STAT 210
Applied Statistics and Data Analysis:
Homework 1

Due on Sept. 11/2022

Question 1

Consider the following system of equations:

$$\begin{aligned} 3x + 2y + 2z + 4w &= 28 \\ 2x + y + z &= 14 \\ 2x + 5z + 5w &= 28 \\ 6x + 2y + 2z + w &= 37 \end{aligned}$$

1. Create a matrix in `R` with the coefficients of the system, and a vector with the constants on the right-hand side of the equations. Call them `mat1` and `vec1`, respectively.
2. Find the inverse of `mat1` and call it `mat2`.
3. Create a list named `list1` having as components `mat1`, `vec1`, and `mat2`. Call these components `item1`, `item2`, and `item3`, respectively.
4. Remove `mat1`, `vec1`, and `mat2` from the working directory.
5. Solve the system of equations and call the solution `vec2`.
6. Verify the solution.
7. Verify that if you multiply the inverse matrix `mat2` by `vec1` you get the solution.
8. Find the eigenvectors of `mat1` and `mat2` and verify that the eigenvectors of `mat2` are the reciprocals of the eigenvectors of `mat1`.

Question 2

Consider the function $f(x) = e^{-|x|}$, for $x \in \mathbb{R}$. We want to use the MonteCarlo method to estimate the value of the integral

$$\int_{-2}^{2} f(x)\, dx$$

1. Plot a graph of this function in the region where you want to calculate the integral.
2. Generate $N = 1000$ random numbers with uniform distribution in the rectangle $[-2, 2] \times [0, 1]$. Count how many points fall below the curve $f(x) = e^{-|x|}$ and estimate the integral using the fraction of these points with respect to the total number of points and the area of the rectangle. Call the estimator I_{1000}
3. Compute analytically the value of the integral and compare with the approximation you obtained in 3. Call I the value of the integral and calculate $|I - I_{1000}|$
4. Repeat for $N = 10^k$ for $k = 4, 5, \ldots, 8$ and compute the deviation $|I - \bar{I}_N|$ from the exact result.

5. Do a log-log plot of the deviation as a function of N. The points should follow approximately a straight line.