Homework 8

Linear Algebra and Probability, Fall 2024

Due Friday, November 8, at 5:00 PM

Read the instructions carefully on this assignment. In some problems, it will be helpful to use a computer. Here is a link to the code from class `https://colab.research.google.com/drive/1E5IcnsBv4zPQHrBZpgfBLqd4jDFwsSZO?usp=sharing`.

1. Let S be the sample space consisting of the 36 possible outcomes of rolling two 6-sided dice with numbers $1, 2, 3, 4, 5, 6$. Define an equiprobable probability function P on this space so that $P(\{(i, j)\}) = \frac{1}{36}$ for all i, j. Find and plot (on paper or computer is fine) the probability mass function (PMF), and cumulative distribution function (CDF) for the following two random variables $X : S \to \mathbb{R}$.

   (a) $X(i, j) = i + j$ (sum of the two dice)

   (b) $X(i, j) = \max(i, j)$ (maximum of the two dice)

2. In the semiconductor industry, the number of visual defects on a semiconductor unit is modeled by a Poisson random variable with rate $\lambda = 3$. Find the following:

   (a) The probability of no defects on the unit.

   (b) The probability that there are more than 3, but less than 6 defects on the unit.

   You should write out exact answers in terms of exponential and factorial functions, the compute them with a computer.

3. Two sports teams, team A and team B, are playing for the championship in a best-of-seven series. Suppose A will beat team B with probability 0.70 in any given game, and the games are independent. What is the probability that team A wins the series? (Note: In practice, teams will stop play after a team wins 4 games. But you can still imagine they will play all 7 games and ask whether team A will win at least 4 games). You should write your answer as a sum of terms involving binomial coefficients, then compute using a computer or calculator.

4. Prove that geometric random variables (with arbitrary parameter $p \in [0, 1]$) have the **memoryless property**. That is, for $k, j \geq 1$,

$$P(X > k + j | X > j) = P(X > k).$$

You may use the fact we showed in class that $P(X > k) = (1 - p)^k$.

The memoryless property can be stated in words as follows: given that you had to wait longer than j trials to see the first success, the probability that you must wait more than $k + j$ trials is the same as having to wait k trials without the conditioning. In other words, knowing how long you have been waiting for the first success does not affect the length of time it will take to see the first success. The geometric distribution and its continuous cousin, the exponential distribution, are the **only** probability distributions with this property.

5. Suppose we have $n = 200$ Bernoulli trials of a random experiment with probability $p = 0.01$. Use the binomial PMF function in Python to compute the probability of getting 4 successes. Also use the poisson pmf function in Python to compute $P(X = 4)$, if $X \sim \text{Poi}(\lambda = np)$ for $n = 200$ and $p = 0.01$. What do you notice about the two numbers you get?

6. Let X be a random variable with

$$P(X = 1) = \frac{1}{6}, \quad P(X = 2) = \frac{1}{6}, \quad P(X = 3) = \frac{1}{3}, \quad P(X = 4) = \frac{1}{3}.$$

Compute the mean, variance, and moment generating function of X. Your answer should be written in terms of sums of fractions, then you may use a calculator or computer to compute the answer.

7. For a Bernoulli random variable

$$P(X = 0) = 1 - p, \quad P(X) = p,$$

compute the variance of X. What value of p gives the largest variance? (Hint: you can use calculus or you can just plot it as a function of p). In loose terms, variance is a measure of how much uncertainty a random variable has. From this perspective of uncertainty, does your answer to which value of p maximizes the variance make sense?

8. Consider a data set with the measurements $10, 15, 16, 17, 18, 21, 13$. Compute the sample mean and unbiased sample variance of this data set (you may use a calculator or computer).

9. If X has moment generating function

$$M_X(t) = \frac{1}{4}e^{-5t} + \frac{1}{2}e^{2t} + \frac{1}{4}e^{7t},$$

find the mean of X.

10. Compute the moment generating function of the Poisson distribution. Use it compute the mean of the Poisson distribution. Hint: Note that $e^{tk}\lambda^k = (e^t\lambda)^k$ and use the formula

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

You will also need to use the chain rule at the end to find the mean.

Optional practice problems (don't turn these in) 2.11.12, 2.11.14, 2.11.16, 2.11.29, 2.11.37