BIOS 234 Final Project. Due: Thursday, March 27th, 2025 @ 11:59pm PDT

1. Consider the following setting: You are working as a statistician in Amgen, California, and they want to design a simple clinical trial for assessing the efficacy of one of their new drugs. The new drug is expected to control LDL (Low Density Lipoprotein) and is the first of its type to be developed "accounting for genetic factors related to cholestrol". Amgen expects this drug, code-named "X" to be a major competitor of Lipitor, marketed by Pfizer. The primary question being asked here is: *How many patients should we collect?*

   Your mentor in the company has heard that you are trained in "Bayesian Statistics". He asks you to present a report to the group about the fundamentals of the Bayesian Sample Size problem. He also wants you to write an R-program that will enable the team to practice what you are preaching: a Bayesian design for a Bayesian analysis.

   In particular, he wants you to focus upon the following simple considerations (at least initially):

   (a) Amgen expects to collect data on LDL levels from patients who have been on Lipitor in California and compare with patients who will be treated with drug "X". For administrative convenience Amgen intends to sample the same number of individuals for both the groups.

   (b) Past analysis on LDL shows that LDL measurements seem to follow a Normal distribution. It is also known from pilot studies that the standard deviation of LDL measurements is usually very robust for such populations, and analysis usually proceeds by assuming that the standard deviation is *known* to be 4.5 units.

   (c) The analysis objective for Amgen is to be able to detect a statistically significant difference in the LDL means from their data when the "true" clinical difference between the LDL means of the Lipitor and X groups are $5, 10, 15$ and 20 units.

   Your group would like to know how to proceed with a Bayesian design and solve the sample size problem in the above setting. Write a report, not exceeding 3 pages, explaining how a Bayesian would perform sample size calculations in the above context. Explain clearly when the Bayesian sample size results would "reproduce" those from a classical perspective as a special case. Prepare an `R` program to implement Bayesian sample size calculations. Carry out some simulations using your program that you would present to your group, explaining how to reach a decision.