Between the 95% interval and the number of repetitions is an inverse proportional relation. To quadruplicate the number of repetitions leads to the bisection of the width of the 95% interval." (Thomas, in his portfolio)

> "This topic was understandable, but with regards to content very extensive. It would have been nice to have more time for deepening this topic." (Anastasia, in her portfolio)

This topic was a more mathematical one and seemed to be quite hard to understand for the students. Like Anastasia, many students expressed a wish for more time with this topic. Together with the next topic, a redesign should be to lengthen this topic.

Excerpts related to topic IV, modeling independence and dependence (two sessions):

> "While working with TinkerPlots I learned in topic IV to realize dependent and independent attributes via different arrangements of the devices in the sampler. I like the visualization of the ramification in the sampler and that I can see how the simulation works. I can prove my hypothesis on a model. I learned to design and to evaluate models, and can view them as pictures of reality." (Kerstin, in her portfolio)

In this topic, students learned to create "factories" (Konold, et al., 2007) with TinkerPlots and most of them appreciated this. For evaluating their models, they were introduced to *p*-values; this seemed not to be an appropriate way of starting hypothesis testing. For a redesign, it must be carefully decided which parts of this topic will be kept and which parts will be integrated in the last topic. Because of the great acceptance of this topic it will be redesigned, rather than erased as was intended after first teaching the topic.

Excerpts related to topic V, hypothesis testing (four sessions):

> "Looking at all topics, I think the last one is the most appealing. There were always elements of everyday life which evoked interest. I could focus on the 'essentials'." (Stella, in her portfolio)

> "Topic V brings the whole course together in a unit, because everything that was taught in topics before was then in topic V and has been deepened somehow. The most important thing was the inferential reasoning at the end of the course. What can I do with simulations and statistics if I can't interpret them or can't draw inferences? If I can't do this, dealing with stochastics is useless. I am glad that we had four weeks for this topic, so we could practice inferential reasoning very often. […]. At some point, my partner and I were at odds about interpreting *p*-values at medium level. Perhaps the interpretation of medium *p*-values can be discussed in more details. Or state clearly that it is not important how to interpret medium *p*-values as long as the interpretation is comprehensible. For me, I decided on the background of the problem how to decide with a medium *p*-value, sometimes like this, sometimes like that. Perhaps this was correct, perhaps wrong. But I think everybody should be able to draw the same conclusion when he looks at the same distribution." (Anton, in his portfolio)

Like Anton says in his reflection, topic V was the topic that brought everything together. Small and large *p*-values were discussed in detail, but as he mentions, there were few discussions on medium *p*-values, that are *p*-values between .05 and .1. In a redesign, there will be more emphasize on these *p*-values.

In summary most reflections were quite positive and students were highly engaged in the course, although it was more weekly work for them than participating in a "normal" course. A detailed review of the reflections is not done yet, but will give valuable hints for redesigning the course (amongst others).

Conclusion

For preservice teachers in Germany, there exist only few courses that build on an introductory statistics course. This course is one attempt to deepen the understanding of probabilities and offers a new approach to probability problems via simulation. The simulation method enables students to solve more complex problems than could be solved by calculation only, and allows them to solve unfamiliar problems. The reflections as being part of the portfolios helped the participants keep track of their knowledge and knowledge gaps, and appeared to be a very valuable instrument for assessment. But it cannot be assumed that the participants of the course fully understood the underlying statistical ideas and sampling that is occurring. In face of much work for the participants, they were highly motivated and worked hard during the whole course. From this first point of view, the design worked well and will be redefined in small details for a future iteration. For further research an interview study was conducted three months after the course and will be analyzed to investigate the impact on the understanding of probabilistic thinking of the participants.

References

Arbeitskreis Stochastik der GDM. (2003). Empfehlung zu Zielen und zur Gestaltung des Stochastikunterrichts. *Stochastik in der Schule, 23*(3), 21-26.

Batanero, C., Biehler, R., Maxara, C., Engel, J., & Vogel, M. (2005). Using simulation to bridge teachers content and pedagogical knowledge in probability. *15th ICMI Study Conference: The Professional Education and Development of Teachers of Mathematics.*

Biehler, R. (1991). Computers in probability education. In: Kapadia, R. & Borovcnik, M. (Eds.). *Chance encounters: Probability in education*. Dordrecht: Kluwer, 169-212.

Cobb, G. (2007). The introductory statistics course: A ptolemaic curriculum? *Technology Innovations in Statistics Education*, 1(1), 1–15. Retrieved from http://escholarship.org/uc/item/6hb3k0nz

Doerr, H. M., & Pratt, D. (2008). The learning of mathematics and mathematical modeling. *Research on technology and the teaching and learning of mathematics: Research syntheses.* Information Age Publishing, 259-285.

Drijvers, P. (2011). Teachers transforming resources into orchestrations. In: *From text to 'lived' resources*. Netherlands: Springer, 265-281.

Garfield, J., Ben-Zvi, D. (2008). *Developing students' statistical reasoning*. Springer Science + Business Media B. V.

Gnanadesikan, M., Schaeffer, R. L., Swift, J. (1987). *The art and techniques of simulation*. Dale Seymour. Teacher's Edition.

Hoyles, C., & Noss, R. (2003). What can digital technologies take from and bring to research in mathematics education?. In: *Second international handbook of mathematics education* (pp. 323-349). The Netherlands: Springer.

Konold, C., Harradine, A. & Kazak, S. (2007). Understanding distributions by modeling them. *International Journal of Computers for Mathematical Learning 12*, 217-230.

Lesh, R. & Doerr, H. M. (2003). Foundations of a Models and Modeling Perspective on Mathematics Teaching, Learning, and Problem Solving. In: Lesh, R. & Doerr, H. M. (Ed). *Beyond constructivism: models and modeling perspectives on mathematics problem solving, learning, and teaching*. Lawrence Erlbaum Associates: Mahwah, NJ, 317–336.

Lesh, R., Hoover, M., Hole, B., Kelly, A., & Post, T. (2000). Principles for developing thought-revealing activities for students and teachers. In: Lesh, R. & Kelly, A. *Handbook of research design in mathematics and science education*, 591-645.

Maxara, C. (2009). *Stochastische Simulation von Zufallsexperimenten mit Fathom. Eine theoretische Werkzeuganalyse und explorative Fallstudie*. Hildesheim: Franzbecker.

Meyfarth, T. (2008). *Die Konzeption, Durchführung und Analyse eines simulationsintensiven Einstiegs in das Kurshalbjahr Stochastik der gymnasialen Oberstufe - Eine explorative Entwicklungsstudie*. Kasseler Online-Schriften zur Didaktik der Stochastik (KaDiSto) Bd. 6. Kassel: Universität Kassel [Online: http://nbn-resolving.org/urn:nbn:de:hebis:34-2006100414792].

Podworny, S. & Biehler, R. (2014). A learning trajectory on hypothesis testing with TinkerPlots - design and exploratory evaluation. In K. Makar; B. de Sousa; R. Gould (Eds.): *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics* (ICOTS9, July, 2014), Flagstaff, Arizona, USA.ISI: Voorburg, The Netherlands.

Prömmel, A. (2012). *Das GESIM-Konzept: Rekonstruktion von Schülerwissen beim Einstieg in die Stochastik mit Simulationen*. Wiesbaden: Springer Spektrum.

Rabardel, P. (2002). *People and technology*. Université paris 8. [Retrieved from: https://hal.archives-ouvertes.fr/hal-01020705].

Voigt, J. (1984). *Interaktionsmuster und -routinen im Mathematikunterricht: Theoretische Grundlagen u. mikroethnographische Falluntersuchungen*. Weinheim: Beltz.

Zawojewski, J. & Lesh, R. (2003). A models and modeling perspective on problem solving strategies. In: Lesh, R. & Doerr, H. M. (Ed). *Beyond constructivism: models and modeling perspectives on mathematics problem solving, learning, and teaching*. New York: Lawrence Erlbaum Associates: Mahwah, 317–336.

Zieffler, A, & Catalysts for Change. (2013). *Statistical thinking. A simulation approach to uncertainty* (second edition). Minneapolis, MN: Catalyst Press.