Problem Set 3:

Problem 2:

Lyft line gets 2 requests every 5 minutes, on average, for a particular route (request times are independent). A user requests the route and Lyft commits a car to take her. All users who request the route in the next five minutes will be added to the car as long as the car has space. The car can fit up to three users. Lyft will make $7 for each user in the car (the revenue) minus $9 (the operating cost).
a. How much does Lyft expect to make from this trip?
b. Lyft has one space left in the car and wants to wait to get another user. What is the probability that another user will make a request in the next 30 seconds?

Solution:

(a) **5**. We can find the expected requests using the Poisson distribution we are given the average requests in given time frame. After finding the expected requests we multiply with 7 to find the amount earned then subtract 9 for operating cost:

$$X \sim Poi(\lambda = 2), \quad E[X] = \lambda = 2, \implies 7 \times 2 - 9 = 5.$$

(b) **~0.165**. We can find the probability with first writing the binomial form, with p being $\frac{\lambda}{m}$(m is the amount of events in 5 min):

$$X \sim Bin(n, \frac{\lambda}{m}) \implies \binom{n}{1}(\frac{\lambda}{m})^1(1 - \frac{\lambda}{m})^{n-1}.$$

Since we are only waiting for one person $k = 1$ and we don't need to divide the time too many times:

$$\binom{30}{1}(\frac{2}{300})^1(1 - \frac{2}{300})^{30-1}$$

```
>>> import scipy.stats as stats
>>> stats.binom.pmf(1, 30, 1/150)
0.16473476739077328
```

Problem 3:

Suppose it takes at least 9 votes from a 12-member jury to convict a defendant. Suppose also that the probability that a juror votes that an actually guilty person is innocent is 0.25, whereas the probability that the juror votes that an actually innocent person is guilty is 0.15. If each juror acts independently and if 70% of defendants are actually guilty

a. Find the probability that the jury renders a correct decision.

b. Determine the percentage of defendants found guilty by the jury.

Solution:

(a) **~0.726**. Can be solved using the binomial distribution:

$$P(9 \leq X \leq 12|Guilty)P(Guilty) = \sum_{i=9}^{12} \binom{12}{i} (0.75)^i (0.25)^{12-i} (0.70),$$
$$P(9 \leq X \leq 12|Not\ Guilty)P(Not\ Guilty) = \sum_{i=9}^{12} \binom{12}{i} (0.85)^i (0.15)^{12-i} (0.30),$$
$$\implies P(9 \leq X \leq 12|Guilty)P(Guilty) + P(9 \leq X \leq 12|Not\ Guilty)P(Not\ Guilty).$$

```
>>> g = stats.binom.cdf(12, 12, 0.75) - stats.binom.cdf(8, 12, 0.75)
>>> ng = stats.binom.cdf(12, 12, 0.85) - stats.binom.cdf(8, 12, 0.85)
>>> g, ng
(0.6487786173820496, 0.9077936688387293)
>>> g*0.70 + ng*0.30
0.7264831328190535
```

(b) **~45.42%**. Similar to part (a):

$$P(9 \leq X \leq 12|Guilty)P(Guilty) = \sum_{i=9}^{12} \binom{12}{i} (0.75)^i (0.25)^{12-i} (0.70),$$
$$P(9 \leq X \leq 12|Not\ Guilty)P(Not\ Guilty) = \sum_{i=9}^{12} \binom{12}{i} (0.15)^i (0.85)^{12-i} (0.30),$$
$$\implies P(9 \leq X \leq 12|Guilty)P(Guilty) + P(9 \leq X \leq 12|Not\ Guilty)P(Not\ Guilty)$$
$$\implies p \times 100.$$

```
>>> g = stats.binom.cdf(12, 12, 0.75) - stats.binom.cdf(8, 12, 0.75)
>>> ng = stats.binom.cdf(12, 12, 0.15) - stats.binom.cdf(8, 12, 0.15)
>>> g, ng
(0.6487786173820496, 5.477914412854723e-06)
>>> percent = (g*0.70 + ng*0.30) * 100
>>> percent
45.41466755417585
```

Problem 4:

To determine whether they have measles, 1000 people have their blood tested. However, rather than testing each individual separately (1000 tests is quite costly), it is decided to use a group testing procedure:

- Phase 1: First, place people into groups of 5. The blood samples of the 5 people in each group will be pooled and analyzed together. If the test is positive (at least one person in the pool has measles), continue to Phase 2. Otherwise send the group home. 200 of these pooled tests are performed.
- Phase 2: Individually test each of the 5 people in the group. 5 of these individual tests are performed per group in Phase 2.
  Suppose that the probability that a person has measles is 5% for all people, independently of others, and that the test has a 100% true positive rate and 0% false positive rate (note that this is unrealistic). Using this strategy, compute the expected total number of blood tests (individual and pooled) that we will have to do across Phases 1 and 2.

Solution:

250 $\pm$ **15**. First we find the expected group amount having no one with measles:

$$X \sim Bin(200, 0.95) \implies E(X) = \sum_{k=0}^{200} \binom{200}{k} (0.95)^k (0.05)^{200-k}$$

```
>>> mean = stats.binom.mean(200, 0.95)
>>> mean
190.0
```

Since there we expect only 10 group to contain at least one person with measles we do 10×5 more tests which will be 250 tests in total. Note that since the standard deviation is given as 3 below the amount of tests done will be within 15 tests of the amount found 65% of the time.

```
>>> std = stats.binom.std(200, 0.95)
>>> std
3.0
```

Problem 5:

The number of times a person's computer crashes in a month is a Poisson random variable with λ = 7. Suppose that a new operating system patch is released that reduces the Poisson parameter to λ = 2 for 80% of computers, and for the other 20% of computers the patch has no effect on the rate of crashes. If a person installs the patch, and has their computer crash 4 times in the month thereafter, how likely is it that the patch has had an effect on the user's computer (i.e., it is one of the 80% of computers that the patch reduces crashes on)?

Solution:

~0.698. We can use the Bayes theorem to find the answer to the problem:

$$P(A|x=4) = \frac{P(A,x=4)}{P(x=4)},$$
$$\text{where,}$$
$$A : \text{patch worked,}$$
$$A^C : \text{patch didn't work,}$$
$$x = 4 : \text{amount of crashes} = 4$$

We can find $P(x=4)$ i.e. denominator like we did in problem 3:

```
>>> a = stats.poisson.pmf(3, 7)
>>> a_c = stats.poisson.pmf(4, 2)
>>> a, a_c
(0.052129252364199796, 0.09022352215774178)
>>> x = a*0.80 + a_c*0.20
>>> x
0.059748106322908195
```

And since we have already calculated the nominator which is a*0.80, we can find the answer:

$$P(A|x=4) = \frac{P(A,x=4)}{P(x=4)} = \frac{a*80}{x}.$$

```
>>> (a*80) / x
0.697987006750877
```

Problem 6:

Let X be a continuous random variable with probability density function:

$$f(x) = \begin{cases} c(2-2x^2) & \text{if } -1 < x < 1, \\ 0 & \text{otherwise.} \end{cases}$$

```
a. What is the value of $c$?
b. What is the cumulative distribution function (CDF) of $X$?
c. What is $E[X]$?
```

Solution:

(a) **0.25**. Since the pdf must sum to 1:

$$\int_{-1}^{1} c(2-2x^2)dx = c(2-\frac{2}{3}x^3)\mid_{-1}^{1} = c(2-\frac{2}{3}) - c(2+\frac{2}{3}) = 4c = 1$$
$$\implies c = \frac{1}{4}$$

(b) $F(x) = \begin{cases} \frac{1}{2x^3} & \text{for } -1 \leq x \leq 1, \\ 1 & \text{for } 1 < x \\ 0 & \text{for } x < -1 \end{cases}$. Since the we have a continuous random variable we integrate:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{2}(1-t^2)dt = \frac{1}{2}(1-\frac{1}{3}t^3)\mid_{-\infty}^{x}$$
$$\implies \frac{1}{2}(\frac{1}{t^3}+\frac{1}{3})\mid_{-\infty}^{x} = \frac{1}{2}((\frac{1}{x^3}+\frac{1}{3})-(\frac{1}{3})) = \frac{1}{2x^3}$$

(c) . To find the expectation we multiply x with $f(x)$ and integrate the whole expression over R:

$$E(x) = \int_{-\infty}^{\infty} x \times \frac{1}{2}(1-x^2)dx = \int_{-\infty}^{\infty} \frac{1}{2}(x-x^3)dx = \frac{1}{2}(\frac{x^2}{2}-\frac{x^4}{4})\mid_{-\infty}^{\infty}$$
$$\implies \frac{x^2}{4}(1-\frac{x^2}{2})\mid_{-\infty}^{\infty}$$