INTUITIVE STRATEGIES FOR TEACHING STATISTICS

Manfred G. Borovcnik
Department of Mathematics, University of Klagenfurt
Sterneckstraße 17, A-9020 Klagenfurt, Austria

1. Trends in teaching stochastics

Statistical inference is a methodology which allows one to generalize findings from data. Thereby data can relate to a small subset, a sample of a bigger population; or they can relate to a data generating process which has been observed for a specific period of time. In both cases it is probability models and the concept of random sample which guarantees that the generalization can be done with some accuracy and with some degree of uncertainty. This is why statistical inference has been entirely based on probability theory, also in teaching - at least in France and in German speaking countries. However, probability is mathematically oriented, difficult, and it takes a long time to learn.

Simplifying probability theory. The increasing demand for skills to apply statistics required this problem to be addressed. Thus in the early 1980s there was an endeavour to simplify access to inference. i) By simplifying probability theorems through simulation, e.g. Watkins (1981) or Travers (1981). Probability is not developed as a mathematical theory for the student; instead those relations needed for inference are motivated by simulating the underlying assumptions. This enables a direct approach from descriptive statistics to inference. ii) By reducing probability theory to samples from a finite population, the nonparametric approach from Noether (1973). Probability is not developed as a theory for the student. The given data are multiplied by systematic permutations to yield the population. The inference then is based on whether the value of the test statistic for the given data is among the extreme values of this statistic for all the possible permutations. iii) By motivating methods through the context to which they are to apply, e.g. Holmes (1983) or Kapadia (1983). Many methods comply with common sense when they are bound to a specific subject matter rather than taught as abstract procedures.

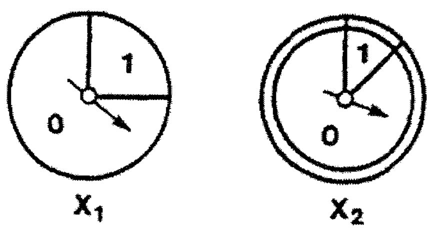
Reduction to data analysis. The late 1980s are signified by a movement back to basics, back to data analysis. The aim of this strand was to eliminate probability theory completely, at least in the earlier stages of teaching. It developed a new type of generalizing results from data which is not based on the random sample argument. The development goes back to Tukey's Exploratory Data Analysis (1977) which investigates data without a theory of probability. Herein, data are analyzed in a detective manner, from various angles, by multiple methods and in an interactive way between the results of intermediate analyses and the analyst's knowledge of the subject matter from which the data originate. Bibby (1983) and the 'Statistics in Society' course from the Open University were first to introduce this approach into Teaching Statistics. The Quantitative Literacy Project combined the exploratory approach with simulation ideas from 1986 onwards (Scheaffer, 1991). Recent developments, however, reduce this approach more and more to exploring data for its own sake: Gordon and Gordon (1992) summarize the state of the art.

Stochastic thinking. There is nothing to say against the exploratory approach; however, there is much to argue against reducing probability and statistics to it and neglecting the probability part of the whole. The psychologists Falk and Konold are the only critical voice in Gordon and Gordon (1992) who remind us that probability establishes a distinct approach towards thinking about reality. This approach is different from logical and causal thinking; it is difficult and at the same time important. Therefore teaching should clarify how it could contribute to solve real problems. The author will focus upon showing why this view on reality is important and how one can teach it successfully.

2. Specific problems of stochastics and its teaching

No learning by trial and error. Trial and error is a fundamental type of learning which accompanies us from our earliest days. This kind of learning is also true for mathematical concepts, even for the more difficult ones. Some theories of learning have integrated this fact. Piaget's operational stage of learning precedes the formal stage of understanding; more refined theories develop a mutual interplay between operations and reflections, see e.g. Dörfler (1984). According to this approach operations upon a material representation of the concept allow for feedback to the individual to revise his intuitions if necessary. Thus more refined forms of the concept may be established in the individual's mind. For example, the notion of number can be established from such concrete operations. It is not until very sophisticated levels that counterintuitive features emerge which are not based on concrete operations. Among such relations are: i) subsets can have the same cardinality as the reference set; ii) the rational numbers are countably infinite even if they have no minimal distance between two numbers and thus are not suitable for counting.

It is the author's thesis that this kind of learning is not possible for probability concepts even from the very basics. Thus, this phenomenon establishes the difficulties which subjects encounter with learning probability. An example similar to one from Falk and Konold (1992) will illustrate the fundamental problem that probability concepts reveal features which allow no direct feedback. In the diagram below are two spinners. The player has to choose one of the spinners and wins if the pointer will stop in the sector marked with a 1. Which wheel of fortune is the better choice? This simple question linked to the very basis of probability has no simple answer in terms of learning by trial and error. The choice of the worse spinner 2 may lead to success whereas the choice of the better spinner 1 may lead to failure. The latter event, however, immediately causes the player to speculate about the 'reasons' behind the failure and may obstruct progress in concept acquisition. Children are led astray by the special margin of spinner 2 (at the age of 7, this author's own experiment). The popularity of astrological predictions shows that such behaviour is not bound to children.



Clearly, longer series of spinning would allow for the feedback needed for concept acquisition. However, such series are usually not the focus of our experience. Furthermore, problems turn up again when one tries to generalize insight from repeated experiments to a single new outcome. It remains an open question how to supply effective experiences from which individuals can learn from trial and error that probability is a concept useful to apply to a single case because, basically, probability is a theoretical and not an experimental concept (see also Steinbring, 1991).

The gap between actions and reflections. There are various approaches to mathematics education which are based on an interplay between actions and reflections starting with Piaget; a modern version being Dörfler

(1984). These may be interpreted as reconstructing the learning process as intelligent trial and error. Actions upon a material representation of the concept in question form the ingredient of a process of reflection during which new objects will emerge from the concrete operations. Hierarchical steps of action - reflection - new object lead to the concept acquisition in an increasingly complex form.

The following simple example shows that probability notions are lacking in such an operational understanding. With coin tossing, the operation is predicting the outcome of the next toss whereas reflection means to evaluate the weight of heads (or tails). The result of the reflection process may yield the relative weights of 1:1, or the probability of 1/2 for heads. However, this result is of no use, at least no direct use for the prediction problem as the operation necessitates an exact and certain prognosis of the next outcome. For empirical evidence of this intuitive conflict see Falk and Konold (1992). Actions and reflections are based on different levels and the feedback between these different levels is missing for probability.

The basic difficulties in the coin tossing context may be best described by a mutual interplay between individual intuitions and theoretical demands. The individual witnesses an intuitive conflict right from the beginning as on the one hand he feels incapable of predicting the next outcome with absolute certainty and on the other hand he deeply desires control of that chaotic situation. At this point mathematics enters the stage and promises to calculate randomness. Any result of this theory will in due course be re-interpreted as a guide for action. Consequently, mathematical statements are either overinterpreted (like the law of large numbers used to derive the prediction of tails after five consecutive heads) or abandoned, reestablishing e.g. causal schemes for the prediction; see Falk and Konold (1992) for their dominance.

Mathematics offers a justification but yields no insight. Conceptual thinking can neither be reduced to mathematics nor to its applications. To *justify* concepts necessitates modes of thought different from those required to *understand* the concept. This might not be a problem as long as one remains within pure mathematics but things are different when one applies or teaches mathematics. Then it has to be brought into the open that a logically consistent argument to justify a concept cannot mirror all of its features. The concept of independence, for example, is mathematically reduced to the multiplication formula and becomes the basic ingredient of theorems about repeated random experiments like the law of large numbers or the central limit theorem. Independence thus gains an important role within theory, but its mathematical definition by no means affects the causal ideas individuals relate to it. In this context, Freudenthal (1983) used the phrase of mathematics as turning the things upside down.

Furthermore, the axiomatic structure of probability is not unique. The classical approach is based on the tendency of an experiment to produce stable relative frequencies whereas the Bayesian approach centres the theory around the idea of probabilistic judgements in the form of a weight of evidence. Thus, different axiomatic theories justify different intuitive ideas and do not cover the ideas of the other position quite well. Conditional probability for example, has a poor representation within classical probability. Now many of the historical puzzles and paradoxes are signified by an intermixture of intuitive ideas of both approaches. While it may be an interesting scientific programme to develop a selfcontained closed theory for one idea of probability, this programme does not yield insight into the true character of probability which remains a mixture of different ideas. With this in mind, Barnett (1973) tries to reconcile the competing positions of classical and Bayesian statistics.

3. Intuitive strategies for teaching statistics

It was argued that an operational understanding is lacking in probability as there is no direct feedback from the level of actions to that of reflections and vice versa. A far more promising approach to teaching statistics is by the interplay between intuitions and theory by Fischbein (1987). According to that, raw primary intuitions are revised by partial theoretical inputs into secondary intuitions whereby intuitions generally form the key to understanding and acceptance of the theory. The approach serves to describe the specific difficulties as well as serving to design teaching activities. For a description of the interplay in the coin tossing context see Borovcnik (1991).

Intuitive reformulation of concepts. Inner mathematical features of a concept are only one side of concept learning. Relating these features to intuitive ideas of individuals will help them to *understand* and *accept* theoretical results which is crucial for teaching to be effective. Some key concepts now double the difficulties as there are inadequate related intuitions *and* their mathematical representation is streamlined to a specific form according to a theory which is far from being intuitive. These difficulties will be illustrated by the Falk problem for the concept of independence and dependence. For intuitive difficulties with these concepts and the use of this problem in empirical research see Borovcnik and Bentz (1991). There are a lot of related puzzles in which teaching the usual mathematical representation will not effectively overcome the misconceptions, as the public discussion around Marilyn vos Savant and the Player's Dilemma shows (Morgan *et al.*, 1991).

The Falk problem deals with a simple urn with two white and two black balls. The urn is thoroughly shaken, two balls are blindly drawn out of it, one after the other without replacement. Consider two questions:

a) What is the probability that the second ball is white, given that the first ball is white?

b) What is the probability for the first ball being white, given that the second is white and the colour of the first is not known?

In what follows, the event W refers to white balls, B to black balls; the index describes the number of the draw. The usual solution, even if supported by tree diagrams will not change wrong causal intuitions because it is too formal (Borovcnik and Bentz, 1991):

$$P(W_1 \mid W_2) = \frac{P(W_1 \wedge W_2)}{P(W_2)} = \frac{P(W_1) \cdot P(W_2 \mid W_1)}{P(W_2)} = \frac{1}{3}.$$

Intuitively there is a big difference between the timely forward direction and the backward direction. A new urn 'causes' the new probability 1/3 in the forward question a) whereas the backward question b) is lacking in such an obvious urn. How to reformulate the representation to make the answer and the concept intuitively more accessible? Conditional probabilities are simplest to introduce via odds, i.e. relative probabilities or weights. The probability 1/3 e.g. amounts to odds of $1/3{:}2/3 \equiv 1{:}2$. A *conditional bet* on the first ball being white after the second is known to be white involves comparing the relative weights for W_1 in the face of the information W_2, i.e.

$$P(W_1 \wedge W_2) : P(B_1 \wedge W_2) = \frac{1}{2} \cdot \frac{1}{3} : \frac{1}{2} \cdot \frac{2}{3} = 1{:}2.$$

As odds of $a{:}b$ yield the probability $\frac{a}{a+b}$, the conditional probability in question b) is

$$P(W_1 \mid W_2) = \frac{P(W_1 \wedge W_2)}{P(W_1 \wedge W_2) + P(B_1 \wedge W_2)} =$$

$$= \frac{P(W_1 \wedge W_2)}{P(W_2)} = \frac{P(W_1) \cdot P(W_2 \mid W_1)}{P(W_2)}.$$

Now the symmetry between forward and backward question a) and b) is obvious. However, the full potential of this representation of conditional probability will not unfold until the Bayes' formula comes to the fore. In its simplest form Bayes' formula deals with two prior hypotheses H_1 and H_2 with prior probabilities $P(H_i)$, an experimental evidence E with known probabilities $P(E \mid H_i)$ under the two hypotheses. Bayes' formula then gives the posterior probabilities

$$P(H_i \mid E) = \frac{P(H_i) \cdot P(E \mid H_i)}{P(H_1) \cdot P(E \mid H_1) + P(H_2) \cdot P(E \mid H_2)}.$$

With odds this reads as

$$\frac{P(H_1 \mid E)}{P(H_2 \mid E)} = \frac{P(H_1)}{P(H_2)} \cdot \frac{P(E \mid H_1)}{P(E \mid H_2)}$$

posterior odds = prior odds × likelihood ratio.

Now one immediately sees that the posterior odds, the new weighting of the hypotheses are linearly dependent on the prior weighting *and* the relative likelihoods of the hypotheses under the empirical evidence. The odds representation of Bayes' formula shows the direct impact of the various inputs to the final probability which dissolves many a related puzzle. Furthermore, it shows that conditional probabilities are neither bound to time direction, nor to causal influence; they merely reflect an indication in summarizing the new state of knowledge in the form of a new probability. The reason for the usual approach to conditional probabilities being so clumsy is that the approach is done *completely within* classical probability whereas the concept is very natural within the Bayesian probability theory; odds re-establish this natural character of the concept. For further details and teaching consequences see Borovcnik (1992).

Teaching as clinical interview. A clinical interview is designed to reveal an individual's thoughts and thereby his/her comprehension of a mathematical concept. This author believes that such an approach is doomed to failure in its pure form as any question, any diagram, or any social interaction will establish a theoretical input for the individual and thus he will *react* to that. The individual will neither reveal everything of his understanding nor will he remain unaffected by an input. This dynamics, however, is not at all crucial in teaching where an intervention is intentional. The difficulty in teaching is, as was stated already, to connect the theoretical input to the pre-existing intuitive world of the individual. It was also argued, that a direct feedback between the operational and reflective level of a concept is missing, which obstructs learning by trial and error. Any teaching in probability is therefore faced with the problem of establishing links between theoretical demands and individual thought.

If such links are not supplied, the teacher might ask on the reflective level whereas the student answers at the operational level. For example the student might correctly evaluate the probability of 1/2 for heads but as this gives no clue for action, he/she might deliberately choose heads (or tails). If the teacher asks for a justification for that choice, the student might be completely confused because in his/her mind there is no such justification. The confusion finally might lead to a complete breakdown of the communication in class. Interviews might overcome this action reflection conflict.

Another link between theory and intuitions are urns and other games. They are a useful reference point which facilitates communication in class. Students may easily read off numerical probabilities, calibrate their individual scale of probability, and get insight into theoretical relations. Urns may therefore be perceived as partial images of a restricted theory which help to clarify intuitions. It is useful to map a real situation onto a suitable urn to deal with the questions. However, empirical investigations do show that intuitions about urns are by no means uncontroversial even in the simplest cases (Borovcnik and Bentz, 1991). If used as a medium between theory and intuitions, these inadequate associations have to be eliminated first. During this phase of teaching, feedback to students' thoughts is crucial and worthwhile. Teaching in the form of an empirical interview might best help with the multifaceted idiosyncratic ideas.

The following variant of the Falk problem should illustrate the potential of the interview approach to teaching. More than 50% of adult students confronted with the backward question b) give the answer 1/2 and justify it with causal arguments. To confront the inadequate causal scheme in this context *and* to establish the idea of new weighting of possibilities when further information is available, the author confronted students first with a third draw and the event second and third ball being white. As drawing is without replacement, it is then logically impossible for the first ball to be white. A first shock for the time dependent causal thinkers - there might be logical reasons to integrate *later* information into the weighting of *earlier* events. The harder causal thinkers needed more; the urn was changed to 4 white and 4 black balls. Again, first ball is unknown, second ball white - the answer was 1/2, the argument causal; second and third ball white - 1/2 again. The urn was changed to 50 white and 50 black balls and subsequently the following events were presented to the students: second ball white, second and third white, ... second to 49th white, second to 50th ball white. In the last case it is still logically possible for the first ball to be white but even the hardest causal thinkers had then become shaky. They recognized that the later information must have an impact on evaluating the first ball even if it is not logically cogent. The next step was to ask *when* this information began to exist - after 45 subsequent white balls, or after 6, or even when only one white ball in the second draw? Now students saw that the conditional probability $P(W_1 \mid W_2)$ has nothing to do with causal deliberations, but merely reflects the new state of information and that there was new information to express in this problem. This insight was facilitated by the interview teaching as neither students nor teacher would otherwise have known of each other's ideas.

Teaching by analogies. A conceptually difficult situation from probability is mapped onto a situation from another context in which the learner is experienced and can deal with problems and restrictions of the results. This can be extended to a teaching strategy; the potential will be illustrated in the context of betting and error theory.

Historically, the development of central concepts in probability come from studies of games of fortune. It is still possible to sharpen individual ideas of probability using this context. Betting is a side strand of today's curriculum and the fair prize of a game is a derived concept, namely the expected value. A basic difficulty with expected value is that it only makes sense for a long series of games. With such a series, the empirical mean will approximate the expectation. However, it is far from uncontroversial that one can deduce what is the fair prize for a *single* bet from that expectation. With odds, it is much easier. Odds are simply another representation of probabilities. If the odds are 1:2 for the winning condition then the stakes should amount to 2:1 which means that one has the possibility to win $2 for a stake of $1, the stake being lost if the winning condition fails to occur. For further details and more intuitive justifications see Borovcnik (1992).

It is important to state that with odds it is very natural to fix the stakes for a *single* game without solving the problem of predicting the outcome of this game. This might overcome the misconception of many individuals who accept probability only in cases where it predicts the exact outcome with certainty. Odds also serve for a new introduction into conditional probability and the Bayes' formula as was already shown here. The idea of the prize of the game makes various relations quite simple: if one plays two games then the prizes should add which yields the additivity of the expected value. For the binomial model situation of playing n times the same game with success probability of p, the single prize of p adds up n times which gives np for the expected value of the binomial distribution. The advantage of such analogies is that mathematical results may be perceived by the learner in advance and mathematics can focus on the issue of how the intuitive situation has to be made precise and which assumptions still have to be considered.

Error theory within geodesy was the second big stream of development for probability theory; the normal distribution and the method of regression are intimately related to it. The Bernoulli law of large numbers deals with the contraction of the distribution of the (theoretical) relative frequencies towards the unknown probability of the binomial experiment. This is proven by Cebychev's inequality with the help of the variance of the binomial distribution. This variance is tricky to calculate and the result is usually not well motivated, a very bad basis for the interpretation and acceptance of such an important theorem by the learners. There remains in the learner's mind a vague idea that the relative frequencies converge to the underlying probability and this forms the basis of misconceptions. An analogue law of large numbers holds for the mean of repeated trials.

In the context of error theory, this result can be intuitively anticipated. The analogy starts with the repeated measurement of a physical quantity a. Too large and too small numbers will eliminate each other if one calculates the mean of the single measurements which leads to the mean being a much better representation than the single measurement. To refer this to the measurement of an unknown probability p, the single measurements are 1 if the event occurs and 0 if the event fails to occur. Despite unusual values however, the mean coincides with the usual fraction of trials with the event 1, which is the relative frequency. Borovcnik (1992) gives details of how to develop a class experiment to get measurements and to study the precision of the repeated measurement. The longer the series the less scattered are the means (the relative frequencies); the procedure for measuring p tends to get more precise if it is based on more data.

This analogy avoids some of the basic misconceptions about the convergence of relative frequencies as it does not investigate a single development of the frequencies with increasing series; it focuses on the distribution of the means (the frequencies) with a fixed number of measurements and only then compares this distribution for different numbers of data. The growing precision is to be expected within error theory. Probability again gets a meaningful interpretation, without the precise prediction of a single outcome as one investigates the variation of measurement values (means) for different sizes of data. The process of making things more precise will lead to the binomial distribution, the central limit theorem, and confidence intervals. Again, it is important that the analogy has enough potential to prepare a way for learners to anticipate mathematical results. For more details about this and other analogies see Borovcnik (1992).

4. Summary

Probability is more a heuristic than a modelling description of reality. There is no direct feedback from reality which is often so helpful for the acquisition of concepts in other branches of mathematics. This is why intuitive ideas gain such an important role for the teaching and learning of probability. Intuitions can prove an obstacle to the acceptance and the comprehension of the concepts. There is an urgent need for a representation of mathematics which relates it to the ideas of the learners so that they can revise their intuitions wherever this is necessary. This paper illustrated three strategies to help an intuitive understanding of probability.

Bibliography

Barnett V. (1973), *Comparative Statistical Inference,* Wiley, New York.

Bibby J. (1983), EDA for 16-19's, in D.R. Grey, P. Holmes, V. Barnett and G.M. Constable (eds.), *Proceedings of the First International Conference on Teaching Statistics,* Teaching Statistics Trust, Sheffield, Vol. 1, pp. 277-281.

Borovcnik M. (1991), A Complimentarity between Intuitions and Mathematics, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics,* International Statistical Institute, Voorburg, Vol. 1, pp. 363-369.

Borovcnik M. (1992), *Stochastik im Wechselspiel von Intuitionen und Mathematik,* Bibliographisches Institut, Mannheim.

Borovcnik M. and Bentz H.-J. (1991), Empirical Research in Understanding Probability, in R. Kapadia and M. Borovcnik (eds.), *Chance Encounters: Probability in Education,* Kluwer, Dordrecht, pp. 73-106.

Dörfler W. (1984), Actions as a Means for Acquiring Mathematical Concepts, in *Proceedings of the Eighth International Conference on Psychology of Mathematics Education,* Sydney, pp. 172-180.

Falk R. and Konold C. (1992), The Psychology of Learning Probability, in F. Gordon and S. Gordon (eds.), *Statistics for the Twenty-First*

Century, MAA Notes, 26, The Mathematical Association of America, pp. 151-154.

Fischbein E. (1987), *Intuition in Science and Mathematics. An Educational Approach*, Reidel, Dordrecht.

Freudenthal H. (1983), *Didactical Phenomenology of Mathematical Structures*, Reidel, Dordrecht.

Gordon F. and Gordon S. (eds.) (1992), *Statistics for the Twenty-First Century*, MAA Notes 26, The Mathematical Association of America.

Holmes P. (1983), Developing Materials for the Schools Council Project in Statistical Education, in D.R. Grey, P. Holmes, V. Barnett and G.M. Constable (eds.), *Proceedings of the First International Conference on Teaching Statistics*, Teaching Statistics Trust, Sheffield, Vol. 1, pp. 269-276.

Kapadia R. (1983), A Practical Approach to Statistics, in D.R. Grey, P. Holmes, V. Barnett and G.M. Constable (eds.), *Proceedings of the First International Conference on Teaching Statistics*, Teaching Statistics Trust, Sheffield, Vol. 1, pp. 169-178.

Kapadia R. and Borovcnik M. (eds.) (1991), *Chance Encounters: Probability in Education*, Kluwer, Dordrecht.

Morgan J.P., Chaganty N.R., Dahiya R.C. and Doviak M.J. (1991), Let's Make a Deal: The Player's Dilemma, *The American Statistician*, 45, N. 4, pp. 284-287.

Noether G. E. (1973), *Introduction to Statistics: A Nonparametric Approach*, Houghton Mifflin Comp., Boston.

Scheaffer R. (1991), The ASA-NCTM Quantitative Literacy Project: An Overview, in D. Vere-Jones (ed.), *Proceedings of the Third International Conference on Teaching Statistics*, International Statistical Institute, Voorburg, Vol. 1, pp. 45-49.

Steinbring H. (1991), The Theoretical Nature of Probability in the Classroom, in R. Kapadia and M. Borovcnik (eds.), *Chance Encounters: Probability in Education*, Kluwer, Dordrecht, pp. 135-168.

Travers K.J. (1981), Using Monte Carlo Methods to Teach Probability and Statistics, in A.P. Shulte and J.R. Smart, *Teaching Probability and Statistics*, NCTM, Reston VA., pp. 210-219.

Tukey J. W. (1977), *Exploratory Data Analysis*, Addison Wesley, New York.

Watkins A.E. (1981), Monte Carlo Simulation: Probability the Easy Way, in A.P. Shulte and J.R. Smart, *Teaching Probability and Statistics*, NCTM, Reston VA., pp. 203-209.