The influence of software support on stochastics courses

Borovcnik, Manfred
University of Klagenfurt, Dept. of Statistics
Universitätsstraße 65-67
9020 Klagenfurt, Austria
E-mail: manfred.borovcnik@uni-klu.ac.at

Introduction

Students in different studies at university bring in widely different pre-knowledge in mathematics and statistics. They also differ in their interest and motivation to do more efforts in learning difficult concepts. This author's experience from academic teaching is that there is a definite order in mathematics aversion, starting from low levels for mathematics students (!), to rising levels amongst mathematics teacher students, students of computer-science, and on top business and administration students. Difficulties with statistics are even higher. In this paper, the author describes some activities he has designed for students of business administration for the second phase of their curriculum. Before, they had passed an introductory course in statistics, which put them at a level comparable to the intended level of the final exam after secondary school. The course content was revised in three cycles of teaching. The number of participants is small (21, 16, and 24 for the third cycle) – the university is a small place and the second phase of statistics is not compulsory.

The course made extensive use of the strategies to cope with the complexity of statistical notions, which are described below in connection with consistent use of spreadsheets in the course. *Dynamic visualizations* to illustrate concepts, and *material embodiment* of calculations to reveal characteristics of concepts, were two of strategies, which proved very successful. Simulation and the exploration of a paradigmatic context (an *analogy*) is another approach used in the course. Below, some excerpts of the approach will be discussed with copies of the spreadsheets, which were developed in cooperation with the students – at the input of the teacher, and later improved in layout for further use. Feedback from the students was quite positive. One has also to take into account that these students are generally not the best at mathematics. The students' motivation was boosted by the use of EXCEL because they already had a good pre-knowledge and they learned many more tricks improving their use of it also for other areas; they learned for the first time that statistical concepts convey interesting statements about the data to be analyzed and the situation in the background. Formulae, which they would have been anxious about, were circumvented by a spreadsheet version of it; i. e. formulae were not simply avoided by retrieving a function from a function menu but were built up step by step by a spreadsheet, with each single step having an obvious rational.

Discussion about the ubiquity of problems to understand the complexity of statistical notions

Exam papers of students over a long period give a clear vision that many have no idea how to interpret such concepts even though they solve numerically the related tasks. It did not matter whether this author had given them extra papers elaborating on such issues, or had extensive discussion about the deeper meaning of confidence intervals, or had identified why wrong interpretations are wrong. The difficulties in learning statistics are not only due to its mathematical ingredients but seem to stem also from highly complex concepts in statistics, which should be illustrated here shortly.

(i) Empirical research in probability understanding now gives persistent "proof" over the decades, starting with the early Kahneman, Slovic and Tversky (1982) studies about elementary misunderstandings, of discrepancies between official concepts and private intuitions and imagination and how this shifts individuals far away from what may be read off from probabilistic statements.

(ii) The mixture between causal interpretations of any random experiment like dice throwing and the search for patterns in the emerging data sequence and probabilistic perceptions about the experiment in the background (like recently shown in Johnston-Wilder 2007) maybe not so important at first sight as dice throwing might not be considered as important for applications. However, for building up mental images about what probability conveys, dice throwing or similar experiments are a substantial didactical tool. For the importance of such didactical objects see a. o. Borovcnik and Peard (1996) – they should balance for the *indirect* knowledge, which is inherent to probability (see Borovcnik 1994).

(iii) The ongoing misinterpretation of the α error of a standard test for significance of any null hypothesis, which is wrongly taken as the probability of the null hypothesis in case of its rejection is but one example. Another one is the inadequate probability interpretation of the result of the confidence procedure for a given set of data. The unknown parameter lies within the realized confidence interval or not, which illustrates according to Callaert (2007) that a probability cannot be associated to it. But in his attempt to give the confidence level a probabilistic meaning by an analogy to a wheel of fortune, he misses the point completely, which highlights the difficulties with an adequate interpretation.

Borovcnik (1992) has elaborated on a mutual relationship between official concepts and imaginations and interpretations, which are feasible for them, and on the other hand private intuitions, which individuals hold prior to education and after passing education. Along the lines of Fischbein (1987), he develops the idea of a dynamic interplay between the “two worlds”, which may be connected only on the level of accompanying private intuitions, which are challenged and revised several times. In order to allow such interplay, more has to be done than just to teach the mathematics.

Strategies to cope with problems to understand the complexity of statistical notions

There are a lot of strategies to help such interplay, all well-known in educational statistics.

Visualization. The EDA approach has initiated an enormous drive towards this general idea. However, visualization is a more general and useful idea to display characteristic features of a concept and not only to depict data in an easily accessible way in order to facilitate interpretation. Special attention may be attributed to dynamic visualization of concepts.

Simulation. There have been a lot of proposals to fill the mathematical gap of key concepts like the central limit theorem by simulation starting from Kissane (1981). One may criticize that many of these simulation proposals miss the point of explaining the concept. However, simulation is a didactical tool now no one can make a roundabout. The question is more how to flexibly integrate simulation into the theoretical development of concepts in class and its discussion with the students.

Embodiment of situations. To give a concrete embodiment of the situation as simple as possible with the main features preserved, is another strategy, advocated and elaborated mainly be the Gigerenzer school (see Krauss e.a. 1999). Their research on the use of “natural frequencies” in the embodiment of a two by two table to deal with Bayes problems (with only two ways classification like in the medical diagnosis context) show considerable progress in successful solving behavior.

Contexts. To refer to paradigmatic contexts does not give only motivation to students but also helps to interpret the concepts in contrast to the context. The medical diagnosis context is a suitable “environment” to clarify the standard situation of a statistical significance test, which elements are usually missing when errors of second type are not considered and that test decisions may never be interpreted fully without reference to this β error under various conditions for the alternative hypothesis.

Analogies. Teaching with the help of analogies is not new; an early reference is Brewer (1989). An analogy thereby is not simply a context. It yields a well-known reference situation; the learners have direct access to key questions and may ask for the necessary concepts to answer “their” questions in the analogue situation. This leads to easily comprehensible models and reduces wrong interpretations of results to a ‘feasible’ level. An analogy lets develop and comprehend the required mathematics step by step.

The list of strategies is by no means complete: e. g. the consistent use of applications and the necessity to interpret theoretical results in the context of a real situation can help to understand more about the restrictions of the used models and the derived results; ‘elementarization’ is another strategy, by which one would teach special cases only, or refer to analogue situations, which are analyzed without the use of mathematics; dynamic visualization is one further strategy to simplify matters. In general, these strategies are intended to substitute for mathematical relations, which are out of the reach of the learners or out of their interest. The target always is to reveal crucial features of notions and procedures without developing the full mathematical background. One may combine the strategies above in introducing a complex concept.

One-dimensional descriptive measures

Introducing measures of location for a set of data could lead to a discussion of the relative merits of mean and median: The mean reacts to all data, while the median is robust as it is not influenced by single data. The list of relative merits might be longer but remains “abstract”. The learner gets a clear vision of the features of mean / median or of standard deviation / interquartile range when a dynamic representation of the data is used in EXCEL (see Fig. 1) and one of the points is moved by a ruler.

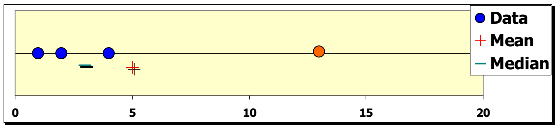

Figure 1

Two-dimensional descriptive measures

The determination coefficient R^2 as well as the correlation coefficient R are keys to understand what is done in regression analysis. The analogy to error theory relates the task to the prediction problem: How to predict the values of an object with respect to the dependent variable if its value of the independent variable is known. By intelligent forecast techniques the error of prediction should be kept small. Supposed, we compare a naïve strategy (#1) of predicting always the mean of the dependent variable with the linear strategy (to predict always the value of the best linear line, #2), which is pre-calculated and used as a black box. The improvement may be judged by variances of errors of predictions. By simple spreadsheet calculation students get an embodiment of the meaning of R^2 as may be seen from the spreadsheet in Fig. 2, $1 - R^2$ being the factor by which the original variance is reduced in the residuals.

data		naive strategy #1		lin. strategy #2		residuals
inde-pendent	dependent = D	prediction F_1	residuals $R_1=D-F_1$	prediction = fit F		R=D-F
x	y	m(y)	y-m(y)	$\hat{y}$	$\hat{y}$ - m(y)	y-$\hat{y}$
1	3	3,17	-0,17	2,16	-1,00	0,84
2	2	3,17	-1,17	2,54	-0,63	-0,54
3	5	3,17	1,83	2,92	-0,25	2,08
6	2	3,17	-1,17	4,05	0,88	-2,05
1	1	3,17	-2,17	2,16	-1,00	-1,16
9	6	3,17	2,83	5,18	2,01	0,82
sum	19,00	19,00	0,00	19,00	0,00	0,00
mean	3,17	3,17	0,00	3,17	0,00	0,00
variance	3,77	0,00	3,77	1,46	1,46	2,31

Figure 2

Test of a binomial parameter

The context of multiple-choice examinations with two answers to each of 20 items leads to questions like: Is the test suitable to check what students have learned? What happens if a student answers the items randomly, how big is the probability to pass the exam (if 11 items are necessary)? What does it mean that the exam is fair? Is it fair? What happens to a student who has learned quite a lot and whose success probability may be compared to 0.7? Is it possible to design a test fulfilling the assumptions of the binomial distribution for the number of correct answers? The context may be visualized by bar diagrams; the success probability may be visualized accordingly. The probabilities may be enlivened by simulating the conditions.

EXCEL lets dynamically simulate these data. Finally the examination by multiple-choice might be evaluated by comparing the distribution of correct answers for those who are randomly answering with those whose solving capacity has a value p (e. g. 0.7 from above). This gives the well-known figure from textbooks illustrating the α and β error (for a special value of the alternative hypothesis). However, with a spreadsheet it is easy to make the figure dynamic (see Fig. 3) and show (i) how the β error varies with a different rule for passing the test (e. g. 8, or at least 14 correct answers are required to pass the test), which is associated to a different α level of the test; or show (ii) how the β error in increased if the solving capacity p of the learners comes closer to the value of 0.5. (Steering the dynamic is again via rulers.)

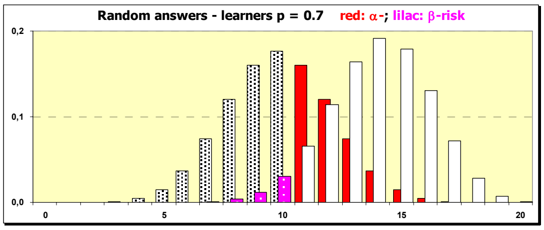

Figure 3

Concluding remarks

Only a few spreadsheets of the course are shown here. Combinatorics may be replaced by recursive approaches (Borovcnik 2007). Resampling yields a new approach towards inferential statistics reducing the learning difficulties (Christie 2004, or Borovcnik 2007). The students were highly engaged in the activities and discussed e. g. fiercely the consequences of alterations to the examination. In the regular evaluation of the course the feedback was raised by 1.6 points on a 5 point scale. Some remarks on competing software conclude this paper: While FATHOM serves more directly the needs of statistical education (e. g. Maxara and Biehler 2007), the effort to learn it establishes a hurdle. The graphic capacity of other software may be more convenient but the students appreciated their progress in using graphs in EXCEL. Needless to say that they use EXCEL in other courses, and that spreadsheets are widely used outside the university. These relative merits of EXCEL are also valid in comparison to other educational software.

REFERENCES

Borovcnik, M.: 1992, Stochastik im Wechselspiel von Intuitionen und Mathematik, BI, Mannheim.

Borovcnik, M.: 1994, Intuitive Strategies for Teaching Statistics, in L. Brunelli, G. Cicchitelli (eds), Proc. 1st Scientific Meeting of the IASE, Universitá di Perugia, Perugia, 355-366.

Borovcnik, M. and Peard, R.: 1996, Probability, in A. Bishop e. a. (eds.), International Handbook of Mathematics Education, part I, Kluwer, Dordrecht, 239-288.

Borovcnik, M.: 2007, Das Sammelbildproblem – Rosinen und Semmeln und Verwandtes: Eine rekursive Lösung mit Irrfahrten. Stochastik in der Schule 27 (2) – to appear.

Borovcnik, M.: 2007, On Outliers, Statistical Risks, and Statistical Inference. CERME 5, WG 'Stochastical Thinking'.

Brewer, J. K.: 1989, Analogies and Parables in the Teaching of Statistics, Teaching Statistics 11(1), 21-23

Callaert, H.: 2007: Understanding Confidence Intervals. CERME 5, WG 'Stochastical Thinking'.

Christie, D. 2004: Resampling with EXCEL. Teaching Statistics 26 (1), 9-14.

Fischbein, E.: 1987: Intuition in Science and Mathematics. An educational Approach, Reidel, Dordrecht.

Johnston-Wilder, P. and Pratt, D.: 2007, The Relationship between Local and Global Perspectives on Randomness. CERME 5, WG 'Stochastical Thinking'.

Kahneman, D., Slovic, P., and Tversky, A. (eds): 1982, Judgement under Uncertainty, Cambridge U. P, Cambridge.

Kissane, B.: 1981, Activities in Inferential Statistics, in A. P. Shulte and J. R. Smart, Teaching Statistics and Probability, NCTM, Reston, Virginia, 182-193.

Krauss, S., Martignon, L., and Hoffrage, U.: 1999, Simplifying Bayesian Inference: The General Case, in L. Magnani e.a. (eds), Model-based Reasoning in Scientific Discovery, Kluwer, New York.

Maxara, C. and Biehler, R.: 2007, Constructing simulations with a computer tool – students' competencies and difficulties. CERME 5, WG 'Stochastical Thinking'.