Comment on Significance Testing in Theory and Practice by Daniel Greco

Howard Seltman
Associate Research Professor
Department of Statistics, CMU
hseltman@stat.cmu.edu
My Chronology

• Two philosophy courses in the early 70’s
• 20 years in medicine, consuming frequentist statistics
• CMU PhD in statistics: the Bayesian light dawns
• 10 years of research, teaching, and consulting at CMU – often masquerading as a frequentist
Key Contributions

• Elucidating the false logic of (typical) frequentists: PMT & weakening evidence

• Elucidating the conditions under which frequentist tests are “reasonable”
  – Appropriate, pre-chosen, plausible H_A
  – Sufficient “smoothness” of H_A
Quibbles

• Steps of Significance Testing
  – Formulate $H_0$
  – Choose (invent) a statistic
  – Find $Pr(S \geq s \mid H_0)$
  – Make a decision

• Justification of smoothness not completely convincing
  – Note: JA Paulos’ Catastrophe Theory of Humor
Valid Frequentist Pedagogy

- $\alpha=0.05$: applies to experiments where $H_0$ is true
- Power=$1-\beta$: applies to experiments where $H_A$ is true
- Above are applicable only when model assumptions are correct
- After an experiment we know “the decision”
- We **want** to know chance we made an error, which has denominators of “all rejects” and “all retentions” of $H_0$
Scenario examination: Feigning omniscience

- E.g., lifetime experience of 1000 experiments with 25% true $H_0$ and the remainder with 40/60/80% power for the “minimally interesting” specific $H_A$
- Reject: 12.5 falsely and 100+150+200 correctly, giving $PVP=97\%$
- Retain: 237.5 correctly and 150+100+50 falsely, giving $PVN=44\%$
Conclusion

• Scenario analysis helps students avoid misinterpretation of frequentist testing
• Daniel’s logic should nicely complement this