Some fundamental ideas in probability

Borovcnik, Manfred
Alps Adria University of Klagenfurt, Department of Statistics
Klagenfurt, Austria
manfred.borovcnik@uni-klu.ac.at

In this panel discussion I want to focus your attention to three – of many – aspects of probability, which influence our individual capacity and the way how we perceive situations and their standard mathematical treatment – these are: thinking probabilistically, weighing the evidence, and the paradox of stabilizing and fluctuating of relative frequencies.

1 Thinking probabilistically

While mathematicians would define thinking probabilistically in terms of adequate use of probabilistic models, individuals are faced with the context of the situations to be modelled. And the ingredients of these situations could lead to directions completely different from standard mathematical models and their solutions. For didactical purpose it has to be clarified in which respects thinking probabilistically could be characterized. Some features are illustrated by context and comprise amongst others:

i. Feedback in probability items is indirect – you may win with the wrong strategy (Figure 1).
ii. Interference with causal perception might lead astray (Figure 2).
iii. Our criteria in probabilistic situations might be completely non probabilistic and emotionally laden (Figure 3).

The spinner with a special margin. Which is the better choice if the spinner is twisted and the amount in the landing sector is paid as prize? While even young children recognize that the left spinner in Figure 1 has a bigger winning sector and thus should be favourable, they might speculate about the special margin of the right spinner (Borovcnik & Peard, 1996).

This might provoke thinking about the deeper reasons of the special design and lead them to use non-rational arguments to justify their choice of the right spinner. Stochastic thinking is unstable even if mathematical prerequisites are available – a switch to magic thinking is quite probable not only with children.

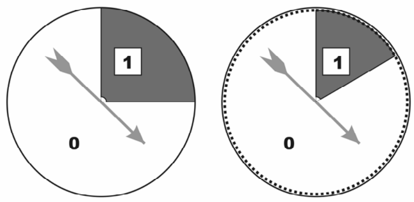

Fig. 1. Feedback is indirect – you may win with the right spinner with the smaller sector

The spinners have still another specialty as also with the best choice of the left spinner one is prone to lose in a single game, which gives a bad feedback on its justification and let the players change their strategy and favour the wrong (right) spinner – and then stay with the second choice (as a reaction of defiance for the frustration of the first choice which was correct but remained unrewarded). Or, worse, if the player won with the right (and a wrong strategy) how to convince this player of the idiosyncrasy of the argument. This irrational behaviour is not bound to children (see Borovcnik & Bentz, 2003, or Borovcnik & Bentz, 1991) as the many awkward justifications for the choice of numbers in the state lotteries show.

The Falk Urn. Draw twice from an urn with two white and two black marbles. The drawn marble is not replaced. Two questions are posed:

i. You see the first marble to be white. How big is the probability that the second marble will be white, too?

ii. The first marble is temporarily hidden from you and laid at the side. Then a second is drawn and proves to be white. What is the probability that the first marble is also white? (Falk & Konold, 1992).

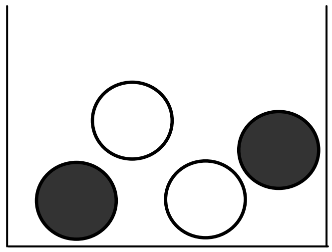

Fig. 2. Falk Urn with two white and two black marbles

While for the first problem there is an obvious urn with 1:2 (odds according to the current composition of the urn) for white whence the probability is 1/3, for the second problem a variety of personal strategies are popular. Especially an interference with causal perceptions is to be observed quite frequently. "As for the first draw there were 2 white and 2 black marbles, the probability has to 2/4 = 1/2 and the later events cannot influence the probability of the earlier." (Borovcnik & Bentz, 2003). This time-bound thinking is characteristic for the causal paradigm, the cause has to be first, the consequence later. The colour of the second draw cannot causally influence the colour of the first.

There is an urgent need of teaching strategies to make it plausible that conditional probabilities have nothing to do with time and causes – they amount to a wider concept to integrate information of whatever kind into the probability judgement of a statement. Such a concept has abundant applications eg., as diagnostic probabilities for medical diagnosing procedures or in judging circumstantial evidence in law.

The three doors – or Monty Hall. Choose one of the three doors (curtains), the main prize is hidden behind one of them (vos Savant, n.d.). Instead of simple guessing the candidates undergo the following procedure in this TV life game:

i. All doors are closed at the beginning;

ii. the candidate chooses on of them;

iii. the moderator talks around personal issues and what the candidate will do with the prize if … and then offers more information in opening an other door behind which a goat is bleating to indicate that the prize still is hidden behind the remaining doors;

iv. finally, the moderator offers the candidate to change the first choice.

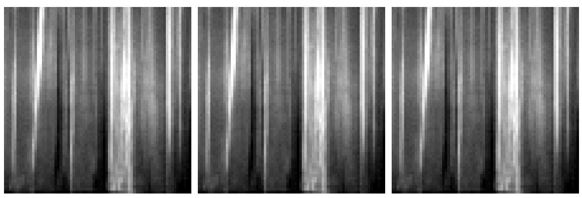

Fig. 3a. Monty Hall's three curtains with two goats and a car as main prize behind

Fig. 3b. Exemplars of a car and a goat to win

There is a vast literature on the topic, an easy solution is found in Gigerenzer (2002). The question is why do people neglect simple mathematical results and react so emotionally in their preference to stay with their first choice and defending it so irrationally while to change the first choice would increase chances to win from 1/3 to 2/3?

One of many explanations is a very personal and deals with being responsible for what one does. The first choice is simply luck or one may also name it as a God's decision to link to the roots of games of chance. The candidate is free of any responsibility for it, it is luck! However, in changing the first choice the candidate's personal involvement is substantial. Now with the personal choice after deliberating all ingredients of the problem personal responsibility enters the fore. Simply compare the two situations: To remain with a choice of a goat first and complain to "have no luck as ever" or to change doors and get the goat, which then is due to a silly decision made by oneself.

In the history of probability, games of chance have often been used to avoid responsibility. If a decision was too difficult and no one could bear responsibility for it, a game was used to outsource responsibility to a higher creature.

2. Weighing the evidence

Accepting that probabilities could help to model a situation, the individual has to estimate the probabilities of various outcomes. In weighing the evidence there is a clash between individual and general perspectives: This may be illustrated by the context of taking out a policy for a comprehensive insurance of one's car.

Taking out an insurance policy. Should one take out a full coverage insurance policy for the car for the next year or not? The focus hereby is to develop a scenario illustrating matters rather than to develop a minute model of the decision situation. Thus we model only two cases, namely "no accident" and "total wreckage" and a damage of 20000 in case of total wreckage and cost of the policy of 1000 (see Borovcnik, 2006, for details and further examples).

Cost [in Euro]		Decision	
		A_1 = insurance yes	A_2 = no insurance
Potential future	T_1 = no accident	1000	0
	T_2 = total wreckage	1000	20000

- The single person has to rely on personal judgement about his/her driving qualities and involved risk. It has also to include utilities of the various outcomes instead of money.
- An insurance company includes frequencies of accidents and the damage in money and not in utilities.

The probabilistic perspective is much easier to use from a general focus than from the individual perspective. Not only because the frequencies of accidents are much easier to estimate than it is to judge the qualitative knowledge about the person's behaviour. The latter is also more engulfed with self-reflection and utility. Here, the elicitation of personal odds for a total crash is difficult and vague and blurred by a variety of personal features (income, debts, risky or risk prone character etc.).

To free of such difficulties one might estimate an interval for the probability instead of an exact value, or to find the break even point for the probability where the decision switches from no insurance to take the insurance policy. In the latter case one would only have to judge whether the personal probabilities are higher or lower than this break even solution.

3. The paradox of stabilizing and fluctuating of relative frequencies

The paradox of stabilizing and fluctuating blurs a rational judgement of information contained in past relative frequencies. This implies that even hard-boiled information about frequencies could be perceived in a way, which it does not imply by fact. At least two related aspects are hard to grasp:

i. Analysis of the timely progression of the relative frequencies of an event gives the impression of stabilizing after a while – especially if it is accompanied by a graphical representation. The attained precision, however, holds only for continuing the series but is completely different in a series starting from anew: After e.g., 600 trials the relative frequencies suggest a fluctuation of 0.5 percentage points. A new experiment shows the same feature but on a different level, which is within roughly ±4% from the first run experiment (see Figure 4a).

ii. While the cumulative relative frequencies stabilize, the next say 5 trials still show their full fluctuation independent of the already attained level. Remarkably, however, the band of fluctuation narrows (roughly to half of the size) if one takes the next 20 trials and narrows generally at a rate of $1/\sqrt{n}$ with series of length n to predict (see Figure 4b).

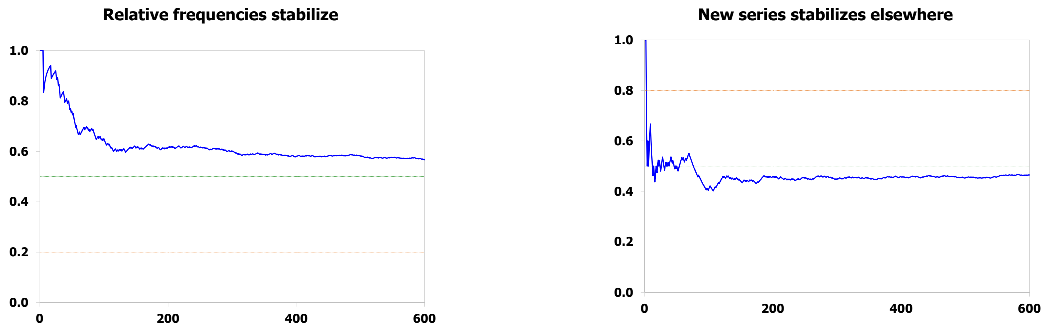

Fig. 4a. Each of the repeated series of experiments show a stabilizing effect for the relative frequencies – however the level is different for each series

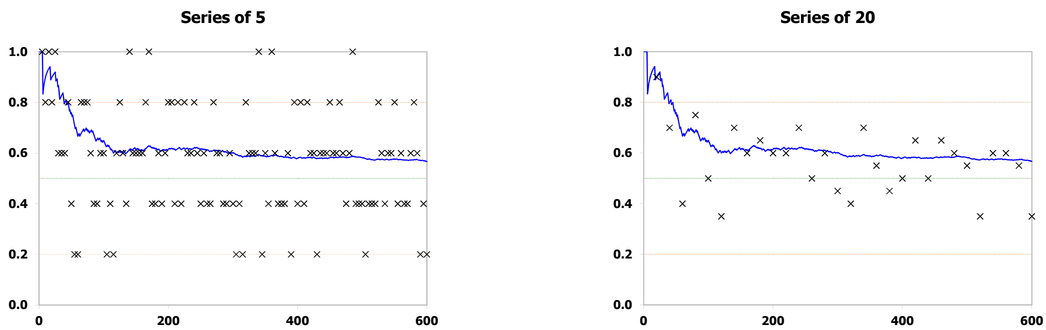

Fig. 4b. Despite the stabilizing effect, single series of next 5 values continue to fluctuate "irregularly" – the fluctuation decreases with predicting the next 20 values

This approach goes back to an idea formulated by Freudenthal (1972) and gives a better view on the law of large numbers. Instead of inspecting or even showing the convergence of the relative frequencies to an obscure limit, the rate of variation for predicting relative frequencies declines with larger size of samples. To be more precise, the rate of fluctuation decreases with the rate of 1 over $\sqrt{n}$.

4. Final remarks

For the final round of discussion, I want to contribute the following thought-provoking theses (see also Borovcnik, 2008):

- Randomness does not exist. It is only a form to think about the world. We have strong interrelations with other kinds of thinking, which might lead us into different directions from those predesigned by probability theory.
- Only in rare cases results from data handling speak for themselves and allow a clear message without referring to probability.
- The peculiarity of probabilistic thinking in contrast to logical, causal, or mystic thinking is important to elaborate on.
- To clarify the abundance of intuitive, personal thought on probability will help to get stable intuitions about probability and its potential.
- To clarify the mutual dependencies between frequentist, Bayesian, and mathematical conceptions and intuitive thought makes the concept flexible and robust.
- Basic notions of expected value, risk, or variability rely on sound conceptions of probability.
- The historical emergence of the concepts was embedded in games of chance. Still they provide a source for understanding and reference models for real situations
- A restricted primitive notion of probability as frequency in the long run neglects other qualitative source of information for weighing the evidence, which are valuable in many cases.
- Any inferential statistical method is intermingled with conditional probabilities and a sound understanding thereof.

Conclusion

Probability is much more often used in the sense of scenarios (Borovcnik, 2006) than it is used as a model, which is aimed to depict the most relevant parts of reality. The objectivistic framework for teaching causes serious problems for understanding (Carranza & Kuzniak, 2008). Following Kapadia & Borovcnik (1991), the author ascertains a need for a reference concept much wider than the frequentist approach can supply. This has to include subjectivist ideas, which base the mathematical approach towards probability on the idea of probability as personal degree of belief. As the numerical values of attributed probabilities get a subjective touch within this approach, it has been dismissed for science, especially as probability has eminent applications within physics.

However, first, the Bayesian school of probability (de Finetti, 1974) resumes any kind of objective information like relative frequencies from comparable past experiments or symmetries of an experiment to validate information about probabilities, second, it can also make use of qualitative information (eg, experience from engineers), which is costly to abandon in applications, and third, the procedures of probabilistic revising – as is done by the Bayesian formula – and statistical inference are quite clumsy and have their gaps in rationality if dealt with completely within an objectivist framework. Barnett (1973) was an attempt to reconcile the schools of objectivist and subjectivist probability but at times the interest of the various schools seemed too self-related to catch up this idea and thus a common approach is still awaiting elaboration. In applications, an informal way of using Bayesian models has been established without careful reflection about the status of the models and the related results.

An agenda for educational probability is Borovcnik (2008), teaching strategies from Borovcnik & Peard (1996) are still relevant. A strategy to use paradoxes to enhance the fundamental ideas like Székely (1986) might be worthy to extend. Vancsó (2008) marks a milestone in this direction.

References

Barnett, V. (1973). *Comparative statistical inference*. New York: Wiley.

Borovcnik, M. (2006). Probabilistic and statistical thinking. In M. Bosch (Ed.), *Proceedings of the Fourth Congress of the European Society for Research in Mathematics Education* (pp. 484-506). Barcelona: European Society for Research in Mathematics Education. Online: http://ermeweb.free.fr/CERME4/CERME4_WG5.pdf.

Borovcnik, M. (2008). A plea for a relatively strong role for probability within stochastic curricula. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. (2 pp.). Online: www.stat.auckland.ac.nz/~iase/publications/rt08/Panel2_Borovcnik.pdf

Borovcnik, M., & Bentz, H.-J. (2003, 1990). *Intuitive Vorstellungen von Wahrscheinlichkeitskonzepten: Fragebögen und Tiefeninterviews*. (Intuitive conceptions of probabilistic concepts: Questionnaire and in-depth interviews.) Technical Reports. Klagenfurt University.

Borovcnik, M., & Bentz, H.-J.: 1991, Empirical research in understanding probability, in R. Kapadia and M. Borovcnik (eds.), *Chance Encounters: Probability in Education*, Kluwer, Dordrecht, 73-106.

Borovcnik, M., & Peard, R. (1996). Probability. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International Handbook of Mathematics Education* (pp. 239-288). Dordrecht: Kluwer.

Carranza, P., & Kuzniak, A. (2008). Duality of probability and statistics teaching in French education. In C. Batanero, G. Burrill, C. Reading, & A. Rossman (Eds.), Joint ICMI/IASE Study: Teaching Statistics in School Mathematics. (5 pp.). Online: www.stat.auckland.ac.nz/~iase/publications/rt08/T1P2_Carranza.pdf.

Freudenthal, H. (1972). 'The empirical law of large numbers' or 'The stability of frequencies'. *Educational Studies in Mathematics*, 4, 484-490.

Gigerenzer, G. (2002). *Calculated risks: How to know when numbers deceive you*. New York: Simon & Schuster.

Falk, R., & Konold, C. (1992). The psychology of learning probability. In F. Sheldon & G. Sheldon (Eds.), *Statistics for the Twenty-First Century*, MAA Notes 26 (pp. 151-164). Washington D. C.: The Mathematical Association of America.

Kapadia, R., & Borovcnik, M. (1991). *Chance encounters: Probability in education*. Dordrecht: Kluwer.

Székely, G.J. (1986). *Paradoxes in probability and mathematical statistics*. Dordrecht/Boston: D. Reidel.

vos Savant, M. (n.d.). *Game show problem*. Online: www.marilynvossavant.com/articles/gameshow.html.

Vanscó, Ö. (2008). Parallel discussion of classical and Bayesian ways as an introduction to statistical inference. In Borovcnik M., IASE at ICME 2008. Research and Development in the Teaching and Learning of Probability. Auckland: International Association for Statistical Education (IASE). Online: www.stat.auckland.ac.nz/~iase/publications/icme11/ICME11_TSG13_16P_vancso.pdf.