CS5960: Nonparametric Methods **Due Wed. 2/4**

Homework 1: Expectation and Simulation

Instructions: Be sure to electronically submit your answers in pdf format for the written part and as an R file for the coding part. You may work together and discuss the problems with your classmates, but write up your final answers entirely on your own.

1 Written Part

1. Let X and Y be continuous, real-valued random variables. Prove the following:
   (a) $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y)$.
   (b) If X and Y are independent, then $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$.
   (c) $E[E[X|Y]] = E[X]$.
2. *Chuck-A-Luck (simplified version).* In the game of Chuck-A-Luck a player tosses three dice and earns a dollar for each 6 that appears. It costs one dollar for each roll of the dice. Let X be the random variable representing the dollar amount won in a single roll of the three dice minus the dollar spent to play. Example: if the player rolls $(6, 2, 6)$, then $X = 1$.
   (a) What is $E[X]$?
   (b) What is $\text{Var}(X)$?
3. Now assume a player starts with \$10 and plays Chuck-A-Luck ten times. Let Y be the random variable representing the total winnings minus the amount paid to play.
   (a) What is $E[Y]$?
   (b) What is $\text{Var}(Y)$?

2 R Simulation

Write R code to simulate the game of Chuck-A-Luck in the following steps. Pay attention to whether the results from the simulation match your answers to written questions above. Use the R command `print` to display your results in parts 2 and 4.

1. Write a function to simulate a roll of Chuck-A-Luck. The function should return the amount won minus the dollar spent (i.e., the random variable X from above.)
2. Run your function 100 times, storing the values in a vector. Compute the sample mean and variance for this simulation. Repeat this experiment for sample sizes of 1,000 and 10,000.
3. Now write a function to simulate the random variable Y above, that is, the amount won starting with \$10 and playing the game ten times.
4. Now generate a random sample of Y of size 100 and compute the sample mean and variance. Repeat for sample sizes 1,000 and 10,000.
5. Plot a histogram of the Y variable. Is the histogram symmetric about the mean?