CS112: Computer System Modeling Fundamentals
Homework 3

Due Tuesday, May 10, 4pm (in class)

Please refer to the course academic integrity policy for collaboration rules, and be sure to make use of the problem solving guidelines covered in section. Remember that you will be graded on both the correctness and the clarity of your solutions.

The problems on this assignment are good practice for the midterm, so you are especially encouraged to start early! Only problems 1b, 7, and 8 require material beyond the first three chapters of the textbook, so save these ones for later and try the rest.

1. The per bit error rate over a certain binary communication channel is 10^{-10}. No other statistics are known about the channel or the data.

   (a) What is the expected number of erroneous bits in a block of 1000 bits?

   (b) Find an upper bound using the Markov Inequality on the probability that a block of 1000 bits has 10 or more erroneous bits.

2. Joe Lucky plays the lottery on any given week with probability p independently of whether he played on any other week. Each time he plays, he has a probability q of winning, again independently of everything else. During a fixed time period of n weeks, let X be the number of weeks that he played the lottery and Y be the number of weeks that he won.

   On an in-class quiz, you were asked to solve parts a-c of this problem. **You do not need to write down the answers to these three parts again, but you should work through them again and understand them before you move on to parts d-f.**

   (a) What is the probability that he played the lottery on any particular week, given that he did not win on that week?

   (b) Find the conditional PMF $p_{Y|X}(y|x)$.

   (c) Find the joint PMF $p_{X,Y}(x, y)$.

   (d) Find the marginal PMF $p_Y(y)$. Hint: One possibility is to start with the answer to part c, but the algebra can be messy. However, if you think intuitively about the procedure that generates Y, you may be able to come up with a simple argument that leads to the answer without digging into the algebra.

   (e) Find the conditional PMF $p_{X|Y}(x|y)$. Do this algebraically using the preceding answers.

   (f) Rederive the answer to part e by thinking as follows: for each one of the $n - Y$ weeks that he did not win, the answer to part a should tell you something.

3. Suppose that X and Y are independent discrete random variables with the same geometric PMF:

$$p_X(k) = p_Y(k) = p(1-p)^{k-1}, \qquad k = 1, 2, \ldots,$$

where p is a scalar with $0 < p < 1$. Show that for any integer $n \geq 2$, the conditional PMF

$$\mathbf{P}(X = k | X + Y = n)$$

is uniform.

4. A wireless data communication service shares available bandwidth with cell phones. The cell phone traffic has priority so that data transmission is slower when there is a lot of cell phone traffic. Suppose that the time to send a data message is exponential with parameter λ_L when there is low cell phone traffic, and is exponential with parameter λ_H when the cell phone traffic is high. Assume that 70% of time the cell phone traffic is high and 30% of time the cell phone traffic is low. Given that a particular message takes time greater than τ to send, what is the probability that the cell phone traffic is high?

5. Let the random variables X and Y have a joint PDF which is uniform over the triangle with vertices at (0,0), (0,1) and (1,0).

   (a) Find the joint PDF of X and Y.

   (b) Find the marginal PDF of Y.

   (c) Find the conditional PDF of X given Y.

   (d) Find $\mathbf{E}[X|Y = y]$, and use the total expectation theorem to find $\mathbf{E}[X]$ in terms of $\mathbf{E}[Y]$.

   (e) Use the symmetry of the problem to find the value of $\mathbf{E}[X]$.

6. **Background:** This is a variation on a problem described in class this past week in which we examined the time it takes to search a singly linked list of size L with entities sorted in ascending order. The list is assumed to be very long, and rather than have positions in the list be discrete, the list is approximated by a continuum from 0 to L.

   In the problem described in class, we had a pointer to a random position in the list. Call that position X. We assumed that X is uniformly distributed between 0 and L. The position of the entity being searched for is a random variable, Y, which is also uniformly distributed between 0 and L.

   The search works as follows: The key of the entity being searched for is compared with the key at the random position X. If the search key is greater than the key at X, then the list is searched linearly starting at location X. Otherwise, the list is searched linearly starting at position 0.

   We were interested in finding the expected length of the search.

   Your Problem: The variation you are to solve in this problem is where we initially start with **two random pointers** into the list. Let:

   - $Y = y$ be a uniform random variable for the position of the key we are searching for.
   - $X_1 = x_1, X_2 = x_2$ be uniform random variables for the position of the two pointers.

   The behavior of our search depends on 3 conditions:

   - If the value at position y is greater than both the values at positions x_1 and x_2, the list is searched linearly starting from the larger of x_1 and x_2.

- If the value at position y is smaller than both the values at positions x_1 and x_2, the list is searched linearly starting from 0.
- If the value at position y lies between the values at x_1 and x_2, the list is searched linearly starting from the smaller of x_1 and x_2. (Note that there are two subcases here, one where the value at x_1 is less than the value at x_2 and vice versa.)

Find the expected value of the search distance.

7. Let X be a random variable and let α be a positive constant. Show that

$$\mathbf{P}(|X| \geq c) \leq \frac{\mathbf{E}\left[|X|^{\alpha}\right]}{c^{\alpha}}, \qquad \text{for all } c > 0.$$

8. Bo assumes that X, the height in meters of any Canadian selected by an equally likely choice among all Canadians, is a random variable with $\mathbf{E}[X] = h$. Because Bo is sure that no Canadian is taller than 3 meters, he decides to use 1.5 meters as a conservative upper bound for the standard deviation of X. To estimate h, Bo uses the average of the heights of n Canadians he selects at random.

   (a) Define H to be Bo's estimate of the of h, obtained by averaging the heights of the n randomly selected Canadians. In terms of h and Bo's 1.5 meter bound for the standard deviation of X, determine the expectation and an upper bound on the standard deviation of H.

   (b) Find as small a value of n as possible such that the standard deviation of Bo's estimator is guaranteed to be less than 0.01 meters.

   (c) Bo would like to be 99% sure that his estimate is within 5 centimeters of the true average height of Canadians. Using the Chebyshev inequality, calculate the minimum value of n that will achieve this objective.

   (d) If we agree that no Canadians are taller than three meters, why is it correct to use 1.5 meters as an upper bound on the standard deviation for X, the height of any Canadian selected at random?