

Comment on “Double Your Variance, Dirtify Your Bayes, Devour Your Pufferfish, and Draw Your Kidstogram,” by Xiao-Li Meng[☆]

ERIC D. KOLACZYK

I enjoyed reading Meng’s article, where we have on full display his characteristic insight and provocativeness laced with humor. Here I’d like to focus briefly on the topic of teaching, which runs like a thread throughout much of the article and is emphasized in the penultimate paragraph, where Meng writes of the importance that one, “[address] the issue of unreliable studies in the most fundamental and sustainable way, i.e. via education.” Indeed, if in being radical our ultimate goal is to truly effect change, then it’s particularly in our teaching that we should look to make inroads. I want to illustrate and support two of the points Meng makes in his article with cameos drawn from my own experiences.

In the last sentence of Section 1, Meng invites the reader to join him in “navigating between generality and particularity to inform and form a collective strategy for communicating and realizing the benefits of each proposal while containing and reducing its negative impact.” To my reading, this is practical statistics and data science in a nutshell. But in my experience, becoming skilled in such practice is something many students find quite challenging. For those that are pursuing undergraduate and higher degrees in statistics and data science, we have a unique opportunity, albeit one that it seems we could leverage better than we do to date. These are the students that can be expected to have both substantial “applied” training and substantial “theory” training. Traditionally, theory and applied have been looked at and taught as separate ends of a spectrum. But, as Michael Jordan has opined recently, this is a “false dichotomy”. Instead, these are two highly interrelated aspects of the same thing, joined together through principle.

In my own teaching, this perspective manifested one day some years ago in an advanced linear modeling course, when I wrote on the whiteboard, “Theory informs principle; principle informs practice.” This was for a class to which students arrived with multiple exposures to the application of linear regression to data, as well as first courses in linear algebra and probability. And the goal of this course was to revisit the topic of linear modeling with all of these experiences on the table, so as to understand “what’s under the hood” of this “car” they had already driven so often. Yet they struggled, compartmentalizing the material and the coursework into

“a theory problem” or an “applied problem” when in fact – as often as I could – the problems frequently asked them to utilize both aspects of their background together. My teaching changed, subtly, when I started emphasizing the role of principle, as informed by theory, in practice. And I’d like to think that student learning changed too (and for the better!) – certainly this phrase has become the most oft-cited that I hear back from my students. Eventually, this perspective went on to inform an entire new masters program, Boston University’s MS in Statistical Practice (MSSP) program. See Kolaczyk, Wright, and Yajima (2021) [1].

I was struck similarly where, in Section 6, Meng writes, “I am fully sympathetic to all deterministically trained minds struggling with stochastic realities.” This reminded me of an experience I had teaching a “Statistics for Poets” type of course at the University of Chicago in the 1990’s. Following design principles originally developed for the course by my colleague, Ron Thisted, and mirroring Chicago’s “Great Books” approach to education in the humanities and social sciences, the students’ entree to topics in probability and statistics was largely through reading and discussion, which then supported exercises in calculations and interpretations. For the first part of the course, I adopted a wonderful little book called *Mathematics and the Unexpected*, by Ivar Eckland. This is a particularly literary book (especially for mathematical exposition!) that attempts to convey something about how we formally represent (un)certainly, touching on a spectrum of relevant topics: from dynamical systems to probability to chaos. While students in general always seemed to enjoy exploring the interplay among these topics, I still recall one year where it became apparent during discussion of the chapter on probability just how shook one student had been by the reading. It turned out that this was truly one of Meng’s “deterministically trained minds” struggling with what appeared to be a first intellectual exposure to “stochastic realities”. And the notion that students could have such a visceral reaction to what I viewed then as purely intellectual topics was an eye-opener. It certainly informed the topic of discussion that day in class! And it has influenced my teaching forever since.

In closing, I’d like to thank Meng for being “radical”, in this article and elsewhere. To be honest, overall I don’t find what he wrote particularly radical, but I suppose that

[☆]Main article: <https://doi.org/10.51387/22-NEJSDS6>.

that perhaps says more about myself than anything else. In the era we find ourselves in particular, as one key part of a number of contributors to the still young and quickly evolving area of data science, it will be important for our arguably conservative field to be more “radical” ... in its admirably principled way!

Accepted 7 September 2022

REFERENCES

- [1] KOLACZYK, E., WRIGHT, H. and YAJIMA, M. (2021). Statistics practicum: placing ‘practice’ at the center of data science education. *Issue 3.1, Winter 2021*.

Eric D. Kolaczyk. Department of Mathematics & Statistics,
McGill University, Canada.

E-mail address: eric.kolaczyk@mcgill.ca