ASSIGNMENT 3
Probability and Statistics

Name: AVISEK MANDAL

Roll Number: 102203700

Batch: 3CO35

Question 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).

OUTPUT:

```
> diff(pbinom(c(6,9), size=12, prob=1/6, lower.tail = TRUE))
[1] 0.001291758
```

Question 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?

OUTPUT:

```
> pnorm(84, mean=72, sd= 15.2, lower.tail=FALSE)
[1] 0.2149176
```

Question 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 $\sim$Poisson($\lambda = 5$). 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 $\sim$ Poisson($\lambda = 5\times10 = 50$). What is the probability that there are between 48 and 50 customers, inclusive?

OUTPUT:

```
> lambda_X <- 5
> prob_X_0 <- dpois(0, lambda_X)
> cat("Probability that no car arrives from 10 AM to 11 AM:", prob_X_0, "\n")
Probability that no car arrives from 10 AM to 11 AM: 0.006737947
>
> # Part 2: Probability that there are between 48 and 50 customers inclusive
> lambda_Y <- 50
> prob_Y_48_50 <- ppois(50, lambda_Y) - ppois(47, lambda_Y)
> cat("Probability that there are between 48 and 50 customers:", prob_Y_48_50, "\n")
Probability that there are between 48 and 50 customers: 0.1678485
```

Question 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).

OUTPUT:

```
> N <- 250    # Total number of processors
> K <- 17     # Number of defective processors
> n <- 5      # Sample size
> X <- 3      # Number of defectives we want in the sample
>
> # Probability of exactly 3 defectives in the sample
> prob_X_3 <- dhyper(X, K, N-K, n)
> prob_X_3
[1] 0.002351153
```

Question 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.

OUTPUT:

```
> n <- 31
> p <- 0.447
>
> # Possible values of X
> x_values <- 0:n
>
> # PMF values
> pmf_values <- dbinom(x_values, n, p)
>
> # Plot the PMF
> plot(x_values, pmf_values, type="h", lwd=2, col="blue",
+      main="Probability Mass Function of X",
+      xlab="Number of Students Using Wikipedia (X)",
+      ylab="Probability")
> points(x_values, pmf_values, pch=16, col="red")
>
> # CDF values
> cdf_values <- pbinom(x_values, n, p)
>
> # Plot the CDF
> plot(x_values, cdf_values, type="s", lwd=2, col="blue",
+      main="Cumulative Distribution Function of X",
+      xlab="Number of Students Using Wikipedia (X)",
+      ylab="Cumulative Probability")
> points(x_values, cdf_values, pch=16, col="red")
>
> # Mean
> mean_X <- n * p
>
> # Variance
> var_X <- n * p * (1 - p)
>
> # Standard Deviation
> sd_X <- sqrt(var_X)
>
> cat("Mean:", mean_X, "\n")
Mean: 13.857
> cat("Variance:", var_X, "\n")
Variance: 7.662921
> cat("Standard Deviation:", sd_X, "\n")
Standard Deviation: 2.768198
```