

Homework 3. Stat 202a. Due Thur, Nov 13, 10:30am.

You must work on the homework INDEPENDENTLY! Collaborating on this homework will be considered cheating. Submit your homework by CCLE. **Late homeworks will not be accepted!** Your homework solution 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.

1. Approximation of an infinite series in C.

It is well known that $1 - 1/2 + 1/3 - 1/4 + 1/5 - 1/6 + \dots = \ln(2)$.

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're off the plot.

2. Kernel density estimates in C.

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 bandwidth b , a vector g of 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 to 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 Jan 1, 2014, from http://www.data.scec.org/eq-catalogs/date_mag_loc.php. Input the data into *R*. (Use maximum magnitude = 9.0, min depth = 0, max depth = 100km, event 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 $\{m_1, m_2, \dots, m_{100}\}$ = 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 using *R*, plot your kernel density estimates $\hat{f}(m_1), \hat{f}(m_2), \dots, \hat{f}(m_{100})$.

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 n ranging up to 1 million.

Figure 2. A plot of your kernel density estimates $\hat{f}(m_1), \hat{f}(m_2), \dots, \hat{f}(m_{100})$ versus m .

After these 2 figures, include all of your C code, followed by all of your *R* code.