Homework 6

Kevin Jin

4/10/2022

Problem 1

```
# Simulate 10 flips of a fair coin
x <- c("T", "H")
sample(x, size = 10, replace = TRUE)
```

```
##  [1] "H" "H" "T" "H" "T" "H" "T" "T" "H" "T"
```

```
# Binomial distribution of Event A (7 or more heads)
p <- dbinom(7, size = 10, prob = 0.5) +
    dbinom(8, size = 10, prob = 0.5) +
    dbinom(9, size = 10, prob = 0.5) +
    dbinom(19, size = 10, prob = 0.5)
round(p, 2)
```

```
## [1] 0.17
```

```
# Binomial distribution of Event B (3 heads or less)
q <- dbinom(4, size = 10, prob = 0.5) +
    dbinom(5, size = 10, prob = 0.5) +
    dbinom(6, size = 10, prob = 0.5) +
    dbinom(7, size = 10, prob = 0.5) +
    dbinom(8, size = 10, prob = 0.5) +
    dbinom(9, size = 10, prob = 0.5) +
    dbinom(10, size = 10, prob = 0.5)
p <- 1 - q
round(p, 2)
```

```
## [1] 0.17
```

Derivation:

Let the sample space Ω be the set of all outcomes of flipping a fair coin 10 times. Let Event A be the probability of observing 7 or more heads. We apply the binomial theorem:

$$
\begin{aligned}
P(A) &= \sum_{k=7}^{10} \binom{10}{k} (\frac{1}{2})^k (\frac{1}{2})^{10-k} \\
&= \frac{120 + 45 + 10 + 1}{1024} \\
&= \frac{11}{64}
\end{aligned}
$$

Let Event B be the probability of observing 3 or less heads. Let B' be the complement of B, that is, the probability of observing 4 or more heads. We apply the binomial theorem again:

$$P(B') = \sum_{k=4}^{10} \binom{10}{k} (\frac{1}{2})^k (\frac{1}{2})^{10-k}$$
$$= \frac{210 + 252 + 210 + 120 + 45 + 10 + 1}{1024}$$
$$= \frac{53}{64}$$

By the complement rule of probability:

$$P(B) = 1 - P(B')$$
$$= 1 - \frac{53}{64}$$
$$= \frac{11}{64}$$

Problem 2

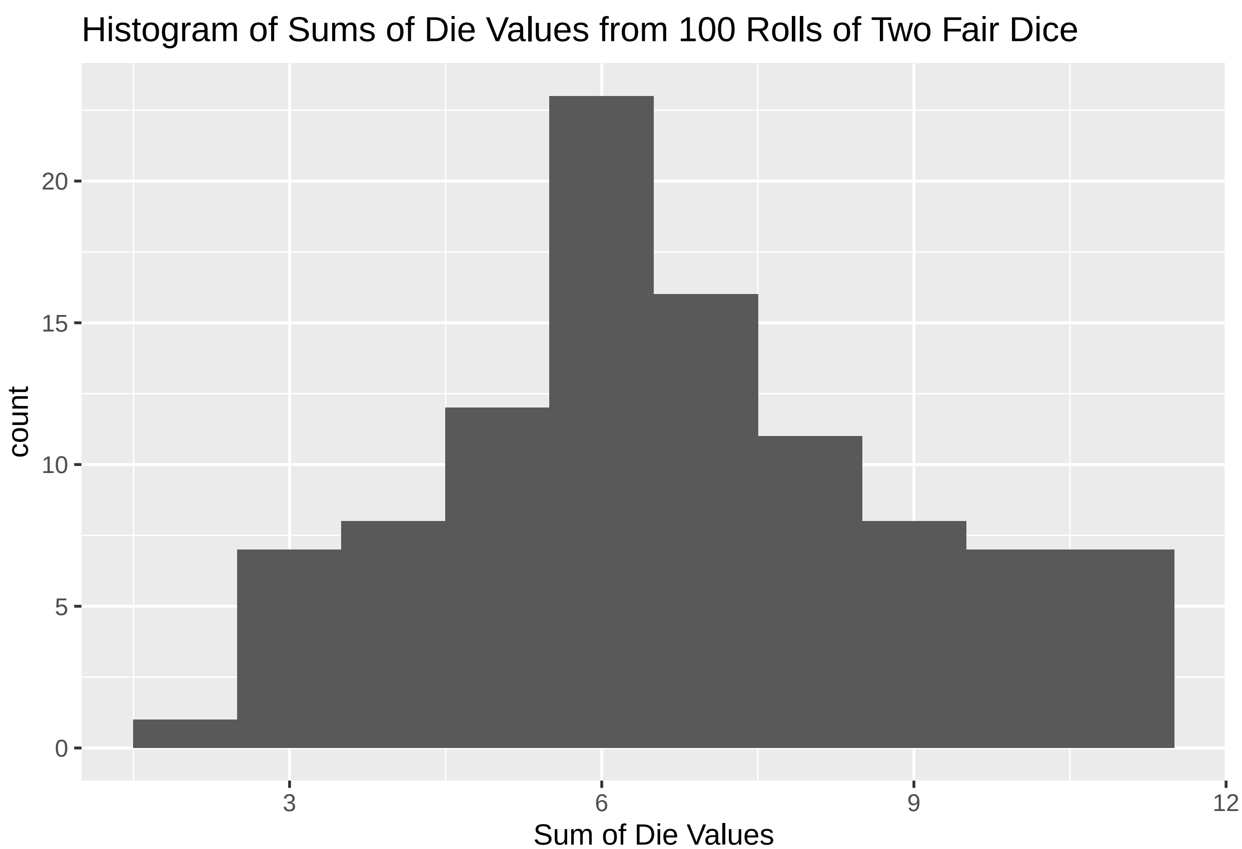
```
library(ggplot2)

# Guarantee reproducibility
set.seed(20220404)

# Simulate 100 rolls of two six-sided fair dice
# Combined sum is the random variable X
X <- sample(1:6, size = 100, replace = TRUE) +
  sample(1:6, size = 100, replace = TRUE)

# Plot distribution of results
n <- length(X)
k <- ceiling(sqrt(n)) # Square root choice

ggplot() +
  geom_histogram(mapping = aes(x = X),
                 bins = k) +
  labs(x = "Sum of Die Values",
       title = "Histogram of Sums of Die Values from 100 Rolls of Two Fair Dice")
```




```r
# Sample mean and sample variance
mean(X)
```

```
## [1] 6.72
```

```r
round(var(X), 2)
```

```
## [1] 4.95
```

Let n be the number of students in the sample. The sample mean $\bar{x}$ is calculated thus:

$$\begin{aligned} \bar{x} &= \frac{\sum_{k=1}^{n}}{n} \\ &= 6.72 \end{aligned}$$

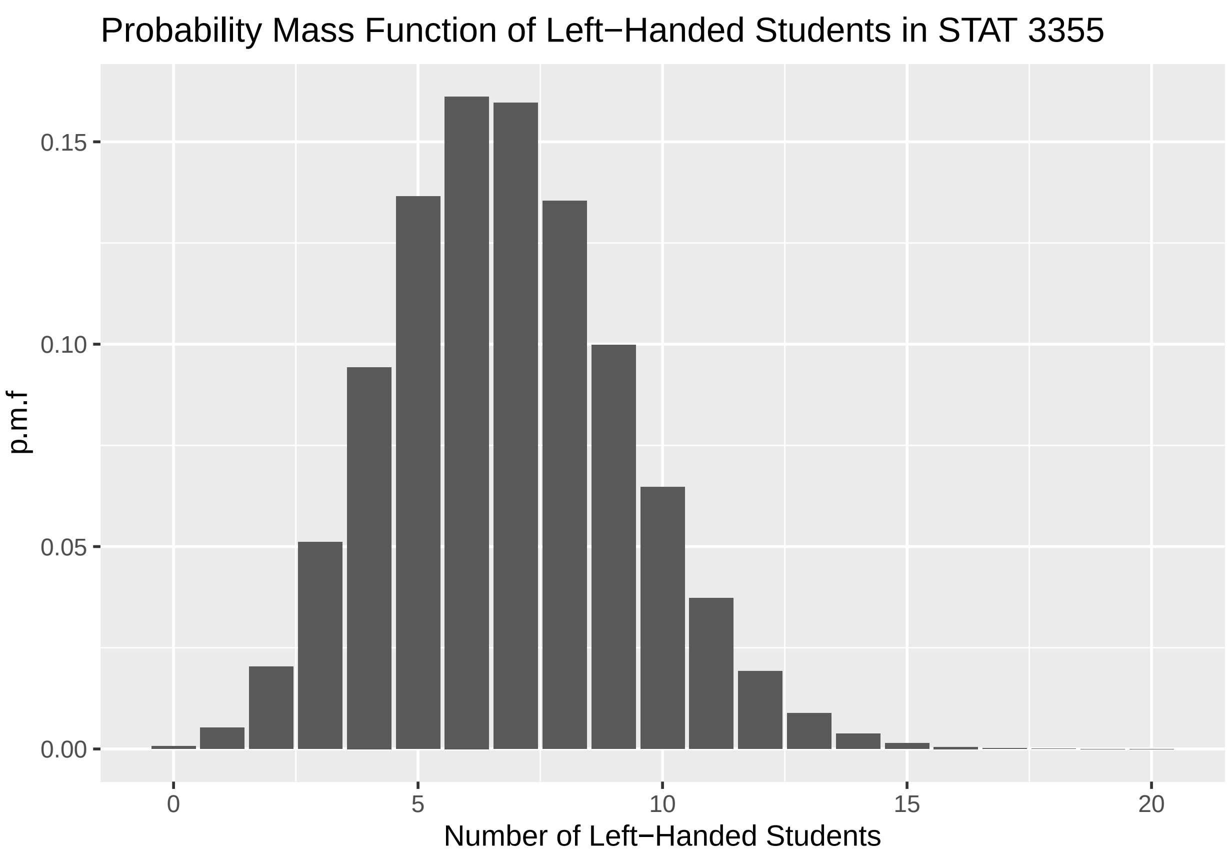
Let x_i be the value of one sample. The sample variance S^2 is calculated thus:

$$\begin{aligned} S^2 &= \frac{\sum_{k=1}^{n}(x_i - \bar{x})^2}{n-1} \\ &= 4.95 \end{aligned}$$

Problem 3

```r
# Probability of 10 or fewer left-handed students in class
y <- pbinom(10, size = 52, prob = 0.131)
```

```r
# Bar chart
x <- 0:20
y <- dbinom(x, size = 52, prob = 0.131)
ggplot() +
  geom_bar(mapping = aes(x = x, y = y),
           stat='identity') +
  labs(title = "Probability Mass Function of Left-Handed Students in STAT 3355",
       x = "Number of Left-Handed Students",
       y = "p.m.f")
```



Problem 4

```r
# Probability of a random cereal box having height of <=10.7 in
pnorm(10.7, mean = 12, sd = 0.5)
```

```
## [1] 0.004661188
```

```r
# Quantiles of the normal distribution
# 25th
round(qnorm(0.25, mean = 12, sd = 0.5), 2)
```

```
## [1] 11.66
```

```r
# Median
round(qnorm(0.50, mean = 12, sd = 0.5), 2)
```

```
## [1] 12
```

```r
# 75th
round(qnorm(0.75, mean = 12, sd = 0.5), 2)
```

```
## [1] 12.34
```