Recent History of a Statement, a Special Issue, and a Controversy

Nicole A. Lazar

Department of Statistics
University of Georgia

Some Early Signs

Ioannidis (2005): Why Most Published Research Findings are False. Other papers follow, that point out misuses and abuses of statistical methodology.

Trafimow (2014): *Basic and Applied Social Psychology* will no longer require null hypothesis significance testing procedures (NHSTP) or inferential statistics more broadly.

Trafimow and Marks (2015): *Basic and Applied Social Psychology* bans the use of p-values, as well as "all vestiges of the NHSTP (p-values, t-values, F-values, statements about 'significant' differences or lack thereof, and so on)."

A Replicability/Reproducibility Crisis?

Concurrently, problems with replicability and reproducibility in various areas of science (neuroscience, social psychology, cancer trials, ...) gain more attention, including in popular press (e.g. *The New Yorker* (2010): The Truth Wears Off).

Open Science Framework Reproducibility Project (2011, psychology) finds low levels of reproducibility in published research. Ongoing project in cancer biology (2019).

Crisis in science? What is the role of statistics – and specifically the use of thresholds such as $p < 0.05$ for publication? P-hacking, researcher degrees of freedom, many signs of abuse of statistical methods in the literature.

ASA Statement on p-Values – 2016

Against this background, ASA convened a committee of experts to discuss use/misuse/abuse of statistics. Statement on p-values was the result.

Approached by ASA Executive Director: Would *The American Statistician* (TAS) publish the ASA Statement on p-values?

Q: How to make this a real TAS paper, not just a policy statement?

A: Invite discussion! Eventually many of the members of the ASA Committee contributed pieces to the online discussion.

Statement and associated discussion brought a greater awareness (to me as well) of general issues of: use and abuse of p-values; replicability crisis; reproducibility crisis.

Symposium on Statistical Inference – 2017

Aim was to continue the conversation started by the ASA Statement.

Brought together statisticians and researchers from a wide range of disciplines, with a wide range of perspectives.

Is there a problem with how statistical inference is commonly conducted? How serious is it, if so? What are the possible solutions?

Symposium on Statistical Inference – 2017

Many differing ideas for how to move forward, including

1. Lower "standard" p-value threshold (for declaring "statistical significance") from 0.05 to 0.005 (for example)
2. Pick and "justify" an α level on a case-by-case basis
3. Bayesian approaches
4. Alternative configurations of p-values
5. And many others ...

A Special Issue of *The American Statistician* – 2019

Following the Symposium, ASA worked with TAS to put together a Special Issue, Statistical Inference in the 21^{st} Century: A World Beyond $p < 0.05$.

Combination of invited papers from participants in Symposium, and papers received through an open call to the community.

Open access, online only, freely available in perpetuity to everyone.

Issue includes 43 papers, with suggestions on how to move forward in: education, publication practices, statistical inference more generally.

A Snapshot in Time . . .

As of May 15, 2019 the ASA Statement on p-values had 315,539 views.

As of May 15, 2019 the lead editorial to the Special Issue had 83,202 views.

Much conversation taking place in the pages of journals, at conferences, among statisticians, between statisticians and scientists, between statisticians and science writers. How do we (collectively) get to a "post $p < 0.05$" world – and do we *want* to? (The controversy . . .)

Stay tuned for more developments!