

From Statistically Significant... to Significantly Statistical

IMS President
Xiao-Li Meng writes
another President's
Column:

We statisticians have successfully—perhaps too successfully—taught everyone that the larger the size, the higher the power to lend

credence to an alternative. This is evident from the 2017 *Nature Human Behaviour's* “Redefine Statistical Significance,” which has over 70 authors, and from the 2019 *Nature's* “Retire Statistical Significance,” with its more than 800 signatories. The statistical community's organized responses regarding the troubled p -value have been led most visibly by American Statistical Association (ASA), via the 2016 ASA's Statement on p -Values, the 2017 ASA Symposium on A World Beyond $p < 0.05$, and the post-symposium special issue in *The American Statistician* (TAS 2019), with its 43 articles on what do to in a world in which p -value has been de-valued.

Given the increased attention to the issue of replicability, what can IMS contribute to the larger conversation? Inspired by a predecessor, I have a somewhat unusual idea, which requires your thoughtfulness in order to be consummated. So please, read on.

If the number 43 is too large for you (because you have taught many that $n=30$ is a good approximation for $n=\infty$ under normal circumstances), the editorial of TAS 2019 by Wasserstein, Schirm, and Lazar is a gentle and humble tour guide. It summarizes the key recommendations by an ATOM: “Accept uncertainty. Be thoughtful, open, and modest.” Indeed, the thoughtfulness and modesty of our profession are well-reflected by the very fact that many statisticians endorse the call to abandon the term “statistical significance.” I have yet to identify another discipline with quite so many members who endorse the idea of abandoning its publicly most-recognized concept.

To a layperson, saying something is “statistically significant” is analogous to saying it is “mathematically proven” or “scientifically valid.” Such colloquial associations are in fact what motivated the call to abandon the term “statistical significance,” because the methods behind it are far less rigorous than mathematical



Photo: Martha Stewart

proofs, and far too simplistic for establishing scientific validity. Yet we should not overlook the epistemological effectiveness of such confidence-inducing terms in promoting and sustaining the public awareness and appreciation of the societal relevance of a discipline (e.g., mathematics) or a collection of them (e.g., science). As Aristotle reminds us, our expectations of absolute exactitude should be qualified when it comes to matters of human opinion and action.

The question, then, is what alternative statistical concept could conceivably maintain the virtues of “statistical significance” without much of its vice? How about we simply drop the word “significance”? Just as we question if a finding is *scientific*, a study is *ethical*, a project is *economical*, an action is *legal*, or a policy is *moral*, we can—and should—ask of any study, “Is it *statistical*?” While the concepts of being *scientific*, *ethical*, *economical*, *legal*, and *moral* are endlessly contested, they have considerable use as yardsticks in both common and specialized parlance. Experts and laypersons alike may ask “Is it X?” with the term “X” signifying what something is or is not. The point is not to lay down incontrovertible definitions but rather to open up questions about what “X” is. Indeed, the lack of such routine questioning would itself be a troubling sign for a society or a historical period.

I dare to suggest that in the light of the dramatically increased societal attention to data science, we should promulgate the use of “statistical” as a yardstick. “Unstatistical” studies can do much harm to our societies in both the short and long term, just as unethical studies or uneconomical projects can. The concept of *being statistical* will not be any more perplexing than any of the concepts mentioned above, and its pithiness will enhance its effectiveness in public discourse and research communications, as well as in private conversations. IMS, as the world's leading learned society in foundational thinking and the building-up of statistics and probability, can play a vital role in framing its core rhetorical components. Indeed, to the best of my knowledge, “Is it statistical?” was first posed by Bernard Silverman, 2000–2001 IMS President (in a private conversation years ago), as a parallel to the question “Is it legal?” or “Is it ethical?”

In the spirit of “casting stones to attract jades” (抛砖引玉 in Chinese), I list below my proposal on the virtues of being statistical, the practice of which should help to reduce the prevalence of irreplicable research findings. I purposefully set the bar high in order to provoke, and hence, I would be happy to praise a study

as being “significantly statistical” if it demonstrates—with due diligence—all of the following virtues, as called for by the purposes and design of the study:

- ☑ Discuss the collection, pre-processing, quality and limitations of the data, and the implications of these;
- ☑ Elucidate, assess, and discuss data analysis and modeling assumptions, as well as their consequences;
- ☑ Investigate and evince a good understanding of selection biases, confounding factors, and when/whether causal conclusions can be drawn;
- ☑ Exhibit coherent probabilistic thinking and treatments of multivariate relationships and distributions;
- ☑ Apply statistical methods with reasonable justifications and acknowledge their shortcomings;
- ☑ Conduct appropriate uncertainty propagation, quantification, and representations;
- ☑ Show good understanding of statistical principles, such as conditioning and the bias–variance trade-off.

A list of virtues can never be exhaustive. There are also other virtues that are critical for data science, but they are not purely or primarily statistical considerations. For example, it is a virtue to understand trade-offs between statistical and computational efficiency, to ensure computational stability and scalability, to consider carefully policy implications, and to describe the essential scientific background, etc.

An invitation to you

My list here is only an invitation for IMS members to contemplate what should be the core considerations of “statistical” or “significantly statistical”. I would greatly appreciate hearing from you. Please either comment on the online version (at <http://bulletin.imstat.org>) or send your thoughts to meng@stat.harvard.edu as I prepare for my IMS Presidential Address at JSM 2019.

Of course, I’d appreciate it most if we all can practice what we preach, by constantly asking ourselves, “Is my study *statistical*?”

OBITUARY: Kimiko Bowman

1927–2019

KIMIKO OSADA BOWMAN, age 91, passed away on January 13, 2019.

Kim immigrated to the USA from her native Japan in 1951. In the course of only five years, she completed an undergraduate degree in mathematics and chemistry at Radford College, and MS and PhD degrees in statistics at Virginia Tech. Many years later, she was also awarded a doctorate in mathematical engineering from Tokyo University. Her close and active collaboration with L.R. Shenton, her PhD advisor at Virginia Tech, focused on the distributional properties of estimators based on non-normal data, and continued for 45 years.

Kim was a member of the scientific staff of Oak Ridge National Laboratory for 50 years, and remained active as a researcher and collaborator with ORNL staff for many years after her retirement in 1994. Kim is fondly remembered by colleagues

at ORNL as a tireless, enthusiastic and dedicated researcher. She authored or co-authored three books and more than 200 articles during her career. She was the recipient of many awards, was a fellow of the American Statistical Association and the American Association for the Advancement of Science, and was an elected fellow of the International Statistical Institute and the Institute of Mathematical Statistics. Her remarkable career was featured in “Statisticians in History,” a special issue of *Amstat News* (September, 2008).

A victim of polio herself, Kim took an active leadership role in advocacy for individuals with disabilities. She served on the National Science Foundation Equal Opportunities for Science and Engineering Advisory Committee, and chaired the NSF Committee on People with Disabilities. She also chaired the Statistical Tracking of



Kimiko “Kim” Bowman

Employment of People with Disabilities Task Force for the President’s Committee on Employment of People with Disabilities.

Kim is survived by a son, Robert Noah Bowman and spouse Cheryl, two grandsons, and one great grandson.

Donations in her name can be made to the Kingwood Church, MAPS Honduras Alliance, 100 Harvest Way, Alabaster, AL 35007.

Max Morris, Iowa State University

This obituary first appeared in the International Statistical Institute’s online “In Memoriam” section. It is reprinted with permission.