ADS2 Practical 11: Power analysis and sample size

Chaochen Wang

Semester 1, 2019/20

Work through this guide alone or in groups. Facilitators are here to help. The time it takes to complete this practical can vary between individuals - this is OK. Do not worry if you do not finish within the session.

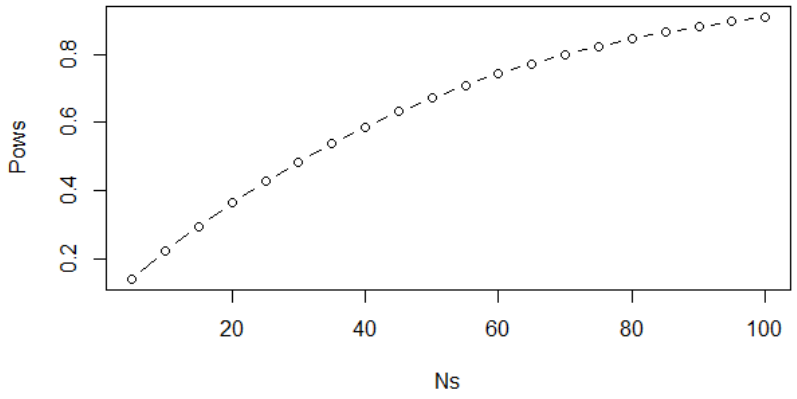
Learning Objectives

- Understand intuition behind power calculations
- Know how to perform power/sample size analysis with formula or in R
- Reveal the relationship among significance level, power, effect size and sample size
- Demonstrate different stages in clinical trails and stopping rules

1. Calculate one-sample test power with simulation in R

Let us assume that the heights of the adult men in a country follow a normal distribution with average 175cm and standard derivation 10cm. A college is providing a free bottle of milk per day to the students. Based on the history of a nearby country, giving a bottle of milk to the college students could improve the average height from 175cm to 178cm. Now, you go to the college and measure the heights of 10 male students, and perform a t-test to see if the students are indeed higher than the average of the country. Simulate the data you get from the students, and record the p-value in the t.test, which is the number? Now repeat the simulation for 10^5 times and look at all the p-values.

1) If you set a cutoff of 0.05, what is the percentage of simulations that you get the p-value larger than 0.05? What does this number mean? What is the power?
2) If you measure 50 students instead of 10 students, then does the power change?
3) The cheating way is use the command power.t.test, try it out to calculate the power number with different Ns (5, 10, 15, 20….100). Try to plot the power versus N. You should see a plot like this.



2. Choose the right sample size

A company is developing a diet pill, which may help people lose weight. In the animal model, the drug could lead to 10% of weight loss. Now the company recruits 20 volunteers and separate them into two groups: placebo group and drug group to perform a trial. Let us assume that the weight in the normal population is 130 pound with standard derivation of 30.

1) If the drug is indeed effective as showed on the animal model, then how possible that they do not see a significant effect of the drug (p-value cutoff = 0.05)? You can simulate it as before.
2) Do you think the power is good? If the company truly believe the effect of the drug and want to be sure that they will not largely miss the effect in the trial (typer II error rate < 0.2), then how many volunteers do they need to recruit?
3) If the company changes their strategy, asking all the volunteers to take the pills and measuring their weight before and afterwards, how many volunteers do they need?
4) Manipulate the p-value cutoff or power number to see the change of needed sample size. You can also plot them out. The output should be similar to these.