Statistical Analysis I (STAT-1301)
Practice Problems with Solutions
Chapter 5

[Question 1]

According to a survey, 35% of employees working at a very large company are happy with their jobs. Suppose that two employees are selected at random from this company. Let X denote the number of employees in this sample of two who are happy with their jobs.

a) Construct the probability distribution table of X.

x	$P(x)$
0	0.4225
1	0.455
2	0.1225
Total	1

- $p(0) = P(\text{None of the two selected employees are happy with their jobs}) =$
  $P(\text{The first selected employee is not happy with his job}) \times \times$
  $P(\text{The second selected employee is not happy with his job}) = 0.65 \times 0.65 = 0.4225.$

- $p(1) = P(\text{One of the two selected employees is happy with their jobs}) =$
  $P(\text{The first selected employee is happy with his job}) \times \times$
  $P(\text{The second selected employee is not happy with his job}) +$
  $+P(\text{The first selected employee is not happy with his job}) \times \times$
  $P(\text{The second selected employee is happy with his job}) = 0.35 \times 0.65 + 0.65 \times 0.35 =$
  $0.455.$

- $p(2) = P(\text{Both of the selected employees are happy with their jobs}) =$
  $P(\text{The first selected employee is happy with his job}) \times \times$
  $P(\text{The second selected employee is happy with his job}) = 0.35 \times 0.35 = 0.1225.$

b) Compute the mean and the variance of the number of employees who are happy with their job in this sample.

- $\mu = \sum xp(x) = 0 \times 0.4225 + 1 \times 0.455 + 2 \times 0.1225 = 0.7.$
- $\sigma^2 = \sum x^2 p(x) - \mu^2 = 0.945 - 0.7^2 = 0.945 - 0.49 = 0.455.$

[Question 2]

The table shows the probability that a person chosen at random from a certain community will have the indicated number of children.

x	Probability
0	0.54
1	0.23
2	$3k$
3	k
4	0.03

a) Determine the value of k.

- $\sum p(x) = 1 \Longrightarrow 0.54 + 0.23 + 3k + k + 0.03 = 1 \Longrightarrow 4k = 1 - 0.8 \Longrightarrow k = \frac{0.2}{4} = 0.05.$

b) Determine the probability that a person chosen at random from this population will have at least one child.

- $P(At\ least\ one) = 1 - p(None) = 1 - 0.54 = 0.46.$

c) Determine the probability that a person chosen at random from this population will have at most 2 children.

- $P(At\ Most\ Two) = 1 - P(None, or\ One\ or\ Two) = 0.54 + 0.23 + 3 \times 0.05 = 0.92.$

[Question 3]

The following table lists the probability distribution of the number of patients entering the emergency room during a 1-hour period at Victorial Hospital.

$\boldsymbol{x}$	$\boldsymbol{p(x)}$	$\boldsymbol{P(X \le x)}$
0	0.45	?
1	0.375	?
2	$2a - 0.3$	?
3	0.025	?

a) Find the value of a.

- $0.45 + 0.375 + 2a - 0.3 + 0.025 = 1 \Longrightarrow 2a = 0.45 \Longrightarrow a = 0.225.$

b) Complete the column of cumulative probability by rewriting the table in your answer sheet.

x	$p(x)$	$P(X \leq x)$
0	0.45	0.45
1	0.375	0.825
2	0.15	0.975
3	0.025	1

c) What is the probability of observing at least 2 patients during that 1-hour period?

- $P(\text{at least 2 patients during that 1} - \text{hour period}) = p(2) + p(3) = 0.15 + 0.025 = 0.175.$

d) What are the mean, the variance and the standard deviation of X?

- $E(X) = \sum x\, p(x) = 0 \times 0.45 + 1 \times 0.375 + 2 \times 0.15 + 3 \times 0.025 = 0.75.$
- $E(X^2) = \sum x^2\, p(x) = 0^2 \times 0.45 + 1^2 \times 0.375 + 2^2 \times 0.15 + 3^2 \times 0.025 = 1.2.$
- $\sigma^2 = E(X^2) - \mu^2 = 1.2 - 0.75^2 = 0.6375.$
- $\sigma = \sqrt{\sigma^2} = \sqrt{0.6375} = 0.7984.$

[Question 4]

The following table lists the probability distribution of the number of refrigerators (denoted by X) owned by all families in a city.

x	$p(x)$
0	0.01
1	0.69
2	0.22
3	?

a) Find $p(3)$.

- $p(3) = 1 - (0.01 + 0.69 + 0.22) = 1 - 0.92 \Longrightarrow 0.08.$

b) What is the probability that a randomly selected family owns at most two refrigerators?

- $P(X \leq 2) = 1 - p(3) = 1 - 0.08 = 0.92.$

c) What are $E(X)$ (also denoted by μ) and $V(X)$ (also denoted by $Var(X)$ or σ^2)?

- $E(X) = 0 \times 0.01 + 1 \times 0.69 + 2 \times 0.22 + 3 \times 0.08 = 0 + 0.69 + 0.44 + 0.24 = 1.37.$
- $Var(X) = 0^2 \times 0.01 + 1^2 \times 0.69 + 2^2 \times 0.22 + 3^2 \times 0.08 - 1.37^2 = 0 \times 0.01 + 1 \times 0.69 + 4 \times 0.22 + 9 \times 0.08 - 1.8769 = 2.29 - 1.8769 = 0.4131.$

[Question 5]

According to a survey, 70% of households said that they have never purchased organic fruits or vegetables. Suppose that this result is true for the current population of households. Suppose a random sample of 10 households is selected.

- First note that this is a binomial distribution because the following four properties are true.

i. X counts the number of households will say that they have never purchased organic fruits or vegetables (success) among 10 households.
ii. Trials are *Bernoulli*. Two possible outcomes for each trial: the selected household has never purchased organic fruits or vegetables (success) or has purchased (failure).
iii. Trials are independent. The sample was randomly drawn from a large population and the results are independent of each other.
iv. The probability p of success (never purchased organic fruits or vegetables) is the same for all households and it is $p = 0.70$.

- Therefore,
  $X \sim Binomial(10,0.70)$ and $p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{10}{x} \times 0.7^x \times 0.3^{10-x}$.

a) Find the probability that exactly 7 households will say that they have never purchased organic fruits or vegetables.
- $p(7) = \binom{10}{7} \times 0.7^7 \times 0.3^{10-7} = 120 \times 0.0823543 \times 0.027 = 0.2668$.

b) Find the probability that none of the selected households will say that they have never purchased organic fruits or vegetables.
- $p(0) = \binom{10}{0} \times 0.7^0 \times 0.3^{10} = 0.0000059 \approx 0$.

c) Find the probability that at most 2 households will say that they have never purchased organic fruits or vegetables.
- $P(X \le 2) = p(0) + p(1) + p(2) = \binom{10}{0} \times 0.7^0 \times 0.3^{10} + \binom{10}{1} \times 0.7^1 \times 0.3^9 + \binom{10}{2} \times 0.7^2 \times 0.3^8$=$0.0000059 + 0.000137781 + 0.0014467 = 0.0016$.

d) Find the average number of households who will say that they have never purchased organic fruits or vegetables.
- $\mu = E(X) = np = 10 \times 0.7 = 7$.

e) Find the variance of this distribution.
- $\sigma^2 = np(1-p) = 10 \times 0.7 \times 0.3 = 2.1$.

f) Find the average number of households who will say that they have purchased organic fruits or vegetables.
- Y counts the number of households will say that they have purchased organic fruits or vegetables (success) among 10 households. Therefore, $Y \sim Binomial(10,0.3)$.
- $\mu = E(Y) = np = 10 \times 0.3 = 3$.

[Question 6]

Based on a study on a population of large size, about 2% of adults have at some point in their life been told that they have hypertension. A random sample of size 20 is selected from the population. Let X be the random variable that denotes the number of adults having hypertension.

- First note that this is a binomial distribution because the following four properties are true.

v. X counts the number of adults having hypertension (success) among $n = 20$ selected adults.

vi. Trials are *Bernoulli*. Two possible outcomes for each trial: the selected individual has hypertension (success) or does not have (failure).

vii. Trials are independent. The sample was randomly drawn from a large population and the results are independent of each other.

viii. The probability p of success (having hypertension) is the same for all individuals and it is $p = 0.02$.

- Therefore,

$$X \sim Binomial(20, 0.02) \quad \text{and} \quad p(x) = \binom{n}{x} p^x (1-p)^{n-x} = \binom{20}{x} \times 0.02^x \times 0.98^{n-x}.$$

a) On average, how many adults having hypertension are found in a sample of 20 adults?

- $\mu = E(X) = np = 20 \times 0.02 = 0.4.$

b) Determine the variance and the standard deviation of X.

- $\sigma^2 = np(1-p) = 20 \times 0.02 \times 0.98 = 0.392.$
- $\sigma = \sqrt{\sigma^2} = \sqrt{0.392} = 0.6261.$

c) Find the probability that a sample of 20 adults has at least 2 adults having hypertension.

- $P(X \geq 2) = 1 - P(X < 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1)$
  $= 1 - {}_{20}C_0 \times 0.02^0 \times 0.98^{20-0} - {}_{20}C_1 \times 0.02^1 \times 0.98^{20-1}$
  $= 1 - \binom{20}{0} \times 0.98^{20} - \binom{20}{1} \times 0.02 \times 0.98^{19} = 1 - 0.98^{20} - 20 \times 0.02 \times 0.98^{19} = 0.0599.$

d) Assume that a new sample of size 10 is drawn from this population. What are the mean, the variance and the standard deviation of the number of adults who do NOT have hypertension in the sample of size 10.

- Y counts the number of adults who do not have hypertension (success) among $n = 10$ selected adults.
- The probability of success (not having hypertension) is 0.98.
- $Y \sim Binomial(10, 0.98).$
- $\mu = E(Y) = np = 10 \times 0.98 = 9.8.$
- $\sigma^2 = np(1-p) = 10 \times 0.98 \times 0.02 = 0.196.$
- $\sigma = \sqrt{\sigma^2} = \sqrt{0.196} = 0.443.$

[Question 7]

The Poisson distribution is used to model the number of patients referred to an oncologist per day. The researchers use a rate of 0.7 patients that are referred to the oncologist per day.

a) Find the mean, the variance and the standard deviation of the number of patients that are referred to the oncologist per day.

- $\mu = 0.7.$
- $\sigma^2 = 0.7.$
- $\sigma = \sqrt{\sigma^2} = \sqrt{0.7} = 0.8367.$

b) Find the probability that in a randomly selected day, the number of patients that are referred to the oncologist will be more than 1 and less than 4.

- X: number of patients that are referred to the oncologist per day. Therefore,

$$X \sim Poisson(0.7) \Longrightarrow p(x) = \frac{e^{-0.7} \times 0.7^x}{x!}.$$

- $P(1 < X < 4) = P(2 \le X \le 3) = P(X = 2) + P(X = 3) = \frac{e^{-0.7} \times 0.7^2}{2!} + \frac{e^{-0.7} \times 0.7^3}{3!} = 0.1550.$

c) Find the probability that in a randomly selected week, the number of patients that are referred to the oncologist will be at least 2.

- Y: number of patients that are referred to the oncologist per week. Therefore,

$$Y \sim Poisson(7 \times 0.7) \Longrightarrow Y \sim Poisson(4.9) \Longrightarrow p(x) = \frac{e^{-4.9} \times 4.9^x}{y!}, \qquad y = 0,1,2,\cdots.$$

- $P(Y \ge 2) = 1 - P(Y < 2) = 1 - P(Y \le 1) = 1 - P(Y = 0) - P(Y = 1)$
  $= 1 - \frac{e^{-4.9} \times 4.9^0}{0!} - \frac{e^{-4.9} \times 4.9^1}{1!} = 0.9561.$

[Question 8]

A university police department receives an average of 6 reports per week of lost student ID cards.

a) Find the probability that at most one such report will be received during a given week by this police department.

- Let X count the number of lost student cards that the university police department receives per week. This situation can be modeled by Poisson distribution as
  - X is a discrete random variable;
  - The occurrences are random;
  - The occurrences are independent.
- Therefore, we have that $X \sim Poisson(4.6)$ and $p(x) = \frac{e^{-4.6} \times 4.6^x}{x!}$, $\; x = 0,1,2,\cdots$
- $P(X \le 1) = P(0) + p(1) = \frac{e^{-6} \times 6^0}{0!} + \frac{e^{-6} \times 6^1}{1!} = e^{-6} \times 7 = 0.0173.$

b) Find the probability that during a given week the number of such reports received by this police department is between 2 and 5, exclusively.

- $P(2 < X < 5) = P(3) + p(4) = \frac{e^{-6} \times 6^3}{3!} + \frac{e^{-6} \times 6^4}{4!} = 0.2231.$

c) Find the probability that during a given period of four days, none of such reports will be received by this police department.

- Y: The number of reports received by this police department during a given period of four days

$$Y \sim Poisson(6 \times 4/7) \Longrightarrow Y \sim Poisson(3.43), \; y = 0,1,2,\cdots.$$

- $P(0) = \frac{e^{-3.43} \times 3.43^0}{0!} = e^{-3.43} = 0.0324.$

d) Find the average number of such reports received during a period of two weeks. Determine also the standard deviation of this distribution.

- V: The number of reports received by this police department during a given period of two weeks
$$V \sim Poisson(6 \times 2).$$
- $E(V) = \mu = 12.$
- $\sigma^2 = 12.$
- $\sigma = \sqrt{12} = 3.464102.$

[Question 9]

The mean number of accidents to occur at a busy intersection during a 24-hour period is 2.10. Assume that the number of accidents to occur at this intersection during a 24-hour period follows a Poisson distribution.

- $X \sim Poisson(2.1) \Longrightarrow p(x) = e^{-2.1} \times \frac{2.1^x}{x!}.$

a) Find the mean and the standard deviation of the number of accidents to occur at this intersection during a 24-hour period.
- $\mu = \lambda = 2.1.$
- $\sigma = \sqrt{\lambda} = \sqrt{2.1} = 1.44914.$

b) Find the probability that the number of accidents to occur at this intersection during a 24-hour period will be between 4 and 7, exclusive.
- $P(4 < X < 7) = p(5) + p(6) = e^{-2.1} \times \frac{2.1^5}{5!} + e^{-2.1} \times \frac{2.1^6}{6!} = 0.04167 + 0.01458 = 0.05626.$

c) Find the probability that no accidents will occur at this intersection during a 36-hour period.
- $Y \sim Poisson(1.5 \times 2.1) \Longrightarrow Y \sim Poisson(3.15) \Longrightarrow p(y) = e^{-3.15} \times \frac{3.15^y}{y!}, \;\; y = 0,1,2,\cdots.$
- $p(0) = e^{-3.15} \times \frac{3.15^0}{0!} = 0.04285.$