1. Section 4.3.1, problem 1
A fair coin is tossed until either a head occurs or 6 tails in a row have occurred. Let X denote the number of tosses. Find the frequency function, mean, and variance of X.

The possible outcomes of this description is:

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>X</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>TH</td>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>TTH</td>
<td>3</td>
<td>0.125</td>
</tr>
<tr>
<td>TTTTH</td>
<td>4</td>
<td>0.0625</td>
</tr>
<tr>
<td>TTTTTH</td>
<td>5</td>
<td>0.03125</td>
</tr>
<tr>
<td>TTTTTTH</td>
<td>6</td>
<td>0.015625</td>
</tr>
<tr>
<td>TTTTTTT</td>
<td>6</td>
<td>0.015625</td>
</tr>
</tbody>
</table>

The frequency function is the table:

<table>
<thead>
<tr>
<th>X</th>
<th>P(X = x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>0.125</td>
</tr>
<tr>
<td>4</td>
<td>0.0625</td>
</tr>
<tr>
<td>5</td>
<td>0.03125</td>
</tr>
<tr>
<td>6</td>
<td>0.015625 + 0.015625 = 0.03125</td>
</tr>
</tbody>
</table>

Mean $E(X) = 1(0.5) + 2(0.25) + 3(0.125) + 4(0.0625) + 5(0.03125) + 6(0.03125) = 1.96875$

$E(X^2) = 1^2(0.5) + 2^2(0.25) + 3^2(0.125) + 4^2(0.0625) + 5^2(0.03125) + 6^2(0.03125) = 5.53125$

Var(X) = $E(X^2) - [E(X)]^2 = 5.53125 - 1.96875^2 = 1.655273$
2. Section 4.3.1
   a. Problem 3
   b. Problem 4

   a. Let $X$ denote the number of spots on a single throw of a fair 6-sided die. Find the mean, variance, and standard deviation of $X$. If you can, relate the mean and variance to the mean and variance in Example 3.

   \[
   \text{Mean} = E(X) = 1(1/6) + 2(1/6) + 3(1/6) + 4(1/6) + 5(1/6) + 6(1/6) = 3.5
   \]

   \[
   E(X^2) = 1^2(1/6) + 2^2(1/6) + 3^2(1/6) + 4^2(1/6) + 5^2(1/6) + 6^2(1/6) = 15.16667
   \]

   \[
   \text{Var}(X) = E(X^2) - [E(X)]^2 = 15.16667 - 3.5^2 = 2.91667
   \]

   In Example 4.3, notice $T = X + Y$, where $X$ is one die (say red) and $Y$ is another die (say green),

   \[
   E(X + Y) = E(X) + E(Y) = 3.5 + 3.5 = 7
   \]

   \[
   \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) = 2.91667 + 2.91667 = 5.8334
   \]

   b. Verify Chebyshev’s inequality for $k = 2$ and $k = 3$ when $X$ is the total number of spots on two rolls of a fair 6-sided die.

   From Example 4.3, $E(X) = 7$ and $SD(X) = 2.415$

   From Chebyshev’s inequality if $k = 2$ then

   \[
   P[|X - 7| > 2(2.415)] \leq \frac{1}{4}
   \]

   $X - 7 < -4.8304$ or $X - 7 > 4.8304$

   $X < 2.1696$ or $X > 11.8304$

   From the example

   \[
   P(X < 2.1696 \text{ or } X > 11.8304) = 1/36 + 1/36 = 0.0555 < .25 \text{ This confirms Chebyshev’s inequality.}
   \]

   From Chebyshev’s inequality if $k = 3$ then

   \[
   P[|X - 7| > 3(2.415)] \leq \frac{1}{9}
   \]

   $X - 7 < 7.245$ or $X - 7 > -7.245$

   $X < -0.245$ or $X > 14.245$

   From the example

   \[
   P(X < -0.245 \text{ or } X > 14.245) = 0 < 1/9 \text{ This confirms Chebyshev’s inequality.}
   \]
3. Section 4.3.1, problem 5  
A biased coin has probability 0.6 of turning up heads. You win $x if a head comes up and you lose $y if a tail comes up. If your expected winnings is $0, what is the relationship between x and y?

\[ 0 = 0.6x + 0.4(-y) = 0.6x - 0.4y \]
\[ 0.4y = 0.6x \]
\[ y = 1.5x \]
4. Section 4.6.1  
   a. Problem 2  
   b. Problem 3  
   c. Problem 4  
   d. Problem 5  

   a. Six people are randomly selected in succession, with replacement, from a class containing 25 men and 20 women. What is the probability of obtaining the sequence 1, 0, 0, 1, 1, where 1 indicates a man was chosen and 0 indicates a woman was chosen?  
   \[ X = \text{a person chosen} \]  
   \[ P(X = 1) = \frac{25}{45} = 0.5556; \ P(X = 0) = \frac{20}{45} = 0.44444 \]  
   \[ P(X = 1 \cap X = 0 \cap X = 0 \cap X = 0 \cap X = 1 \cap X = 1) = 0.55556^3 \cdot (0.44444)^3 = 0.015 \]  
   This is possible because of “with replacement” thus we have independence.  

   b. Write down all the other outcomes of this sequential sampling experiment that lead to 3 men and 3 women being chosen. What are their probabilities?  
   
   All possible outcomes:  
   
   | 1, 1, 1, 0, 0, 0 | 1, 0, 1, 0, 1, 0 | 0, 1, 0, 0, 1, 1 | 0, 1, 1, 1, 0, 0 | 1, 1, 0, 1, 0, 0 | 1, 0, 1, 0, 0, 0 | 0, 1, 0, 1, 0, 1 | 0, 0, 1, 1, 1, 0 | 0, 1, 0, 0, 1, 1 | 0, 1, 1, 0, 1, 1 | 0, 0, 1, 0, 1, 1 | 0, 0, 0, 1, 1, 1 | 1, 0, 0, 1, 1, 1 | 0, 1, 0, 1, 1, 0 | 0, 0, 1, 1, 0, 1 | 0, 1, 0, 1, 1, 1 | 0, 0, 1, 0, 1, 1 | 0, 0, 0, 1, 1, 1 | 1, 0, 0, 0, 1, 1 | 0, 1, 0, 0, 1, 1 | 0, 0, 0, 1, 1, 1 | 1, 0, 0, 0, 1, 1 | 0, 1, 0, 0, 1, 1 | 0, 0, 0, 1, 1, 1 |

   There should be a total C(6,3) = 20 outcomes, probability of each outcome is the same at 0.015  

   c. What is the probability that 3 men are chosen in the sampling experiment?  
   \[ P(X = 3) = 20(0.015) = 0.3 \]  

   d. What is the probability that 2 or more women are chosen?  
   Let \( Y = \text{women are chosen}, \ n = 6, \ p = 0.4444 \)  
   \[ P(Y \geq 2) = 1 - P(Y \leq 1) = 1 - \text{phinom}(1,6,0.444) = 0.8294166 \]
5. Section 4.6.1, problem 13
Use R’s ”pbinom” function to verify Chebyshev’s inequality for k = 2 and k = 3 when X \sim Binomial(50, 0.4).

mean = 50(0.4) = 20
standard deviation = \sqrt{50(0.4)(0.6)} = 3.4641

If k = 2, \( P(|X – 20| > 2(3.4641)) \leq \frac{1}{4}, \)
\( X – 20 < -6.928 \) or \( X – 20 > 6.928 \)
\( X < 13.0718 \) or \( X > 26.928 \)
Thus we can find \( P(X \leq 13 \text{ or } X \geq 27) = pbinom(13,50,0.4) + 1 - pbinom(26,50,0.4) = 0.05939392 \)
Which is less than \( \frac{1}{4} = 0.25 \), thus Chebyshev’s inequality holds.

If k = 3, \( P(|X – 20| > 3(3.4641)) \leq \frac{1}{9}, \)
\( X – 20 < -10.3923 \) or \( X – 20 > 10.3923 \)
\( X < 9.6077 \) or \( X > 30.3923 \)
Thus we can find \( P(X \leq 9 \text{ or } X \geq 30) = pbinom(9,50,0.4) + 1 - pbinom(29,50,0.4) = 0.0041 \)
Which is less than \( \frac{1}{9} = 0.1111 \), thus the Chebyshev’s inequality holds.
6. Section 4.8.3
   a. Problem 1
   b. Problem 2
   c. Problem 3

   a. A club has 50 members, 10 belonging to the ruling clique and 40 second-class members. Six members are randomly selected for free movie tickets. What is the probability that 3 or more belong to the ruling clique?
   Let \( X \) = the number of members selected from the ruling clique, then we can use Hypergeometric with \( m = 10 \), \( n = 40 \) and \( k = 6 \)
   \[ P(X \geq 3) = 1 - P(X \leq 2) = 1 - \text{phyper}(2,10,40,6) = 0.0856 \]

   b. What are the mean and variance of the number in the sample that belong to the ruling clique?
   Mean = \( 6 \times \frac{10}{50} = 1.2 \)
   Variance = \( 6 \times \frac{10}{50} \times \frac{40}{50} \times \frac{1 - 5/49}{49} = 0.862 \)

   c. What are the answers to (1) and (2) if the club has 50,000 members, 10,000 in the ruling clique and 40,000 second-class members.
   Here, \( m = 10000 \), \( n = 40000 \) and \( k = 6 \)
   \[ P(X \geq 3) = 1 - P(X \leq 2) = 1 - \text{phyper}(2,10000,40000,6) = 0.09887 \]
   Mean = \( 6 \times \frac{10000}{50000} = 1.2 \)
   Variance = \( 6 \times \frac{10000}{50000} \times \frac{40000}{50000} \times \frac{1 - 5/49999}{49999} = 0.9599 \)

Note: \( p = .2 \) if we use binomial, \( P(X \geq 3) = 1 - P(X \leq 2) = 1 - \text{pbinom}(2,6,0.2) = 0.09888 \)
Which is the same answer in part c.
7. Section 4.8.3, problem 5
Huck and Jim are waiting for a raft. The number of rafts floating by over intervals of time is a Poisson process with a rate of \( \lambda = 0.4 \) rafts per day. They agree in advance to let the first raft go and take the second one that comes along. What is the probability that they will have to wait more than a week?

Hint: If they have to wait more than a week, what does that say about the number of rafts in a period of 7 days?

Let \( X = \# \) rafts in a period of 7 days.

The probability that they will have to wait for more than a week is the same as the probability of having no raft or only 1 raft passing by (because they decide to let the 1st raft go) in a period of 7 days.

The parameter \( \mu = 0.4 \times 7 = 2.8 \)

\[
Pr(X = 0) + Pr(X = 1) = e^{-2.8} \times \frac{2.8^0}{0!} + e^{-2.8} \times \frac{2.8^1}{1!} = 0.2311 \text{ or in R studio ppois(1,2.8)}
\]
8. Section 4.8.3
   a. Problem 7
   b. Problem 8

   a. Fire ant colonies occur according to a spatial Poisson process with a rate of 1.5 colonies per acre. What is the probability that a 10 acre plot of land will have 10 or fewer fire ant colonies?
      We are looking at 10 acres, mean = 10(1.5) = 15
      \[ P(X \leq 10) = \text{ppois}(10,15) = 0.1185 \]

   b. What are the mean and variance of the number of colonies on a 10 acre plot?
      \[ \text{Mean} = 15, \text{Variance for Poisson is the same as mean} = 15 \]
9. Of all customers purchasing automatic garage door openers, 75% purchase chain-driven model. Let \( X \) = the number among the next 15 purchasers who select the chain-driven model.

a. What is the frequency function (pmf) of \( X \)?
b. Compute \( P(X > 10) \).
c. Compute \( P(6 \leq X \leq 10) \).
d. Compute \( \mu \) and \( \sigma^2 \).
e. If the store currently has in stock 10 chain-driven models and 8 shaft-driven models, what is the probability that at least 7 out of the 15 customers select a chain-driven model from this stock?

a. This is binomial with \( n = 15, p = 0.75 \), the frequency function is:
   \[
   f(x) = C(15, x)(0.75)^x(1 - 0.75)^{15-x}
   \]
b. \( P(X > 10) = 1 - P(X \leq 10) = 1 - \text{pbinom}(10,15,0.75) = 0.6865 \)
c. \( P(6 \leq X \leq 10) = P(X \leq 10) - P(X \leq 5) = \text{pbinom}(10,15,0.75) - \text{pbinom}(5,15,0.75) = 0.3127 \)
d. \( \mu = 15(0.75) = 11.25, \sigma^2 = 15(0.75)(0.25) = 2.8125 \)
e. This switches to a hypergeometric because the size of the population is only 18. We have \( m = 10, n = 8 \) and \( k = 15 \)
   \( P(X \geq 7) = 1 - P(X \leq 6) = 1 - \text{phyper}(6,10,8,15) = 1 \)
10. Suppose we have a frequency function for two variables $X$ and $Y$, $f(x,y) = \frac{x+y}{30}$, for $x = 0, 1, 2$ and $y = 0, 1, 2, 3$.

a. Determine the marginal distributions of $X$ and $Y$.

b. Determine $E(X)$ and $E(Y)$.

c. Determine $E(X + Y)$.

d. If $Z = 2X + 10$, determine $E(Z)$.

e. Determine $E(XY)$.

f. Determine $\text{cov}(X, Y)$.

g. Are $X$ and $Y$ independent? Justify your answer.

a. Put the function into the following table

<table>
<thead>
<tr>
<th>$X$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0333</td>
<td>0.0667</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>0.0667</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.1333</td>
<td>0.3</td>
</tr>
<tr>
<td>3</td>
<td>0.1333</td>
<td>0.1667</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Marginal distribution of $X$:

<table>
<thead>
<tr>
<th>$X$</th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(X = x)$</td>
<td>0.2</td>
<td>0.3333</td>
<td>0.4667</td>
</tr>
</tbody>
</table>

Marginal distribution of $Y$:

<table>
<thead>
<tr>
<th>$Y$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(Y = y)$</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
</tr>
</tbody>
</table>

b. $E(X) = 0(0.2) + 1(0.3333) + 2(0.4667) = 1.2667$

$E(Y) = 0(0.1) + 1(0.2) + 2(0.3) + 3(0.4) = 2$

c. $E(X + Y) = E(X) + E(Y) = 1.2667 + 2 = 3.2667$

d. $E(Z) = E(2X + 10) = 2(1.2667) + 10 = 12.533$

e. $E(XY) = 0.0667 + 2(0.1) + 2(0.1) + 4(0.1333) + 3(0.1333) + 6(0.1667) = 2.4$

f. $\text{cov}(X,Y) = E(XY) - E(X)E(Y) = 2.4 - (1.2667)(2) = -0.1334$

g. No, for any $x$ and $y$, $f(x,y) \neq f(x)f(y)$.

Proof

$P(X = 0) = 0.2$, $P(Y = 0) = 0.1$, $0.2(0.1) = 0.02 \neq 0$