Ethics Part II - Week 4 Lab

Question Sheet

Zak Varty

This lab will not focus on the practical analysis of randomised control trial data or of A/B test outcomes, though the latter of these will be covered as introductory material in the learning agents course. Instead, we focus of the ethical aspects of these study designs and what data scientists can learn from the ethical design of clinical trails.

Task 1: Outline the similarities and differences between Randomised Control Trials and A/B test. In doing so you might want to think about:

- The statistical problem being addressed
- How and where these techniques are used
- The ethical implications of each approach.

Task 2: In a clinical trail an RCT might compare a new treatment to either no intervention, a placebo (a dummy treatment that is known to have no effect) or to the current standard treatment. Briefly describe the ethical implications of each approach and how these same concepts might translate to A/B testing.

Task 3: In both RCTs and A/B testing there are different types of tests that we might want to conduct. Investigate and explain the differences between the following, giving an example of when each might be used:

- An equivalence trail
- A non-inferiority trial
- A superiority trial.

Task 4: When conducting an A/B test or RCT it is important that we consider enough individual cases that we have the statistical ability to detect a difference between groups, if such a difference really exists. (In the live session we will work through such a sample size calculation)

- What aspects of the test will influence the number of observations needed?
- What are the statistical consequences of failing to meet this sample size?
- What are the ethical consequences of failing to meet this sample size?

- What might make obtaining this sample size difficult?

Task 5: (Extension - suggest revisiting after live session.)

An online retailer is proposing a change to the location of the 'Buy Now' button on their website. The aim of this change is to increase the proportion of page visits that are converted to a sale by at least some amount $\Delta > 0$. Let p_1 and p_2 respectively denote the proportion of page visits that lead to sales under the current and proposed layouts.

To assess the benefits of the change, the retailer is planning an AB test. This will be applied to n customers, where half of these customers are shown each layout. You have been asked to help formalise this AB test and to advise on the required sample size to meet the power requirements for this trial.

(a) State both in words and mathematically the null and alternative hypotheses for this superiority trial.

(b) Let N_1 and N_2 denote the number of sales made under each layout within the AB test, while $n_1 = n/2$ and $n_2 = n/2$ denote the number of customers directed to each layout. State the sampling distributions of N_1 and N_2. Hence find expressions for the mean and variance of D, the change in the sample proportion of customers who make a purchase when the proposed layout is used rather than the current layout.

(c) By applying a Gaussian approximation to the sample sales counts, state the approximate sampling distribution of D under the null hypothesis and the most pessimistic version of the alternative hypothesis.

(d) Calculate an expression of the power of this AB test for given values of the type 1 error rate α, the type 2 error rate β, the sample size n, and the smallest meaningful difference δ. Hence find an expression for the minimum required sample size to achieve this power.

(e) The retailer currently expects around 7% of page visits to convert to sales and would consider an increase of 0.5% or greater to be a meaningful improvement in sales conversion. Calculate the minimum required sample size for this AB test, given that the retailer will accept a 10% chance of falsely identifying a meaningful improvement and a 5% chance of failing to identify a meaningful improvement when one truly exists.