STATS 116: Homework 3

Due: Thursday, July 20, 2023 at 10:00 pm PDT on Gradescope

There are 7 problems on this assignment, each worth 8 points, although subparts within a problem may not be equally weighted. Credit will be assigned primarily based on reasoning and work, not the final answer. You do not need to simplify arithmetic expressions unless otherwise noted. While you may discuss the problems on this assignment other students, you must write up your own solutions. As per the syllabus, you may occasionally use the Internet or other public resources to clarify concepts with citation when this information is used as part of your own solution to a homework problem. However, you may not search for direct solutions to any problems assigned for homework or exams. For example, you can ask ChatGPT to clarify a particular concept from lecture that may be related to a problem, but you cannot feed it any part of a course assignment or a substantively similar version.

1. (BH 3.3) Let X be a random variable with CDF F, and $Y = \mu + \sigma X$, where μ and σ are real numbers with $\sigma > 0$. Find the CDF F_Y of Y, in terms of F.

2. (BH 3.20) Suppose that a lottery ticket has probability p of being a winning ticket, independently of other tickets. A gambler buys 3 tickets, hoping this will triple the chance of having at least one winning ticket.

   (a) What is the distribution of how many of the 3 tickets are winning tickets?

   (b) Show that the probability that at least 1 of the 3 tickets is winning is $3p - 3p^2 + p^3$ in two different ways: by using inclusion-exclusion, and by taking the complement of the desired event and then using the PMF of a certain named distribution.

   (c) Show that the gambler's chances of having at least one winning ticket do not quite triple (compared with buying only one ticket), but that they do approximately triple if p is small.

3. Suppose X and Y are random variables for which $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ for all real x and y. In class, we mentioned that this implies X and Y are independent, i.e. $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ for all subsets A and B of $\mathbb{R}$. In this problem, we will show a weaker conclusion. For simplicity, you may assume X and Y are discrete and both supported on the set of integers $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$, though this is not actually necessary for these statements to hold.

   (a) Show that $P(X > x, Y > y) = P(X > x)P(Y > y)$ for any x, y.

   (b) Show that $P(X = x, Y = y) = P(X = x)P(Y = y)$ for any x, y.

4. (BH 3.30) A company with n women and m men as employees is deciding which employees to promote.

   (a) Suppose for this part that the company decides to promote t employees, where $1 \leq t \leq n + m$, by choosing t random employees (with equal probabilities for each set of t employees). What is the distribution of the number of women who get promoted?

   (b) Now suppose that instead of having a predetermined number of promotions to give, the company decides independently for each employee, promoting the employee with probability p. Find the distributions of the number of women who are promoted, the number of women who are not promoted, and the number of employees who are promoted.

5. (BH 4.26) Suppose Nick is flipping a nickel with probability p_1 of Heads and Penny is (independently) flipping a penny with probability p_2 of Heads. Let $X_1, X_2, \ldots,$ be Nick's results and $Y_1, Y_2, \ldots$ be Penny's results, with $X_i \sim \text{Bern}(p_1)$ and $Y_j \sim \text{Bern}(p_2)$.

   (a) Find the distribution and expected value of the first time at which they are simultaneously successful, i.e., the smallest n such that $X_n = Y_n = 1$. Hint: Define a new sequence of Bernoulli trials (coin flips).

   (b) Find the expected time until at least one has a success (including the success). Hint: Define a new sequence of Bernoulli trials.

   (c) For $p_1 = p_2$, find the probability that their first successes are simultaneous, and use this to find the probability that Nick's first success precedes Penny's.

6. The game of **wizards** is played with a standard deck of 52 cards, augmented with two jokers and six "wizard" cards for a total of 60 cards. At the beginning of the first round, the cards are shuffled a single card is dealt to each of three players, including you. The cards are played in sequence, starting with the first player, which is you. The rule is that the highest ranked card of the same suit as the card played by the first player (you) wins the round (ranks go from 2 at the lowest up to Ace at the highest), with the following exceptions:

   - The first player to play a wizard always wins
   - Spades are a "trump" suit, meaning that if the card played by the first player (you) is a diamond, heart, or club, then other players can win if they play a spade. If there are spades (but no wizards), the highest spade wins
   - Anyone who plays a joker cannot win the round

   Each player must declare (without seeing the other players' cards) whether they think they will win. A player gets -10 points if they declare incorrectly. They get 20 points if they correctly declare they will win, and 10 points if they correctly declare they will lose. You are dealt the King of Hearts and you'd like to maximize your expected score. Should you declare that you will win the trick?

7. The probability generating function of a random variable X supported on the nonnegative integers is denoted by G_X and given by $G_X(s) = \mathbb{E}[s^X]$ for all values of s where the expectation exists.

   (a) Suppose $X \sim \text{Bern}(p)$. Find $G_X(s)$ for any real s.

   (b) Now suppose $Y \sim \text{Pois}(\lambda)$. Find $G_Y(s)$ for any real s.

   (c) Show that $p_X(0) = G_X(0)$ where p_X is the PMF of X. Hint: Recall $0^0 = 1$.

   (d) Show that $\mathbb{E}[X] = G'_X(1)$, where $G'_X(s)$ is the derivative of the function G_X evaluated at s.