

- (4 points) A grocer sells an average of  $\lambda = 2.8$  pineapples per day, and we will model the demand using a Poisson process. Suppose that the grocer sells them for \$1.00 each. To be safe, the grocer orders 6 pineapples each day, and that each pineapple costs the grocer \$0.40 - and that any unsold pineapples are thrown away at the end of each day. What is the expected profit (or loss) per day for the grocer from the pineapples?

*Solution.* Let  $X$  correspond to the demand for pineapples in a day. Then,  $X \sim \text{Poisson}(2.8)$ . The cost of ordering 6 pineapples is  $6 \cdot 0.40 = 2.40$ . The demand for pineapples can be broken up with the following table:

Demand ( $x$ )	Revenue	Probability	Revenue $\cdot$ Probability
0	\$0	0.608	\$0
1	\$1	0.1703	\$0.17
2	\$2	0.2384	\$0.48
3	\$3	0.2225	\$0.67
4	\$4	0.1557	\$0.62
5	\$5	0.0872	\$0.44
$\geq 6$	\$6	0.0651	\$0.39
<b>Total</b>			<b>\$2.77</b>

Accounting for the upfront cost of ordering, we get an expected profit of

$$2.77 - 2.40 = \$0.37.$$

2. Suppose an average of 3 people enter a store over an hour. We will use a Poisson process to determine the probability that exactly 2 customers enter during a two-hour long period:

- (a) (2 points) First, if we set out Poisson time interval to the full two hours, explain why we should use  $\lambda = 6$ .

*Solution.* The problem says that over the course of an hour, an average of 3 people enter the store. Since each occurrence is independent of the previous, then we can just add the averages together.

- (b) (2 points) Use this to determine  $P(\text{exactly 2 customers})$ .

*Solution.* If we let  $X$  count the number of customers, then  $X \sim \text{Poisson}(6)$ . Hence,

$$P(X = 2) = 0.0446.$$

- (c) (1 point each) But, we can also model this as two separate and independent one-hour periods, now with  $\lambda = 3$ . For a time interval of a single hour, determine:

- i.  $P(0 \text{ customers})$

*Solution.*  $P(X = 0) = 0.0498$

- ii.  $P(1 \text{ customer})$

*Solution.*  $P(X = 1) = 0.1494$

- iii.  $P(2 \text{ customers})$

*Solution.*  $P(X = 2) = 0.2240$

- (d) (2 points) Briefly, explain why in our pair of one-hour time frames, there are three distinct ways for 2 people to come over two total hours. [Hint: What if no customers come in the first hour? What if one does?]

*Solution.* The three distinct time frames are as follows:

- i. No one comes in the first hour, and two come in the second hour.
- ii. One comes in the first hour, and one comes in the second hour.
- iii. Two come in the first hour, and no one comes in the second hour.

These outcomes are possible because of the independent nature of how the hour intervals interact with each other.

- (e) (2 points) For each of these possibilities, determine their probability, and then sum of their probabilities. What do you notice, relative to your answer from part (b) above?

*Solution.* Let each item in the following list correspond to the previous part's time frames (just so I don't have to write them twice). Since the two trials are independent, we multiply the probabilities together.

- i.  $P(X = 0) \cdot P(X = 2) = 0.0498 \cdot 0.2240 = 0.0112$ .
- ii.  $P(X = 1) \cdot P(X = 1) = 0.1494 \cdot 0.1494 = 0.0223$ .
- iii.  $P(X = 2) \cdot P(X = 0) = 0.2240 \cdot 0.0498 = 0.0112$ .

Summing these probabilities,  $(0.0112 \cdot 2) + 0.0223 = 0.0446$ , we get the exact answer from part (b).

3. (2 points each) Suppose that  $X \sim \text{Binomial}(500, 0.02)$ .

(a) Determine  $E[X]$  and  $\text{Var}[X]$ .

*Solution.* For the mean:

$$E[X] = np = 500 \cdot 0.02 = 10.$$

For the variance:

$$\text{Var}[X] = np(1 - p) = 500 \cdot 0.02 \cdot 0.98 = 9.8.$$

(b) Determine  $P(X \leq 9)$ .

*Solution.*  $P(X \leq 9) = 0.4567$ .

4. (2 points each?) Let  $Y \sim \text{Poisson}(10)$ .

(a) Determine  $E[Y]$  and  $\text{Var}[Y]$ .

*Solution.* The mean and variance are both equal to  $\lambda$ :

$$E[Y] = \lambda = 10 = \text{Var}[Y].$$

(b) Determine  $P(Y \leq 9)$ .

*Solution.*  $P(Y \leq 9) = 0.4579$ .

5. (2 points) Compare the previous two problems. [Just write a sentence or two here.]

*Solution.* For both distributions, we get around the same expected value, variance, and cumulative probability for  $X, Y \leq 9$ . Maybe there is a connection when  $\lambda = np$ ?

6. (4 points) We will now justify the result of the previous page. Suppose that  $n$  is very large and  $p$  is very small. Let  $\lambda = np$ . We want to explain why, if  $B$  and  $L$  are random variables so that  $B \sim \text{Binomial}(n, p)$  and  $L \sim \text{Poisson}(\lambda)$ , then  $P(B = x) \approx P(L = x)$  for small  $x$ .

You have two options - you only need to do one, but are welcome to do both; if both are well done, I might give some bonus points.

- (a) Make an intuitive argument - talk about how, while Poisson has an infinite support, when  $\lambda$  is smallish, the effective support (i.e., the set of values that  $L$  could take on with reasonable probabilities) is finite. Think about how this might relate to having a fixed number of trials in a binomial-type experiment.

*Solution.* I think it's easier to think about this problem from the binomial perspective first. When we have a very large  $n$  with a very small  $p$ , we are essentially breaking up a large interval into many partitions and attributing a small probability to each partition. For example, consider the probability of winning the lottery. For many people, say  $n$  of them, they each have a small probability of winning,  $p$ . That is why we only have a couple of winners, even though so many people are playing.

Now, for the perspective of the Poisson distribution, we can look at the problem from a similar lens. However, instead of treating each person as a trial, we can instead look at the relative amount of winners,  $\lambda$ . This hits at the main point of having an infinite support. While the count of winners could theoretically consist of any non-negative integer, it is more probable that this count will be a very small number.

Therefore, since both the binomial (with very large  $n$  and small  $p$ ) and the Poisson (with small  $\lambda$ ) concentrate nearly all their probability mass on the same small set of outcomes  $\{0, 1, 2, \dots\}$ , the calculations for  $P(X = x)$  for those small values of  $x$  are nearly identical.

- (b) Explain - using calculations and at least pointing at the limits - why the pmfs for  $B$  and  $L$  are approximately equal for smallish  $x$ . Two hints: You might rewrite the pmf for  $B$  replacing any  $p$  with  $\lambda/n$  and should recall from Calculus I that  $\lim_{n \rightarrow \infty} (1 - \frac{\lambda}{n})^n = e^{-\lambda}$ .

*Solution.* I'm skipping this one.

7. (3 points each) Suppose that  $X$  is a random variable which is geometrically distributed, with probability of success  $p$  and failure  $q = 1 - p$ .

- (a) Suppose that  $a$  is a positive integer. Show that  $P(X > a) = q^a$ .

*Solution.* From the CDF, we know:

$$F(a) = P(X \leq a) = 1 - (1 - p)^a = 1 - q^a.$$

Hence,

$$P(X > a) = 1 - P(X \leq a) = 1 - (1 - q^a) = q^a.$$

- (b) Now, let  $b$  be another positive integer. Show that  $P(X > a + b \mid X > a) = P(X > b)$ . That is, show that the probability of needing more than  $b$  failures before a success is equal to the probability that, given we have already seen at least  $a$  failures that we'll see at least  $b$  more.

*Solution.* The conditional probability can be written as

$$P(X > a + b \mid X > a) = \frac{P((X > a + b) \cap (X > a))}{P(X > a)}.$$

Now, we need to find  $P((X > a + b) \cap (X > a))$ . If the first success occurs after trial  $a + b$ , then it has certainly occurred after trial  $a$ . This means that any outcome in set  $\{X > a + b\}$  is also in set  $\{X > a\}$ . Therefore,  $\{X > a + b\} \subseteq \{X > a\}$ . Then, we know that when we take the intersection of a set with one of its subset, it results the subset itself:  $\{X > a + b\} \cap \{X > a\} = \{X > a + b\}$ .

Now we can use the result from part (a) and solve the equation:

$$P(X > a + b \mid X > a) = \frac{P((X > a + b) \cap (X > a))}{P(X > a)} = \frac{P(X > a + b)}{P(X > a)} = \frac{q^{a+b}}{q^a} = q^b = P(X > b).$$

8. Suppose that  $X$  is a discrete random variable. Define the value  $m(X)$  to be the smallest  $x$  so that

$$P(X \leq x) \geq 0.5.$$

- (a) (3 points) Briefly, explain why  $m(X)$  must exist, and must be in the support of  $X$  - i.e.,  $f(x) > 0$ . [Hint: You can assume that there is a smallest  $x$  in the support overall.]

*Solution.* The CDF of  $X$  must increase from 0 to 1 over its support. To do so, the CDF must pass the 0.5 mark. The first point where it crosses is the value  $m(X)$  by definition. Since the CDF only increases in regions of non-zero probability,  $m(X)$  must also have non-zero probability. Therefore,  $m(X)$  is in the support of  $X$ .

- (b) (2 points) Briefly, explain how you can use trial and error to find  $m(X)$  for any random variable whose CDF the calculator has built in.

*Solution.* Start with  $x = 1$ , and keep inputting larger integer values for  $x$  in the formula `geometcdf( $p, x$ )` until you find a value that is closest to 0.5.

- (c) (5 points) Fill in the following table:

$X$ is distributed	$E(X) = 1/p$	$m(X)$
$X \sim \text{geometric}(0.6)$	1.67	1
$X \sim \text{geometric}(0.3)$	3.33	2
$X \sim \text{geometric}(0.1)$	10	7
$X \sim \text{geometric}(0.05)$	20	14

- (d) (3 points) As  $p$  gets smaller, what can you say about the values of  $E(x)$ ,  $m(X)$ , and their relative positions?

*Solution.* As  $p$  gets smaller, both  $E(X)$  and  $m(X)$  get larger, and both values tend to double when  $p$  halves. In terms of their relative positions,  $E(X)$  will always be larger than  $m(X)$ .