STUDENTS' PROBABILSTIC SIMULATION AND MODELING COMPETENCE AFTER A COMPUTER-INTENSIVE ELEMENTARY COURSE IN STATISTICS AND PROBABILITY

Carmen Maxara and Rolf Biehler
University of Kassel, Germany
carmen@mathematik.uni-kassel.de

Modeling and simulation with the software Fathom *has become an important part of an introductory course on probability and statistics for future mathematics teachers at our institution. We describe our conception of modeling and simulation competence that students are supposed to acquire. We use various means such as modeling guidelines, simulation plan and a guidebook with examples for simulations to support students' learning processes. We report on results of empirical studies that made us change and extend our initial educational approach.*

INTRODUCTION

The role of simulation in an elementary course on probability and statistics is dependent on the goals one intends to achieve in the course. Many studies address the role of simulation for developing the concept of sampling distribution and for preparing concepts and methods of inferential statistics (delMas, Garfield, and Chance, 1999). We focus more on the role of simulation in elementary probability. In this respect, we build on ideas developed by Gnanadesikan, Scheaffer and Swift (1987) and Konold (1994) and on ideas in the German tradition where more emphasis is given to modeling and simulation in elementary probability.

Ideally the use of simulation can serve two different pedagogical purposes (Biehler, 1991). Students are to develop "modeling competence" related to elementary probability, and simulation competence can be part of it. Instead of setting up a formal model and determining unknown probabilities mathematically they can set up a model, implement it in a computer and develop estimates of unknown probabilities by means of simulation. In this sense simulation replaces mathematics. A different purpose is to use simulation for making probabilistic situations more experiential. The observed phenomenon can be used as a reference for concept development, for elaborating probabilistic intuitions, and for overcoming misconceptions. In this sense simulation replaces the real world, which is to be modeled.

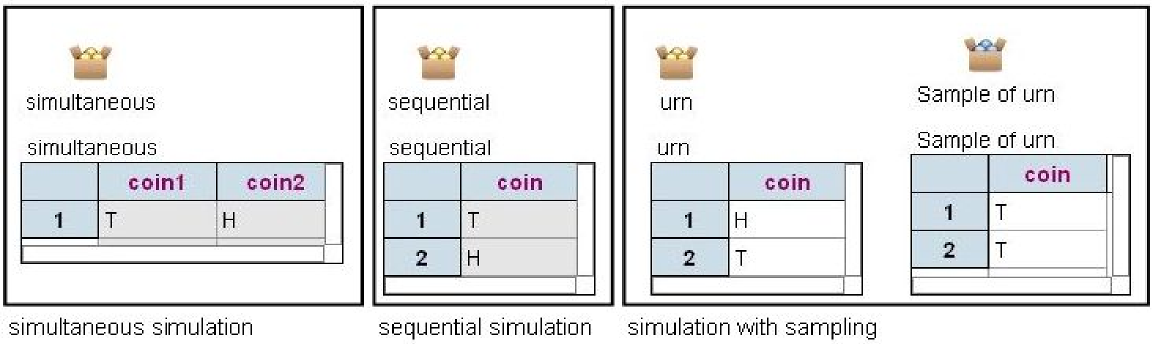
An appropriate software tool is essential for putting such an approach into practice. Students should work actively on analyzing data, simulating, analyzing methods and building models. The software *Fathom* achieves best our criteria for a software tool that supports both the learning and the use of probability and statistics in problem solving (Biehler, 1997).

At the department of mathematics and informatics in our University we are in the process of redesigning our introductory course on stochastics (probability and statistics) for future mathematics teachers. This is obligatory for student teachers that are going to teach in grade 5 to 10 (pupils' ages 11 to 16 years). In the last years the software *Fathom* was used in this course. The course starts with exploratory data analysis and descriptive statistics. In the first part of the course the students learn how to work with *Fathom* to analyze real data. The students can use these capabilities to analyze simulated data in the later parts of the course. Simulation as a method is introduced in parallel to the concept of probability. Situations are modeled mathematically and by simulation and results are compared.

A PRIORI DIDACTICAL ANALYSIS OF *FATHOM*'S SIMULATION CAPABILITIES

Developing *Fathom* as students' cognitive tool for modeling probabilistic situations implies that students develop a kind of mental model of the tool with regard to simulation problems. It also implies that they are to develop a certain style of working with the tool. Our a priori analysis aimed at supporting this process of instrumental genesis (Lagrange, 1999). As a result we constructed several concepts and notions that we communicated to the students. We distinguish three types of simulations (simultaneous, sequential, and sampling based), we developed the notion of a step-wise "simulation plan," and we had to develop the traditional concepts of event and random variable as bridging concepts between the world of probability and the world of *Fathom*. In other words these traditional concepts had to gain new meanings in the software context but we were aware that meaning conflicts may arise (Godino and Batanero, 1998).

- *Three types of simulations.* In a simultaneous version each particular step of a multi-step random experiment is represented by a separate attribute. One realization of the experiment is represented by one row. In a sequential simulation each step of the random experiment is represented by one case of the same attribute. In the third, sampling-based version, the simulation of the random experiment is realized by drawing a sample from a collection of data (an urn) where the number of cases in the sample represents the number of steps. The three different versions of simulations in *Fathom* are visualized in the following picture. The example is the simulation of a double throw of a coin.



It is tempting for a student to choose among the three different possibilities according to the similarity to the real situation. The simultaneous simulation seems to be most adequate if two coins are thrown simultaneously and the sequential one is adequate if one coin is repeatedly thrown. However, due to the specificities of the software, an optimal choice has to take into account the purpose of the simulation. Depending on the type of simulation, some events or random variables can be defined only in a complicated way or they cannot be defined at all. Moreover the possibilities of generalizing or modifying a simulation depend on the type. For instance, the waiting time until the first occurrence of a "T" can be easily defined by little modifications of the sampling-based simulation but not at all with the other two types. We regard this as a source of problems for the students.

- *Guidebook "An introduction into simulation with Fathom."* As part of our a priori analysis we developed a comprehensive overview of possible simulations in *Fathom*. We constructed a set of working environments with simulation components (about 200). Based on these results, a Guidebook "Introduction into Simulation with *Fathom*" with many examples in *Fathom* was developed. This guidebook was especially prepared for the students in our elementary statistic course and contains only a selection of examples and options. The guidebook contains further aspects of our a priori analysis that we describe below.

WORKING STYLES WITH SIMULATIONS: MODELING GUIDELINES AND THE 'SIMULATION PLAN'

The simulation in the course and in the guidebook recommends a three step process for a stochastic simulation: setting up a stochastic model, writing a *simulation plan* and the realization in *Fathom*. This three step process is used as a guideline for stochastic modeling.

In the first step, setting up a stochastic model, the students should build a model of a real situation with a random outcome by describing the situation for example by a concrete urn-model or in a more abstract way. They have to specify the set of possible results (the sample space), the probability distribution, the number of steps of the experiment and the interesting events or random variables. In the second step the students are to transfer the constructed model in a plan for a simulation of a random experiment in *Fathom*. The simulation plan was an important aspect in the conception of the courses. It aims at supporting students to reflect about the simulation process before (or after) the realization of their simulation in *Fathom*.

We became aware of the need for such a "modeling guideline" when observing our first generation students at work and when doing interviews with them. The students often directly jumped into using the software without enough stochastic thinking beforehand. At least after finishing the simulation in *Fathom* the modeling guideline and the simulation plan becomes relevant for them, when they have to communicate and document their simulation.

The simulation plan is always composed of five steps. We will illustrate these components with the following famous example:

A math course consists of 23 students. What is the probability that the birthday of at least two students are on the same day of the year? We assume that every day of the year is equally likely as a birthday. Estimate the probability by a simulation.

Stochastic Modeling

Concrete model. We're modeling the problem with sampling from an urn with 365 consecutively numbered balls (1-365). From this urn we're sampling 23 balls with replacement. Every drawn ball represents the birthday of one student. If two or more numbers are equal at least two students' birthdays are at the same day.

Formal model of single experiment. We describe the set of results of one step of the random experiment with $\Omega_1 = \{1, 2, 3, ..., 364, 365\}$ with an uniform probability distribution $P(1) = P(2) = ... = P(365) = \frac{1}{365}$.

Formal model of compound experiment. The whole random experiment consists of 23 steps, so the set of results can be described by $\Omega = \underbrace{\Omega_1 \times \Omega_1 \times ... \times \Omega_1}_{23-times} = \{(x_1, x_2, ..., x_{23}) \mid x_1, ..., x_{23} \in \Omega_1\}$ with the probability distribution $P(\{(\omega_1, \omega_2, ..., \omega_{23})\}) = P(\{\omega_1\}) \cdot P(\{\omega_2\}) \cdot ... \cdot P(\{\omega_{23}\}) = \left(\frac{1}{365}\right)^{23}$, $\omega_1, \omega_2, ..., \omega_{23} \in \Omega_1$. Alternatively students can leave the model for the compound experiment implicit and work and think with more concrete representations of results of compound experiments as paths in a tree diagram.

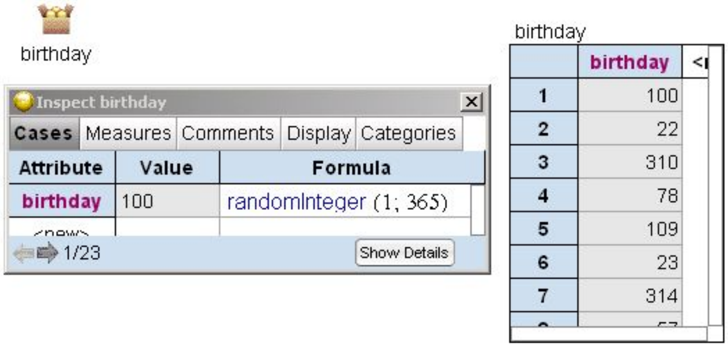
Plan of Simulation

Step:	Example:	General:
1	We create a new collection "birthday" with an attribute "birthday" and intend to represent every student by a case (sequential simulation).	Defining a collection and one or more attributes; choice of the type of the implementation in *Fathom* (simultaneous, sequential, or sample).
2	We choose the random function *randomInteger(1,365)* to define the attribute. The function randomly generates an integer representing the birthday of a student. We add 23 cases to the collection. Each case stands for the birthday of one student.	Choice of an appropriate random machine for the probability space and for the simulation of the random experiment.
3	For the event E: "the birthdays of at least two students are on the same day" we define a measure "E" with the formula: *uniqueValues(birthday)* $\leq$ *22*. If this expression is true the event occurs.	Defining events and random variables as attributes or *measures*. (A measure is a technical concept in *Fathom* that will be described later.)
4	We collect 1000 measures. (The simulation is repeated 1000 times.)	Realization and repetition of the simulation.
5	We analyze the results with a summary table and a bar chart.	Statistical data analysis of the simulated results.

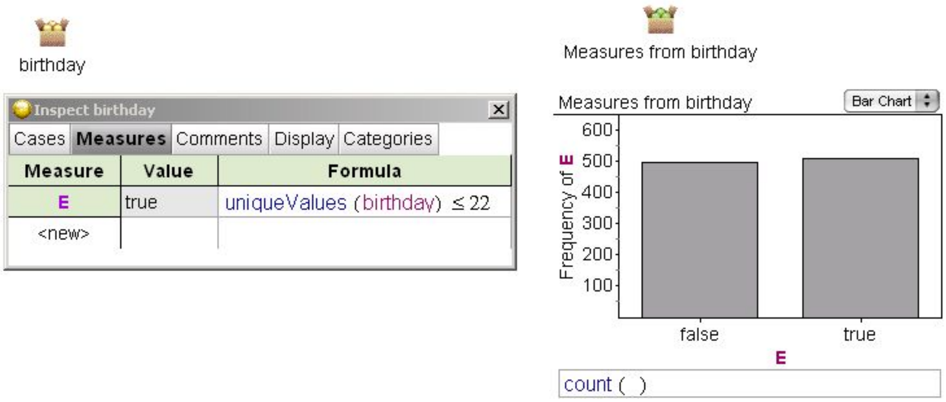
Realization in Fathom

The realization of the simulation plan in *Fathom* is the third step in the overall process. This is the most practical part of the activity. The simulation plan is getting real. In addition the students can experiment with their assumptions, with the sample size, with the number of repetitions of an experiment and so on.

The steps easily transfer one to one in *Fathom*. Each particular step corresponds to a certain action in *Fathom*. This is an advantage of the software *Fathom* because this supports the problem solving use of the software tool.



The simulation of random experiments consisting of several steps may be realized in one of the three mentioned versions. We realize the example of the birthday-problem in a sequential simulation, in which each particular step of the random experiment is represented in one case.



Most important is *Fathom*'s notion of a *measure*. A measure can be defined by a formula. It refers to a collection as a whole. The inspector window of a collection contains a panel for defining measures. The mental model about measures that students have to develop is that measures are properties of a collection as a whole. In this case, a measure is a random property of the compound random experiment. If we rerandomize the values of the attributes in the collection the value of the measure will vary accordingly. In this sense, a measure in *Fathom* is a new reification of the abstract concept of a random variable or event, and we expected new learning opportunities from this aspect. On the other hand we expected several semiotic conflicts (Godino and Batanero, 1998) at this point. In everyday language, the occurrence of an event is either true or false and the *Fathom* notion of an event as a Boolean expression directly corresponds to this notion. On the other hand the notion of an event as a subset of the sample space is not directly supported by this reification. The notion of a random variable as a function on the compound sample space is supported. However the distinction between events and random variables may become blurred, because both can be defined as measures. Moreover, the above problem could be solved by observing the random variable X defined as uniqueValues(birthday), and in the next step reading off from a histogram of the distribution of X the relative frequency that $X \leq 22$ occurred – without explicitly using the concept of event.

SIMULATION AND A NEW CONTENT STRUCTURE IN THE COURSE

In the first course with an emphasis on simulation, the students had to work with the three step design: stochastic model, simulation plan, realization in *Fathom*. In the lectures the simulation was continuously integrated in the different versions (simultaneous, sequential, with sampling), but not so explicitly in these three steps. The concept of probability was initially defined by elementary events in discrete probability spaces following a step by step extension to events, random variables and results in random experiments consisting of several steps. The concept of a random variable was introduced about the same time as the concept of event. The introduction of combinatorial-mathematical methods for calculating probabilities was always done in parallel to simulation methods. Further mathematical contents were among others the empirical law of large number, probability distribution of random variables, the mean, the expected value, fair games, (in)dependence, the path rule, binomial coefficient, binomial distribution, hypergeometric distribution, the variance, the standard variation, $1/\sqrt{n}$ – law. In working sessions and homework assignments students had to solve problems in several textual contexts, particularly with stochastic contexts like urns, dies, and game contexts.

CHANGES OF OUR CONCEPTION OF MODELING COMPETENCE AS A RESULT OF OUR EMPIRICAL STUDIES

After the courses we performed interviews with a selection of students, we observed them while working in pairs on modeling tasks and we analyzed written homework assignments. We cannot go into details of these studies in this paper but we will summarize some changes in our didactical conception of simulation and modeling competence and changes in the content and orientation of the course as a consequence of these studies.

The modeling guidelines, the guidebook and the concept of simulation plan proved its worth. Nevertheless, we saw several limitations in the students' competence.

- The status of the modeling guideline was more emphasized particularly the stage of stochastic modeling. We had observed unsatisfactory limitations in students' knowledge and language, which often was close to the software processes but not to the probabilistic concepts. More emphasis was put on the explanatory aspect of the simulation plan instead of being just a technical advice how to run the simulation.
- We introduced more and new tasks that required more thinking and idealization in the first stage of modeling. In the first generation we concentrated too much on the interface "stochastic problem – software" and lost sight of the interface "real world – probability model." Initial situations should be idealized and simplified by students themselves. Students were asked to make assumptions themselves. They should decide whether a model fits a situation or not. For example, we used subtasks as:
  (1) Express the following situations as multilevel random experiments. (2) Would you classify the levels of the random experiments dependent or independent of each other? Justify your decision. Give real conditions under which independence is likely to be an adequate model. (3) In which cases you think is an empirical test with data necessary to support or reject the assumption of (in)dependence?
  The goals of such new problem types were to improve the modeling competence, the competence of giving reasons, a deeper understanding of important terms (in this case: independence), and more reflective activities.
- We added examples, where a model was compared to real data in order to exercise model validation.
- We found that the students were not really sufficiently reflecting on the usefulness and relative advantages of using simulation as compared to theoretical mathematical solutions. We added more tasks, where we used both methods in combination. The accuracy and the limits of simulation methods became a larger topic. We used simulation for developing hypotheses that were explained and proved by theoretical analysis and we used simulations to check the validity of theoretical solutions.
- We found that intuitive conceptions and misconceptions persisted or co-existed more than expected. Although the reality of the simulation well showed that intuitions are wrong, simulation as such often did not help to improve. Therefore we addressed intuitive expectations explicitly in the course as "intuitive theories" that should be expressed before starting simulation and modeling. We think it's important that students better reflect about their intuitions

before they begin to model or try to compute something. So we asked "What do you think about this problem? What's your estimate?" For example, estimates of the students regarding the birthday problem were - as expected - much too low. To the question "What's the probability for at least two birthdays at the same day of 42 random picked persons?", the median of the students' estimates was < 10%. Such wrong or inadequate intuitions are taken up again after the simulation and theoretical analysis of the problem. The contradictions between the intuitive theory, the mathematical theory and simulation have to be discussed. This discussion is essential, because the simulation in itself doesn't explain the contradictions and doesn't necessarily improve stochastically understanding. We had been aware of this problem, which is well discussed in the literature, but still underestimated the problems of the students.

- We found the need for new tasks formats in order to stimulate students' flexible modeling and simulation competence. There is a long tradition with tasks in elementary probability that require from students more or less the determination of single probabilities or expected values of random variables instead of a more general exploration of the situation by means of the model. We found it difficult to formulate open subtasks that direct the students' attention to general aspects of the model and to further explore and vary the model. Similar to guidelines for problem solving activities we found it helpful to add a step to our simulation and modeling plan, namely the step "Asking further questions." The questions and the model should be varied. In the birthday problem possible further questions are: "What are the probabilities for 10, 20, 300 persons?", "How many persons must be there for a probability > 50% of at least one double birthday?", "What happens if the equal probability assumption is not true?"

CONCLUSIONS

In the process of redesigning the course we will expand our modeling guidelines. The initial three step design is now expanded in a six step process:

1. Intuitive theories and expectations, 2. Building a stochastic model, 3a. Generating a simulation plan, 3b. Theoretical analysis of the problem, 4. Comparing simulation and theoretical analysis, 5. Comparing the intuitive theory with theoretical analysis and/or simulation; debugging of misconceptions, 6. Exploring further questions: Varying assumptions, contexts and questions.

We think that explicitly extending the course that was supposed to be "only" about probability and modeling in the direction of probabilistic thinking and more subjective aspects of modeling is an important step.

REFERENCES

Biehler, R. (1991) Computers in probability education. In R. Kapadia and M. Borovcnik (Eds.), *Chance Encounters - Probability in Education*, (pp. 169-221). Dordrecht: Kluwer.

Biehler, R. (1997). Software for learning and for doing statistics. *International Statistical Review* 65(2), 167-189.

delMal, R. C., Garfield, J., and Chance, B. L. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. *Journal of Statistics Education* , 7(3).

Fathom Dynamic Statistics™ Software. Emeryville, CA: Key Curriculum Press.

Gnanadesikan, M., Scheaffer, R. L., and Swift, J. (1987). *The Art and Techniques of Simulation*. Palo Alto: Dale Seymour.

Godino, J. D. and Batanero, C. (1998). Clarifying the meaning of mathematical objects as a priority area of research in mathematics education. In A. Sierpinska and J. Kilpatrick (Eds.), *Mathematics Education as a Research Domain: A Search for Identity*, (pp. 177-195). Dordrecht: Kluwer.

Konold, C. (1994). Teaching probability through modeling real problems. *The Mathematics Teacher*, 87(4), 232-235.

Lagrange, J.-B. (1999). Complex calculators in the classroom: Theoretical and practical reflections on teaching pre-calculus. *International Journal of Computers for Mathematical Learning*, 4, 51-81.