Probabilistic Scoring

Client: Alan Malik—Patient Tool

Lee Panter

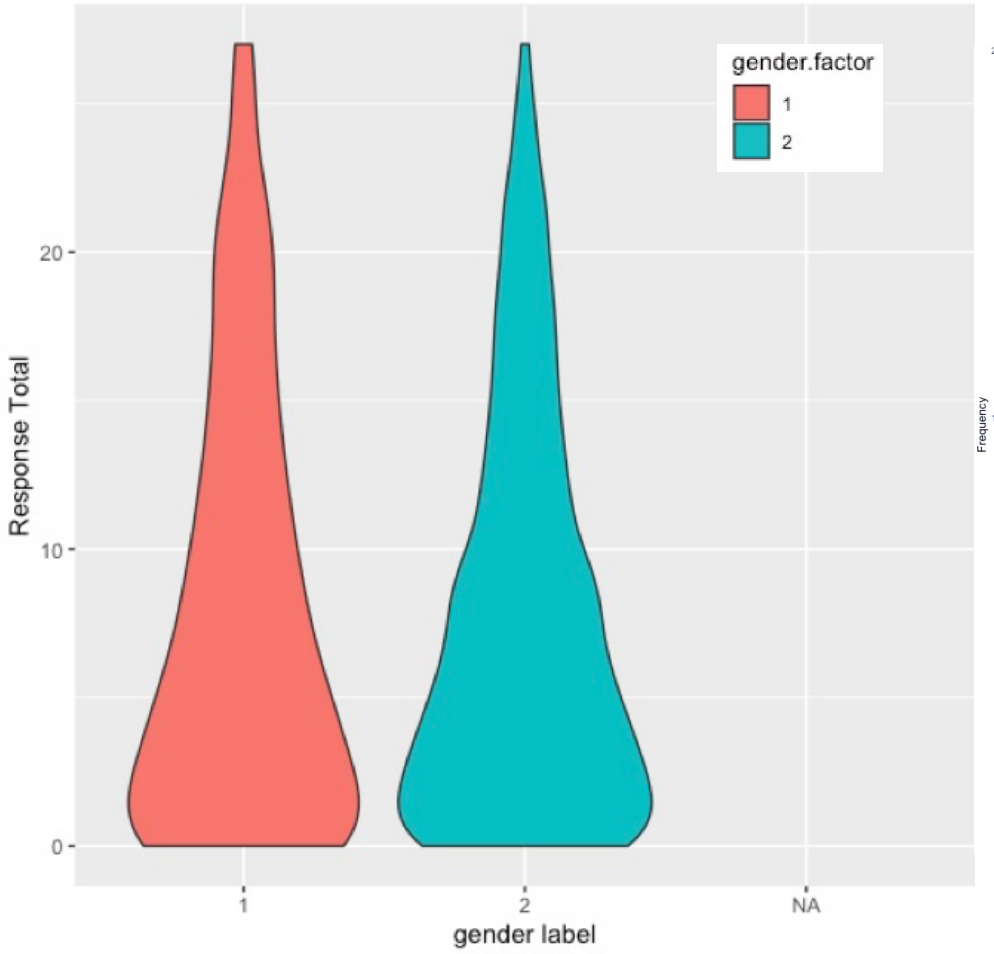
Updates: Ideal Outcomes

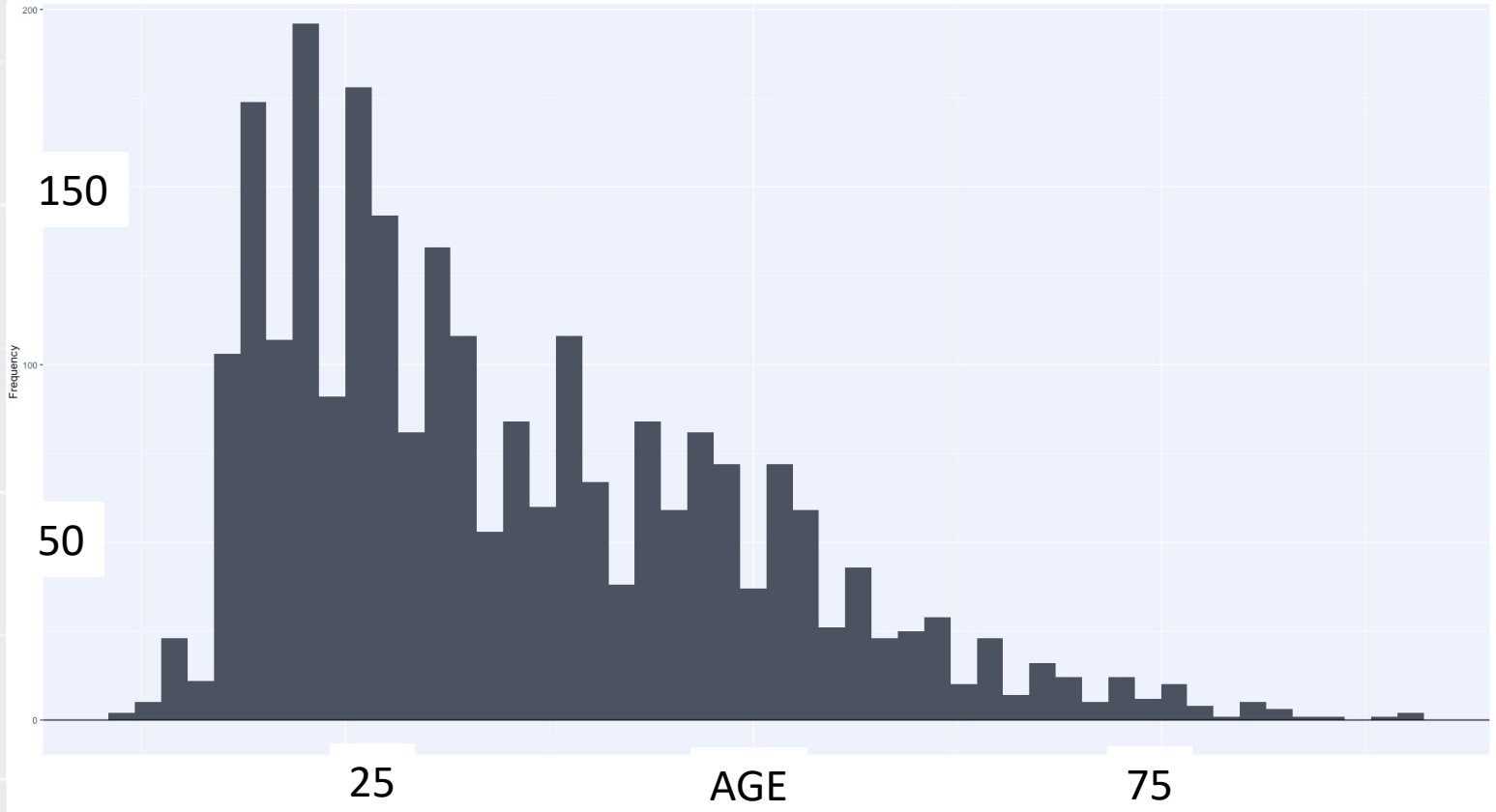
- Show that probabilistic scoring is superior in accuracy to conventional-linear scoring by using existing data from PHQ9 research
- Show that Probabilistic Scoring accuracy is a function of sample size, and that larger sample sized imply better classification accuracy
- Show that Probabilistic Scoring converges to the best estimate of “Baseline Truth” in higher Sample Numbers

Data

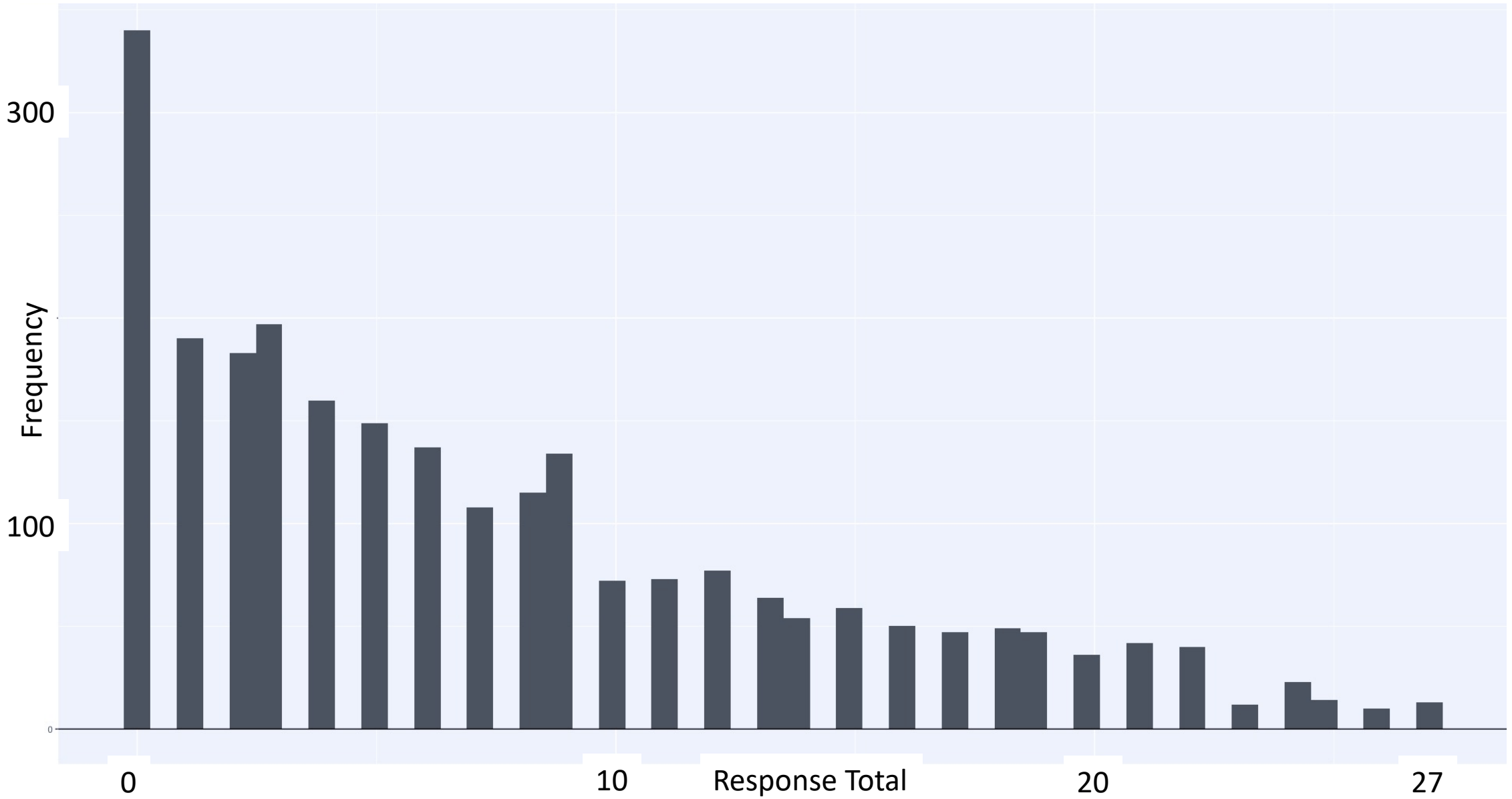
- Full data is 286 variables, 2495 observations
  - PHQ2, PHQ9, QDP
  - Data dictionary needed if further analyses will be expected

- Lots of calculated and ”nested” variables
- Demographic data: Age & Gender





So how crazy am I?



Immediate Goals

- Replicate analysis or get Alan to walk me through an analysis

- Cross-validation sample-size convergence:
  - Generate algorithm for calculating Evidence Coefficients, and accuracy measures as a function of sample size
  - Create cross-validation test-training data sets based upon an increasing sample-size
  - Compare average (minimum) accuracy across CV sets as sample size increases for Probabilistic Score and Conventional Score

- Investigate stability of convergence
  - Are some points converging faster? Not at all?
  - Can the limit be determined? Can a generalized proof of convergence be obtained?

THANK YOU!

Suggestions, comments, concerns: Lee.panter@ucdenver.edu