DS3: Week 3 - Homework 2

Brian Quistorff

2017-06-27

1. Examine one of the experiments that was a major ethical violation (e.g. those listed in the notes).
   (a) Write out what the experimental goal was.
   (b) Describe an new ethical experiment that could learn some or all of the information desired. Write what would be the treatment(s) and control, how it would be randomized, and what data would be collected.
   (c) Would there be any other way to learn this information non-experimentally?
2. We are going to build a tool for "power calculations". One of the key trade-offs with an experiment is that the larger (and more costly) the experiment the smaller of an effect that one can measure, which is called the minimum detectable effect (MDE). We want a tool to relate these to help with experimental design.
   (a) Build a base function takes N, Cohen's d (effect divided of standard deviation of the data), $M = 100$, and $\alpha = 0.05$ (numbers in parentheses are defaults). For each iteration of M loops, generate data for N units, where half are treated ($W = 1$ vs $W = 0$), so that $y = d \cdot W + e$ where $e \sim N(0, 1)$. The function then reports the proportion of loops where the estimated $\hat{d}$ was significant at the level of α.
      i. What is the power for $N = 1000$, $d = 0.5$, and $\alpha = 0.05$?
      ii. What is the power for $N = 1000$, $d = 0.5$, and $\alpha = 0.01$?
      iii. What is the relation between significance level and power?
   (b) Build a higher-level function that takes d and $power = 0.80$ that outputs the minimum N required. This will call the base function. Do a straight-forward linear search starting from $N = 4$ and incrementing by 2. What is the minimum N for $d = 0.5$?
   (c) Build a higher-level function that takes N and $power = 0.80$ that outputs the minimum d. This will call the base function. Use the command `optim(...,method='BFGS')` and pass in a user supplied function that relates the difference between the power for a particular value of d, N and the target power. What is the minimum d for $N = 100$?