STAT3355(HW-5)

Tulasi Janjanam

2024-11-04

```
{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

Problem 1

```r
## Given probability of a defective bulb
p <- 3 / 75
q <- 1 - p

# Probability of finding the first defective bulb on the 6th trial
k <- 6
P_X_6 <- q^(k - 1) * p
round(P_X_6, 3)
```

```
## [1] 0.033
```

LaTeX:

$$P(X = 6) = (1 - p)^{6-1} \cdot p$$

```r
# Probability of taking at least four trials
P_X_less_4 <- sum(q^(0:2) * p)  # P(X = 1) + P(X = 2) + P(X = 3)
P_X_geq_4 <- 1 - P_X_less_4
round(P_X_geq_4, 3)
```

```
## [1] 0.885
```

LaTeX:

$$P(X >= 4) = 1 - (P(X = 1) + P(X = 2) + P(X = 3))$$

```r
# Probability of taking at most 10 trials
P_X_leq_10 <- sum(q^(0:9) * p)  # Summing probabilities from X = 1 to X = 10
round(P_X_leq_10, 3)
```

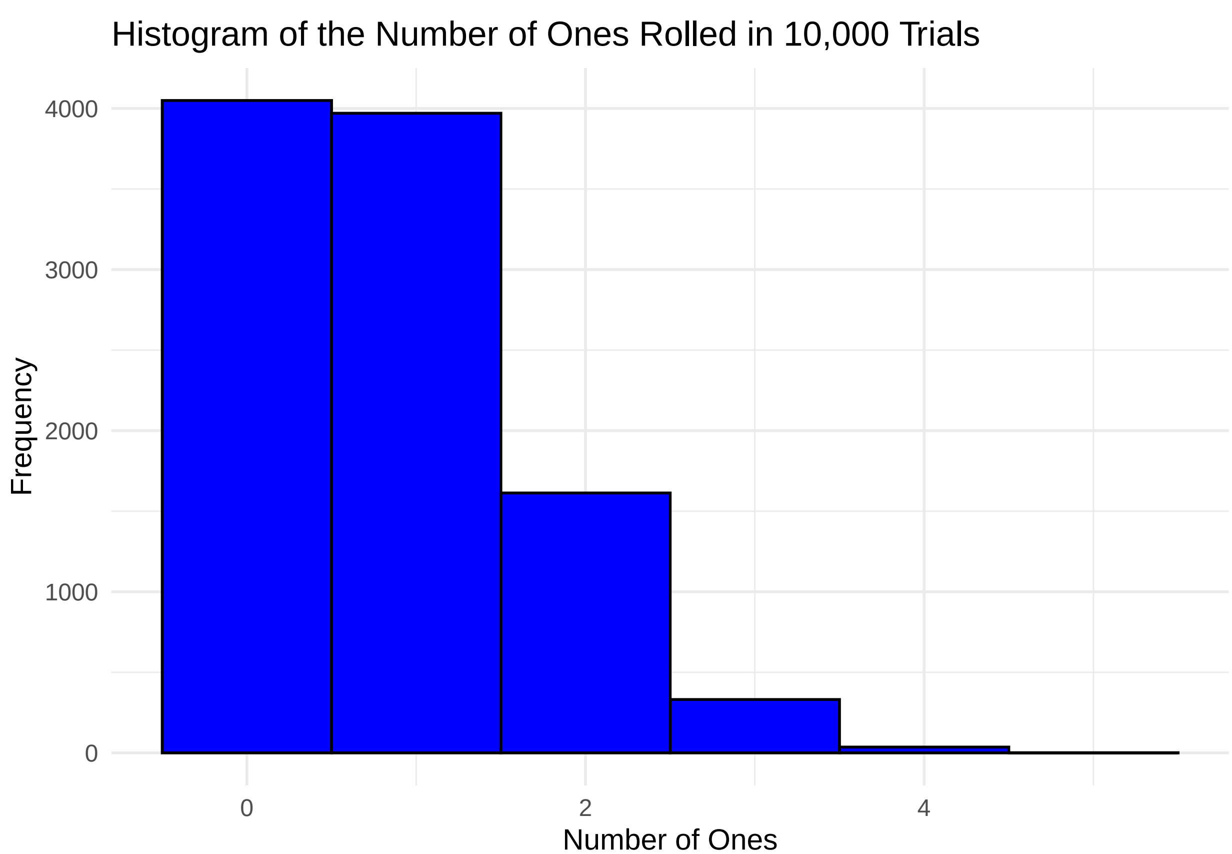
```
## [1] 0.335
```

LaTeX:

$$p(X = k) = (1 - p)^{k-1} * p$$

Problem 2

```r
set.seed(20220707)
n_trials <- 10000

simulate_rolls <- function() {
  sum(sample(1:6, 5, replace = TRUE) == 1)
}

X <- replicate(n_trials, simulate_rolls())
library(ggplot2)
X_data <- data.frame(X = X)

# Plot the histogram
ggplot(X_data, aes(x = X)) +
  geom_histogram(binwidth = 1, fill = "blue", color = "black") +
  labs(title = "Histogram of the Number of Ones Rolled in 10,000 Trials",
       x = "Number of Ones",
       y = "Frequency") +
  theme_minimal()
```



Calculate sample mean and sample variance
```r
# Calculate sample mean and sample variance
sample_mean <- mean(X)
sample_variance <- var(X)

cat("Sample Mean:", round(sample_mean, 3), "\n")
```

```
## Sample Mean: 0.834
```

```r
cat("Sample Variance:", round(sample_variance, 3), "\n")
```

```
## Sample Variance: 0.705
```

LaTeX:

Sample Mean:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Sample Variance:

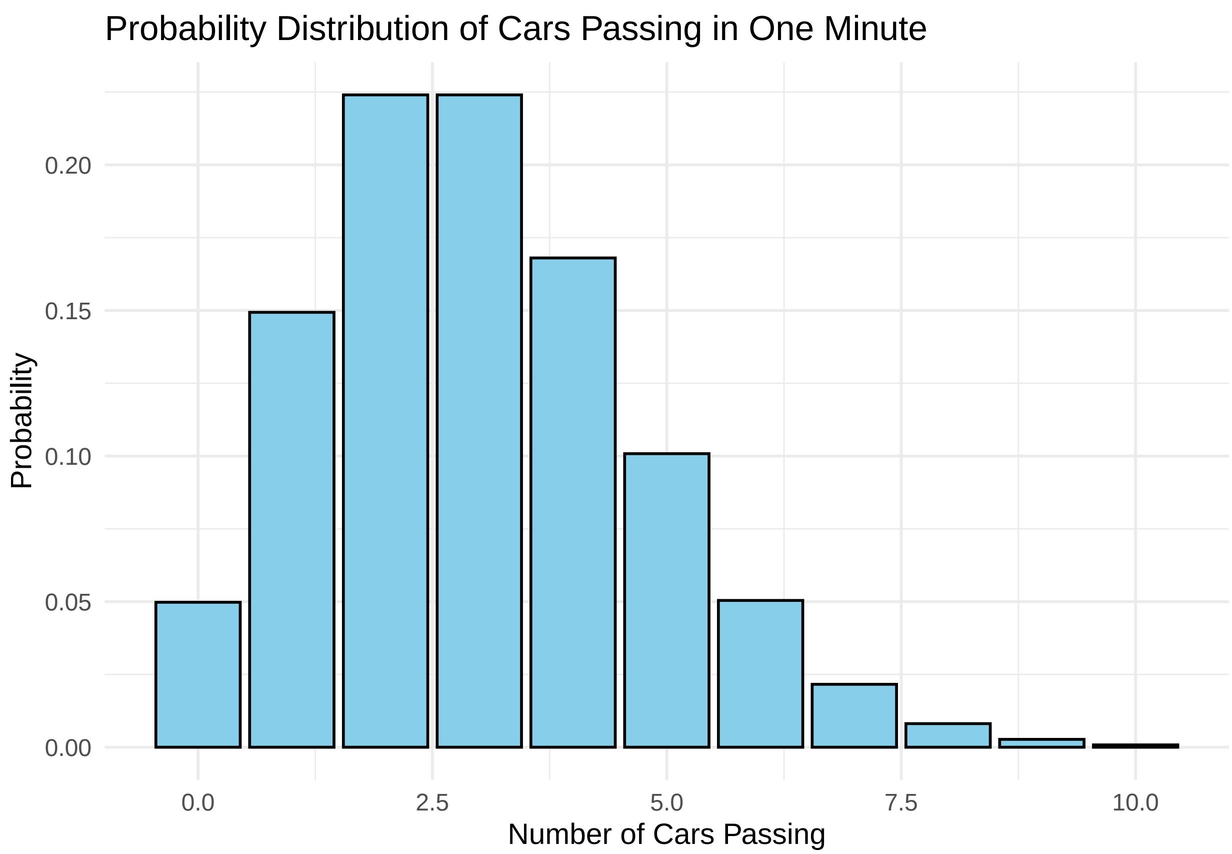
$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

Problem 3

```r
# Define the average rate (lambda) for cars per minute
lambda <- 180 / 60

# (a) Calculate the probability that congestion will occur: P(X > 5) = 1 - P(X <= 5)
prob_congestion <- 1 - ppois(5, lambda)
cat("Probability of congestion (more than 5 cars in one minute):", round(prob_congestion, 3), "\n")
```

```
## Probability of congestion (more than 5 cars in one minute): 0.084
```

```r
# (b) Plot a bar chart
library(ggplot2)
x_values <- 0:10
probabilities <- dpois(x_values, lambda)
data <- data.frame(Cars = x_values, Probability = probabilities)

ggplot(data, aes(x = Cars, y = Probability)) +
  geom_bar(stat = "identity", fill = "skyblue", color = "black") +
  labs(title = "Probability Distribution of Cars Passing in One Minute",
       x = "Number of Cars Passing",
       y = "Probability") +
  theme_minimal()
```



LaTeX:

$$P(X > 5) = 1 - P(X \leq 5)$$

Problem 4

```
# Given mean and standard deviation
mean_score <- 500
sd_score <- 100

# (a) Probability of scoring 585 or less
prob_585_or_less <- pnorm(585, mean = mean_score, sd = sd_score)
cat("Probability of scoring 585 or less:", round(prob_585_or_less, 3), "\n")
```

```
## Probability of scoring 585 or less: 0.802
```

```
# (b) Lower quartile (25th percentile), median (50th percentile), and upper quartile (75th percentile)
lower_quartile <- qnorm(0.25, mean = mean_score, sd = sd_score)
median <- qnorm(0.50, mean = mean_score, sd = sd_score)
upper_quartile <- qnorm(0.75, mean = mean_score, sd = sd_score)

cat("Lower quartile (25th percentile):", round(lower_quartile, 3), "\n")
```

```
## Lower quartile (25th percentile): 432.551
```

```r
cat("Median (50th percentile):", round(median, 3), "\n")
```

```
## Median (50th percentile): 500
```

```r
cat("Upper quartile (75th percentile):", round(upper_quartile, 3), "\n")
```

```
## Upper quartile (75th percentile): 567.449
```

LaTeX:

(a)

$$P(X \leq 585) = \Phi\left(\frac{585 - 500}{100}\right)$$

(b) Quartiles:

$$Q_1 = \Phi^{-1}(0.25), Q_2 = \Phi^{-1}(0.5), Q_3 = \Phi^{-1}(0.75)$$

Problem 5

```r
# Number of trials and probability of heads
n <- 10
p <- 0.5

# (a) Probability of observing seven or more heads (Event A)
prob_A <- 1 - pbinom(6, n, p)
cat("Probability of observing seven or more heads (Event A):", round(prob_A, 2), "\n")
```

```
## Probability of observing seven or more heads (Event A): 0.17
```

```r
# (b) Probability of observing three or fewer heads (Event B)
prob_B <- pbinom(3, n, p)
cat("Probability of observing three or fewer heads (Event B):", round(prob_B, 2), "\n")
```

```
## Probability of observing three or fewer heads (Event B): 0.17
```

```r
# Determine which event is more likely
if (prob_A > prob_B) {
  cat("Event A is more likely to happen.\n")
} else if (prob_B > prob_A) {
  cat("Event B is more likely to happen.\n")
} else {
  cat("Both events are equally likely to happen.\n")
}
```

```
## Event B is more likely to happen.
```

LaTeX for calculations:

(a)

$$P(X \geq 7) = 1 - P(X \leq 6)$$

(b)

$$P(X \leq 3) = \sum_{k=0}^{3} \binom{10}{k} (0.5)^k (0.5)^{10-k}$$

Problem 6

```
# Given probabilities
P_G <- 0.95   # Probability that the cab is green
P_R <- 0.05   # Probability that the cab is red
P_S_R_given_R <- 0.8   # Probability witness says red given cab is red
P_S_R_given_G <- 0.2   # Probability witness says red given cab is green

# Calculate P(S_R)
P_S_R <- (P_S_R_given_R * P_R) + (P_S_R_given_G * P_G)

# Calculate P(R | S_R) using Bayes' Theorem
P_R_given_S_R <- (P_S_R_given_R * P_R) / P_S_R

# Display the result rounded to 3 decimal places
cat("Probability that the cab was red given the witness said it was red:", round(P_R_given_S_R, 3), "\n
```

```
## Probability that the cab was red given the witness said it was red: 0.174
```

LaTeX for calculations:

$$P(\text{Red} \mid \text{Witness says Red}) = \frac{P(\text{Witness says Red} \mid \text{Red}) \cdot P(\text{Red})}{P(\text{Witness says Red})}$$

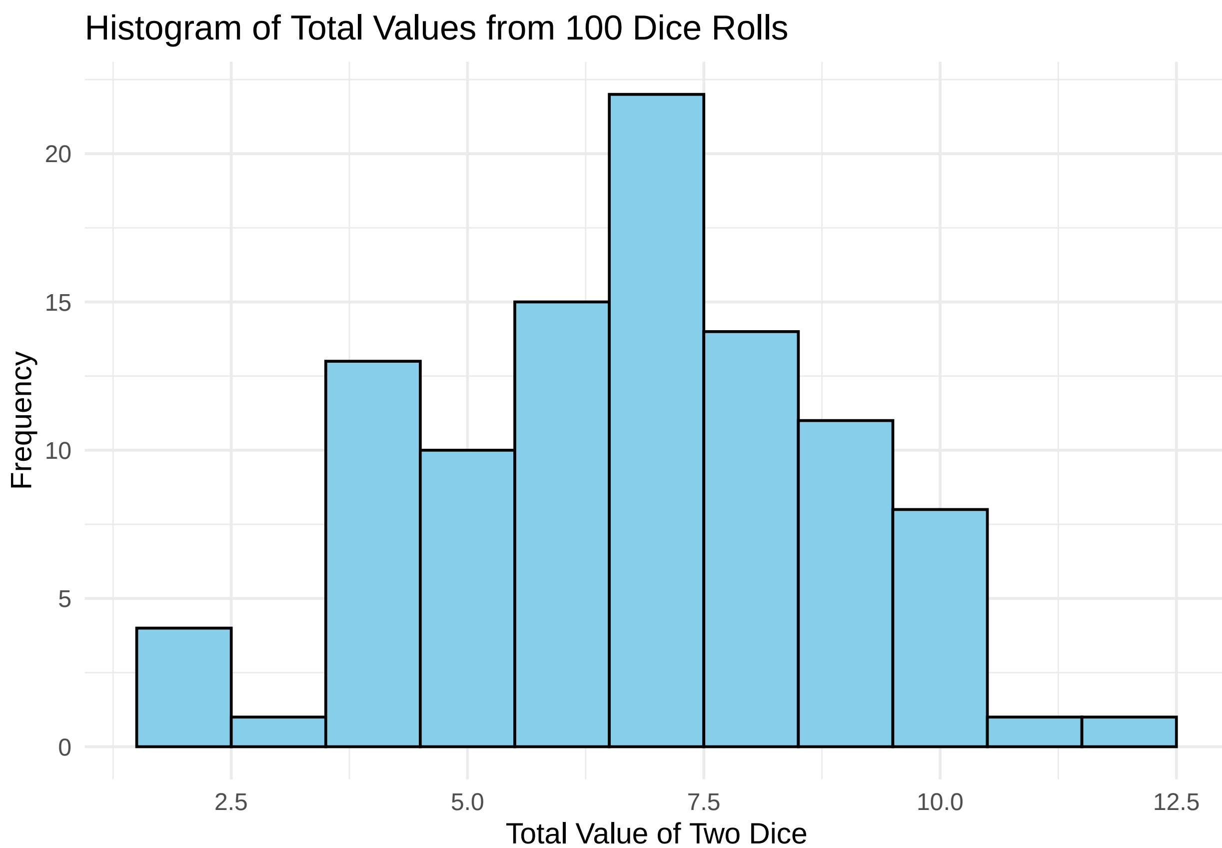
where

$$P(\text{Witness says Red}) = P(\text{Witness says Red} \mid \text{Green}) \cdot P(\text{Green}) + P(\text{Witness says Red} \mid \text{Red}) \cdot P(\text{Red})$$

Problem 7

```
set.seed(20191031)
n_rolls <- 100
X <- replicate(n_rolls, sum(sample(1:6, 2, replace = TRUE)))

library(ggplot2)
X_data <- data.frame(X = X)

# Plot the histogram
```

```r
ggplot(X_data, aes(x = X)) +
  geom_histogram(binwidth = 1, fill = "skyblue", color = "black") +
  labs(title = "Histogram of Total Values from 100 Dice Rolls",
       x = "Total Value of Two Dice",
       y = "Frequency") +
  theme_minimal()
```




```r
sample_mean <- mean(X)
sample_variance <- var(X)

# Print sample mean and variance
cat("Sample Mean:", round(sample_mean, 2), "\n")
```

```
## Sample Mean: 6.71
```

```r
cat("Sample Variance:", round(sample_variance, 2), "\n")
```

```
## Sample Variance: 4.55
```

LaTeX:

Sample Mean:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Sample Variance:

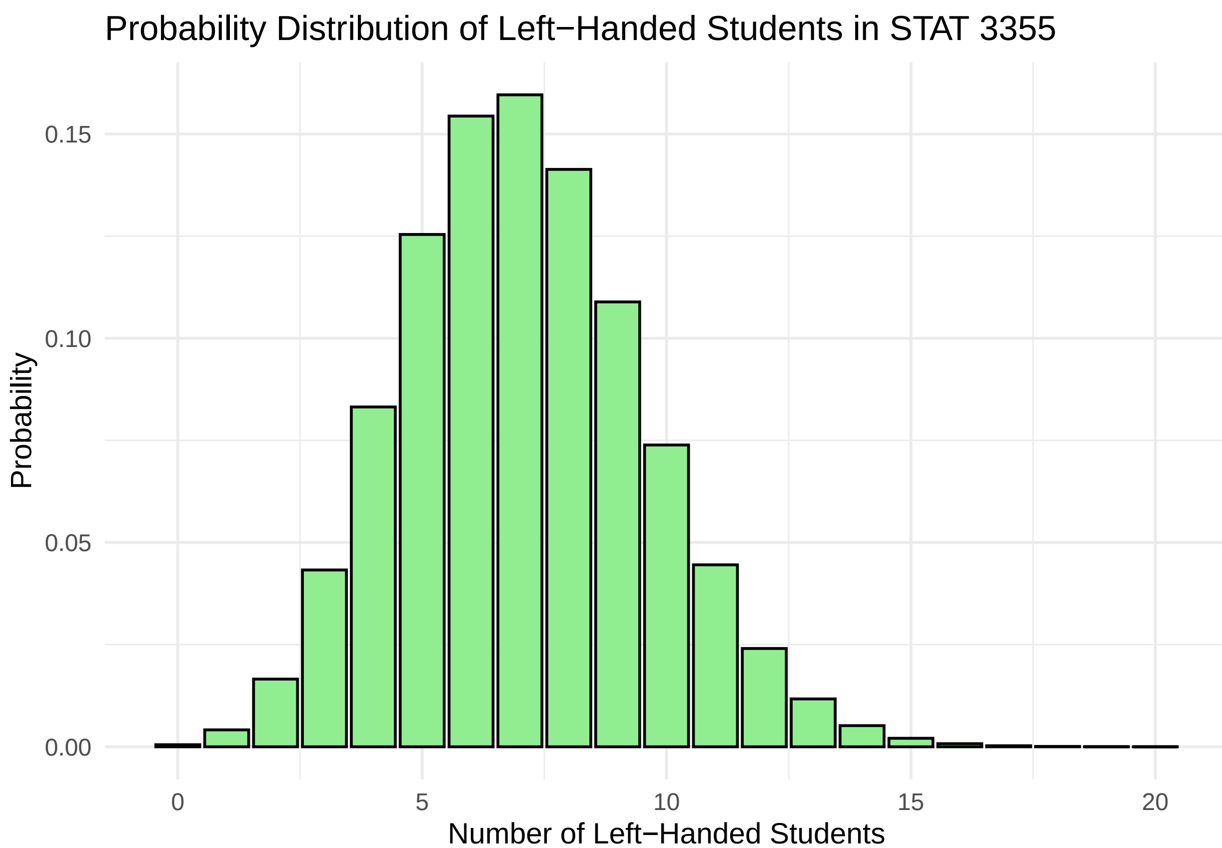
$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Problem 8

```r
# Number of students and probability of being left-handed
n_students <- 54
p_left_handed <- 0.131

# (a) Probability that 10 or fewer students are left-handed
prob_10_or_fewer <- pbinom(10, n_students, p_left_handed)
cat("Probability of 10 or fewer left-handed students:", round(prob_10_or_fewer, 3), "\n")
```

```
## Probability of 10 or fewer left-handed students: 0.911
```

```r
# (b) Plotting the probability distribution for 0 to 20 left-handed students

library(ggplot2)
x_values <- 0:20
probabilities <- dbinom(x_values, n_students, p_left_handed)
data <- data.frame(Left_Handed_Students = x_values, Probability = probabilities)

# Plot the bar chart
ggplot(data, aes(x = Left_Handed_Students, y = Probability)) +
  geom_bar(stat = "identity", fill = "lightgreen", color = "black") +
  labs(title = "Probability Distribution of Left-Handed Students in STAT 3355",
       x = "Number of Left-Handed Students",
       y = "Probability") +
  theme_minimal()
```



LaTeX:

$$P(X \le 10) = \sum_{k=0}^{10} \binom{54}{k} (0.131)^k (1 - 0.131)^{54-k}$$

Problem 9

```
# Parameters for the normal distribution
mu <- 12          # Mean
sigma <- 0.5      # SD

# (a) Probability of height
height_limit <- 10.7
probability <- pnorm(height_limit, mu, sigma)
cat("Probability that a randomly chosen cereal box has height 10.7 inches or less:", round(probability,
```

```
## Probability that a randomly chosen cereal box has height 10.7 inches or less: 0.0047
```

```
# (b) Calculate quartiles
Q1 <- qnorm(0.25, mu, sigma)  # Lower quartile
median <- qnorm(0.5, mu, sigma)  # Median
Q3 <- qnorm(0.75, mu, sigma)  # Upper quartile

# Print quartiles
cat("Lower Quartile (Q1):", round(Q1, 2), "\n")
```

```
## Lower Quartile (Q1): 11.66
```

```r
cat("Median (Q2):", round(median, 2), "\n")
```

```
## Median (Q2): 12
```

```r
cat("Upper Quartile (Q3):", round(Q3, 2), "\n")
```

```
## Upper Quartile (Q3): 12.34
```

LaTeX:

(a)

$$P(X \leq 10.7) = \Phi\left(\frac{10.7 - 12}{\sqrt{0.52}}\right)$$

(b) Quartiles:

$$Q_1 = \Phi^{-1}(0.25), Q_2 = \Phi^{-1}(0.5), Q_3 = \Phi^{-1}(0.75)$$