

Betting for Noise

Contribution to the Discussion Meeting for:
Glenn Shafer (2020) Testing by Betting: A strategy for statistical and
scientific communication

RSS, 9th September 2020

Barbara Osimani, Univpm

Working through some issues

Psychologists have long wrestled with the problem of replicability. Bayesian inference and a break with the deterministic model of science might offer some respite, says **Andrew Gelman**



NLshop/iStock/Thinkstock

Case Study

Durante, Arsena, Griskevicius (2013) The fluctuating female vote: Politics, religion and the ovulatory cycle. *Psychological Science*, 24, 1007-1016.

Predictors:

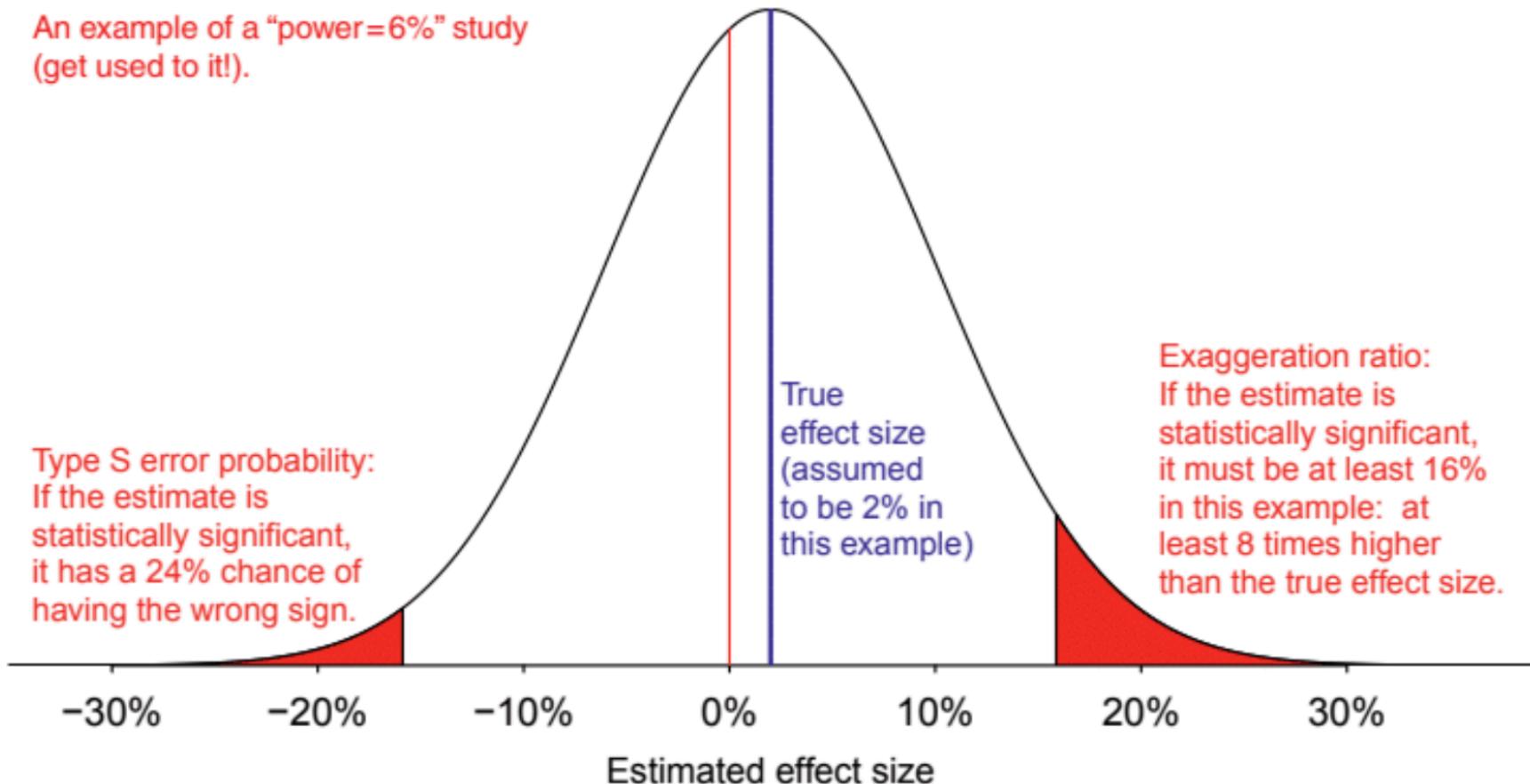
- 1) ovulation &
- 2) marriage status (single vs. married, respectively)

→ Effects:

- More liberal vs. more conservative
- Less religious vs. more religious
- Vote for Obama vs. Mitt Romney

Weak signal and high noise

An example of a “power=6%” study
(get used to it!).



The bell curve reproduces the probability distribution of differences between the two groups (married vs. single women), that could be hypothetically be observed if the effect size were just 2% (instead of the presumed 20%)

Because of the high variability in the population, the spread of the distribution is rather extended across the set of values of the variable. Within 6 SD it covers a range which goes from -30% to +30% difference between the groups.

Higher Order Evidence

“One direction for statistical analysis that appeals to me is Bayesian inference, an approach in which data are combined with prior information (in this case, **the prior expectation that newly studied effects tend to be small**, which leads us to downwardly adjust large estimated effects in light of the high probability that they could be coming largely from noise)”.

Gelman (2015)

Higher Order Evidence

“... but these steps will not be easy because they move away from the usual statistical paradigm in which each scientific study stands alone.
To resolve the replication crisis in science we may need to consider each individual study in the context of an **implicit meta-analysis...**”

(hyper-prior – hierarchical bayes)

Gelman (2015)

Questions

1. Is there any opportunity to accommodate Gelman's proposal in your approach?
2. More specifically, is it possible to model bets for noise? May we think of a testing protocol in that sense? Can Sceptic make money by betting against too precise hypotheses? How?

Questions

One explanation for the “reproducibility crisis” is that noise is strategically exploited in order to “sample to a foregone conclusion” — that is, one selects and discloses evidence opportunistically

One way to do this is precisely to pretend to have stopped earlier than one actually did.

3. How can your approach cope with this sort of phenomena?

Critical points and suggestions

How can we make the entire framework more palatable to Doctors, or scientists more generally?

Suggestion:

I think we need to contextualize your framework in the context of information theory if we don't want doctors jump from an apparatus they use without understanding it (standard hypothesis testing) to another one they can understand even less.