Homework 3

Stat 345 - Spring 2020

Name: ______________________________

Problem 1

Let X be a random variable with the following probability mass function:

$$p(x) = \frac{3x+1}{22} \qquad x = 0, 1, 2, 3$$

a) Verify that this pmf is a valid probability mass function.

b) Find $P(1 \leq X < 3)$ using pmf.

c) Determine the cumulative distribution function (cdf) of X.

d) Determine expected value and variance for X.

Problem 2

Let X be a random variable with probability density function

$$f(x) = c(1 - x^2) \qquad -1 < x < 1$$

and $f(x) = 0$ otherwise

a) What is the value of c?

b) What is the cumulative distribution function of X?

c) Find $P(-0.5 \leq X < 0.2)$.

d) Determine expected value and variance for X.

Problem 3

Suppose there are 3.4 millions flights every month worldwide. The probability that the commercial airplane will crash is 10^{-6}. What is the probability that there will be

a) at least 2 such accidents in the next month?

b) at most 1 accident in the next month?

Problem 4

If the probability of hitting a target is 0.2, and 11 shots are fired independently. Define the random variable X and its distribution.

a) What is the probability of the target being hit at least 3 times?

b) What is the probability of scoring more hits than misses?

c) What is the expected value and variance of the number of hits in 11 shots?

d) What is the conditional probability that the target is hit at least 3 times, given that it is hit at least once?

e) Suppose the person has to pay $5 to enter the shooting range and he gets $3 dollars for each hit. Find the expectation and the variance of the profit.

Problem 5

The life of a semiconductor laser at a constant power is normally distributed with a mean of 7000 hours and a standard deviation of 600 hours.

a) What is the probability that a laser fails before 5000 hours?

b) What is the life in hours that 95% of the lasers exceed?

c) If four lasers are used in a product and they are assumed to fail independently, what is the probability that all four are still operating after 7000 hours?

d) *(bonus (5 pts))* Given that a laser already lasts 5000 hours, what is the probability that it lasts at least another 3000 hours?

Problem 6

Insects are expected to be attracted to the roses. A commercial insecticide is advertised as being 98% efective. Suppose 2000 insects infest a rose garden where the insecticide has been applied, and let X=number of surviving insects. Evaluate the probability that fewer than 35 insects survive.

a) Write the probability mass function using the Binomial distribution. Write the expression for the probability that fewer than 35 insects survive. Do not calculate.

b) Write the probability mass function using the Poisson distribution. Write the expression for the probability that fewer than 35 insects survive. Do not calculate.

c) Use the normal approximation to the probability that fewer than 35 insects survive. Calculate the probability.

Problem 7

Imagine that subway trains in New York City always arrive exactly on time and run every day 24 hours a day, with the time between successive trains fixed at 12 minutes. Bob arrives at the train stop at a uniformly random time on a certain day (the time that Bob arrives is independent of the train arrival process).

a) What is the distribution of how long Bob has to wait for the next train? What is the average time that he has to wait?

b) Given that the train has not yet arrived after 7 minutes, what is the probability that Bob will have to wait at least 2 more minutes?

c) Bob moves to Los Angeles where subway system is less organized. Now, when any train arrives, the time until the next train arrives is an Exponential random variable with mean 12 minutes. Bob arrives at the train stop at a random time. What is the distribution of his waiting time for the next train? What is the average time that Bob has to wait?

d) When Bob complains to a friend how much worse transportation is in Los Angeles, the friend says that Bob arrives at a uniform instant between the previous train arrival and the next train arrival. The average length of that interval between trains is 12 minutes, but since Bob is equally likely to arrive at any time in that interval, his average waiting time is only 6 minutes. Explain what is wrong with his friend's reasoning.

e) *(bonus (5 pts))* Tired of the poor public transportation in Los Angeles, Bob bought the car. Eventually, in a few years Bob moves back to New York City and wants to sell his car. He decides to sell it to the first person to offer at least \$13,000 for it. Assume that the offers are independent Exponential random variables with mean \$10,000. Find the expected number of offers Bob will have.

Problem 8

A truncated discrete distribution is one in which a particular class cannot be observed and is eliminated from the sample space. In particular, X has range $0, 1, 2, ...$ and 0 class cannot be observed (as is usually the case), the zero-truncated random variable T has pmf

$$P(T = x) = \frac{P(X = x)}{P(X > 0)}, \quad x = 1, 2, ...$$

a) Find pmf of the zero-truncated Poisson random variable T if $X \sim Poisson(\lambda)$

b) *(bonus (5 pts))* Find the expected value of the zero-truncated Poisson random variable T.

c) Find pmf of the zero-truncated Negative Binomial random variable T if $X \sim NegativeBinomial(r, p)$

d) *(bonus (5 pts))* If we let $r \to 0$ in c), we get an interesting distribution, called the *logarithmic series distribution*. A random variable Y has a logarithmic series distribution with parameter p if

$$P(Y = y) = \frac{-(1-p)^y}{y log(p)}, \quad y = 1, 2, ..., \quad 0 < p < 1$$

Find the expected value of Y.

Interesting facts: *Zero-truncated discrete distributions* (there are many of them) can model the count data when the number of instances or individuals falling into the zero-category class cannot be determined. For example, zero-truncated discrete distributions can model the length of hospital stay in days (the minimum of at least one day is required) or model the numbers of traffic violation for drivers during a certain period (there will be no record of those who have received no tickets). The *logarithmic series distribution* has proven useful in modeling how rare a species is relative to other species in a certain location. It is also can be used to model the distribution of numbers of items of a product purchased by a customer during a certain time period.