WILL THE REAL BAYESIAN PROBABILITY PLEASE STAND UP!?

Egan J Chernoff
University of Saskatchewan, Saskatoon, Canada
egan.chernoff@usask.ca

Whether examined from a mathematical perspective or especially from a philosophical perspective, as Bayesian inference (probability) enters school mathematics curricula worldwide it definitely brings with it "some baggage." On its way to school, Bayesian inference (probability) will also pick up some baggage from the field of mathematics education (e.g., informal inference, "subjective" probability). Further adding to this baggage, mathematics and school mathematics are not one and the same. Through examining Bayesian inference (probability) from the perspectives of philosophy, mathematics, mathematics education and school mathematics (and popularization), this session will, ultimately, try to better establish what is meant by Bayesian inference (probability)...as it goes to school. In other words, "Will the real Bayesian probability please stand up!?"

Painting with a broad brush, the teaching and learning of mathematics (not in a "Math Wars" sense) is under attack. Public denunciations of the necessity of the teaching and learning of mathematics (e.g., Algebra II if you live in the United States of America) have been recently published by: political scientist Andrew Hacker (2012) in *The New York Times*, eminent biologist E. O. Wilson (2013) in the *Wall Street Journal*, and novelist/essayist Nicholson Baker in *Harper's Magazine*. However, perhaps to the surprise of those (too) close to the subject (e.g., mathematicians, mathematics educators and mathematics teachers), public condemnations of the necessity of the teaching and learning of mathematics are nothing new. As detailed by Baker, William McAndrew, in the 1900s; Henry C. Morrison, in the 1910s; John Franklin Bobbitt, in the 1920s; Arthur Dean, in the 1930s; Underwood Dudley, in the 1980s; Paul E. Burke, also in the 1980s; Michael Smith in the 1990s; Derek Stolp in the 2000s; Michael Wiener, also in the 2000s; and others, have all openly questioned various aspects of the teaching and learning of mathematics. Whether the criticism is new or old, the result is the same.

The mathematics (if you will) is simple: criticism equals ostracism. For example, the more recent condemnations have resulted in (with mere clicks of mice) individuals, such as Hacker, Wilson and (to a lesser extent because his article was behind a paywall) Baker being denounced by the mathematics (education) community through a flurry of social media activity (e.g., blogs, Twitter, and other social media platforms). We must, however, keep in mind that cyber ostracism is a relatively new phenomenon. Nevertheless, as detailed in Baker's (2013) article, Michael Smith, after his condemnation of the teaching and learning of mathematics, was on the receiving end of the silent treatment from his math department — the entire department. Beyond cold-shouldering, those who denounce the necessity of the teaching and learning of mathematics or the necessity of certain aspects of the teaching and learning of mathematics (e.g., Bobbitt, Burke, Dean, Dudley, McAndrew, Morrison, Smith, Stolp, Wiener and others) are often accused of biting the hand that feeds them – among many other accusations.

The main contention of my invited paper (6A1) for Session 6A: 'Bayesian inference (probability) goes to school: meanings, tasks and instructional challenges' – part of "Topic 6: Innovation and reform in teaching probability within statistics" at the 9th International Conference on Teaching Statistics (ICOTS9) – will read to many as yet another, albeit different, denunciation of certain aspects of the teaching and learning of mathematics. Yes, I am well aware of the potential ramifications (as I have detailed above): ostracism from the mathematics education community (online and otherwise), opposition and cold-shouldering from colleagues, being accused of "biting the hand that feeds me" and other allegations. Given, however, that the intended audience for my contention is individuals with a heavily vested interest in the teaching and learning of statistics, I take solace in the fact that my contention will not (necessarily) read as a denunciation of certain aspects of the teaching and learning of mathematics, which would result in ostracism, but, rather, as a contention worthy of discussion during both the formal and informal activities that comprise ICOTS9 and after the conference is over.

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

I contend, given the (past and) more recent denunciations of the teaching and learning of the mathematics, coupled with the recent move of statistics (further) into the mainstream (e.g., Nate Silver, *Moneyball*, “Big Data”), that calculus is currently perilously perched at the top of Mount School-Mathematics, its days are numbered and statistics is waiting in the wings as its replacement. Although not a new contention for me (see, for example, Reimann, 2013) or others (e.g., Benjamin, 2009), I wish, with this article, to further contend that as “Bayesian inference (probability) goes to school” it could play a pivotal role in the teaching and learning of statistics and statistics education, respectively, coming to (further) prominence in elementary, secondary and tertiary classrooms around the world and in the mathematics education research community. However, whether examined from a mathematical perspective or especially from a philosophical perspective, as Bayesian inference (probability) enters school mathematics curricula worldwide it definitely brings with it "some baggage." On its way to school, Bayesian inference (probability) will also pick up some baggage from the field of mathematics education (e.g., informal inference, "subjective" probability), the chasm that exists between mathematics and school mathematics and the popular notion of Bayesian inference (probability).

Inspired by Caleb Gattegno, particularly, “only awareness is educable,” and the potential for Bayesian probability, as it goes to school, to help supplant calculus with statistics as the “darling” of the school mathematics classroom, I, in the remainder of this article, detail (what I consider) the baggage associated with Bayesian inference (probability), hereafter referred to as Bayesian probability, as it “goes to school.” My presentation of different “pieces” of baggage, from a variety of different perspectives, is, ultimately, designed to represent the subject, that is, Bayesian probability, in a greater context. Baggage from the domains of philosophy, mathematics, mathematics education, psychology, school mathematics and popular culture are now commented on in turn.

PHILOSOPHY

The field of philosophy contributes baggage as Bayesian probability heads to school, with three (potential) particular philosophical pieces of baggage are now highlighted. First, as well-documented elsewhere (e.g., Gillies, 2000; Hacking, 1975), Bayesian probability is but one particular interpretation of probability. Further, the Bayesian interpretation of probability represents one of two dominant interpretations of probability – the frequentist interpretation of probability being the other. Thus, amongst what Gillies calls a “wide divergence of opinions about the philosophy” (p. 1) of probability, Bayesians and the frequentists represent the “Janus-faced nature” (Hacking, 1975, p. 12) of probability. Second, and compounding the baggage stemming from multiple interpretations of probability, there are different interpretations of probability within the Bayesian interpretation. Thus, as Bayesian probability goes to school, it has baggage stemming from whether or not we are discussing Bayesian probability, as Gillies notes, as a *general classifier* or a *specific theory*. Alternatively stated, Bayesian probability can be used as a general classifier for specific theories (e.g., Bayesian probability encompasses both the logical and subjective interpretations and others) or for a specific theory (e.g., subjective probability or logical or others). Should Bayesian probability, as it heads to school, denote a specific theory (as opposed to a general classifier) then, according to Chernoff (2008) and Chernoff and Russell (2014), a third piece of baggage, nomenclatural issues, are along for the ride. For example, each of subjective, Bayesian, belief-type and others names are all currently used as descriptors for Bayesian probability. (Will the real Bayesian probability please stand up!?) Loaded with philosophical Bayesian baggage, we now look for a reprieve in the field of mathematics where, despite any philosophical differences, “an almost complete consensus and agreement exists about the mathematics” (Gillies, 2000, p. 1).

MATHEMATICS AND STATISTICS

Despite a consensus about the mathematics, the same, unfortunately, cannot be said for mathematicians and statisticians, which results in further Bayesian probability baggage. In other words, the debate over different interpretations of probability in the field of philosophy is (rightly or wrongly) represented by a feud in the field of mathematics and statistics. Whether or not the feud really did exist or exists, at this point, is immaterial. Stories of offices in university mathematics and statistics departments being divided along party lines, whether true or not, have been cemented into the folklore of “Bayesians versus frequentists” and will, as Bayesian probability goes to school, definitely come along for the ride (case in point: some individuals may take umbrage with me not capitalizing Frequentist). Mathematics teachers, those who take the opportunity to humanize the subject for their students (e.g., Gauss in elementary school adding up the number from 1 to 100), are likely, once Bayesian probability is established in school, to draw upon the many stories associated with the feud between Bayesians and frequentists. With major advances in computing power, the fields of mathematics and statistics are rapidly moving away from classical statistics to the ability to conduct large-scale data analysis. Different approaches to dealing with “big data,” however, only further cement differences between different interpretations of and approaches to probability. For example, objective Bayesians (i.e., Bayesians whose priors stem from frequentist means) and others will head to school, which adds to the baggage. Fortunately, according to Chernoff and Russell (2014), a similar feud, despite prominent individuals’ implicit support for one interpretation over another, is not currently found in the field of mathematics education – where individuals advocate for an approach that utilizes the classical, frequentist and *subjective* interpretations of probability.

MATHEMATICS EDUCATION

Probability, over the past few years, has further cemented itself as a burgeoning area of research in the field of mathematics education (e.g., Biehler & Pratt, 2012; Borovcnik & Kapadia, 2009; Burrill & Elliott, 2006; Chernoff & Sriraman, 2014c). Inevitably, as Bayesian probability heads to school, particular researchers in the field of mathematics education will be provided with a unique opportunity to further investigations into probabilistic thinking and the teaching and learning of probability and to contribute to the current cannon of literature in the field (e.g., Borovcnik & Peard, 1996; Fischbein, 1975; Jones, 2005; Kapadia & Borovcnik, 1991; Piaget & Inhelder, 1975; Shaughnessy, 1992; Shulte & Smart, 1981). However, said investigations are not immune from Bayesian probability baggage.

For example, certain aspects of Bayesian probability, especially the modeling of beliefs (to establish priors), will provide researchers in the area probability education with a unique opportunity to connect different areas of research in mathematics education (e.g., modeling, affect, probability and others). However, Chernoff & Sriraman (2014ab) noted that mathematics education research topics, when utilized in research investigating probabilistic thinking, are used with (academic) impunity, that is, they are exempt from the academic scrutiny applied in other areas of research in the field of mathematics education. Thus, should the research to be conducted not adhere to the traditional academic scrutiny found in other areas of research in mathematics education when discussing, for example, modeling or affect, then Bayesian probability will gather even more baggage on its way to school.

A second piece of Bayesian probability baggage stemming from the field of mathematics education stems from the use of the term “subjective probability.” Worthy of note, subjective probability has been a staple of formative reviews of research in the field (e.g., Garfield & Ahlgren, 1988; Hawkins & Kapadia, 1984; Jones, Langrall & Mooney, 2007; Jones & Thornton, 2005; Shaughenssy, 1992). Chernoff (2008), in an examination of the state of probability theory specific to the field of mathematics education, concluded that the state of the term “subjective probability” is subjective and includes the inconsistent use of multiple terms, such as “subjective,” “Bayesian,” “intuitive,” “personal,” “individual,” “epistemic,” “belief-type,” “epistemological” and others. Moreover, Chernoff, working from Gillies (2000) distinction, also determined that the term subjective probability in mathematics education has dual meaning – “subjective probability” is concurrently used as a general classifier and as a specific theory. In an attempt to rectify the dual usage of the term, Chernoff suggested subjective probability remain the general classifier and a further distinction be made, more aligned with the distinction found in probability theory, for the specific theory. In further examining parallels between mathematics education and probability theory, Chernoff contended that subjective probability (the specific theory not the general classifier) in the field of mathematics education trended toward the “personal” (or “subjective”) interpretation found in probability theory rather than the “logical” interpretation. As a result he suggested that “subjective probability” be used as a general classifier while “personal probability” (or other terms) could be used for the specific theory.

Chernoff and Sriraman (2014a) suggested that the next period of probability education research may, one day, be defined as a renaissance period for psychological research in the field of mathematics education. This contention was derived from: research in the field of psychology having played a foundational role in probability education (e.g., Tversky & Kahneman, 1974); recent recognition of major developments in the field of cognitive psychology (e.g., Gilovich, Griffin, & Kahneman, 2002), which are associated with the original heuristics and biases research program of Kahneman and Tversky (e.g., Kahneman, Slovic, & Tversky, 1982); and, the more recent recognition and adoption, by researchers in the field of mathematics education, of the research on heuristics and risk conducted by Gerd Giegerenzer and colleagues (e.g., Gigerenzer, Todd, & the ABC Research Group, 1999). Should future research head in this particular direction, there is little doubt that Bayesian probability, as it goes to school, has and will play a central role in this research thread. In addition, should research head in this particular direction, further pieces of baggage will be accumulated. For example, Kahneman and Tversky (1972) explicitly declared (my italics):

> We use the term ‘subjective probability’ to denote any estimate of the probability of an event, which is given by a subject, or inferred from his behavior. These estimates are not assumed to satisfy any axioms or consistency requirements. We use the term “objective probability” to denote values calculated, on the basis of stated assumptions, according to the laws of the probability calculus. *It should evident that this terminology is noncommittal with respect to any philosophical view of probability*.

Based on the above caveat, coupled with the subjective probability baggage stemming from the fields of philosophy, mathematics and statistics and mathematics education, it appears, at this particular point, difficult to pin down what is meant when one uses the term subjective probability, which is eponymously used with the term Bayesian probability.

SCHOOL MATHEMATICS

Probability has, just recently, made its way into worldwide mathematics curricula (Jones, Langrall & Mooney, 2007). Thus, as Bayesian probability goes to school, we must be reminded that the objective and classical interpretations of probability are still getting their footing. Specific to Bayesian probability, the use of technology in the school mathematics classroom is a major piece of baggage that must be addressed. Recent developments in technology have resulted in huge advantages in computation in mathematics, which has allowed Bayesians (and others) to conduct large-scale data analysis and tackle previously “unanswerable” questions. However, the huge advances in computation in mathematics does not necessarily translate to school mathematics where many individuals are still engaged in a debate over whether or not calculators should be allowed in the school mathematics classroom. While certain individuals are clinging to the “being able to calculate the unit price in the grocery store” example as a reason for why calculators should not be allowed in the school mathematics classroom, other individuals, such a Conrad Wolfram, are advocating for a “Computer-Based Math” approach, which would allow teachers to stop teaching calculating and start teaching mathematics. For example, in his TED Talk, Wolfram (2010) argues that certain concepts in calculus are “amenable to a much younger age group” through computer-based mathematics. In a similar vein, adopting a computer-based approach to Bayesian probability, as it goes to school, would allow for a reduction in the current chasm that exists between mathematics and school mathematics and more closely align statistics with school statistics. Given the current technological climate that exists in a number of school mathematics classrooms, perhaps the real Bayesian probability will not be able to stand up.

DISCUSSION

The argument for supplanting calculus with statistics is based on the current zeitgeist. [Insert joke about the difference between statisticians and data scientists here.] For example, certain individuals such as, Nate Silver, David Spiegelhalter, Hans Rosling and others have, essentially, reached rock star status. The daily weather, sporting events and other regular occurrences are now discussed from an ever increasingly statistical perspective. Further, advances in technology are allowing individuals to “crunch” numbers to provide new perspectives. Slowly, but surely, statistics has become a staple of popular culture. A similar transition to the mainstream could, potentially, take place in school mathematics. However, for statistics to rightly take its place as the pinnacle of school mathematics, it would help if the real Bayesian probability identified itself as it makes its way to school mathematics. As demonstrated, though, the fields of philosophy, mathematics, mathematics education and school mathematics all provide Bayesian probability baggage that must be addressed. Then again, we all have baggage. Alternatively stated, Bayesian probability going to school is an opportunity. Bayesian probability going to school is an opportunity for researchers in the field of mathematics education to open new areas of investigation. Bayesian probability going to school is an opportunity for researchers in the field of mathematics education to reconnect with the psychological roots of their field. Bayesian probability going to school is an opportunity for teachers to continue to discuss Bayes’ theorem in their classroom. While I am not able to predict, based on the baggage that could be accumulated, if the real Bayesian probability will, in fact, stand up, I am confident that, ultimately, *a* version of Bayesian probability will go to school…how’s that for a prior…

ACKNOWLEDGEMENTS

I would like to acknowledge Professor Per Nilsson for the opportunity to be one of the invited speakers for Session 6A, ‘Bayesian inference (probability) goes to school: meanings, tasks and instructional challenges’. I am humbled and excited to be able to speak at ICOTS9. Thank you.

REFERENCES

Baker, N. (2013, September). Wrong answer: The case against algebra II. *Harper’s Magazine*, 31–38.

Benjamin, A. (2009, June). *Arthur Benjamin: Teach statistics before calculus* [Video file]! Retrieved from http://www.ted.com/talks/arthur_benjamin_s_formula_for_changing_math_education.html

Biehler, R., & Pratt, D. (Eds.) (2012). Probability in reasoning about data and risk [special issue]. *ZDM—The International Journal on Mathematics Education, 44*(7), 819–952.

Borovcnik, M., & Kapadia, R. (Eds.) (2009). Research and developments in probability education [special issue]. *International Electronic Journal of Mathematics Education, 4*(3). Retrieved from: http://www.iejme.com/032009/main.htm

Borovcnik, M., & Peard, R. (1996). Probability. In A. Bishop, K. Clements, C. Keitel, J. Kilpatrick, & C. Laborde (Eds.), *International handbook of mathematics education* (Part 1, pp. 239–288). Dordrecht: Kluwer.

Burrill, G. F., & Elliott, P. C. (Eds.) (2006). *Thinking and reasoning with data and chance: Sixty-eighth yearbook*. Reston, VA: National Council of Teachers of Mathematics.

Chernoff, E. J. (2008). The state of probability measurement in mathematics education: A first approximation. *Philosophy of Mathematics Education Journal, 23*. Retrieved from http://people. exeter.ac.uk/PErnest/pome23/index.htm

Chernoff, E. J., & Russell, G. L. (2014). Preface to Perspective I: Mathematics and Philosophy. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking: Presenting Plural Perspectives* (pp. 3-6). Berlin/Heidelberg: Springer Science.

Chernoff, E. J., & Sriraman, B. (2014a). Commentary on Probabilistic Thinking: Presenting Plural Perspectives. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking: Presenting Plural Perspectives* (pp. 721-728). Berlin/Heidelberg: Springer Science.

Chernoff, E. J., & Sriraman, B. (2014b). Introduction to Probabilistic Thinking: Presenting Plural Perspectives. In E. J. Chernoff & B. Sriraman (Eds.), *Probabilistic Thinking: Presenting Plural Perspectives* (pp. xv-xviii). Berlin/Heidelberg: Springer Science.

Chernoff, E. J., & Sriraman, B. (Eds.) (2014c). *Probabilistic Thinking: Presenting Plural Perspectives* (Volume 7 of Advances in Mathematics Education Series). Berlin/Heidelberg: Springer Science.

Fischbein, E. (1975). *The intuitive sources of probabilistic of probabilistic thinking in children*. Dordrecht, The Netherlands: Reidel.

Garfield, J. B., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implication for research. *Journal for Research in Mathematics Education, 19*, 44–63.
Gigerenzer, G., Todd, P. M., & the ABC Research Group (Eds.) (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.
Gillies, D. (2000). *Philosophical theories of probability*. New York: Routledge.
Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). New York: Cambridge University Press.
Hacker, A. (2012, July 29). Is algebra necessary? *The New York Times*. Retrieved from http://www.nytimes.com
Hacking, I. (1975). *The emergence of probability: A philosophical study of early ideas about probability induction and statistical inference*. Cambridge: Cambridge University Press.
Hawkins, A. S., & Kapadia, R. (1984). Children's conceptions of probability—a psychological and pedagogical review. *Educational Studies in Mathematics*, *15*, 349–377.
Jones, G. A. (Ed.) (2005). *Exploring probability in school: Challenges for teaching and learning*. New York: Springer.
Jones, G. A., Langrall, C. W., & Mooney, E. S. (2007). Research in probability: Responding to classroom realties. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909–955). New York: Macmillan.
Jones, G. A., & Thornton, C. A. (2005). An overview of research into the learning and teaching of probability. In G. A. Jones (Ed.), *Exploring probability in school: Challenges for teaching and learning* (pp. 65–92). New York: Springer.
Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology, 3*, 430-454.
Kapadia, R., & Borovcnik, M. (Eds.) (1991). *Chance encounters: Probability in education*. Dordrecht, The Netherlands: Kluwer.
Piaget, J., & Inhelder, B. (1975). The origin of the idea of chance in students (trans: L. Leake Jr., P. Burrell, & H. D. Fischbein). New York: Norton. (Original work published 1951.)
Riemann, S. (2013, January). Math-frolic interview #11: Egan Chernoff aka MatthewMaddux. *The Math-Frolic Interviews*. Retrieved from: http://math-frolic.blogspot.ca
Shaughnessy, J. M. (1992). Research in probability and statistics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465–494). New York: Macmillan.
Shaughnessy, J. M. (2003). Research on students' understanding of probability. In J. Kilpatrick, W. G. Martin & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 216–226). Reston, VA: National Council of Teachers of Mathematics.
Shulte, A. P., & Smart, J. R. (Eds.) (1981). Teaching statistics and probability: 1981 yearbook. Reston, VA: National Council of Teachers of Mathematics.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1131.
Wilson, E. O. (2013, April 5). Great scientist ≠ good at math. *The Wall Street Journal*. Retrieved from http://onlinewsj.com
Wolfram, C. (2010, July). *Conrad Wolfram: Teaching kids real math with computers* [Video file]. Retrieved from http://www.ted.com/talks/conrad_wolfram_teaching_kids_real_math_with_computers.html