Name (Last, First): Kunjangada Arun Sushma

Student ID: 002473132

Assignment 7

NEU_COE_INFO6105_Fall2024

Instructions:

1. For answering **programming questions**, please use Adobe Acrobat to edit the pdf file in two steps **[See Appendix: Example Question and Answer]**:
   a. Copy and paste your R or python code as text in the box provided (so that your teaching team can run your code);
   b. Screenshot your R or python console outputs, save them as a .PNG image file, and paste/insert them in the box provided.
   c. Show all work - credit will not be given for code without showing the code in action by including the screenshot of R or python console outputs.
2. To answer **non-programming questions**, please type or handwrite your final answers clearly in the boxes. Show all work - credit will not be given for numerical solutions that appear without explanation in the space above the boxes. **You're encouraged to use R or python to graph/plot the data and produce numerical summaries; please append your code and screenshot of the outputs at the end of your pdf submission.**
3. **[Total 111 pts = 21 + 15 + 30 + 42 pts + 3 Extra Credit pts]**

Grading Rubric

Each question is worth 3 points and will be graded as follows:

3 points: Correct answer with work shown

2 points: Incorrect answer but attempt shows some understanding (work shown)

1 point: Incorrect answer but an attempt was made (work shown), or **correct answer without explanation (work not shown)**

0 points: Left blank or made little to no effort/work not shown

Reflective Journal [3 pts]

(Copy and paste the link to your live Google doc in the box below)

https://docs.google.com/document/d/1ptEhnYHniNtT1yxDPcvXK7LpJaPGzi80BCSGZZhom7Y/edit?usp=sharing

Part I. Discrete and Continuous Random Variables (7 x 3 = 21 pts)

1) **(18 pts)** A random sample of adults was taken and asked the question "Last week, how many days did you sit down and eat at a restaurant?" X represents the number of days they responded with. Here is the resulting probability model:

Days (X)	0	1	2	3	4	5	6	7
Probability	0.40	0.28	0.16	0.05	0.04	0.03	0.02	0.02

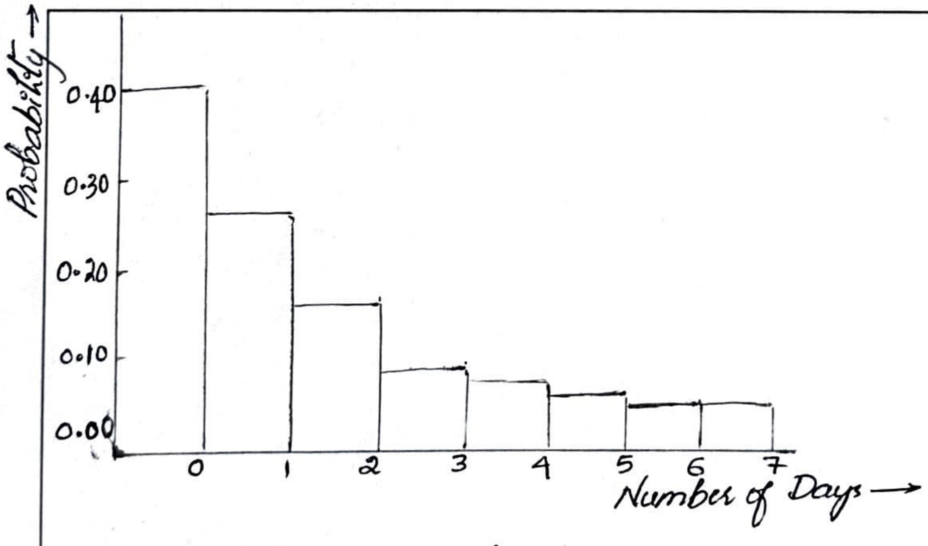
(a) Show that the probabilities in this distribution add up to 1.

Answer:

To show that the probabilities in This distribution adds up to 1, we'll add the given probabilities.

$0.40 + 0.28 + 0.16 + 0.05 + 0.04 + 0.03 + 0.02 + 0.02$
$= \underline{1}$

(b) Make a histogram of the probability distribution. Describe the shape of the distribution.



• The distribution is right skewed.
• It starts with the highest probability at 0 days and decreases as the number of days increases, with a longer tail extending to the right.

(c) Find and interpret P(X < 4).

$P(x<4) = P(x=0) + P(x=1) + P(x=2) + P(x=3)$

$= 0.40 + 0.28 + 0.16 + 0.05$

$= \underline{\underline{0.89}}$

There is an 89% chance that a randomly selected person ate out fewer than 4 times in the past week.

(d) Express the event "Ate out at least twice last week" in terms of X. What is the probability of this event?

Ate out atleast twice last week $= X \geq 2$

$P(X \geq 2) = P(X=2) + P(X=3) + P(X=4) + P(X=5) + P(X=6) + P(X=7)$

$= 0.16 + 0.05 + 0.04 + 0.03 + 0.02 + 0.02 = 0.32$

There is a 32% chance that a randomly selected person ate out atleast twice last week.

(e) Find and interpret the mean.

$\mu_x = \Sigma x_i \cdot p_i$

$\mu_x = (0 \times 0.40) + (1 \times 0.28) + (2 \times 0.16) + (3 \times 0.05) + (4 \times 0.04) + (5 \times 0.03) + (6 \times 0.02) + (7 \times 0.02)$

$= 0 + 0.28 + 0.32 + 0.15 + 0.16 + 0.15 + 0.12 + 0.14 = \underline{\underline{1.32 \text{ days}}}$

On average, a person ate out about 1.32 days in the past week.

(f) Find the standard deviation.

$\sigma_x = \sqrt{\Sigma (x_i - \mu_x)^2 \cdot p_i}$

$= \sqrt{(0-1.32)^2 \times 0.40 + (1-1.32)^2 \times 0.28 + (2-1.32)^2 \times 0.16 + (3-1.32)^2 \times 0.05 + (4-1.32)^2 \times 0.04 + (5-1.32)^2 \times 0.03 + (6-1.32)^2 \times 0.02 + (7-1.32)^2 \times 0.02}$

$= \sqrt{2.743868} \simeq \underline{1.65 \text{ days}}$

2) **(3 pts)** A new type of gambling game uses two dice and the sum of the two dice determines your winnings. It costs $5 to play the game. If the sum is a 2 or a 12, the player will win $50. If the sum is a 7, the player will win $10. Any other sum and the player will lose and not get anything. If X represents the winnings (with -5 representing a loss, since a player will lose that amount of money), create a table to represent the probability distribution. Using the mean of this distribution, do you believe the game is in favor of the player or the "house"?

Answer:

Probability of Sum = 2 or 12: There is only 1 way to get 2 (1+1) & 1 way to get 12 (6+6)
Probability of Sum = 7: There are 6 combinations (1+6, 2+5, 3+4, 4+3, 5+2, 6+1) out of 36 possible combinations.

Winnings	$50	$10	-$5
Probability	2/36	6/36	28/36

$\mu_x = \Sigma x_i p_i$

$= 50 \times \frac{2}{36} + 10 \times \frac{6}{36} + (-5) \times \frac{28}{36}$

$= \frac{100}{36} + \frac{60}{36} - \frac{140}{36} = \frac{20}{36} \simeq \underline{\underline{0.56}}$

- Since the expected value is positive ($0.56), the game is in favour of the player.
- On average, a player can except to win about 56 cents per game.

Part II. Combining Random Variables (5 x 3 = 15 pts)

1) **(12 pts)** In the 2021 baseball season, the Brewers runs per game, B, varied from game to game but had $\mu_B = 3.7$ and $\sigma_B = 1.1$.The Cubs runs per game, C, were $\mu_C = 4.9$ and $\sigma_C = 1.6$.

Find the following (show your work): $\mu_B = 3.7$, $\sigma_B = 1.1$, $\mu_c = 4.9$, $\sigma_c = 1.6$

μ_{B+C}	$\mu_{B+C} = \mu_B + \mu_c$ $= 3.7 + 4.9$ $= 8.6$ runs per game	μ_{B-C}	$\mu_{B-C} = \mu_B - \mu_c$ $= 3.7 - 4.9$ $= -1.2$ runs per game
σ_{B+C}	$\sigma_{B+C} = \sqrt{1.1^2 + 1.6^2}$ $= \sqrt{1.21 + 2.56} = \sqrt{3.77}$ $\simeq 1.94$ runs per game.	σ_{B-C}	$\sigma_{B-C} = \sqrt{1.1^2 + 1.6^2}$ $= \sqrt{1.21 + 2.56} = \sqrt{3.77}$ $\simeq 1.94$ runs per game

2) **(3 pts)** At our town's ice cream shop, if you order a cone, you can determine the number of scoops you would like. Let X represent the number of scoops ordered and the probability that a random customer would ask for that number of scoops.

X	1	2	3	4	5
P(X)	0.20	0.38	0.28	0.11	0.03

It costs \$2.50 for the cone and \$0.75 for each scoop of ice cream. Find the amount of money the store expects to make on a randomly selected customer.

Answer:

Cost of cone = \$2.50 Cost per scoop = \$0.75

Revenue for each number of scoops:

For $X=1$: Revenue $= 2.50 + 0.75 \times 1 = 2.50 + 0.75 = 3.25$

For $X=2$: Revenue $= 2.50 + 0.75 \times 2 = 2.50 + 1.50 = 4.00$

For $X=3$: Revenue $= 2.50 + 0.75 \times 3 = 2.50 + 2.25 = 4.75$

For $X=4$: Revenue $= 2.50 + 0.75 \times 4 = 2.50 + 3.00 = 5.50$

For $X=5$: Revenue $= 2.50 + 0.75 \times 5 = 2.50 + 3.75 = 6.25$

Expected revenue $= (3.25 \times 0.20) + (4.00 \times 0.38) + (4.75 \times 0.28) + (5.50 \times 0.11) + (6.25 \times 0.03)$

$= 0.65 + 1.52 + 1.33 + 0.605 + 0.1875$

$= 4.2925$

$\simeq 4.29$

The expected amount of money the store makes on a randomly selected customer is approximately \$4.29

Part III. The Binomial Distribution (10 x 3 = 30 pts)

Directions: For the situations below, define a random variable X for the situation and then decide if they follow a binomial distribution model by commenting on the four requirements. **(12 pts)**

1) You roll a DnD dice (20-sided), 20 times, and record the number that shows on the dice.

- Fixed Number of Trails: Yes, $n=20$
- Two Possible Outcome: No, The outcome is a number between 1 and 20, not just 2 possible outcomes.
- Constant Probability of Success: Not applicable due to more than two outcomes.
- Independence of Trails: Yes, each roll is independent.

Conclusion: This does not follow a binomeal distribution.

2) A basketball player can make 60% of their free throws. The coach plans on having a free throw shooting competition and the player will be shooting 100 shots.

- Fixed Number of Trails: Yes, $n=100$
- Two Possible Outcomes: Yes, each shot can either be a "make" (success) or "miss" (failure).
- Constant Probability of Success: Yes, the probability $p=0.60$ is the same for each shot.
- Independence of Trails: Yes, each sot is independent.

Conclusion: This follows a binomial distribution.

3) From a standard deck of cards, you pull out a card, record the suit, put it back, and reshuffle. You continue until you get two spades in a row.

- Fixed Number of Trails: No, because the number of draws is not predetermined, we stop once we acheive two spades in a row.
- Two Possible Outcomes: Yes, each draw can result in a "spade" or "not a spade".
- Constant Probability of Success: No, because the probability of drawing a spade changes as cards are drawn.
- Independence of Trails: No, the draws are dependent because cards are not replaced.

Conclusion: This doesn't follow a binomial distribution.

4) You are conducting a survey at your school to see how many students own smartphones. The probability that a student will own a smartphone is 0.85. You plan on having 200 students participate in your survey.

- Fixed Number of Trails: Yes, $n=200$
- Two possible Outcomes: Yes, each student either owns a smartphone (success) or does not (failure).
- Constant Probability of Success: Yes, The probability $p=0.85$ is assumed constant.
- Independence of Trails: Yes, each student's ownership is independent of the others.

Conclusion: This does follow binomial distribution.

5) **(18 pts)** Zach is an 85% field goal kicker. Each goal he kicks is independent of the next kick. At practice today, he will be doing 30 kicks. $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$ $p=0.85$ $(1-p)=0.15$

For the following table, fill in the missing pieces.

Situation	***Probability Notation***	***Formula***	***Value***
What is the probability Zach will make exactly 25 kicks?	$P(X=25)$ dbinom(25, 30, 0.85)	$P(X=25)=\binom{30}{25}(0.85)^{25}(0.15)^{30-25}$ $=\frac{30!}{5!(30-5)!}\times(0.85)^{25}\cdot(0.15)^{30-25}$ $=142506(0.85)^{25}(0.15)^{30-25}$	$\simeq 0.1861$
What is the probability Zach will make exactly 20 kicks?	$P(X = 20)$ dbinom(20,30,0.85)	$P(X=20)=\binom{30}{20}(0.85)^{20}(0.15)^{30-20}$ $=\frac{30!}{10!(30-10)!}\times(0.85)^{20}(0.15)^{10}$ $=30045015(0.85)^{20}(0.15)^{10}$	$\simeq 0.0067$
What is the probability Zach will make no more than 20 kicks?	$P(X\leq 20)$ pbinom(20, 30, 0.85)	$P(X\leq 20)=\sum_{k=0}^{20}P(X=k)$ $=P(X=0)+P(X=1)+P(X=2)$ $+\dots+P(X=19)+P(X=20)$	$\simeq 0.0097$
What is the probability Zach will make no more than 24 kicks?	$P(X \leq 24)$ pbinom(24, 30, 0.85)	$P(X\leq 24)=\sum_{k=0}^{24}P(X=k)$ $=P(X=0)+P(X=1)+P(X=2)$ $+\dots+P(X=23)+P(X=24)$	$\simeq 0.2894$
What is the probability Zach will make exactly 28 kicks?	$P(X=28)$ dbinom(28, 30, 0.85)	(30 28)$0.85^{28}0.15^{2}$	$\simeq 0.1034$
What is the probability Zach will make 27 or more kicks?	$P(X\geq 27)$ 1-pbinom(26, 30, 0.85)	$P(X = 27) + P(X = 28) + P(X = 29) + P(X = 30)$	$\simeq 0.3217$

Part IV. The Geometric Distribution (14 x 3 = 42 pts)

1) **(18 pts)** For the following situations, decide if it is a binomial setting, a geometric setting, or neither. Explain your answer.

a) You keep drawing cards out a deck, without replacement, until an ace is drawn.

• Neither binomial or geometric.

• This situation does not fit either distribution because the trails are not independent (drawing without replacement affects the probabilities).

b) You roll a dice 20 times and record the number of sixes you have rolled.

• Binomial Distribution.

• There is a fixed number of trails (20 rolls), two possible outcomes (rolling a six or not), a constant probability of success (1/6 for each roll), and the rolls are independent.

c) Crest claims that 40% of Americans use their toothpaste. You take a random sample of 50 Americans and count how many use Crest toothpaste.

• Binomial Distribution.

• There is a fixed number of trails (50 Americans), two outcomes (use Crest or do not use Crest), a constant probability of success (0.40), and the selection of individuals is independent.

d) You flip a coin until you get tails.

• Geometric Distribution.

• This is a geometric setting because you continue flipping until the first success (getting tails). Each flip is independent with the same probability of success (0.5 for tails).

e) 5% of the tomatoes at a farmer's market have imperfections on them. You randomly choose one tomato at a time until you find one with an imperfection.

• Geometric Distribution.

• This represents a geometric distribution because you continue choosing tomatoes until the first success (finding an imperfect tomato). Each choice is independent, with a constant probability of finding an imperfect (0.05).

f) 5% of the tomatoes at the farmer's market have imperfections on them. You randomly choose 20 tomatoes and count the number of imperfections on them.

• Binomial Distribution.

• This is a binomial setting because there is a fixed number of trials (20 tomatoes), two outcomes (imperfect or not imperfect), a constant probability of success (0.05), and the choice are independent.

• Binomial Distribution: Involves a fixed number of trails, each with two possible outcomes (success or failure), a constant probability of success for each trail, and trails are independent.

• Geometric Distribution: Involves a process where trails continue until first success is achieved, with each trail being independent and having the same probability of success.

2) **(12 pts)** In a recent Pew study, it was found that 28% of all homes have at least one dog. Let X = the number of houses you look at before finding a house with a dog. Find and interpret the following probabilities.

a) $P(X = 3)$ There is a 14.5% chance of finding the first house with a dog on the third house checked.

$P(X=k) = (1-p)^{k-1} \cdot p$

$P(X=3) = (1-0.28)^{3-1} \times (0.28)$

$= (0.72)^2 \times 0.28$

≈ 0.1452

b) $P(X < 5)$ There is a 73.13% chance of finding a house with a dog within first 4 houses.

$P(X<5) = P(X=1)+P(X=2)+P(X=3)+P(X=4)$

$P(X=1) = (0.72)^0 \times (0.28) = 0.28$

$P(X=2) = (0.72)^1 \times (0.28) = 0.2016$

$P(X=3) = (0.72)^2 \times 0.28 = 0.145152$

$P(X=4) = (0.72)^3 \times 0.28 = 0.10450944$

$P(X<5) \approx 0.7313$

c) $P(X > 4)$

$P(X>4) = 1 - P(X<5)$

$= 1 - 0.7313$

≈ 0.2687

There is a 26.87% chance it will take more than 4 houses to find ~~to~~ one with a dog.

d) $E(X)$

$E(X) = \frac{1}{p}$

$= \frac{1}{0.28} \approx 3.57$

On average, you would need to look at about 3.57 houses before finding one with a dog.

3) **(12 pts)** Zach's teammate, Taylor, is the quarterback for the football team. She can complete 65% of her passes.

For the following table, fill in the missing pieces.

Situation	***Probability Notation***	**Formula**	***Value***
What is the probability that it takes Taylor 3 incomplete passes before she has a completion?	$P(X=4)$	$P(X=k)=(1-p)^{k-1}p$ $P(X=4) = (1-0.65)^{4-1} \times (0.65)$ $= (0.35)^3 \times 0.65$	≈ 0.0279
What is the probability it takes Taylor fewer than 3 incomplete passes befor a completion?	$P(X < 3)$	$P(X=3) = P(X=1)+P(X=2)$ $P(X=1) = (0.35)^0 \times 0.65 = 0.65$ $P(X=2) = (0.35)^1 \times 0.65 = 0.2275$ $P(X=3) = 0.65 + 0.2275$	$= 0.8775$
What is the propability that it takes Taylor 4 incomplete passes before she has a completion?	$P(X=5)$	$(0.35)^4(0.65)$	≈ 0.0097
What is the probability Taylor fails to complete a pass in 5 attempts?	$P(X>5)$ or $1-P(X\le 5)$	$1 - (P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5))$	≈ 0.0052

THE END