EE 313 Homework 1: Central Limit Theorem

Let A be a random variable which can be represented as the sum of a group of random variables.

$$A = K_1 + K_2 + ... + K_n, \tag{1}$$

where $K_1, K_2, ..., K_n$ are independent and identically distributed. Therefore, we can write

$$E[K] = E[K_1] = E[K_2] = ... = E[K_n], \tag{2}$$

and

$$Var(K) = Var(K_1) = Var(K_2) = ... = Var(K_n). \tag{3}$$

1 Find the general expressions for $E[A]$ and $Var[A]$ by using $E[K]$ and $Var[K]$.

2 Let $K_1, K_2, ..., K_n$ be a sequence of uniform random variables with the uniformity range of $[-5, 5]$.

- By using MATLAB generate 10000 samples of A for each n value in $\underline{n} = \{1, 2, 3, 5, 10\}$. Take the normalized histogram of each sample vector and plot them separately side by side. (*You are encouraged to use subplot here to keep your plots organized.*) Calculate the parameters of Gaussian PDF for each case with the help of the expected value and variance expressions you have found earlier and plot them onto their respective histogram plots. How do the plots change with increasing n? Do the PDFs and histograms match? Comment on it.
- Let us keep $n = 10$. Generate another 100000 samples of A, take the histogram and plot the Gaussian PDF onto the histogram. What do increasing the number of samples achieve? Comment on it.
- Let us keep $n = 10$. Extend the uniformity range from $[-5, 5]$ to $[-10, 10]$, generate 10000 samples and take the histogram. Change the parameters of the Gaussian PDF accordingly and plot it onto the histogram. What changes? Comment on it.

3 Let $K_1, K_2, ..., K_n$ be a sequence of exponential random variables where $\lambda = 1$.

- By using MATLAB generate 10000 samples of A for each n value in $\underline{n} = \{1, 2, 3, 5, 10\}$. Take the normalized histogram of each sample vector and plot them separately side by side. (*You are encouraged to use subplot here to keep your plots organized.*) Calculate the parameters of Gaussian PDF for each case with the help of the expected value and variance expressions you have found earlier and plot them onto their respective histogram plots. How do the plots change with increasing n? Do the PDFs and histograms match? Comment on it.
- Let us keep $n = 10$. Generate another 100000 samples of A, take the histogram and plot the Gaussian PDF onto the histogram. What do increasing the number of samples achieve? Comment on it.
- Let us keep $n = 10$. Change the parameter of λ from 1 to 2. Generate 10000 samples and take the histogram. Change the parameters of the Gaussian PDF accordingly and plot it onto the histogram. What changes? Comment on it.

4 Let $K_1, K_2, ..., K_n$ be a sequence of Bernoulli random variables where $p = 0.5$.

- By using MATLAB generate 10000 samples of A for each n value in $\underline{n} = \{1, 2, 3, 5, 10\}$. Take the normalized histogram of each sample vector, **take the CDF of the histograms,** and plot the histogram CDFs separately side by side. (*You are encouraged to use subplot here to keep your plots organized.*) Calculate the parameters of Gaussian CDF for each case with the help of the expected value and variance expressions you have found earlier and plot them onto their respective histogram CDF plots. How do the plots change with increasing n? Do the CDFs and histograms match? Comment on it.
- Let us keep $n = 10$. Generate another 100000 samples of A, take the histogram CDF and plot the Gaussian CDF onto the histogram. What do increasing the number of samples achieve? Comment on it.
- Let us keep $n = 10$. Change the parameter of p from 0.5 to 0.75. Generate 10000 samples and take the histogram CDF. Change the parameters of the Gaussian CDF accordingly and plot it onto the histogram CDF. What changes? Comment on it.

5 Let $K_1, K_2, ..., K_n$ be a sequence of Poisson random variables where $\lambda = 1$.

- By using MATLAB generate 10000 samples of A for each n value in $\underline{n} = \{1, 2, 3, 5, 10\}$. Take the normalized histogram of each sample vector, **take the CDF of the histograms,** and plot the histogram CDFs separately side by side. (*You are encouraged to use subplot here to keep your plots organized.*) Calculate the parameters of Gaussian CDF for each case with the help of the expected value and variance expressions you have found earlier and plot them onto their respective histogram CDF plots. How do the plots change with increasing n? Do the CDFs and histograms match? Comment on it.
- Let us keep $n = 10$. Generate another 100000 samples of A, take the histogram CDF and plot the Gaussian CDF onto the histogram. What do increasing the number of samples achieve? Comment on it.
- Let us keep $n = 10$. Change the parameter of λ from 1 to 2. Generate 10000 samples and take the histogram CDF. Change the parameters of the Gaussian CDF accordingly and plot it onto the histogram CDF. What changes? Comment on it.

Notes

- Although you can use for loops or similar loops to achieve what it is wanted, it is shorter and more efficient to do vector and matrix operations on MATLAB. Most of the MATLAB functions are optimized to do these operations as well. Do not forget, not only your results but also your methods to achieve those results will be considered when your homeworks are evaluated.
- In the 4th and 5th part you will deal with CDFs, not PDFs. Be careful to not confuse them.
- You will have 7 plots for each part between 2 and 5. First 5 plots in each part should be in subplots to keep the plots organized.
- Keep your explanations short, preferably 1 or 2 sentences. Please do not over-explain the results, over-explaining will not get you any extra points.
- Do not forget to add labels, captions and other important details to your plots.

Functions

- Functions you may find a good use of: *rand, exprnd, binornd, poissrnd, histogram, normpdf, normcdf, cumsum, subplot*
- You may also check "A Brief Note on MATLAB/Octave Functions for Probability Distributions" link on the Moodle page for alternative functions.
- You do not have to use these functions, you are free to use other functions not listed here.