EE 131A | Matlab project
Probability | Monday, December 1, 2014
Instructor: Professor Roychowdhury | Due: Friday, December 19, 2014 at 5:00 pm

In this project we will further analyze random variables and their various properties. We will also investigate how are random variables used to model practical systems. Each part will have a combination of MATLAB programming, mathematical analysis and technical writing. You will be graded on all three components.

When producing your plots **clearly indicate** the x-axis, the y-axis and what is being plotted (using legends, title etc.). You may need to rescale x-axis to ensure that your plot is showing the right quantity.

Make sure to attach in the appendix of your project report **all MATLAB programs** that you used to generate the data.

1. *Generation of binomial RVs.*

   (a) Write a MATLAB program to simulate a $X \sim \text{binomial}(n, p)$ random variable for $(n, p) = (100, 0.2)$. Use $t = 10000$ trials to generate X, and plot pmf.

   (b) What value of k maximizes $P(X = k)$ for a general binomial(n, p) ?

   (c) What value of k maximizes $P(X = k)$ in your simulations ? Contrast the result in (b) with your experimental findings from part (a).

2. *Generation of Gaussian.*

   (a) Write a MATLAB program to generate $t = 10, 10^4, 10^7$ realizations of standard normal. Plot cdf and pdf in each case, and determine sample mean and sample variance. What fraction of values exceeds 1, 2, and 3 for each choice of t ? How does this fraction compare to Q values from the table in the textbook?

   (b) Super-impose the analytical expression for $f(x)$ over your sample pdf from part (a). Comment on the (dis)agreement.

   To plot sample pdf and sample cdf you should use `hist` command with bin-width 0.2 with the suitably chosen ranges.

3. *Simulation of Game of craps.* A player rolls two fair dice. If the sum of the dice is either 2, 3, or 12, the player loses; if the sum is either a 7 or an 11, he or she wins. If the outcome is anything else, the player continues to roll the two dice until he or she rolls either the initial outcome or a 7. If the 7 comes first, the player loses; whereas if the initial outcome reoccurs before the 7, the player wins.

   (a) Write a MATLAB program to generate sum of two fair dice using $t = 10000$ samples. Compute the probability of winning on the first roll based on your samples?

   (b) Suppose that the first roll of the two dice yields a sum of 5. we know probability of two dice yields a sum of 5 is $\frac{4}{36}$ and probability of two dice yields a sum of 7 is $\frac{6}{36}$.

Use these information to write a MATLAB program to compute the probability that the player wins based on your samples? Try $t = 10000$ samples.

(c) Calculate pmf, use the information to write a MATLAB program to compute the winning probability of a player at the game of craps. Try $t = 10000$ samples.

4. *Simulation of binary transmission systems* A binary transmission system sends a "0" bit using a -2 voltage signal and a "1" bit by transmitting a +2. The received signal is corrupted by noise N that has a Laplacian distribution with parameter α. Assume that "0" and "1" bits are equiprobable.

   (a) Write a MATLAB program to generate Laplacian distribution with parameter $\alpha = 2, 0.5$ respectively. Use the transformation method (refer to Selection 4.9 in textbook) based on $t = 10000$ samples of the unit-interval uniform random variable. Plot the resulting Laplacian cdf and pdf in each case.

   (b) Assume the received signal Y is given by $Y = X + N$. Suppose that the receiver decides a "0" was sent if $Y < 0$, and a "1" was sent if $Y \geq 0$. Write a MATLAB program to simulate the transmission of 10,000 bits under the SNR = 0dB for this channel and compute the empirical error probability.

   (c) Derive analytically the expression for the error probability under SNR = 0dB. How does the analysis compare to your simulation from part (b) ?

   For the above, SNR denotes the Signal-to-Noise Ratio and SNR(dB)=$10\log_{10}(\frac{1}{\sigma^2})$. You need to express the standard deviation of Laplacian distribution σ in terms of its parameter α.

5. *Central Limit Theorem* Let X_1,X_2,... be a sequence of iid random variable with finite mean μ and finite variance σ^2, and let S_n be the sum of the first n random variables in the sequence:

$$S_n = X_1 + X_2 + ... + X_n.$$

   (a) Let X_i be a continuous uniform random variable taking values in the interval $(1, 6)$. Write a MATLAB program to plot the pdf and cdf of S_n. Consider $n = 1, 5, 10, 50$ and compare your results.

   (b) Calculate analytically the mean and the variance of X_i and of S_n in part (a).

   (c) Write a MATLAB program to generate the standard Gaussian random variable with the same mean and variance as S_n. Superimpose this plot with the plots from part (a).
   Use t = 10,000 samples in the above.