Distributions for Parameters

Nancy Reid University of Toronto

Monash University

Feb 21 2020



Reproducibility and Statistical theory



David Spiegelhalter @d_spiegel y

This paper motivates the call for the end of significance. A 25% mortality reduction, but because P=0.06 (two-sided), they declare it 'did not reduce' mortality. Appalling. jamanetwork.com/journals/jama/...

Research

JAMA | Original Investigation | CARING FOR THE CRITICALLY ILL PATIENT

Effect of a Resuscitation Strategy Targeting Peripheral Perfusion Status vs Serum Lactate Levels on 28-Day Mortality Among Patients With Septic Shock The ANDROMEDA-SHOCK Randomized Clinical Trial

Monash Feb 2020

- comparing two treatments for septic shock
- randomized clinical trial
- estimated hazard ratio 0.75 [0.55, 1.02]
- 2-sided p-value 0.06

after adjusting for confounders

34.9% vs 43.4% unadjusted

- Discussion: " a peripheral perfusion-targeted resuscitation strategy did not result in a significantly lower 28-day mortality when compared with a lactate level-targeted strategy"
- Abstract: "Among patients with septic shock, a resuscitation strategy targeting normalization of capillary refill time, compared with a strategy targeting serum lactate levels, did not reduce all-cause 28-day mortality."

• 2014: Basic and Applied Social Psychology published an editorial banning p-values

actually "null hypothesis significance testing"

• "prior to publication, authors will need to remove all vestiges of the NHSTP ... *p*-values, ... , statements about 'significant differences' or lack thereof, and so on"

"confidence intervals are also banned"

- 2014: *Nature* published a News Feature by R. Nuzzo: "*p*-values, the gold standard of statistical validity, are not as reliable as many scientists assume"
- 2016: American Statistical Association released a public statement on statistical significance and *p*-values



732 North Washington Street, Alexandria, VA 22314 • (703) 684-1221 • Toll Free: (888) 231-3473 • www.amstat.org • www.twitter.com/AmstatNews

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND P-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative

Science March 7, 2016

Monash Feb 2020

• 2017: Another *Nature* article *p* < 0.005

• Articles solicited for special issue of American Statistician

comment

Redefine statistical significance

We propose to change the default P-value threshold for statistical significance from 0.05 to 0.005 for claims of new discoveries.

David L Ingensen, Larenco Canego L Agenco L Advances, Biala A, Nang L C, Yangamustan, Sharing M, Chang M, Lang M, Lang L, Lang M, Lang

... a recent timeline

- 2019: American Statistician publishes special issue 43 articles; 400 pages
- Editorial introduction advises "abandon 'statistical significance' "
- Nature publishes a letter agreeing with this
- "we are not advocating a ban on P values, confidence intervals or other statistical measures – only that we should not treat them categorically
- "This includes dichotomization as statistically significant or not, as well as categorization based on other statistical measures such as Bayes factors."







"Lawyers and judges pay close attention to standards, guidances, and consenus statements from respected and recognized professional organizations."

"Despite the fairly clear and careful statement of principles, legal actors did not take long to misrepresent the ASA principles." 2016

"distorted into strident assertions that statistical significance was unnecessary for scientific conclusions."



HARVARD DATA SCIENCE REVIEW

P-Values on Trial: Selective Reporting of (Best Practice Guides Against) Selective Reporting

outlines a 2018 Supreme Court case appealing a conviction for wire fraud, based on misleading investors Harkonen v. United States 13-180

the fraud centered on *p*-hacking the results of a Phase III trial of a drug

by Deborah Mavo

marketed by Harkonen

in the appeal "his defenders argued that the ASA guide provides compelling new evidence that the scientific theory upon which petitioner's conviction was based [that of statistical significance testing] is demonstrably false" Monash Feb 2020

Monash Feb 2020

11

What to do?

- report actual p-value, not "*", p < 0.05, etc. to sensible number of decimal points
- supplement *p*-value with sample size, estimated power, etc.
- clarify 'exploratory' and 'confirmatory' p-values
- · report effect sizes and estimated standard errors
- report confidence intervals
- pre-register trials, specifying primary and secondary outcomes
- pre-specify data analysis
- provide a *p*-value function
- or some analogous distribution

Spiegelhalter 2017

NEJM

significance function Baves posterior