



Research and Development in the Teaching and Learning of Probability

A Report on the Topic Study Group 13 at ICME 11 – Monterrey, Mexico – July 6-13, 2008

Manfred Borovcnik, University of Klagenfurt

The series of the ICME meetings are probably the biggest congresses dedicated to mathematics education in the world. The 11th convention of this kind took place in Monterrey, Mexico. This report focuses on the results of the topic study group 13 on probability issues in education. The rapporteur will classify the contributions alongside the topics, which were contained in the call for papers prior to the congress.

The importance of research in probability teaching

Probability and statistics education are relatively new disciplines. Both have only recently been introduced into the main stream school curricula in many countries. While the application-oriented statistics is undisputed in its relevance, discussion about probability is more ambivalent. Reduction of probability to the classical conception, mainly based on combinatorics, or its close connection to higher mathematics, are sometimes used as arguments to abandon it in favour of the statistics part. However, there are key reasons for a strong role for probability within stochastics curricula:

1. Misconceptions on probability affect people's decision in important situations, such as medical tests, jury verdict, investment, assessment, etc.
2. Probability is essential to understand any inferential procedure of statistics.
3. Probability offers a tool for modelling and "creating" reality. For example, modern physics cannot be formulated without reference to probability concepts. The concepts of risk (not only in financial markets) and reliability are closely related to and dependent upon probability.

Thus the challenge is to teach probability in order to let the students understand it. The focus has to be on creating approaches to probability that are more accessible and motivating. Additionally, the frequentist and subjectivist views of probability, and connections of probability to practical applications should be taken into account. Simulation is one such strategy, as is visualization of abstract concepts; there are more. The use of technology also enables to reduce the calculation technicalities and focus the learner on the concepts instead. The world of personal attitudes and intuitions is another source for success or failure of teaching probability. These challenges of course could not be met by a single working group at a conference. However, some valuable contributions may be noticed, which will be clear from this report. This follows the main themes emerging in the sessions, which nevertheless did overlap: individuals' corner; impact of technology; teachers' corner; conditional probability and Bayes' theorem; fundamental ideas; panel discussion; future perspectives.

Individuals' Corner

This topic was intended to comprise “Individual understanding and misconceptions of notions – Intuitive concepts of children and adults about probability”. Quite a few of the papers may be attributed hereto. There has been a shift away from misconceptions, which may be changed by suitable teaching, towards pre-conceptions, which should be taken up and refined in teaching. Such a shift of focus in research may be traced throughout all papers, which reported on empirical research.

K. Rolka (Germany, joint work with S. Prediger) presented a study of a group of 12 year olds who were observed playing a game of chance with tokens moved forward on a playing board by the result of a die. The icosahedron used had more red sides than sides of any other colour whence it favoured the red token. The fierce discussion amongst the children leads to a *jointly* agreed best strategy. Views as “there are more red sides on the die” and “the red token wins more often” were advocated likewise. The common struggle for a strategy seems to generate a better understanding of the value of their tactics – by their discussion they seem to be much more aware of the risk that a token of a different colour could win; much more than they might have been aware by simply *observing* the game. The social situation of the class with the children interacting in their discussion is a feature of Rolka’s investigation – the situation is thought also to be exemplary for later teaching.

With D. Abrahamson (US), the experiment is signified by a single child (Li, 11 years) and the interviewer interaction, an in-depth interview after a teaching phase with an urn experiment replaced gradually by the computer environment. Abrahamson is not only interested in the personal understanding of the child but in the learning trajectory of it and how the interaction of the representation of the notions by different media could influence learning positively and speed it up. He seems to be cautious to let the interplay of the different embodiments of the same notion work as the main input for learning (and not the teacher or interviewer). The experiment involves the binomial distribution.

Four blocks serve as unifying element and change their appearance from *spoons* to *scoop samples* to the *building elements* of the combination tower of all possibilities, and finally build up the *histogram* of repeated samples. As a side effect, the histogram gets a *greenish* impression resembling the proportion of green marbles in the urn.

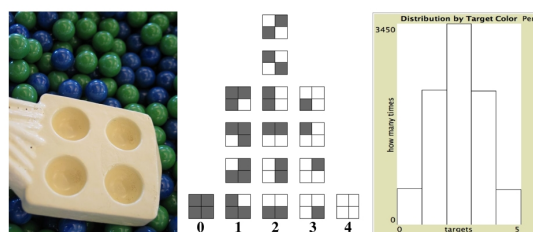


Figure 1. Abrahamson’s four blocks as unifying elements

The empirical studies by F. Chiesi & C. Primi (Italy) and L. Zapata (Colombia/US) deal with heuristics in the paradigm of Kahneman & Tversky. The Italian group studies the development of *negative* and *positive recency* by age. They compare 9, 11, and 25 years olds to imitate a longitudinal survey.

From a bag with blue and green marbles a marble is drawn repeatedly with replacement. The result “all marbles of the same colour” is presented (the marbles are not actually drawn). The numbers of both colours is known – they are varied from same numbers to a strong bias to either of the colours.

According to *negative recency*, people choose the change: with 4 blue they would predict a green one for the fifth draw. With *positive recency*, they would predict that colour, which continues the series “observed”. How frequently are these heuristics used, and are they independent of the colour composition of the bag? Interestingly, the study shows an increase of the normative solution first (from age 9 to 11), with a drop down later (age 25). To the same extent, *negative recency* decreases (from age 9 to 11) first and then increases (age 25). With *positive recency* there is a decrease (9 to 11) and it remains then (25) at this level. More in-depth investigations are needed to clarify what has happened here, or whether such a *development* can be confirmed (of course it is only fictionally longitudinal).

On the outset of L. Zapata's investigations there are well-known tasks from the Kahneman & Tversky school: The conjunction of two statements is frequently assumed to be more probable than each of the single statements – this may be due to the fact that the conjunction resembles much more authentic whence it is more “plausible” and thus judged to be more likely. (Remember the term *plausibility* for the likelihood function, which uses the similarity in connotation between the two.) Other tasks relate to the laws of large and small numbers reflecting that people are widely unaware of the huge influence of sample size on variability of results observed. New in Zapata's study are single interviews with *teachers* with the target to clarify what may be learned from more experienced teachers: Are they better in anticipating difficulties and can they offer suitable media or representations to cope with?

One result of an investigation of K. Lysø (Norway) with student teachers is the documentation of an inclination to reformulate tasks, which is not much known about up to now: The students quite often reformulated two-stage experiments into one-stage tasks getting a wrong answer – but remained unwilling to see why their reconstruction is misleading:

Anna has three red, two green and one blue pencil in her pen case. She asks Maria to pick out two pencils without looking. Anna thinks that the probability that both of the pencils are red is $1/5$ but Maria thinks that the probability is $1/3$. Does either of them have the correct answer?

“I agree with Maria. Argument: There are a total of 6 pencils and Maria picks out 2 pencils; this leads to the probability $2/6 = 1/3$ for red pencils.” (More than 20%)

“No, one has to divide the red pencils, 3 pieces, on the number of the total number of pencils, 6 pieces, $= 3/6 = 50\%$ chance for a red pencil.” (About 15%)

S. Anastasiadou (Greece) develops a battery of simple items to research relations between algebraic and graphic skills in student teachers. Using similarity diagrams she detects a widespread lack of skill to change between different representations of a task or a notion. Missing conceptual flexibility hinders a deeper comprehension. She explains the result by an inclination to learn compartmentalized concepts, i.e., students learn the concepts together with the representation. With different representations, they learn *different* concepts – they fail to recognize the same notion in different shapes.

Impact of Technology

There has not been a systematic evaluation of the possibilities and limitations of new media in the group. However, as far as the presenters focused towards teaching or teacher in-service, they freely used various kinds of software in a substantial way. Spreadsheets (EXCEL), Fathom, or Tinkerplot were used for efficient calculations but also for illustrating key ideas (concept of distribution, law of large numbers, etc) e.g., by S. Inzuna (Mexico) or R. Peard (Australia).

New media indirectly form the backbone of the elaboration of D. Pratt (UK) on shaping the experience of naïve probabilists. By means of intentional sequences of the programme ChanceMaker, he supplies new and challenging experiences to learners in order to shape their intuitions and strategies. Pratt talks of new challenges for designers of software and teachers using this software alike. In a fusion of control over the initial parameters (via randomness) and representations of results (histograms for the distribution of data or statistics like the mean), he seeks to prepare *new* insights into randomness, which should widen and refine intuitive notions, which might have been too narrow (and thus from time to time misleading) previously. Software offers more efficient, graphically orientated possibilities to supply (in fast motion) and order experience with randomness. His essay could be attributed likewise to the heading “Fundamental Ideas”: How is it possible to form such basic ideas – with the support of new media?” According to Pratt, a new world of up-to-date unknown intuitions might emerge, which would affect onto concepts and its understanding, too.

Teachers' Corner

This topic embraces pre- and in-service education as well teachers' conceptions of teaching and of probabilistic notions. Some of the contributions fit to the heading of fundamental ideas as well.

K. Lysø (Norway) presented an innovative starter for the elementary probability course for teacher students. He starts with an empirical investigation with the students. He wins the students' motivation and can build a bridge between their intuitions and the mathematical concepts. He can refer to these items and their discussion later in the course with success. The author of this report confirms that this way is rewarded by the students' attention and success. The items used are not really decisive – Lysø uses a battery of quite normal tasks covering the main primitive concepts including some items referring to two-stage experiments. It depends mainly on how the items are dealt with after the written test. It may have a lasting effect on how one opens the discussion about which solutions are feasible, or which reconstruction of the task would make sense and therefore lead to a sensible solution even if it does not coincide with the “normative” solution.

L. Zapata (Columbia/US) tries to derive meta-knowledge for teachers from her in-depth interviews with teachers. Wherein do teachers with experience differ from novices? Which learning problems may be anticipated by teaching expertise? Which strategies to counter such problems would expert teachers apply? Surprisingly, or better, unsurprisingly, novice teachers witness the same misleading intuitive conceptions as their students and thus are not really able to help them. Possibly this result is another argument to include (at least) one didactical course on the subject in teacher education at university and not restrict it to a mathematical course of the subject. Probability is much more prone to such difficulties than other topics.

J. Watson (Australia, together with S. Ireland) reported about the results of in-depth interviews covering issues on the relations between empirical and theoretical probabilities. The interviews build on two course units with 12 year olds, which covered first coin tossing and its tabulation of the results by the joint effort of the class and second the same experiments done by Tinkerplot (a didactic software, which is getting more and more popular). Using the software has widened the students' experience, which is reflected by their relative success. Some questions remain open for further scrutiny: Can the computer really generate randomness? How to deal with diagrams from the software (how to help to read them correctly, e.g., their scale)? How to ensure that the children have sufficient experience in proportional thinking – is fractions without reference to probabilities sufficient?

V. Kataoka (Brazil) reports about a series of workshops in teacher in-service. Teachers' difficulties – as noted in the presentation – seem to be surprisingly similar everywhere. The interactive approach, however, could well count to the more innovative approaches worldwide. In the rapporteur's memory is one special experiment used in the workshop, which illustrates sustainably the importance of suitable models and data sampled by randomness (when data really comes from *random* samples?)

We break a stick randomly into three pieces. After that, the subjects are asked to form a triangle of the three pieces. Finally the success rate is determined with which triangles actually could be formed. Try it with spaghetti – without explaining in advance what you intend.

Success rates of 75% are not rare. In contrast to it, there are (at least) two models for randomly breaking the stick (with 25% and ca 19% success probabilities). The obvious discrepancy between theory and model let us gradually start to doubt whether we could break the stick truly randomly into 3 parts. Conclusion: Relative frequencies are sometimes of no value to estimate an unknown probability. There may be a great potential in the clash of notions. Clearing the cloud and the tension are eagerly awaited in the class. Analogous examples are less emotionally laden than the spaghetti.

An interesting extra-mural activity is presented by H. Trevethan (Mexico) who describes a project in the context of a science fair. The concept of such fairs envisages that a group of students work on a project together with the aim of presenting this project to a wider public at the fair. Aside of the interaction with the public, a jury of experts evaluates the projects including the performance of the groups at the fair. There are awards in various disciplines to win. Of course, the activity aims at students with a special interest in the subject – it is not for the usual teaching in class. However, there are many advantages of the approach: The autonomous activity of the learners, their own responsibility, presenting in public etc. The students must be well prepared to face the imponderabilities of a live presentation. In the case at hand it deals with a game of chance “Shut the box”, which is certainly open to (stochastic) strategies. To elaborate on these strategies, to play against people from the audience with success, to finally explain these strategies to the curious audience, that was the task of the group. Changing the role from an (too often passive) learner to one who (actively) explains and is responsible for what goes on – that *is* of lasting effect onto learners. This authentic (and not faked) transfer of responsibility could well be more often also taken up in teaching in class. Mathematically, conditional probabilities and Bayes’ theorem are the key concepts to develop the winning strategy.

Conditional probability and Bayes’ Theorem

Conditional probability and Bayesian inference are important ingredients also of university teaching, including courses for non-mathematical studies. Many different types of errors have been investigated in isolation. According to C. Batanero and C. Diaz (Spain), a synthesis is missing. There is neither a study investigating connections between various types of misconceptions, nor an analysis whether misconceptions are related to mathematical knowledge, i.e., whether they decrease with better achievement in mathematics. Consistently, they develop a questionnaire with – familiar – items, test it and apply it to university students. Data are analyzed by means of factor analysis. They describe some phenomena, which remain with higher mathematics education, but in general a significant decrease in misconceptions is found with higher mathematics level. For interrelations between several misconceptions, the result is less optimistic as these misconceptions seem to be quite isolated with not many relations in between. As a consequence of this investigation, endeavour in mathematics education in probability has to be fostered while the types of misconceptions still have to be singly put to the fore in teaching again and again in order to extend students’ experience with them.

Most empirical studies into understanding conditional probability are done within the context of cognitive or psychological research. For example, the influence of the data format in the items on students’ solution behaviour is analysed.

The problem of diagnosis (especially within medical context) is one of the most often used: With known probabilities of positive and negative diagnostic findings for the disease and for being healthy in the population, one asks for the probability of having the disease in case one is tested positively.

P. Huerta (Spain) criticizes it as a serious flaw of research if it neglects the structure of the posed problems into account. Then one cannot generalize the found results to all kinds of problems with conditional probabilities. He describes the structure of “ternary problems of conditional probabilities” and classifies 20 different types of which only one subclass (and from it mainly one type of tasks) has been used in research. A graph with all problems is used to determine the grade of difficulty of a special problem at hand. By this structural analysis, Huerta develops a plan for future empirical research to cover all types of conditional probability problems to enhance the insight, which might be gained from it. In later stages of research he plans to extend the research from an analysis of single probands to classroom analysis in order to evaluate teaching interventions for their relative success.

Conditional probabilities are considered to be difficult; Bayesian inference is no less difficult. However, their role in a successful curriculum in stochastics is undisputed.

L. Martignon and S. Krauss (Germany) present a class experiment on this very topic with 10 (!) year olds. With the help of Wason cards, they manage to initiate learning steps in the children.

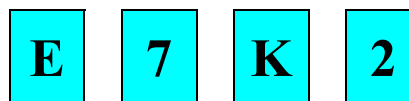


Figure 2. The original Wason cards

Which cards do you necessarily have to turn around in order to check, whether the following rule holds for the set of 4 cards? “If one side of a card exhibits a vowel, its other side must exhibit an odd number”

While within this “logical” context children solve the task badly, they improve with contexts closer to everyday tasks. Martignon and Krauss use 32 cards (equal numbers for all types) and let the children turn the “right” cards. In this statistical variant the children act quite successfully. The researchers move on to represent the Wason cards by tinker cubes, which can be put together to build towers. With two colours they represent the cards with the advantage that both sides may be seen simultaneously. In using such media they favour learning steps in proportional thinking, right from the beginning in connection to probabilities. They report on encouraging results from their pilot projects.

We all know about the pitfalls of the interpretation of results from statistical tests or from confidence intervals. These originate from the reduction of the interpretation of probability to situations, which may be repeated independently in the same manner on and on. On this issue there has been a debate not only in the foundations of statistics but also in the didactical community. Taking this as a starter, Ö. Vancsó (Hungary) decided to develop a parallel course in classical *and* Bayesian statistics. He tried his ideas and refined them in several cycles in teacher pre-service and reports about his positive experiences with the approach: “Now I have really understood what is meant by confidence intervals” one of his students expresses it.

Vancsó compares the debate, which approach is the better, with the geometry debate in the 19th century. We forget, however, the progress of modern mathematics, which is no longer devoted to the analysis of *absolute* true statements. Therefore, it is a false dichotomy, to teach *either* classical statistics *or* Bayesian statistics. Both theories offer a consistent theory of probability – the choice comes only with the application. His didactical motto is “You will understand a theory much better if you contrast it to another – especially to another, which is quite different”. Accordingly he works on a conception and materials for teaching both schools of statistics in *parallel*, without favouring either of the approaches.

Fundamental Ideas

With this complex of topics, the organizers intended to attract papers devoted to probabilistic ideas like “random variable – distribution – expectation”, or to the convergence of relative frequencies, or to the central limit theorem. A further topic on “Revising probabilities – Bayes’ theorem – independence” has been the target of research of many contributions that it formed a separate group.

R. Peard (Australia) in his approach to probability goes back to the roots motivating concepts by problems and questions from games of fortune. Interestingly, however, he does not argue with the potential to explain the notions much better by the context, in which they emerged. Games of chance have meanwhile been discredited by their closeness to combinatorics (which is not always easy to understand) and by their artificiality (we want and have to teach real applications to our students!). Instead he argues by the relevance of this topic for applications as games of chance have spread so

much nowadays (at least in Australia) that they developed to one of the most important business sector, which grows still at fast speed. It would therefore be necessary to get students to be familiar with this business and clarify wherein their chances really lie.

R. Kapadia (UK) presented tasks from the national attainment target tests and concludes from the bad achievement of students therein that teaching compared to 20 years behind has not really improved. This may be rooted back to recent trends as statistics, mainly simple data handling is favoured at the cost of probability in the curricula. And, as far it concerns probability, teaching still focuses too much and narrowly on equal likelihood and experimental probability, there is not enough work on misconceptions, risk as a concept related to probability is hardly anywhere tangled. Clearly, if people have to judge probabilities, they will have a strong bias towards “equal probabilities”, especially when they are (or feel to be) confronted with two possibilities. The fundamental idea of judging – subjectively coloured and supported by qualitative and objective information – probabilities and risk has still not found a sustainable form of teaching.

The panel discussion was devoted to “Fundamental Ideas in Probability Teaching at School Level”.

Y. Wu presented a panoramic view of the Chinese situation. School reality is centrally organized; only two (!) textbooks are in use in the whole country. Despite the clash in cultures, the approach towards probability is surprisingly similar to what we promote in Western countries. For details, the reader should look at the presentation, which is available from the homepage of the group.

R. Kapadia (UK) deepened his elaborations from the presentation and enriched the discussion with a list of behaviour, which should be explicit topics of instructional endeavour:

- People use their experience in order to judge probabilities incompletely and – even worse – in a haphazard manner.
- People have difficulties to judge very small and very high probabilities especially if these are connected to adverse consequences
- People are inclined to attribute *equal* chances to the – given or seen – possibilities, especially if there are two.
- People attribute probabilities and process these into new ones neglecting even the most basic rules (e.g., all probabilities sum to 1).

In her panel thoughts, J. Watson (Australia) focused on the relation between expectation (probability) and variation as well as on including risk in teaching probability:

- Variation is more fundamental than expectation: Experience about variation is earlier than an appreciation of expectation. To build up a language about randomness has also to include suitable terms for variation. Ideas of variation lead more directly to the concept of distribution.
- Links across the curriculum should be established from risk to decision making as risk is fundamental but badly understood: Risk is negatively connotated (eg., bad life style). Its perception is biased, especially small risks raise serious problems. Is clarifying risks really intended?

M. Borovcnik (Austria) outlines some peculiarities of stochastic thinking, which make it so different from other approaches:

- There is no direct control of success with probabilities – the rarest event may occur and “destroy” the best strategy.
- Interference with causal re-interpretations may lead a person completely astray.
- Our criteria in situations with uncertainty may stem from “elsewhere” and may be laden with emotions – Probability and divination have a common source in ancient Greece.

Considering such features of stochastic thinking, paradoxes like the stabilizing of relative frequencies while at the same time new events have full-fledged variability may not seem special. Faced with uncertainty, we tend to rely on personal views rather than rationality. Many aspects influence our perception of the situation and what we are prepared to accept as mathematical solution of it. One difficulty may also lie in a primitive attribution of an ontological character of probability to situations: Probability does not exist – it is only *one of many views* to reflect on phenomena of the real world.

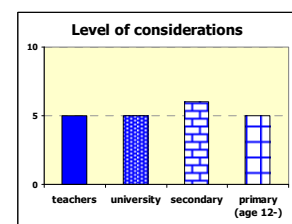
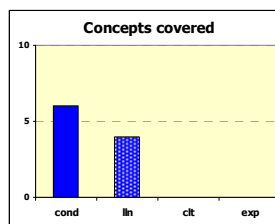
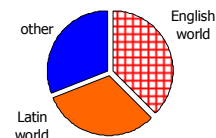
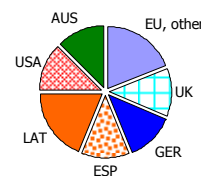
A general description of the contributions and a perspective for the future

The group could offer the possibility to present their results to 17 researchers selected out of 32 submitted papers after careful examination. The sessions at ICME were ordered to these topics:

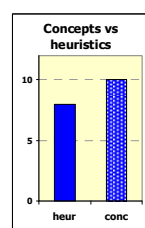
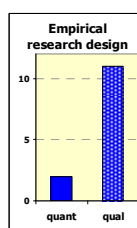
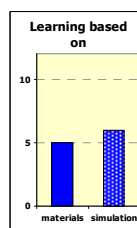
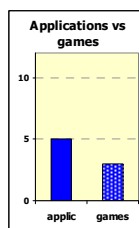
- Issues in Probability Teaching and Learning;
- Informal Conceptions;
- Conditional probability and Bayes' theorem.

The panel discussion completed the programme. The authors came from Europe, USA, Australia and Latin America, the English, the Spanish world, and the “rest” are distributed “evenly”. Some graphs illustrate the variety of approaches in the accepted papers.

Papers and PowerPoints of the presentations are available from the conference website. The hope is that ICME will continue to organize topic study groups on probability and statistics separately. By this “strategy” we did in fact split the potential audience as all the study groups are held at the same time slots. However, the great interest in our group on probability as well as the number of persons who attended the parallel statistics group confirm that we can attract many more people to our topic by two separate groups. The split into the two groups allowed also for a more convenient focus of the pertinent presentations and discussions.



Cond: conditional probability and Bayes; *lln*: law of large numbers; *clt*: central limit theorem; *exp*: expected value



Links

Topic Study Group 13 on “Probability” at ICME 11: <http://tsg.icme11.org/tsg/show/14>

Topic Study Group 14 on “Statistics” at ICME 11: <http://tsg.icme11.org/tsg/show/15>

Joint ICME/IASE study: http://www.ugr.es/~icmi/iase_study/

Author's address

Prof. Dr. Manfred Borovcnik
 Institut für Statistik, Alpen-Adria-Universität Klagenfurt
 9020 Klagenfurt, Austria
manfred.borovcnik@uni-klu.ac.at