

# Notes on “Predictive Inference and Scientific Reproducibility” by Dean Billheimer

Gabe Harris

November 3, 2021

## Abstract

- hypothesis tests and parameter estimation for inference are “noble, but misguided.”
- “observables are fundamental”
- prediction of future observables given current data and understanding is the goal of statistical modeling
- Advantages of predicting future observables:
  - “an interpretable numerical summary of a quantity of direct interest to current and future researchers”
  - “a calibrated prediction of what’s likely to happen in future experiments”
  - “a prediction that can be either ‘corroborated’ or ‘refuted’ through experimentation”
  - “avoidance of inference about parameters; quantities that exist only as convenient indices of hypothetical distributions”
- “Adoption of this paradigm would improve our rigor for scientific accuracy and reproducibility by shifting our focus from ‘finding differences’ among hypothetical parameters to predicting observable events based on our current scientific understanding.”
- Elaborate on how prediction improves “rigor for scientific accuracy and reproducibility”? Concrete examples?

*“The only useful function for a statistician is to make predictions, and thus provide a basis for action.”*

W. Edwards Deming

# 1 Introduction

- “current disciplinary dilemma over  $p$ -values,” conventional hypothesis tests. Conventional ways of doing statistics compromises reproducibility **elaborate**
- misguided to try to “fix” hypothesis testing
- “larger and deeper inferential problems”
  - “dichotomization of results as either ‘significant’ or ‘nonsignificant’”
  - “the failure to incorporate the consequences of different types of testing errors”
  - “unclear definition of what it means to ‘reproduce’ a result”
- scientists without expertise in statistics are frustrated with follow-on experiments that do not provide ‘significant’ results, despite indications from traditional hypothesis testing and confidence intervals. This is a failure of statistics.
- “A focus on prediction allows the comparison of competing theories according to the quality of predictions they make, thus leading to better scientific inference.”
- reliability and validity in educational and psychometric testing theory
  - reliability: stability of measurement
  - validity: meaningfulness of the measurement
  - “Probabilistic prediction ... constitutes a type of ‘measurement reliability’ for inferences”
  - “focus on observable events, rather than unobservable parameters, ensures the ‘meaningfulness’ of results for future investigators”

Additional advantages

- focus on important (to the study) quantities, avoiding less informative inferential summaries “that arise historically because of their convenient sampling properties”
- “interpretation of ‘causes’” (e.g. regression parameter estimates) “based on the entire distribution of observables, thus encouraging the identification of relevant changes” **elaborate**
- “the practical significance conferred to individuals...avoiding paradoxes of practical vs. statistical significance.” **elaborate? examples?**
- interpretation of experimental results as sequential and progressive, with later results related to and influenced by earlier ones → better reproducibility **so I’m thinking reproducibility is enhanced progressively through the process of iterated experiments with updated prior information**
- “full accounting of the variability contributing to, or inherent in, the observed data, avoiding uncertainty laundering (Gelman 2016)” **What is “uncertainty laundering”?**

More fundamental advantage: “predictions of observable events is understandable to humans.”