Homework 1

MA 506 Probability and Statistical Inference

Maximum points: 100

Due: September 28 (Wednesday), 11:59pm

Question 1: (15 points)

Fibonacci sequence (F_n), is defined as a sequence of numbers such that each number is the sum of two preceding ones starting from 0 and 1 ($F_0 = 0$ and $F_1 = 1$). Hence, we have the following value at n^{th} position

$$F_n = F_{n-1} + F_{n-2}$$

Write a python function that takes the number of elements the user wants in the generated Fibonacci sequence as a parameter. The function should return the generated Fibonacci sequence when called. Hence,

- For user input n = 0, function should return [ ]
- For user input n = 1, function should return [0]
- For user input n = 2, function should return $[0, 1]$
- For user input n = 3, function should return $[0, 1, 1]$ .. and so on

Question 2: (15 points)

Write a function in python that takes a list of numbers from the user and returns a different list which only contains those numbers from the original list that are within the range $[\mu - \sigma, \mu + \sigma]$. Here μ and σ are the mean and standard deviation of numbers in the original list.

Question 3: (25 points)

Write a python function that takes an integer (n) as an input and:

1. (**5 points**) If the integer is less than 1, the function should print an error message and return nothing.
2. for other cases:

(a) (**10 points**) Returns a numpy array with a checkerboard pattern. For example if user provides n = 5 to the function, then it should return the 5x5 array

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

(b) (**10 points**) Plots the checkboard pattern as an image with grayscale colormap.

Question 4 : (45 points)

We want to understand the application of Central limit theorem to the dice throw experiment for a biased dice. For this:

1. (**5 points**) Generate a population of 1 million dice throws with each face probability as follows

Face	1	2	3	4	5	6
Probability	1/3	1/12	1/12	1/6	1/6	1/6

2. (**5 points**) Assuming a sample size of $n = 10000$, draw $m = 10000$ samples from the population with replacement.

3. (**15 points**) Generate the histogram for 10000 realizations of the random variable

$$Z_m = \frac{S_m - \mu}{\sigma}$$

where S_m is the average of all the dice faces included in a sample, μ and σ are the mean and standard deviation of the population for the dice throw experiment. Additionally, overlay a pdf of a standard gaussian distribution ($\mu = 0, \sigma = 1$) on the plotted histogram.

4. (**15 points**) Fixing the sample size as 10000, consider 4 different cases of different number of samples: [100 5000, 10000, 20000]. For each of these 4 cases, make a similar histogram as you made in part 3 and plot these histograms in 2x2 matplotlib subplot grid.

5. (**5 points**) Explain in your own words what you observe in the plots in part 4 above.