ON THE DELICATE RELATION BETWEEN INFORMAL STATISTICAL INFERENCE AND FORMAL STATISTICAL INFERENCE

Rolf Biehler
University of Paderborn, Germany
biehler@math.upb.de

Informal inferential reasoning has become a topic of concern in statistics education. Laying intuitive grounds for more formal procedures or shaping the reasoning of younger students for whom formal inferential procedures will be too difficult to understand can be found among the aims. However, formal inferential reasoning as such is controversial itself: Bayesian and non-Bayesian ways of reasoning, Neyman-Pearsonian and Fisherian ways of hypothesis testing, confidence intervals. And: real applications of these procedures are context dependent. This raises questions with regard to which view of formal statistical inference we design preparatory informal inference activities for. The paper will critically discuss several approaches.

INFORMAL STATISTICAL REASONING AS A FIRST STEP TO FORMAL INFERENCE

Informal statistical reasoning has become an important topic in statistics education research (Makar & Rubin, 2009; Zieffler, Garfield, delMas, & Reading, 2008). When informal statistical reasoning or inference is the first step, an important question is what is the next step. Or, in other words, to which type of "formal" inference is informal inference the first step. Rossman (2008) expressed this concern from the perspective of statistics.

DIFFERENT APPROACHES TO FORMAL INFERENCE AND REAL INFERENCES

Statistical inference is controversial in itself. We can distinguish Bayesian and non-Bayesian ways of reasoning. Among the approaches to hypothesis testing we have to distinguish between Neyman-Pearsonian and Fisherian ways among the non-Bayesian approaches. The Neyman-Pearson approach interprets hypothesis testing as decision making, where as hypothesis testing in the Fisherian sense interprets low p-values as "evidence" against a null hypothesis. The interpretation and use of tests in experimental research is also debated: we face a significance test controversy (Batanero, 2000). Statisticians make a distinction between inference based on random samples versus inference based on randomization. High quality inference is only possible if specific conditions for the data collection had been taken into account (random sample, randomization). In practice, statistical methods are also applied even if these strong conditions are not met. To which extend this is legitimate is also debated. Zieffler, Harring, and Long (2011) discuss these different scenarios and what kind of inferences can be drawn by using randomization or permutation tests. Misinterpretations of test results are abundant (Gigerenzer, Krauss, & Vitouch, 2004; Haller & Krauss, 2002).

Another distinction is necessary. Ideal inference procedures are different from real inferences in practice, where contextual knowledge is important. It is seldomly the case that one hypothesis test in experimental science will provide a definite answer and decision. Usually the evidence provided by a test is combined with other evidences, and inference beyond the population from which the sample was drawn cannot be justified by statistics alone, but has to be tentative and theory-based.

FUNDAMENTAL IDEAS AS A GUIDELINE FOR PREPARING FORMAL INFERENCE

The idea of a spiral curriculum that is guided by "fundamental ideas" is an important design principle for curriculum design and should also guide the question how to prepare for formal inference. Burrill and Biehler (2011) developed an approach to fundamental ideas and distinguished fundamental ideas from the perspective of statistics from the perspective of applied probability (called stochastics in the European tradition). The latter approach was developed by Heitele (1975) and updated by Biehler (2014b), where as the companion paper (Biehler, 2014a) discusses fundamental ideas from the perspective of statistics.

In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014), Flagstaff, Arizona, USA.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

SOME DESIGN PROBLEMS

Before being confronted with ideas of inference, students will usually have made experiences in two contexts, in (exploratory) data analysis and in probability. The question is how we build on such previous experiences. When starting with EDA it is necessary to consciously introduce the distinction between sample and population and introduce the idea of “growing samples” (Bakker, 2004; Gil & Ben-Zvi, 2010). From the probabilistic perspective the following is important, as Burrill and Biehler (2011, p. 61) formulate: “Probability should not be taught ‘data-free’ but with a view towards its role in statistics. Probability models should be introduced as models to predict real data from random experiments and how empirical data may randomly differ from the theoretical distribution even if this distribution is assumed to be true.”

Starting from this twofold perspective, the long version of this paper will describe and review different learning trajectories that can finally lead to different types of formal inference. The thesis will be put forward that it might be better to postpone kinds of informal inference in earlier stages to later one in order to avoid the establishment of premature kinds of reasoning that cannot be build on in a later stage.

REFERENCES

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht: CD-β Press.

Batanero, C. (2000). Controversies around the role of statistical tests in experimental research. *Mathematical Thinking and Learning, 2*(1-2), 75-97. doi: 10.1207/S15327833MTL0202_4

Biehler, R. (2014a). Leitidee Daten und Zufall - fundamentale Ideen aus Sicht der Statistik. In H. Linneweber-Lammerskitten (Ed.), *Fachdidaktik Mathematik - Grundbildung und Kompetenzaufbau im Unterricht der Sekundarstufe I und II* (pp. 69-92). Seelze: Klett Kallmeyer.

Biehler, R. (2014b). Leitidee Daten und Zufall – Fundamentale Ideen aus Sicht der angewandten Stochastik. In H. Linneweber-Lammerskitten (Ed.), *Fachdidaktik Mathematik - Grundbildung und Kompetenzaufbau im Unterricht der Sekundarstufe I und II* (pp. Downloadbereich). Seelze: Klett Kallmeyer.

Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill & C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher tducation - A joint ICMI/IASE study: The 18th ICMI Study* (pp. 57-69). Dordrecht: Springer.

Gigerenzer, G., Krauss, S., & Vitouch, O. (2004). The null ritual - What you always wanted to know about significance testing but were afraid to ask. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social science* (pp. 391-408). Thousand Oaks, CA: Sage.

Gil, E., & Ben-Zvi, D. (2010). *Emergence of reasoning about sampling among young students in the context of informal inferential reasoning*. Paper presented at the Eighth International Conference on Teaching Statistics (ICOTS 8), Ljubljana, Slovenia.

Haller, H., & Krauss, S. (2002). Misinterpretations of significance: A problem students share with their teachers? *Methods of Psychological Research Online 2002 (Institute for Science Education), 7*(1), 1-20.

Heitele, D. (1975). An epistemological view on fundamental stochastic ideas. *Educational Studies in Mathematics, 6,* 187 - 205.

Makar, K., & Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal, 8*(1), 82-105.

Rossman, A. (2008). Reasoning about informal statistical inference: One statistician's view. *Statistics Education Research Journal, 7*(2), 5-19.

Zieffler, A., Garfield, J., delMas, R. C., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal, 7*(2), 40-58.

Zieffler, A., Harring, J., & Long, J. D. (2011). *Comparing groups: Randomization and bootstrap methods using R*. Hoboken, N.J.: Wiley.