Homework 3. Stat 202a. Due Thu, Nov 11, 11:59pm.

You must work on the homework INDEPENDENTLY! Collaborating on this homework will be considered cheating. Submit problems 1 and 2 via CCLE and submit the extra-credit by email to frederic@stat.ucla.edu. **Late homeworks will not be accepted!** Questions 1 and 2 will be worth a combined 10 points, and the extra-credit is worth a maximum of 2 points. Your homework solution to 1 and 2 should be a single PDF document. The first pages should be your *output* from the problems above. After that, on subsequent pages, include all your *code* for these problems. For the extra-credit problem, your submission should be text only, in the body of the email, not as an attachment, and should contain only your python function.

1. Approximation of an infinite series in C.

It is well known that 1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 +/- ... = ln(2).
Write a C function called *alt2(n)* that computes the first *n* terms in this series, as a function of *n*. Call your C function from *R* to evaluate *alt2(n)* for various *n*. Using *R*, plot *alt2(n)* vs. *n*, for *n* ranging from some small number up to 1 million. You may set up your range of the y-axis in a way that you feel is appropriate. You do not need to show alt2(n) for all values of n and should not plot alt2 for very small values of n if they are off the plot.

2. Kernel density estimation in C and plotted in R.

Write a C function to compute a Gaussian kernel density estimate for univariate data. The inputs to the function should be two integers, m and n, a vector gof m gridpoints at which to calculate the estimates, a vector x consisting of the n observed data points, and a vector y of length m which will contain the resulting density estimates.

Gather data on all earthquakes of magnitude at least 3.0 in the longitude range -122.0 to -118.0 and latitude range 34.0 to 38.0, from Jan 1, 1960 to Oct 1 2021, from http://service.scedc.caltech.edu/eq-catalogs/date_mag_loc.php . Input the data into R. Use minimum magnitude = 3.0, maximum magnitude = 9.0, min depth = -5km, max depth = 100km, event type = earthquake, geographic type = local. Take this vector of earthquake magnitudes, and use your C function to make a kernel density estimate of the earthquake magnitudes, using a Gaussian kernel with bandwidth selected using the rule of thumb suggested by Scott (1992). You may calculate this bandwidth in R. Let $\{m_1, m_2, ..., m_{100}\}$ = a vector of 100 equally spaced magnitudes spanning the observed range of magnitudes in your dataset, compute your kernel estimates on this grid using the C function, and plot your kernel density estimates $\hat{f}(m_1), \hat{f}(m_2), ..., \hat{f}(m_{100})$.

Output: Your output for this assignment should be a pdf document containing the following, in this order.

Figure 1. A plot of alt2(n) versus n, for several values of n ranging up to 1 million.

Figure 2. A plot of your kernel density estimates $\hat{f}(m_1)$, $\hat{f}(m_2)$, ..., $\hat{f}(m_{100})$ versus m.
After these 2 figures, include all of your C code, followed by all of your *R* code.

3. Writing a function in Python.

For extra-credit, write a function in Python that takes no arguments, and when executed, allows the user to play a game. Your function should start by printing instructions so the user knows what the rules of your game are and how to play. Your game must somehow involve the words "202", "kernel", and "statistics" in it. Your code should consist with a function that, when called, runs the game. If you have other helper functions too, put them lower down, so that the function to play your game is at the very top of your code. Send this Python code directly to frederic@stat.ucla.edu . Your email should contain text only.