Purpose: Explore multi-dimensional optimization.

Assignment: Each problem involves mathematical and computational parts. For the computational parts, please turn in the simplest/shortest algorithm you can write that accomplishes the task. You may discuss the problems with other students, but all work must be your own. If you consult a book or online resource, cite it.

Your code: For each computational problem, you need to show your code and demonstrate via sample output that it works.

Built-in functions: `R` provides several built-in functions for numerical solving and optimization such as `optim`, `uniroot`, and `nlm`. You may use any of these to check your results, but you must write your algorithms using only basic `R` functionality. If you are not sure whether you can use a certain function, please ask.

Problem 1: Seasonal influenza

Influenza is more common during certain times of the year. Consider a region with a large number of people and a single hospital. Suppose we model the number of cases of influenza in day t as

$$X_t \sim \text{Poisson}(\lambda_t)$$

for $t = 1, \ldots, 365$. The rate λ_t is *seasonal*:

$$\lambda_t = \lambda \left[1 + a \cos\left(\frac{2\pi t}{365} - \phi \right) \right],$$

where $\lambda > 0$ is the overall rate, $a \in (0, 1)$ is the amplitude, and $\phi \in (0, 2\pi)$ is the seasonal offset, in radians.

a. Set $\lambda = 103$, $a = 0.8$, $\phi = 3\pi/10$ and plot λ_t for $t = 1, \ldots, 365$. Write code to simulate $X_1, \ldots, X_{365}$ and plot the simulated data over the curve for λ_t. Turn in your code and the plot.

b. Fix $\lambda = 103$. Create a heatmap or contour plot of the log-likelihood for your simulated data on the set $(a, \phi) \in [-1, 1] \times [0, 4\pi]$. What do you see? Explain.

c. Derive and implement an algorithm to estimate λ, a, and ϕ by maximum likelihood. Explain your derivation clearly, turn in your code, and confirm that your algorithm gets the correct MLE. Plot your iterates on the heatmap/contour plot.

Extra Credit: A randomized controlled trial with missing data

At a certain school, there are 200 students. All are non-smokers. Researchers randomly select exactly 100 students to receive anti-smoking education. These are the "treated" students. The "untreated" (control) students receive no anti-smoking education. One year later, researchers conduct an *anonymous* survey of all 200 students in the school. The survey is anonymous because the researchers want to encourage truthful responses. The survey asks only one question: "Have you smoked one or more cigarettes in the last week?". Since the survey is anonymous, researchers do not know which survey responses come from treated and untreated students. From the results of this survey, researchers observe that 47 students report smoking at least one cigarette in the last week.

Let $Z_i = 1$ if student i received the anti-smoking education, and let $Z_i = 0$ otherwise. Let $Y_i = 1$ if student i reports no cigarettes smoked in the last week, and let $Y_i = 0$ if the student reports smoking at least one cigarette. Assume students' survey responses are truthful. Define the "average treatment effect" as

$$\mu = \mathbb{E}[Y_i|Z_i = 1] - \mathbb{E}[Y_i|Z_i = 0].$$

An estimator $\hat{\mu}$ of μ is obtained by substituting sample proportions for the expectations above.

Immediately after the survey is conducted, a smoking cessation support group is established for students who smoke, and 24 students *voluntarily choose* to join the group. All of these students answered "Yes" to the survey question about smoking. The researchers learn that *at least* 13 students in the support group received the anti-smoking education treatment, and *at least* 1 student in the support group did not. Let $W_i = 1$ if student i participates in the smoking cessation support group. Define the "average effect of treatment on smokers joining the support group" as

$$\delta = \mathbb{E}[W_i|Y_i = 0, Z_i = 1] - \mathbb{E}[W_i|Y_i = 0, Z_i = 0].$$

An estimator $\hat{\delta}$ of δ is obtained by substituting sample proportions for the expectations above.

Answer the questions below by considering the observed data, and make no assumptions about the relationship between treatment and response.

a. What are the minimum and maximum values $\hat{\mu}$ could have in this trial? Can researchers rule out the possibility that the anti-smoking treatment has no effect, $\hat{\mu} = 0$?

b. What are the minimum and maximum values $\hat{\delta}$ could have in this trial? Can researchers rule out the possibility that the anti-smoking treatment has no effect on smokers' joining the support group, $\hat{\delta} = 0$?

c. Plot every value that the pair $(\hat{\mu}, \hat{\delta})$ can take in this trial.

d. Consider the event that the anti-smoking education treatement effect is positive but the treatment effect on smokers joining the support group is negative, $(\hat{\mu} > 0, \hat{\delta} < 0)$ for *this population of 200 students*. What can you say about this event?