Name: Shreeya Chatterji
Class: CO16
Roll No: 102103447

Probability and Statistics (UCS410)

Experiment 3: Probability distributions

(1) Roll 12 dice simultaneously, and let X denotes the number of 6's that appear. Calculate the probability of getting 7, 8 or 9, 6's using R. (Try using the function pbinom; If we set S = {get a 6 on one roll}, P(S) = 1/6 and the rolls constitute Bernoulli trials; thus X ~ binom(size=12, prob=1/6) and we are looking for P(7 ≤ X ≤ 9).

CODE:

```
#Q1.
# Number of dice rolls
n <- 12
# Probability of getting a 6 on one roll
p_success <- 1/6

# Calculate the cumulative probabilities using binomial distribution
cum_prob_6 <- pbinom(6, size = n, prob = p_success)
cum_prob_9 <- pbinom(9, size = n, prob = p_success)

# Calculate the probability of getting 7, 8, or 9 sixes
prob_7_to_9 <- cum_prob_9 - cum_prob_6

cat("Probability of getting 7, 8, or 9 sixes:", prob_7_to_9, "\n")
```

OUTPUT:

```
> prob_7_to_9 <- cum_prob_9 - cum_prob_6
>
> cat("Probability of getting 7, 8, or 9 sixes:", prob_7_to_9, "\n")
Probability of getting 7, 8, or 9 sixes: 0.001291758
```

CODE(USING dbinom):

```
###USING DBINOM
# Number of dice rolls
n <- 12
# Probability of getting a 6 on one roll
p_success <- 1/6

# Calculate the probabilities using binomial distribution (PDF)
prob_7 <- dbinom(7, size = n, prob = p_success)
prob_8 <- dbinom(8, size = n, prob = p_success)
prob_9 <- dbinom(9, size = n, prob = p_success)

# Calculate the probability of getting 7, 8, or 9 sixes
prob_7_to_9 <- prob_7+prob_8+prob_9

cat("Probability of getting 7, 8, or 9 sixes:", prob_7_to_9, "\n")
```

OUTPUT:

Name: Shreeya Chatterji
Class: CO16
Roll No: 102103447

```
>
> # Calculate the probability of getting 7, 8, or 9 sixes
> prob_7_to_9 <- prob_7+prob_8+prob_9
>
> cat("Probability of getting 7, 8, or 9 sixes:", prob_7_to_9, "\n")
Probability of getting 7, 8, or 9 sixes: 0.001291758
>
```

(2) Assume that the test scores of a college entrance exam fits a normal distribution. Furthermore, the mean test score is 72, and the standard deviation is 15.2. What is the percentage of students scoring 84 or more in the exam?

CODE:

```
#Q2
# Mean and standard deviation
mean_score <- 72
standard_deviation <- 15.2

# Score threshold
threshold <- 84

# cumulative distribution function (CDF)
probability_above_threshold <- 1 - pnorm(threshold, mean = mean_score, sd = standard_deviation)

# probability to percentage
percentage_above_threshold <- probability_above_threshold * 100

cat("Percentage of students scoring 84 or more:", percentage_above_threshold, "%\n")
```

OUTPUT:

```
NAs introduced by coercion
> #(2)
> outcomes <- c("Success", "Failure")
> probab <- c(0.9, 0.1)
> # Generate a sample space for 10 surgical procedures
> sample_space <- sample(outcomes, size = 10, replace = TRUE, prob = probab)
> # Display
> cat("Sample space for next 10 Procedures:\n", sample_space)
Sample space for next 10 Procedures:
 Success Success Success Success Success Success Success Success Success Failure
```

```
>
> cat("Percentage of students scoring 84 or more:", percentage_above_threshold, "%\n")
Percentage of students scoring 84 or more: 21.49176 %
>
```

(3) On the average, five cars arrive at a particular car wash every hour. Let X count the number of cars that arrive from 10AM to 11AM, then X ~Poisson(λ = 5). What is probability that no car arrives during this time. Next, suppose the car wash above is in operation from 8AM to 6PM, and we let Y be the number of customers that appear in this period. Since this period covers a total of 10 hours, we get that Y ∼ Poisson(λ = 5×10 = 50). What is the probability that there are between 48 and 50 customers, inclusive?

CODE:

Name: Shreeya Chatterji
Class: CO16
Roll No: 102103447

```r
#Q3.
# library for Poisson distribution calculations
library(stats)

# Parameters for the Poisson distribution
lambda_x <- 5   # Average number of cars from 10AM to 11AM
lambda_y <- 50 # Average number of customers from 8AM to 6PM

# Probability that no car arrives during 10AM to 11AM
prob_x <- dpois(0, lambda = lambda_x)

# Probability of having between 48 and 50 customers (inclusive) from 8AM to 6PM
prob_y <- sum(dpois(48:50, lambda = lambda_y))

cat("Probability that no car arrives during 10AM to 11AM:", prob_x, "\n")
cat("Probability of having between 48 and 50 customers from 8AM to 6PM:", prob_y, "\n")
```

OUTPUT:

```
>
> cat("Probability that no car arrives during 10AM to 11AM:", prob_x, "\n")
Probability that no car arrives during 10AM to 11AM: 0.006737947
> cat("Probability of having between 48 and 50 customers from 8AM to 6PM:", prob_y, "\n")
Probability of having between 48 and 50 customers from 8AM to 6PM: 0.1678485
>
```

(4) Suppose in a certain shipment of 250 Pentium processors there are 17 defective processors. A quality control consultant randomly collects 5 processors for inspection to determine whether or not they are defective. Let X denote the number of defectives in the sample. Find the probability of exactly 3 defectives in the sample, that is, find P(X = 3).

CODE:

```r
#Q4
# Parameters for hypergeometric distribution
total_processors <- 250
defective_processors <- 17
sample_size <- 5

# p(x=3)
prob_3 <- dhyper(3, m = defective_processors, n = total_processors - defective_processors, k = sample_size)

cat("Probability of exactly 3 defective processors in the sample:", prob_3, "\n")
```

OUTPUT:

```
e)
>
> cat("Probability of exactly 3 defective processors in the sample:", prob_3, "\n")
Probability of exactly 3 defective processors in the sample: 0.002351153
>
```

(5) A recent national study showed that approximately 44.7% of college students have used Wikipedia as a source in at least one of their term papers. Let X equal the number of students in a random sample of size n = 31 who have used Wikipedia as a source.

(a) How is X distributed?

(b) Sketch the probability mass function.

(c) Sketch the cumulative distribution function.

(d) Find mean, variance and standard deviation of X.

Name: Shreeya Chatterji
Class: CO16
Roll No: 102103447

CODE:

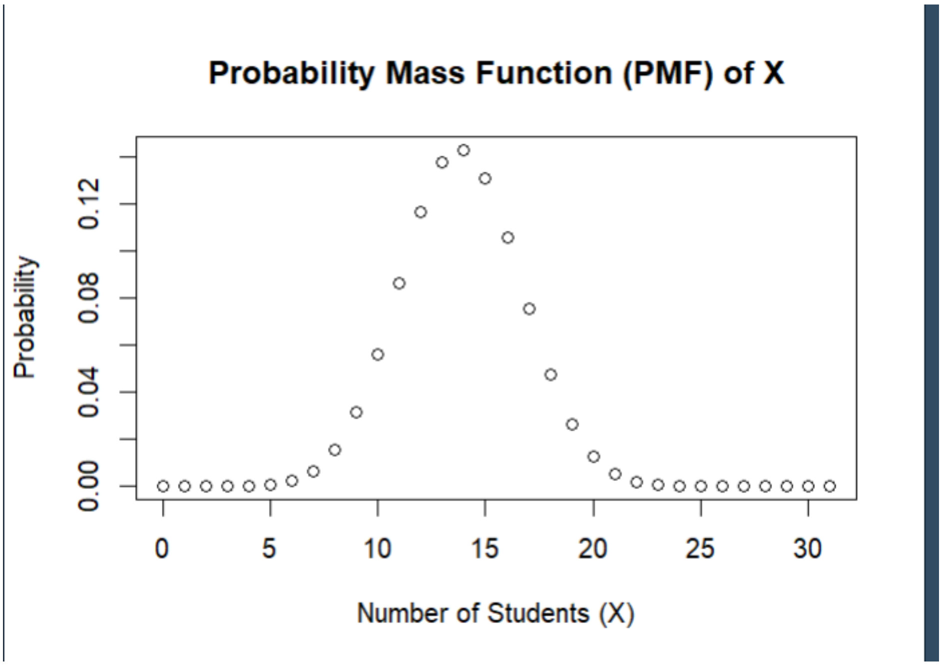
```r
#Q5
# Given probability
p_success <- 0.447

# Sample size
n <- 31

# (a) X is distributed as a binomial distribution
# (b) Sketch the probability mass function (PMF)
xx <- seq(0,31,1)
pmf_values <- numeric()
cdf_values <- numeric()

for(i in 1:length(xx))
{
  pmf_values[i] = dbinom(xx[i],n,p_success)
}

plot(xx,pmf_values)
# (c) Sketch the cumulative distribution function (CDF)

  for(i in 1:length(xx))
  {
    cdf_values[i] = pbinom(xx[i],n,p_success)
  }
# (d) Mean, variance, and standard deviation
mean_x <- n * p_success
variance_x <- n * p_success * (1 - p_success)
std_dev_x <- sqrt(variance_x)

# Print the results
cat("Mean of X:", mean_x, "\n")
cat("Variance of X:", variance_x, "\n")
cat("Standard Deviation of X:", std_dev_x, "\n")

# Plot PMF and CDF
plot(xx, pmf_values, xlab = "Number of Students (X)", ylab = "Probability", main = "Probability Mass Function (PMF) of X")
plot(xx, cdf_values, xlab = "Number of Students (X)", ylab = "Cumulative Probability", main = "Cumulative Distribution Function (CDF) of X")
```

OUTPUT:

```
> # Print the results
> cat("Mean of X:", mean_x, "\n")
Mean of X: 13.857
> cat("Variance of X:", variance_x, "\n")
Variance of X: 7.662921
> cat("Standard Deviation of X:", std_dev_x, "\n")
Standard Deviation of X: 2.768198
```

PLOTS:

PDF: plot(xx, pmf_values, xlab = "Number of Students (X)", ylab = "Probability", main = "Probability Mass Function (PMF) of X")



Name: Shreeya Chatterji
Class: CO16
Roll No: 102103447

CDF: plot(xx, cdf_values, xlab = "Number of Students (X)", ylab = "Cumulative Probability", main = "Cumulative Distribution Function (CDF) of X")