

Quantitative Understanding in Biology

Problem Set 1: Understanding Type I Error Rates

21 September, 2023

You start a new rotation in a lab that studies the effect of different compounds on the ability of mice to form new memories. You learn that the lab has been following a protocol that involves testing each compound on 10 mice, and using another 10 mice treated with a placebo as a control. If the first analysis doesn't yield a p-value passing the statistical significance threshold of 0.05, the standard practice in the lab is to add another 10 mice to each group and perform a second analysis with 20 animals per group.

Part I

Show, using a simulation in R, that when a compound really has no effect, the Type I error rate of the first analysis is 5%.

Part II

Using a simulation in R, show that the Type I error rate for the overall protocol that the lab is using is increased by approximately 70%.

What percentage of the time will this lab conclude that ineffective compounds have an effect?

Part III

What alpha value should the lab be using to maintain an overall Type I error rate of 5% when following their protocol?

Logistics

Answers are due by 11:59 PM on Thursday, 28th September, 2022. E-mail your solution to dag4010@med.cornell.edu, jbanfelder@rockefeller.edu, and las2017@med.cornell.edu

This Problem Set should be completely individually.

Your submission should be generated using R Markdown. Include both the script needed to reproduce your results and generate any figures in your write-up, as well as the compiled HTML file.