9.S918: Statistical Inference for Brain and Cognitive Sciences, Pset 3
due 7 May 2024

30 April 2024

1 Analyzing pilot data, power analysis, and preregistering an experiment

It's long been known that the wording of public opinion surveys can affect how people respond to those surveys. A particularly interesting example was first documented by Rugg (1941). Compare the two closely related questions in (1) below:

(1) a. Should the US government allow public speeches against democracy?
b. Should the US government forbid public speeches against democracy?

Perhaps surprisingly, the proportion of people who respond "yes" to the *allow* version of the question is lower than the proportion of people who respond "no" to the *forbid* version of the question.

One might also speculate that participants' responses are modulated by characteristics of their personality, such as the Big Five "openness to experience" dimension, which has been reported to be associated with tendencies toward right-wing authoritarianism (Butler, 2000).

The file

https://rlevy.github.io/statistical-inference-spring-2024/assets/assignments/pset_3/forbid_allow_openness_pilot_data.csv

contains hypothetical pilot data from 100 participants (50 in each condition) for a study attempting to reproduce Rugg's original results and also to look at the relationship between openness to experience (operationalized as a standardized continuous numeric variable where higher values indicate more openness to experience).

Tasks:

1. Consider the following scientific questions:

(a) Does Rugg's original generalization—that the choice of *allow* vs. *forbid* wording affects participant response patterns—still hold today?

(b) Do response patterns vary with openness to experience?

(c) Does the effect of wording choice (if it exists) vary with openness to experience?

Fit a logistic regression model to the provided pilot data such that answers to these three questions are provided by estimates and CIs on model parameters. What can we conclude regarding these questions from the pilot data?

2. Use your fitted model to conduct a power analysis for each of the three questions. Assuming that participants are randomly assigned to wording condition (in 50/50 proportions) and that the distribution of openness to new experience in the pilot data is reflective of the overall distribution in the population from which you will draw participants for your full experiment, plot statistical power curves for each of the three questions as a function of the total number N of new participants you will recruit in your full experiment.

3. Should you really divide your participant pool 50/50 between the two wording conditions, or would an uneven split improve statistical power? Answer this question using your intuitions. Then, choose a value of N and generate a power curve as a function of the proportion of participants assigned to the *allow* condition.

References

Butler, J. C. (2000). Personality and emotional correlates of right-wing authoritarianism. *Social Behavior and Personality: an international journal*, *28*(1), 1–14.

Rugg, D. (1941). Experiments in wording questions: Ii. *Public Opinion Quarterly*, *5*(1), 91.