Name: Ivan Leung

Student ID: 67485846

STAT 67

DUE: Date posted on Canvas

Instructions

Please read the following instructions carefully:

1. You must demonstrate your work/show your thinking to receive full credit.
2. Please write your answers on the exam in the appropriate space below each question. If you are writing on your own paper, please label neatly.
3. Answers must be rounded to an appropriate number of decimal places, when in doubt, round to 4 decimal places.
4. Please submit the PDF to gradescope by 6:00 PM PST on the due date.

The table below will be used for grading purposes.

Question	Points	Score
1	11	
2	15	
3	10	
4	17	
5	17	
6	13	
7	17	
Total:	100	

Distribution	pmf	Distribution	pdf
Bernoulli	$f(x) = \begin{cases} p & x = 1 \\ 1-p & x = 0 \end{cases}$	Exponential	$\lambda e^{-\lambda x}$ for $x > 0$
Binomial	$f(x) = \binom{n}{x} p^x (1-p)^{n-x}$ for $x = 0, 1, \cdots$	Uniform	$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & else \end{cases}$
Geometric	$f(x) = (1-p)^{x-1} p$ for $x = 1, 2, 3, \cdots$	Normal	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$ for $x \in (-\infty, \infty)$
Poisson	$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ for $x = 0, 1, 2, \cdots$		

1. A sample of 200 individuals is selected, and the individuals are classified according to their smoking habits and the smoking habit of their parents. The results are given in the contingency table below.

Individual's smoking habit	Parents' smoking habit			Total
	Both parents smoke	Only one parent smokes	Neither parent smokes	
Smokes	36	48	16	100
Does not Smoke	30	42	28	100
Total	66	90	44	200

(a) (3 points) What is the probability that an individual has at least one parent who smokes?

$$P(\text{at least one parent}) = \frac{66+90}{200} = \boxed{0.78}$$

(b) (4 points) What is the probability that an individual smokes or has parents who both smoke?

$$P(\text{smoke} \cup \text{Both parents}) = P(\text{Smoke}) + P(\text{Both parents}) - P(\text{Smoke} \cap \text{Both parents})$$

$$= \frac{100}{200} + \frac{66}{200} - \frac{36}{200} = \frac{130}{200} = \boxed{0.65}$$

(c) (4 points) Given that an individual in the sample does not smoke, what is the probability that neither of the individual's parents smokes?

$$P(\text{neither parents} \mid \text{not smoke}) = \frac{28}{100} = \boxed{0.28}$$

2. The number of days ahead of time that travelers purchase their airline tickets can be modeled by an exponential distribution with $\lambda = \frac{1}{14}$

(a) (4 points) What is the *expected value* of the number of days ahead of time a traveler will purchase an airline ticket? What is the *variance* of the number of days ahead of time a traveler will purchase an airline ticket? What is the *standard deviation* of the number of days ahead of time a traveler will purchase an airline ticket? $E(X) = \frac{1}{\lambda} = \boxed{14 \text{ days}}$

$$\text{Var}(X) = \frac{1}{\lambda^2} = \boxed{196 \text{ days}^2}$$

$$\sigma = \sqrt{\text{Var}(X)} = \sqrt{196} = \boxed{14 \text{ days}}$$

(b) (5 points) Find the probability that the number of days ahead a traveler will purchase an airline ticket is between 10 and 16 days.

$$P(10 < X < 16) = P(X < 16) - P(X < 10)$$

$$= \text{pexp}(16, \tfrac{1}{14}) - \text{pexp}(10, \tfrac{1}{14})$$

$$= \boxed{0.1706}$$

(c) (6 points) Given that the number of days ahead of time a traveler will purchase an airline ticket is more than 7 days, what is the probability that the number of days the traveler purchases the ticket is more than 9 days?

$$P(X > 9 \mid X > 7) = P(X > 2) = 1 - P(X < 2)$$

$$= 1 - \text{pexp}(2, \tfrac{1}{14})$$

$$= \boxed{0.8669}$$

3. In the statistic department of UCI, there are 60 graduate students in total, 45 PhD students and 15 master students. Among master students, 40% of them are international students. Among PhD students, 20% are international students.

   (a) (5 points) What is the proportion of international students in the statistic department?

   $$P(\text{international students}) = \frac{0.40(15) + 0.20(45)}{60} = \boxed{0.25}$$

   (b) (5 points) What is the proportion of master students among all domestic students.

   $$P(\text{master} \mid \text{domestic}) = \frac{15}{0.6(15) + 0.8(45)} = \boxed{0.3333}$$

4. You have joined the development team of a new search engine. You are in charge of analyzing the time taken to respond to a search query. The search time of the current algorithm is normally distributed, with mean 0.79 seconds and standard deviation 0.21 seconds.

   (a) (2 points) Let X be the algorithm's search time. Describe the distribution of X.

   $$\boxed{X \sim \text{Normal}(0.79, 0.21)}$$

   (b) (2 points) According to the empirical rule, where do approximately 68% of the values fall?

   upper: $0.79 + 0.21 = 1.00$

   lower: $0.79 - 0.21 = 0.58$

   Approximately 68% of the values fall between 1 second and 0.58 seconds

   (c) (4 points) What is the probability that a random search takes more than 1.21 second to be responded?

   $$P(X > 1.21) = 1 - P(X \leq 1.21)$$
   $$= 1 - \text{pnorm}(1.21, 0.79, 0.21)$$
   $$= \boxed{0.0228}$$

   (d) (4 points) What is the probability that a random search is responded in more than 1.063 but less than 1.126 seconds?

   $$P(1.063 < X < 1.126) = P(X < 1.126) - P(X < 1.063)$$
   $$= \text{pnorm}(1.126, 0.79, 0.21) - \text{pnorm}(1.063, 0.79, 0.21)$$
   $$= \boxed{0.0420}$$

   (e) (5 points) What is the value x^* of X such that 70 per cent of the searches take longer than x^*?

   $$x_c = z_c \sigma + \mu$$
   $$= \text{qnorm}(0.70, 0.79, 0.21)$$
   $$= \boxed{0.9001 \text{ seconds}}$$

5. You are playing Super Mario Bros together with 2 of your friends. You got to level 4, where you encounter your nemesis Bowser. Bowser is very strong, and he is defeated only 41% of the times. Each of you will play level 4 one time. $n = 3 \quad p = 0.41$

   (a) (2 points) Let X be the total number of times that Bowser is defeated. What is the distribution of X?

   $\boxed{X \sim \text{Binomial}(3, 0.41)}$

   (b) (3 points) What is the probability that only 1 of you defeats Bowser?

   $P(X=1) = \binom{3}{1}(0.41)^1(1-0.41)^2 = \text{dbinom}(1, 3, 0.41)$
   $= \boxed{0.4282}$

   (c) (2 points) You want to understand how likely it is to correctly predict the number of times Bowser is defeated. What is the variance of X?

   $Var(X) = npq = 3(0.41)(1-0.41)$
   $= \boxed{0.7257}$

   (d) (1 point) What is the probability that you beat Bowser - regardless of whether your friends beat him or not? $n = 1 \quad p = 0.41$

   Since it has only one trial, it is Bernoulli Distribution.

   $P(X=1) = p = \boxed{0.41}$

   Suppose that, after your friends are gone, you decide to play level 4 until you beat Bowser. Let Y be the number of times you play level 4.

   (e) (3 points) What is the distribution of Y?

   $\boxed{Y \sim \text{Geometric}(0.41)}$

   (f) (3 points) What is the probability that you play less than 3 times?

   $P(X<3) = P(X \leq 2) = 1 - (1-0.41)^2 = \text{pgeom}(2-1, 0.41)$
   $= \boxed{0.6519}$

   (g) (3 points) What is the expected number of times that you play?

   $E(X) = \frac{1}{p} = \frac{1}{0.41} = \boxed{2.439 \text{ plays}}$

6. Suppose 15% of people do not own a calculator, 40% of people own one calculator, 40% own two calculators, and the remaining 5% own three calculators. Let X be the number of calculators that a randomly selected person owns.

x = number of calculators	0	1	2	3
$P(X = x) = f(x)$	0.15	0.40	0.40	0.05

(a) (3 points) What is the probability a person owns fewer than 3 calculators?

$$P(X<3) = P(X=0)+P(X=1)+P(X=2)$$
$$= 0.15 + 0.4 + 0.4$$
$$= \boxed{0.95}$$

(b) (3 points) What is the expected number of calculators that a person owns?

$$E(X) = 0(0.15) + 1(0.40) + 2(0.40) + 3(0.05)$$
$$= \boxed{1.35 \text{ calculators}}$$

(c) (3 points) What is E[X^2]?

$$E(X^2) = 0^2(0.15) + 1^2(0.40) + 2^2(0.40) + 3^2(0.05)$$
$$= \boxed{2.45 \text{ calculators}^2}$$

(d) (4 points) What is the standard deviation for the number of calculators that a person owns?

$$\sigma = \sqrt{Var(X)} = \sqrt{E(X^2) - (E(X))^2}$$
$$= \sqrt{2.45 - 1.35^2}$$
$$= \boxed{0.7921 \text{ calculators}}$$

7. A tunnel opens at 7am and on average 27 red trucks enter this tunnel from 7am to 10am on Monday mornings. Suppose the red trucks arrive independent of each other and at a constant rate.

(a) (1 point) Let X be the number of red trucks that pass through the tunnel between 7am and 10am over the next Monday. What is the distribution of X?

$$\boxed{X \sim \text{Possion}\left(\lambda = 27 \frac{\text{red trucks}}{3 \text{ hrs}}\right)}$$

(b) (2 points) Again let X be the number of red trucks that pass through the tunnel between 7am and 10am next Monday. How many red trucks would you expect to pass through the tunnel between 7am and 10am next Monday?

$$E(X) = \lambda = \boxed{27 \text{ red trucks}}$$

(c) (5 points) What is the probability that 8 red trucks pass through the tunnel between 8am and 9am? **State the appropriate distribution** and any parameter values for any random variable(s) you use to model the situation. Write the probability statement and show your work in order to solve the problem.

$$Y \sim \text{Possion}\left(\lambda = 9 \frac{\text{red trucks}}{1 \text{ hrs}}\right)$$

$$P(Y = 8) = \frac{e^{-\lambda}\lambda^{y}}{y!} = \frac{e^{-9}(9)^8}{8!} = \text{dpois}(8, 9)$$

$$= \boxed{0.1318}$$

(d) (4 points) Suppose it takes a half hour for a red truck to pass through the tunnel. If there are no red trucks in the tunnel when it enters the tunnel at 7:35am on a Monday, what is the probability it will be the only red truck in the tunnel the whole time it spends in the tunnel? **State the appropriate distribution** and any parameter values for any random variable(s) you use to model the situation. Write the probability statment and show your work to receive full credit.

The amount of time until the next red truck arrives

$$Z \sim \text{Exponential}(9)$$

$$P(Z > 0.5) = 1 - P(Z \leq 0.5) = 1 - (1 - e^{-(9 \times 0.5)} = \boxed{0.0111}$$

(e) (5 points) Let W represent the amount of time in hours it takes for the 9^{th} red truck to arrive at the tunnel on a Monday morning. What ***time*** do you expect the 9^{th} red truck to arrive at the tunnel on a Monday morning (to the nearest 10 minutes)? Recall the tunnel opens at 7am. Your final answer should be a ***time***.

Lets break up W into the sum of 9 random varicbles

Let W_1 = the waiting time until the first red truck

W_2 = the waiting time between the arrival of the first and second trucks

$\vdots$

W_9

$$W_i \sim \text{Exponential}\left(\lambda = 9 \frac{\text{red trucks}}{\text{hour}}\right) \qquad W = W_1 + W_2 + \ldots + W_9$$

$$E(W) = E(W_1 + W_2 + \ldots + W_9)$$

$$= E(W_1) + E(W_2) + \ldots + E(W_9)$$

$$= \sum_{i=1}^{9} E(W_i) = 9E(W_1) = 9\left(\tfrac{1}{9}\right) = 1 \text{ hour}$$

7am + 1 hour

$$= \boxed{8 \text{ am}}$$