<u>Introduction to Probability</u>
Batch 23 (Class of 2025 Winter) Foundation Term
<u>Individual Assignment</u>
<u>Answers</u>
By
Vridhi Wadhawan (12420090)

Q2.

Given:

X is the outcome of the first die & Y is the outcome of the second die.
Then, X+Y is the sum of the 2 dice
$S = \{(X,Y) : X,Y \in \{1,2,3,4,5,6\}\}$

Answer:

(a) $\{X+Y=4\} = \{(1,3),(2,2),(3,1)\}$
(b) $\{X+Y=9\} = \{(3,6),(4,5),(5,4),(6,3)\}$
(c) $\{Y=3\} = \{(1,3),(2,3),(3,3),(4,3),(5,3),(6,3)\}$
(d) $\{X=Y\} = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$
(e) $\{X>2Y\} = \{(6,1),(6,2),(5,1),(5,2),(4,1),(3,1)\}$

Q3.

Given: Sample Space = S = {(1,2,3,4)}

Known:

For the probability of any outcome on the given sample space to be true it must follow the following 2 conditions to fulfill:

1) The probability of every outcome has to be greater than 0
   P(outcome)>0
2) The sum of the probabilities of all the outcomes should be equal to 1

Answer:

(a) P(1) = 0.6, P(2) = 0.05, P(3) = 0.4, P(4) = 0.2
All values are greater than 0.
Sum of all probabilities = P(1) + P(2) + P(3) + P(4) = 0.6 + 0.05 + 0.4 + 0.2 = 1.25 > 1
Conclusion : Since the sum of all probabilities is greater than 1, it is not a valid probability function.

(b) P(1) = 0.5, P(2) = 0.2, P(3) = 0.2, P(4) = 0.1.
All values are greater than 0.
Sum of all probabilities = P(1) + P(2) + P(3) + P(4) = 0.5 + 0.2 + 0.2 + 0.1 = 1
Conclusion : Since all the probabilities are greater than 0 and the sum of all probabilities is equal to 1, it is a valid probability function.

(c) P(1) = 0.15, P(2) = 0.3, P(3) = 0.1, P(4) = 0.45.
All values are greater than 0.
Sum of all probabilities = P(1) + P(2) + P(3) + P(4) = 0.15 + 0.3 + 0.1 + 0.45 = 1
Conclusion : Since all the probabilities are greater than 0 and the sum of all probabilities is equal to 1, it is a valid probability function.

(d) P(1) = 0.3, P(2) = 0.3, P(3)=−0.2, P(4) = 0.6.
Since P(3) = -0.2, which is less than 0 violates the first condition of probability.
Sum of all probabilities = P(1) + P(2) + P(3) + P(4) = 0.3 + 0.3 - 0.2 + 0.6 = 1
Conclusion : Even though the sum of all the probabilities is 1 but since P(3) = -0.2 violates the first condition of probability being that the probability of all outcomes must be more than 0. Thus, it is not a valid probability function.

Q4.

Given:

Total number of students in the program = 20
Let x be the number of students possible in a sub-group = 3, 4, 5, 6
x = 3, 4, 5, 6

To Find:

a. How many subgroups of students are possible ?
b. Find the probability that three most senior students attend the conference.

Answer:

a. The number of ways to choose sub-groups is given by $^{n}C_{r}$
Total number of possible sub groups = $^{20}C_3 + {}^{20}C_4 + {}^{20}C_5 + {}^{20}C_6$
= 1140 + 4845 + 15504 + 38760
= 60249

Thus, there are 60249 possible subgroups of students that might attend the conference.

```
Console  Terminal  Jobs
~/
> #Q4. Part (I)
> #The number of ways to choose sub-groups is given by nCr
> #Total number of possible sub groups =  20C3 + 20C4 + 20C5 + 20C6
> a <-choose(20,3)
> a
[1] 1140
> b <-choose(20,4)
> b
[1] 4845
> c <-choose(20,5)
> c
[1] 15504
> d <-choose(20,6)
> d
[1] 38760
> result <- a+b+c+d
> result
[1] 60249
> print(paste("Total number of possible sub groups are ", result))
[1] "Total number of possible sub groups are  60249"
```

b. Funding is fixed for 6 students and selection is to be made from 20 students
Total ways to choose 6 students students from 20 students = $^{n}C_{r}$ = $^{20}C_6$ =38760

For all 3 senior students to be chosen for the 6 spots they must be included in the 6 selected students out of 20 which means only 3 spots (6 spots - 3 senior students = 3 remaining spots) remain for the rest of the 17 students (20 students - 3 senior students already selected =17 students remaining).

Remaining 3 spots are to be filled by the 17 remaining students= $^{n}C_{r}$ = $^{17}C_{3}$ = 680
Probability that all 3 senior students are chosen =

$$P\text{ (3 seniors)} = {}^{17}C_3 \,/\, {}^{20}C_6 = 680 / 38760 = 1 / 57 = 0.01754386$$

Thus, the probability that the 3 most senior students attend when 6 students are chosen at random is 1 / 57 = 0.01754386

```
Console   Terminal ×   Jobs ×
~/
> #Q4. Part (II)
> #Total ways to choose 6 students students from 20 students = nCr = 20C6
> d <-choose(20,6)
> d
[1] 38760
> #Remaining 3 spots are to be filled by the  17 remaining students= nCr = 17C3
> e <-choose(17,3)
> e
[1] 680
> #Probability that all 3 senior students are chosen = P (3 seniors) = 17C3 / 20C6
> f <- e/d
> f
[1] 0.01754386
> print(paste("Probability that all 3 senior students are chosen ", f))
[1] "Probability that all 3 senior students are chosen  0.0175438596491228"
```

Q5.

A) Given: A & B are mutually exclusive

$P(A) = 0.30$ & $P(B) = 0.60$

Then, $P(A \cap B) = 0$ & $P(A \cup B) = P(A) + P(B) = 0.30 + 0.60 = 0.90$

(a) At least one of the two events occurs

P(At least one of the two events occurs) = P(A) + P(B)
= 0.3 + 0.6
= 0.90

(b) Both of the events occur

$P(A \cap B) = 0$

(c) Neither event occurs

P(Neither of the events occurs) = 1 - P(A) - P(B) - $P(A \cap B)$
= 1 - 0.3 - 0.6 -0
= 1 - 0.9
= 0.10

(d) Exactly one of the two events occurs

P(Exactly one of the two events occurs) = P(A) + P(B) - $P(A \cap B)$
= 0.3 + 0.6 - 0
= 0.90

B)

See the assignment of probabilities to the Venn diagram in the figure.

(a) P(No events occur) = h

(b) P(Exactly one event occurs) = a + c + f

(c) P(Exactly two events occur) = b + d + e

(d) P(Exactly three events occur) = g
(e) P(At least one event occurs) = a + b + c + d + e + f + g
(f) P(At least two events occur) = b + d + e + g
(g) P(At most one event occurs) = a + c + f + h
(h) P(At most two events occur) = a + b + c + d + e + f + h

Q6.

Given:

$X \in \{0, 0.01, 0.02, ..., 0.99, 1\}$
All outcomes are equally likely
Total number of outcomes for X = 101 (i.e, Including 0)

(a) $P(X \leq 0.33)$

Values of $X \leq 0.33 = \{0, 0.01, 0.02, ..., 0.33\}$
Total Number of values of X in the above set = 33 +1 = 34 (i.e, Includes 0)
Thus, $P(X \leq 0.33) = 34/101 = 0.3366337$

(b) $P(0.55 \leq X \leq 0.66)$

Values of $0.55 \leq X \leq 0.66 = \{0.55, 0.56, 0.57, ..., 0.66\}$
Total Number of values of X in the above set = 11+1 = 12
(i.e, Includes the starting and the ending values)
Thus, $P(0.55 \leq X \leq 0.66) = 12/101 = 0.1188119$

```
Console  Terminal ×  Jobs ×
~/
> #Q6
> # Create the sequence of outcomes
> x <- seq(0, 1, by = 0.01)
> x
  [1] 0.00 0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08 0.09 0.10 0.11 0.12 0.13 0.14 0.15 0.16 0.17 0.18
 [20] 0.19 0.20 0.21 0.22 0.23 0.24 0.25 0.26 0.27 0.28 0.29 0.30 0.31 0.32 0.33 0.34 0.35 0.36 0.37
 [39] 0.38 0.39 0.40 0.41 0.42 0.43 0.44 0.45 0.46 0.47 0.48 0.49 0.50 0.51 0.52 0.53 0.54 0.55 0.56
 [58] 0.57 0.58 0.59 0.60 0.61 0.62 0.63 0.64 0.65 0.66 0.67 0.68 0.69 0.70 0.71 0.72 0.73 0.74 0.75
 [77] 0.76 0.77 0.78 0.79 0.80 0.81 0.82 0.83 0.84 0.85 0.86 0.87 0.88 0.89 0.90 0.91 0.92 0.93 0.94
 [96] 0.95 0.96 0.97 0.98 0.99 1.00
> # Total number of outcomes
> n <- length(x)
> n
[1] 101
> # (a) P(X <= 0.33)
> prob_a <- sum(x <= 0.33) / n
> prob_a
[1] 0.3366337
> print(paste("P(X <= 0.33) =", prob_a))
[1] "P(X <= 0.33) = 0.336633663366337"
> # (b) P(0.55 <= X <= 0.66)
> prob_b <- sum(x >= 0.55 & x <= 0.66) / n
> prob_b
[1] 0.1188119
> print(paste("P(0.55 <= X <= 0.66) =", prob_b))
[1] "P(0.55 <= X <= 0.66) = 0.118811881188119"
```

Q7.

Let X be the number of infected people in a group of 1000 (Fixed number of trials), where each person is infected Independently with probability 0.01 (Identically Distributed).

$X \sim \text{Binomial}(n = 1000, p = 0.01)$

(a) The probability that exactly 10 people are infected

$P(X = 10) = {}^{n}C_r\,(p)^r\,(1\text{-}p)^{n-r} = {}^{1000}C_{10}\,(0.01)^{10}\,(0.99)^{990} = 0.1257402$

(b) That probability that at least 16 people are infected

$P(X \geq 16) = 1 - P(X \leq 15) = 1 - P(X = 15) - P(X = 14).....\text{-}P(X = 0) = 0.04787059$

(c) The probability that between 12 and 14 people are infected

$P(12 \leq X \leq 14) = P(X = 12) + P(X = 13) + P(X = 14)$

$= {}^{1000}C_{12}\,(0.01)^{12}\,(0.99)^{988} + {}^{1000}C_{13}\,(0.01)^{13}\,(0.99)^{987} + {}^{1000}C_{14}\,(0.01)^{14}\,(0.99)^{986}$

$= 0.2202376$

(d) The probability that someone is infected

$P(X \geq 1) = 1 - P(X = 0) = 1 - (0.99)^{1000} = 0.9999568$

```
Console   Terminal   Jobs
~/
> #Q7
> # Given
> m <- 1000
> p <- 0.01
> # (a) Probability exactly 10 people are infected
> prob_exactly_10 <- dbinom(10, m, p)
> prob_exactly_10
[1] 0.1257402
> print(paste("Probability exactly 10 are infected:", prob_exactly_10))
[1] "Probability exactly 10 are infected: 0.125740211126207"
> # (b) Probability at least 16 people are infected
> prob_at_least_16 <- 1 - pbinom(15, m, p)
> prob_at_least_16
[1] 0.04787059
> print(paste("Probability at least 16 are infected:", prob_at_least_16))
[1] "Probability at least 16 are infected: 0.0478705857579471"
> # (c) Probability between 12 and 14 people are infected (inclusive)
> prob_between_12_14 <- sum(dbinom(12:14, m, p))
> prob_between_12_14
[1] 0.2202376
> print(paste("Probability between 12 and 14 are infected:", prob_between_12_14))
[1] "Probability between 12 and 14 are infected: 0.220237594340806"
> # (d) Probability that at least one person is infected
> prob_someone_infected <- 1 - dbinom(0, m, p)
> prob_someone_infected
[1] 0.9999568
> print(paste("Probability that at least one person is infected:", prob_someone_infected))
[1] "Probability that at least one person is infected: 0.999956828752589"
```

Q9.

A. In each room the number of phone calls made by the guest follows a Poisson distribution with parameter $\lambda = 2$.

$X \sim \text{Poisson}(2)$

Then,

Probability that more than 6 calls are made in 1 room =

P(More than 6 calls are made in 1 room) = $P(X > 6)$

$= 1 - P(X \leq 6) = 0.004533806$

B. The 100 rooms (Fixed number of trials) in the hotel are independent

$$Y \sim \text{Binomial}(n = 100, p)$$

Where,

p = P(X > 6) is the probability that a given room has more than 6 calls.

Then,

Probability that in at least three rooms more than six calls are made =

P(At least 3 rooms had more than 6 calls are made) = $P(Y \geq 3) = 1 - P(Y \leq 2)$

$= 0.01086967$

```
Console   Terminal ×   Jobs ×
~/
> #Q9
> #(a): Probability that more than six calls are made in one room
> lambda <- 2
> prob_more_than_6 <- 1 - ppois(6, lambda)
> prob_more_than_6
[1] 0.004533806
> print(paste("Probability that more than 6 calls are made in one room:", prob_more_than_6))
[1] "Probability that more than 6 calls are made in one room: 0.00453380552624882"
> #(b): Probability that in at least 3 rooms more than 6 calls are made
> no_rooms <- 100
> # Y ~ Binomial(no_rooms, prob_more_than_6)
> prob_at_least_3 <- 1 - pbinom(2, no_rooms, prob_more_than_6)
> prob_at_least_3
[1] 0.01086967
> print(paste("Probability that in at least 3 rooms more than 6 calls are made:", prob_at_least_3))
[1] "Probability that in at least 3 rooms more than 6 calls are made: 0.0108696679123289"
```

Q8.

(a) $P(X = 3) = 1 - P(X = 0) - P(X = 1) - P(X = 2) = 1 - 0.4 - 0.3 - 0.2 = 0.1$

(b) Probability that a single home has no violations = $P(X = 0) = 0.4$

For two independent homes, probability that both have no violations =

P(Both 0) = $(0.4)^2 = 0.16$

(c) Since 2 homes are selected independently, there are two possible orders

First home has at most 1 and second at least 2, or vice versa

At most 1 violation i.e., X = 0 or X = 1 $\quad P(X \leq 1) = P(0) + P(1) = 0.4 + 0.3 = 0.7$

At least 2 violations i.e, $X \geq 2$ $\quad P(X \geq 2) = P(2) + P(3) = 0.2 + 0.1 = 0.3$

Hence, the required probability that one home has at most 1 violation and other at least 2 violations = P(One home has at most 1 violation and other at least 2 violations) =

$= 2 \times (0.7) \times (0.3)$

$= 2 \times 0.21$

$= 0.42$

(d) Let Y be the number of homes (out of 5) that have at most 1 violation.

$P(X \leq 1) = 0.7$

Thus, Y follows a binomial distribution with parameters n = 5 and p = 0.7.

Probability that fewer than 3 of the homes have at most 1 violation i.e (Y < 3)

$= P(Y < 3) = P(Y \leq 2) = P(Y = 0) + P(Y = 1) + P(Y = 2) = 0.16308$

```
Console   Terminal ×   Jobs ×
~/
> #Q8
> # Given probabilities for X (violations per home)
> prob_0 <- 0.4
> prob_1 <- 0.3
> prob_2 <- 0.2
> prob_3 <- 1 - (prob_0 + prob_1 + prob_2)
> # (a)Print the completed probability for X = 3
> prob_3
[1] 0.1
> print(paste("P(X = 3) =", prob_3))
[1] "P(X = 3) = 0.1"
> # (b) Two homes: both have no violations
> prob_both_none <- prob_0^2
> prob_both_none
[1] 0.16
> print(paste("Probability both homes have no violations =", prob_both_none))
[1] "Probability both homes have no violations = 0.16"
> # (c) Two homes: one has at most 1 violation and the other has at least 2 violations
> prob_at_most1 <- prob_0 + prob_1
> prob_at_most1
[1] 0.7
> prob_at_least2 <- prob_2 + prob_3
> prob_at_least2
[1] 0.3
> prob_mixed <- 2 * prob_at_most1 * prob_at_least2
> prob_mixed
[1] 0.42
> print(paste("Probability one home has at most 1 and the other at least 2 violations =", prob_mixed))
[1] "Probability one home has at most 1 and the other at least 2 violations = 0.42"
> # (d) Five homes: fewer than 3 have at most 1 violation
> n <- 5
> p <- prob_at_most1
> prob_fewer_than_3 <- pbinom(2, n, p)
> prob_fewer_than_3
[1] 0.16308
> print(paste("Probability that fewer than 3 of 5 homes have at most 1 violation =", prob_fewer_than_3))
[1] "Probability that fewer than 3 of 5 homes have at most 1 violation = 0.16308"
```

Q10.

Given:

Elevators weight capacity = 1000 pounds

3 Men and 3 women are travelling in the elevator

Weights are normally distributed

Mean weight of an adult male = 172 pounds

Mean weight of an adult female = 143 pounds

Standard Deviation of weight of an adult male = 29 pounds

Standard Deviation of weight of an adult female = 29 pounds

Known: (Standard Deviation)^2= Variance

Answer:

Let the weight of the 3 men in the elevator be M_1, M_2 and M_3

M_i ~ N(172,29^2), where 172 is the average weight and 29^2 is the variance

Let the weight of the 3 women in the elevator be F_1, F_2 and F_3

W_i ~ N(143,29^2), where 143 is the average weight and 29^2 is the variance

Total Weight in the elevator = W = $M_1 + M_2 + M_3 + F_1 + F_2 + F_3$

W = 172 X 3 + 143 X 3 = 516 + 429 = 945 pounds

Individual variance of weight = (Standard Deviation)^2 = (29)^2 = 841

Total variance of weight = 3 X 841 + 3 X 841 = 6 X 841 = 5046

Total standard deviation of weight = sqrt (50460) = 71.0352

Probability that total weight exceeds 1000 pounds =

P(W > 1000) = 1 - P(W ≤ 1000)

P(W ≤ 1000) = Elevator Capacity - Total Mean / Total Standard Deviation

$= 1000 - 945 / 71.03552 = 55 / 71.03552 = 0.7742605$

Thus , $P(W > 1000) = 1 - 0.7742605 = 0.2257395$

Hence, the probability that the total weight of the passengers exceeds the elevator's capacity is approximately 0.22 or 22%.

Console Terminal Jobs

```
~/
> #Q10
> # Given:
> elevator_capacity <- 1000
> #mean for men and women
> mean_men <- 172
> mean_women <- 143
> #standard deviaton for men and women
> sd_men <- 29
> sd_women <- 29
> # Number of people in the elevator
> num_men <- 3
> num_women <- 3
> # Total mean and variance and standard deviation
> total_mean <- num_men * mean_men + num_women * mean_women
> total_mean
[1] 945
> total_variance <- num_men * sd_men^2 + num_women * sd_women^2
> total_variance
[1] 5046
> total_sd <- sqrt(total_variance)
> total_sd
[1] 71.0352
> # The probability that the total weight exceeds 1000 pounds
> prob_exceed <- 1 - pnorm(elevator_capacity, mean = total_mean, sd = total_sd)
> prob_exceed
[1] 0.2193873
> print(paste("Probability that the total weight exceeds 1000 pounds:", prob_exceed))
[1] "Probability that the total weight exceeds 1000 pounds: 0.219387344258883"
```

Q1.

(a) Ravi believes there's a 40% chance of rain tomorrow. Since there are only two possible outcomes—rain (success, X = 1) or no rain (failure, X=0) and there is only one trial—this scenario is modeled by a Bernoulli distribution.

Distribution : Bernoulli Distribution

Parameter:

Probability of success (p) = 0.4

Probability Mass Function (PMF)

$P(X = 1) = 0.4$ $\quad P(X = 0) = 1 - P(X = 1) = 1 - 0.4 = 0.6$

(b) When Priyanshu spends his free afternoons watching ship traffic in the harbor—where approximately 4 large ships dock each hour—the Poisson distribution is an appropriate model. This distribution is suitable because at any point there is no upper limit of ships that can dock in a harbour. Thus, it describes the number of events occurring in a fixed time interval at a constant rate.

Distribution : Poisson Distribution

Parameter:

Average number of ships per hour (λ) = 4

Probability Mass Function (PMF)

$P(X = k) = e^{-4}4^k / k!$ $\quad k=0,1,2,\dots$

(c) When Priya and Praveen play a strategy game, both are equally likely to win, and they play 10 matches. Each match is an independent trial where Priya has a win probability of 0.5. Consequently, the outcome of one match does not affect the outcome of another, the probability of success remains constant across trials, and the number of trials (10 matches) is fixed. Thus, this scenario meets all the criteria of a binomial distribution.

Distribution : Binomial Distribution

Parameters**:**

Number of trials (n) = 10

Probability of Priya winning a single match (p) = 0.5

Probability Mass Function (PMF)

$P(X = k) = {}^{10}C_k\,(0.5)^k\,(0.5)^{10-k}$ k = 1, 2, ….9, 10

(d) When Neha plays a board game, her income per turn is determined by rolling a standard 6-sided die, whose outcomes are 1, 2, 3, 4, 5 and 6, which are all equally likely. The outcome of the die is multiplied by 100 to determine Neha's income which can take 100, 200, 300, 400, 500 and 600 only, whose probability is again 1/6. As each outcome of the die is equally likely, her income follows a discrete uniform distribution, where every outcome has the same probability.

Distribution : Discrete Uniform Distribution

Support (Possible Values of X) = {100, 200, 300, 400, 500, 600}

Parameters:

Minimum Value (a) = 100

Maximum Value (b) = 600

Number of possible outcomes (n) = 6

Probability of each value on the die = 1/6

Probability Mass Function (PMF)

$P(X = x) = 1/6 \;\; x \in \{100, 200, 300, 400, 500, 600\}$

(e) Each day, Shyam has a 10% chance of receiving no mail. Since today is Monday, the number of days he receives mail (success) over five days (number of trials) where each day is independent, with a constant probability of receiving mail at 0.9 (1 − 0.11 - 0.11 − 0.1 = 0.9). Since the focus is on receiving mail, we use 0.9 rather than 0.1. Here there are a fixed number of trials, i.e, 5 days, the probability of success remains constant,i.e, 0.9, and each trial is an independent event as the success and failure of one doesn't affect the other. Thus, this scenario meets all the criteria of a binomial distribution.

Distribution : Binomial Distribution

Parameters:

Number of days (n) = 5

Probability of receiving mail on a single day (p) = 0.9

Probability Mass Function (PMF)

$P(X = k) = {}^{5}C_k\,(0.9)^k\,(0.1)^{5-k}$ k = 1, 2, 3, 4, 5