Homework #1
Due Feb 25, 2019, 11:59pm

Problems:

1. Consider the data in the files a100.csv, b100.csv, c100.csv and d100.csv.
   a. Determine the underlying probability distributions (and its parameters) of each data set, by creating a histogram and over-plotting with the most similar probability distribution, until the agreement is acceptable. Create a label with the name of the distribution, and its parameter values on the plot.
   b. Create a new series from each data set through the formula

   \[ y_i = \sum_{j=0}^{K-1} x_{i+j}, \]

   i.e. each new number is the sum of \( K \) adjacent elements of the original series. Determine the probability distribution and its parameter for each sequence for \( K = 5, 20 \) and \( 80 \). Calculate the mean and variance of the original distributions and compare to the derived (summed) series.

2. The files noise01.csv to noise10.csv contain a random noise from a real instrument, measuring the intensity of light as a function of the voltage on a light source. The voltage goes from 0.1V to 1.0V, encoded in the filename. (0.1V, 0.2V, 0.3V, 0.4V, 0.5V, 1.0V). Prove that the noise is due to the Poisson distribution of the discrete photons using iPython.

3. Fit a linear relationship to the data in the files *fit.csv. Use an iPython notebook for the fitting (bfit.csv, cfit.csv, dfit.csv, efit.csv).

4. A die is rolled 24 times. Use the Central limit theorem to estimate the probability that
   a. The sum is greater than 84
   b. The sum is equal to 84
   c. Perform a hundred numerical realizations to illustrate the result