

***P*-values are really quite nifty**

The P-Value Controversy: Where Do We Go from Here?

2019 Joint Statistical Meetings

Denver, Colorado

Philip B. Stark

July 2019

University of California, Berkeley

What's a P -value?

Suppose X is a random variable whose distribution is dominated by the uniform distribution if the null hypothesis is true:

$$\mathbb{P}_0\{X \leq p\} \leq p \text{ for all } p \in [0, 1].$$

Then the observed value of X is a P -value.

Where's the problem?

- Irrelevant/silly null.
 - Too abbreviated
 - Ignore design

Where's the problem?

- Irrelevant/silly null.
 - Too abbreviated
 - Ignore design
- $\mathbb{P}_0\{X \leq p\} > p$.
 - Selective inference
 - Ignore design

Where's the problem?

- Irrelevant/silly null.
 - Too abbreviated
 - Ignore design
- $\mathbb{P}_0\{X \leq p\} > p$.
 - Selective inference
 - Ignore design
- Misinterpretation

Where's the problem?

- Irrelevant/silly null.
 - Too abbreviated
 - Ignore design
- $\mathbb{P}_0\{X \leq p\} > p$.
 - Selective inference
 - Ignore design
- Misinterpretation
- Education and communication
 - “Disservice courses”

Where's the problem?

- Irrelevant/silly null.
 - Too abbreviated
 - Ignore design
- $\mathbb{P}_0\{X \leq p\} > p$.
 - Selective inference
 - Ignore design
- Misinterpretation
- Education and communication
 - “Disservice courses”
- Software makes it easy to commit silliness

Where's the problem?

- Irrelevant/silly null.
 - Too abbreviated
 - Ignore design
- $\mathbb{P}_0\{X \leq p\} > p$.
 - Selective inference
 - Ignore design
- Misinterpretation
- Education and communication
 - “Disservice courses”
- Software makes it easy to commit silliness
- *Cargo-Cult Statistics* and *Quantifauxcation*
 - Editorial policies
 - Disciplinary practice & culture

Why not just confidence intervals or confidence sets?

Not all hypotheses are about parameters. E.g., 2-sample problem.