IMPROVING RESEARCH IN STATISTICS EDUCATION

Djordje Kadijevich
Megatrend University and Mathematical Institute SANU, Serbia
djkadijevic@megatrend.edu.rs

Researchers in education tend to operate within their own narrow research circles ignoring the work of others in the field. Aimed at improving research in statistics education, this paper presents two requirements that should be respected by new research studies. These requirements deal with promoting the human face of statistics and utilizing research in mathematics education in research in statistics education. Each of them is briefly described and justified in general terms and illustrated with concrete examples.

INTRODUCTION

Researchers in mathematics education tend to operate in their own narrow research circles with an ignorance of the work of others in the field (Kilpatrick, 1987). One reason can be found in the field's high tolerance toward discursive diversity that has usually been misinterpreted as a license for ignoring other researchers' work (Sfard, 2005). Guy Brousseau once estimated that about 80% of the research in mathematics education "is reorganizing, reformulating, and problematizing work that has already been done" (Gjone, 1999, p. 51). Although the fraction may not be that large, high-quality research studies are still rather rare in education (Burkhardt & Schoenfeld, 2003). To improve the matters, some standards of educational research may be used (see Kadijevich, 2005, for basic standards in research in mathematics education).

The author of this study has followed the research activities and outcomes of various research circles in mathematics and statistics education (e.g., his papers have so far been accepted for a number of ICMI studies on diverse topics) and has witnessed, to a smaller of a larger extent, the above mentioned findings concerning the research practice and its outcomes. Aimed at improving research in statistics education, this paper presents two requirements that should be respected by new research studies. These requirements are briefly described and justified in general terms and illustrated with concrete examples.

TWO REQUIREMENTS

Promoting the human face of statistics

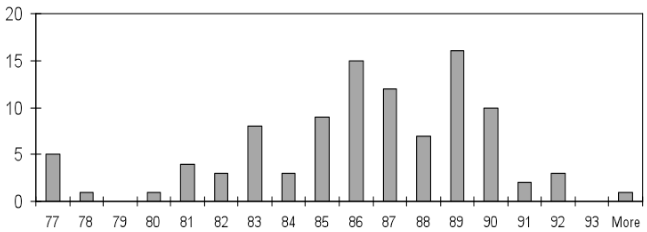
From the late 1980s mathematical teaching at the upper secondary level in Denmark has been required to include historical, structural and applicative issues in covered mathematical topics, which, when implemented appropriately, would, to use Phil Davis's words, display a human face of mathematics rather than an algorithmic, or computerized, or other view of the subject (Kadijevich, 1998). Despite differences in the features of mathematical and statistical knowledge, the learning of the latter has similar difficulties to that of the former (Kadijevich, Kokol-Voljc, & Lavicza, 2008). It is thus appropriate to require research to promote the human face of statistics by studying the historical, structural and applicative issues of statistical knowledge. Such an integrated approach to research in statistics education has not been applied to the author's knowledge. Mostly, structural issues and learning problems they generate have been widely studied (see, for example, Shaughnessy, 2007, for the state of this research).

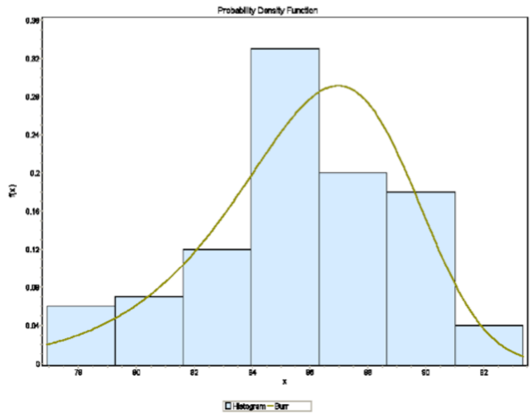
In order to illustrate how this integrated approach should be applied, let us consider the notion of distribution, keeping in mind that other important statistical constructs and procedures should be approached in that way.

First, undoubtedly, it is useful to know in what way the concept of distribution had been developed from the theory of errors (Bakker, 2004). However, it may be more important to know in what way the probability density function for normal distribution has been found (Stahl, 2006) and can be found today by an interested learner. Second, as regards the application of the concept of distribution, two issues are critical: (1) how one can apply an instance of distribution (i.e., the concrete distribution), and (2) how one can determine an instance of distribution to be used for some non-deterministic modelling. If one knows that the price of an item at the market varies

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

according to a normal distribution with M = 9.50$ and SD = 0.25$ (fictional data), it is very wise (the learner should find to what extent) to buy this item in a large quantity when its price is 9.10$ and sell all these items when the price in question is 9.90$. But, frequently, an appropriate instance of distribution is to be find first, and this is particularly relevant to non-deterministic modelling based upon simulations. Although many of the author's students have confidently use *MS Excel* for curve fitting (fitting the data with linear or non-linear curve), just few of them have been able to generalize this process to distribution fitting (finding an appropriate distribution for given data). Figure 1 shows a part of search for an appropriate distribution by using *EasyFitXL* (an *MS Excel* add-in available at www.mathwave.com). The things are more complex if a mixture of distribution is present.





Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
2	Burr	0.06531	1	0.42256	1	8.5299	27
60	Weibull (3P)	0.06667	2	0.45162	3	8.3519	19
37	Log-Logistic (3P)	0.06673	3	0.9481	10	11.939	35
31	Kumaraswamy	0.06777	4	0.45637	4	8.3421	18
59	Weibull	0.06925	5	0.43171	2	9.3606	29

$$f(x) = \frac{\alpha k \left(\frac{x}{\beta}\right)^{\alpha-1}}{\beta \left(1 + \left(\frac{x}{\beta}\right)^{\alpha}\right)^{k+1}}$$

k = 9.5333 , α = 30.878, β = 93.688

Figure 1. A part of search for a proper distribution from a detailed, puzzling histogram

Utilizing research in mathematics education in research in statistics education

As the previous subsection shows, research in statistics education may benefit from research in mathematics education by promoting the human face of statistics. Many outcomes of research in mathematics education may be used to improve research in statistics education. The following two paragraphs presents concrete examples.

Research in mathematical modelling has uncovered critical activities in a transition from one modelling stage to the next (Galbraith & Stillman, 2006), which enables teachers to help modelers overcome typical difficulties in the process of modelling. Both steps in *modelling cycle* (e.g. problem, model, solution, interpretation, evaluation; Galbraith & Stillman, 2006) and *statistical investigation cycle* (e.g., problem, plan, data, analysis, and conclusion; Makar, 2008) can be viewed as an instance of the following four steps: formulating, representing, operating and interpreting (the last three are according to Peschek and Schneider, 2001, the three basic steps in doing mathematics). Having in mind these four steps, researcher may uncover critical activities concerning a transition from one step to the next regarding his/her specific context (e.g. solving tasks with computer algebra system; Kokol-Voljc, Kadijevich, & Haapasalo, 2009).

According to Kadijevich, Kokol-Voljc and Lavicza (2008), the task-technique-theory is suitable for the learning of statistics, which should be examined in terms of task, technique, theory, and learner's profile, each influenced by instructional context (needs, values, learning support offered, etc.). *Task* involves objects to be learned, *technique*, which usually has both pragmatic and epistemic values, stands for technique for solving task (paper & pencil and, if available, its software variant), whereas *theory* deals with statistical theory learned, typical learner's misconceptions as well as mathematical knowledge built into technology applied (Artigue, 2002).

Consider, for example, task "What sample size should one use to predict the outcome of the forthcoming presidential election with an error of 1% or less?" to be solved with the *Fathom* software. Analyze techniques (paper & pencil and computer-based) and theoretical issues to emerge (have emerged) from solving this task for different learner's profiles (extrapolated from Kieran & Drijvers, 2006). Figure 2 presents an experimental approach with *Fathom* that should reveal that $SD_{max}= 0.5$ for $p = 0.5$ as well as that $N = 5{,}000$ is likely to give an error of about 1.4% provided that the student knows that function *stdDev()* computes the standard error (*SE*) of the mean and that $SD = SE\sqrt{N}$.

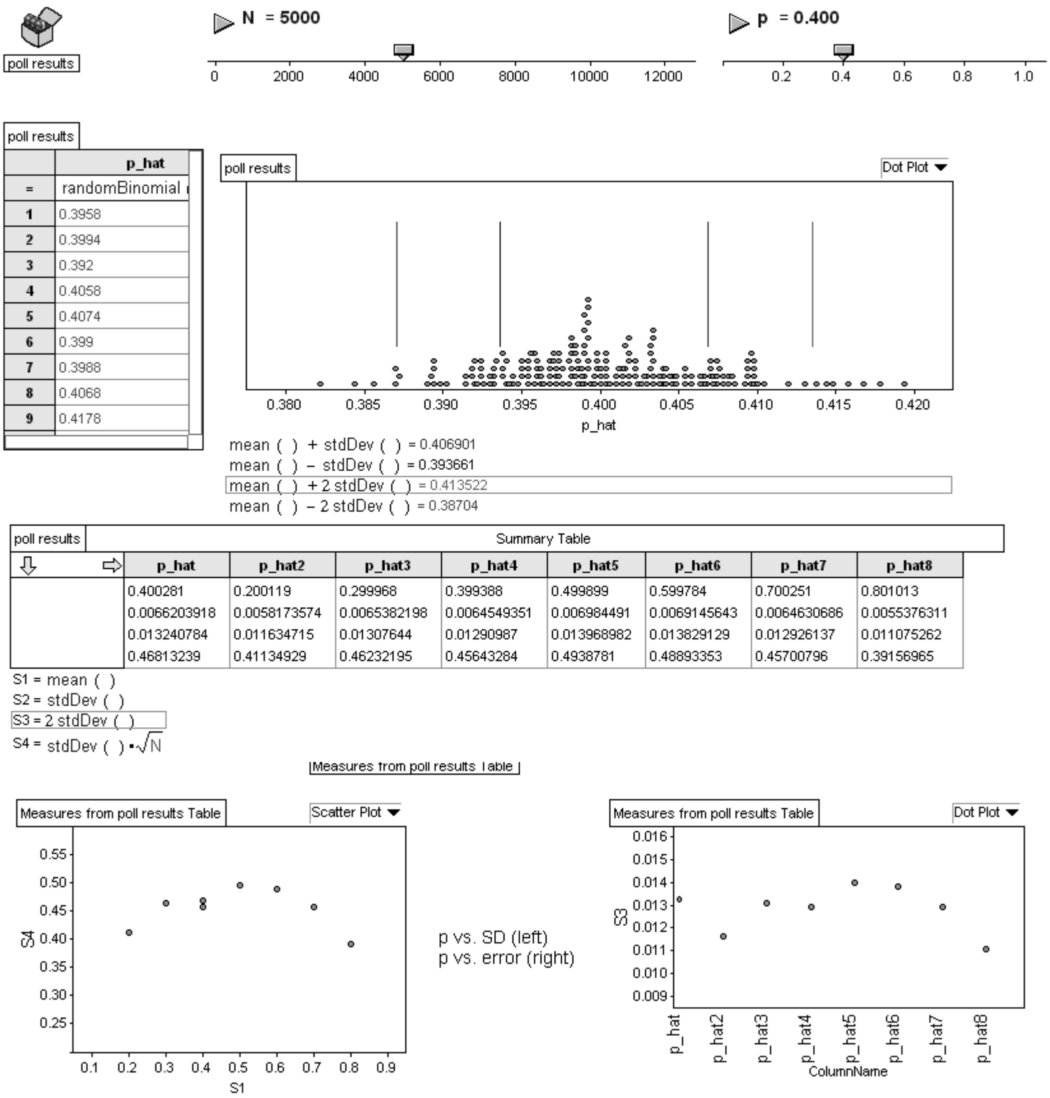


Figure 2. An experimental approach to polling voters problem

The previous two examples show how research in statistics education may be improved by using findings of research in mathematics education. The latter research may also beneficial from findings of the former (e.g., main components of construct's genuine understanding are interconnected with a complex web, Garfield & Ben-Zvi, 2005; understand both data and model used and, if needed, question and improve their quality). In general, relating research in mathematics education and statistics education would be beneficial for both fields for appropriately chosen content to transfer.

CLOSING REMARKS

Findings of educational research that (with rather small modifications) hold true in different research areas give more strength to these findings and the theory behind them.

[Mathematics advances its knowledge when, among other ways, a theorem proved for space *X* is then proved for different space *Y*, for example.] Requirements to promote the human face of statistics and utilize research in mathematics education in research in statistics education give good research directions towards improving research in statistics education and attaining a less fragmented picture of educational research as a whole.

ACKNOWLEDGEMENT

This paper results from the first author's work on projects 144032D and 144050A funded by the Serbian Ministry of Science.

REFERENCE

Artigue, M. (2002). Learning mathematics in a CAS environment: the genesis of a reflection about instrumentation and the dialectics between technical and conceptual word. *International Journal of Computers for Mathematical Learning*, *7*(3), 245-274.

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

Burkhardt, H., & Schoenfeld, A. H. (2003). Improving education research: Toward a more useful, more influential, and better-funded enterprise. *Educational Researcher*, *32*(9), 15–21.

Galbraith, P., & Stillman, G. (2006). A framework for identifying student blockages during transitions in the modelling process. *Zentralblatt für Didaktik der Mathematik*, *38*(2), 143-162.

Garfield, J., & Ben-Zvi, D. (2005). A framework for teaching and assessing reasoning about variability. *Statistics Education Research Journal*, *4*(1), 92-99.

Gjone, G. (1999). [Review of the book *Theory of didactical situations in mathematics: Didactique des mathematiques, 1970–1990*]. *Nordic Studies in Mathematics Education, 7*, 47–52.

Kadijevich, Dj. (1998). Promoting the Human Face of Geometry in Mathematical Teaching at the Upper Secondary Level. *Research in Mathematical Education*, *2*(1), 21-39.

Kadijevich, Dj. (2005). Towards basic standards for research in mathematics education. *The Teaching of Mathematics*, *8*(2), 73-81. Online: http://elib.mi.sanu.ac.rs/journals/tm/.

Kadijevich, Dj., Kokol-Voljc, V. & Lavicza, Z. (2008). Towards a suitable designed instruction on statistical reasoning: Understanding sampling distribution with technology. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Proceedings of the ICMI Study 18 Conference and IASE 2008 Round Table Conference, ITESM, Monterrey, México* (on CD). International Commision on Mathematical Instruction & International Association for Statistics Education.

Kieran, C., & Drijvers, P. (2006). The co-emergence of machine techniques, paper-and-pencil techniques, and theoretical reflection: a study of CAS use in secondary school algebra. *International Journal of Computers for Mathematical Learning*, *11*(2), 205 -263.

Kilpatrick, J. (1987). Editorial. *Journal for Research in Mathematics Education, 18*, 82.

Kokol-Voljc, V., Kadijevich, Dj., & Haapasalo, L. (2009). CAS-based task requirements and critical activities in completing them. In Dj. Kadijevich & R. M. Zbiek (Eds.), *Proceedings of the 6th CAME Symposium* (pp. 31-34). Belgrade: Megatrend University. On line: www.megatrend.edu.rs/came6.htm.

Makar, K. (2008). A model of learning to teach statistical inquiry. In C. Batanero, G. Burrill, C. Reading & A. Rossman (Eds.), *Proceedings of the ICMI Study 18 Conference and IASE 2008 Round Table Conference, ITESM, Monterrey, México* (on CD). International Commision on Mathematical Instruction & International Association for Statistics Education.

Peschek, W., & Schneider, E. (2001). How to identify basic knowledge and basic skills? Features of modern general education in mathematics. *The International Journal of Computer Algebra in Mathematics Education*, *8*(1), 7-22.

Sfard, A. (2005). What could be more practical than good research? On mutual relations between research and practice of mathematics education. *Educational Studies in Mathematics, 58*, 393-413.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1010). Greenwich, CT: Information Age Publishing, Inc., and NCTM.

Stahl, S. (2006). The evolution of the normal distribution. *Mathematics Magazine, 79*(2), 96-113.