MATH 343 (SPRING 2021, PHILLIPS): SOLUTIONS TO SELECTED PROBLEMS IN WRITTEN HOMEWORK 7

The statements of the problems are not identical to what is in the book.

Problem 1 (Problem 5.12 in the book). In one toss of a fair coin, we expect heads, give or take heads or so.
   In 100 tosses of a fair coin, we expect heads, give or take heads or so.
   In 400 tosses of a fair coin, we expect heads, give or take heads or so.
   Note that the standard deviation is proportional to the square root of the number of tosses.

Solution. For \( n \) tosses, the expected number of heads is \( \frac{n}{2} \) and the standard deviation is
\[
\sqrt{n \left( \frac{1}{2} \right) \left( 1 - \frac{1}{2} \right)} = \frac{\sqrt{n}}{2}.
\]

For \( n = 1 \) this gives \( \frac{1}{2} \) and \( \frac{\sqrt{1}}{2} = \frac{1}{2} \).

For \( n = 100 \) this gives \( \frac{100}{2} = 50 \) and \( \frac{\sqrt{100}}{2} = 5 \).

For \( n = 400 \) this gives \( \frac{400}{2} = 200 \) and \( \frac{\sqrt{400}}{2} = 10 \).

Therefore:
   In one toss of a fair coin, we expect \( \frac{1}{2} \) heads, give or take \( \frac{1}{2} \) heads or so.
   In 100 tosses of a fair coin, we expect 50 heads, give or take 5 heads or so.
   In 400 tosses of a fair coin, we expect 200 heads, give or take 10 heads or so. \( \square \)

Problem 2 (Problem 5.14 in the book). A multiple choice quiz consists of five questions, each with four possible answers. You randomly guess each answer. What is the probability of getting each of the following?

1. Exactly one right.
2. Exactly three right.
3. More right than wrong.
4. At least one wrong.

Solution. For convenience, let \( X \) be the random variable whose value is the number of correct answers. It is a binomial random variable because the guesses are independent. For each problem, the probability of a correct guess is \( \frac{1}{4} \).

Date: 26 May 2021.
Solution to (1): This is
\[ \Pr(X = 1) = \binom{5}{1} \left( \frac{1}{4} \right)^1 \left( 1 - \frac{1}{4} \right)^4 = 5 \left( \frac{3^4}{4^5} \right) = \frac{405}{1024}. \]
(This is approximately 0.395508.)

Solution to (2): This is
\[ \Pr(X = 3) = \binom{5}{3} \left( \frac{1}{4} \right)^3 \left( 1 - \frac{1}{4} \right)^2 = \frac{5 \cdot 4 \cdot 3 \cdot 3^2}{1 \cdot 2 \cdot 3 \cdot 4^5} = \frac{90}{1024} = \frac{45}{512}. \]
(This is approximately 0.0878906.)

Solution to (3): Using the result from Part (1) at the second step, this is
\[ \Pr(X \geq 3) = \Pr(X = 3) + \Pr(X = 4) + \Pr(X = 5) \]
\[ = \frac{90}{1024} + \frac{5 \cdot 4 \cdot 3 \cdot 3^2}{1 \cdot 2 \cdot 3 \cdot 4^5} + \left( \frac{1}{4} \right)^5 \]
\[ = \frac{90 + 5 \cdot 3 + 1}{1024} = \frac{53}{512}. \]
(This is 0.103515625.)

Solution to (4): This is
\[ \Pr(X \leq 4) = 1 - \Pr(X = 5) = 1 - \left( \frac{1}{4} \right)^5 = \frac{1023}{1024}. \]
(This is approximately 0.999023.)

\[ \square \]

Problem 3 (Problem 7.42 in the book). Sixteen ounce boxes of shredded wheat cereal packaged automatically by machine are sometimes overweight and sometimes underweight. The actual weight in ounces over or under 16 is a random variable \(X\) whose probability density function is
\[ f_X(x) = \begin{cases} 
\frac{3}{32} (4 - x^2) & -2 < x < 2 \\
0 & \text{otherwise.}
\end{cases} \]

Negative values of \(x\) correspond to ounces under 16. Find the probability of each of the following.

1. A box of cereal is more than one ounce underweight.
2. A box of cereal is neither underweight nor more than one ounce overweight.
3. A box of cereal is exactly 16 ounces.

Solution. Solution to (1):
\[ \Pr(X < -1) = \int_{-\infty}^{-1} f_X(x) \, dx = \int_{-2}^{-1} \frac{3}{32} (4 - x^2) \, dx = \left. \frac{3}{32} \left( 4x - \frac{x^3}{3} \right) \right|_{-2}^{-1} = \frac{5}{32}. \]
(This is 0.15625.)

Solution to (2):
\[ \Pr(0 \leq X \leq 1) = \int_{0}^{1} f_X(x) \, dx = \int_{0}^{1} \frac{3}{32} (4 - x^2) \, dx = \left. \frac{3}{32} \left( 4x - \frac{x^3}{3} \right) \right|_{0}^{1} = \frac{11}{32}. \]
(This is 0.34375.)

Solution to (3): \( \Pr(X = 16) = 0 \) because \( X \) is a continuous random variable; \( \Pr(X = x) = 0 \) for any real number \( x \).
Problem 4 (Problem 8.2 in the book). Consider a normal population with $\mu = 1.6$ and $\sigma = 2.1$. Convert the following values to $z$-scores.

1. $3.7$.
2. $1.6$.
3. $-1.3$.
4. $1.8$.
5. $0$.

Solution. Recall that for a given value $x$, if the mean is $\mu$ and the standard deviation is $\sigma$, then the corresponding $z$-score is

$$z = \frac{x - \mu}{\sigma}.$$  

Solution to (1):

$$z = \frac{3.7 - 1.6}{2.1} = 1.$$

Solution to (2):

$$z = \frac{1.6 - 1.6}{2.1} = 0.$$

Solution to (3):

$$z = \frac{-1.3 - 1.6}{2.1} \approx -1.380952.$$

Solution to (4):

$$z = \frac{1.8 - 1.6}{2.1} \approx 0.0952381.$$

Solution to (5):

$$z = \frac{0 - 1.6}{2.1} \approx 0.7619048.$$

□

Problem 5 (Problem 8.13 in the book). Serum cholesterol levels in 17 year olds have a normal distribution with mean 176 mg/dl and standard deviation 30 mg/dl. What percentage of 17 year olds have the following cholesterol levels?

1. Below 166 mg/dl.
2. Above 166 mg/dl.
3. Below 186 mg/dl.
4. Above 220 mg/dl.
5. Above 260 mg/dl.

Solution. Let $X$ be the random variable whose value is the serum cholesterol level of a randomly chosen 17 year old. Recall that for a given value $x$, if the mean is $\mu$ and the standard deviation is $\sigma$, then the corresponding $z$-score is

$$z = \frac{x - \mu}{\sigma}.$$  

Let $Z$ be a standard normal random variable. Then $\Pr(X \leq x) = \Pr(Z \leq z)$.

Solution to (1): We have

$$z = \frac{166 - 176}{30} \approx -0.333333.$$  

Using the table inside the front cover of the book for $z = -0.33$, I got

$$\Pr(X < 166) \approx \Pr(Z < 0.33) \approx 0.37070.$$

A more accurate value, obtained from an online calculator, is

\[ \Pr(X < 166) \approx \Pr(Z < -0.3333333333333333) \approx 0.369. \]

Therefore about 37% (more accurately, about 36.9%) of 17 year olds have serum cholesterol level below 166 mg/dl.

Solution to (2): Using the result from Part (1), about \([100 - 37] = 63\%\) (more accurately, about 63.1%) of 17 year olds have serum cholesterol level above 166 mg/dl.

Solution to (3): Since a normal distribution is symmetric about its mean, this has the same answer as Part (2): about 63\% (more accurately, about 63.1\%) of 17 year olds have serum cholesterol level below 186 mg/dl.

Solution to (4): We have

\[ z = \frac{220 - 176}{30} \approx 1.466667. \]

Using the table inside the front cover of the book for \(z = 1.47\), I got

\[ \Pr(X > 220) = 1 - \Pr(X < 220) \approx 1 - \Pr(Z < 1.47) \approx 1 - 0.92922 = 0.07078. \]

Using an online calculator which gives only three digits, I got

\[ \Pr(X > 220) = 1 - \Pr(X < 220) \approx 1 - \Pr(Z < 1.4666666666666667) \]

\[ \approx 1 - 0.929 = 0.071. \]

Therefore about 7.1% of 17 year olds have serum cholesterol level above 220 mg/dl.

Solution to (5): We have

\[ z = \frac{260 - 176}{30} = 2.8. \]

Using the table inside the front cover of the book for \(z = 2.8\), I got

\[ \Pr(X > 260) = 1 - \Pr(X < 260) = 1 - \Pr(Z < 2.8) \approx 1 - 0.99744 = 0.00256. \]

Therefore about 0.256% of 17 year olds have serum cholesterol level above 260 mg/dl. (The calculator website I used gives only three digits, so is no improvement here.)

\(\square\)

**Problem 6** (Problem 8.24 in the book). You roll a fair die 120 times. Use the normal approximation to approximate the probability of rolling a 1 exactly 20 times.

**Solution.** Let \(Z\) be a standard normal random variable.

We must use the continuity correction, since otherwise we get the answer zero. Let \(X\) count the number of 1’s. Its mean and standard deviation are

\[ E(X) = \frac{120}{6} = 20 \quad \text{and} \quad \sigma_X = \sqrt{120 \left( \frac{1}{6} \right) \left( 1 - \frac{1}{6} \right)} = \frac{5\sqrt{6}}{3} \approx 4.082483. \]

This gives \(z\)-scores for 19.5 and 20.5 of approximately

\[ -0.122474 \quad \text{and} \quad 0.122474. \]

Using the table inside the front cover of the book for \(z = 0.12\) and \(z = -0.12\), my approximation is

\[ \Pr(Z < 0.12) - \Pr(Z < -0.12) \approx 0.54776 - 0.45224 = 0.09552. \]
A more accurate value, obtained from an online calculator, is
\[
\Pr(\frac{Z}{0.1224744871391589} < -0.1224744871391589) - \Pr(\frac{Z}{-0.1224744871391589} < -0.549 - 0.451 = 0.098.
\]
(The calculator website I used gives only three digits.)

\[\square\]

**Problem 7** (Problem 8.25 in the book). Use the normal approximation to approximate the probability of getting heads 8, 9, or 10 times in 15 tosses of a fair coin. Then use the binomial distribution to find the exact probability.

**Solution.** We do the normal approximation first. The mean is \(\frac{15}{2} = 7.5\) and the standard deviation is
\[
\sqrt{15 \left( \frac{1}{2} \right) \left( \frac{1}{2} \right)} = \frac{\sqrt{15}}{2}.
\]
(This is about 1.93649.)

Let \(Z\) be a standard normal random variable. We need \(\Pr(z_1 < Z < z_2)\) with
\[
z_1 = \frac{7.5 - 7.5}{\sqrt{15}/2} = 0 \quad \text{and} \quad z_2 = \frac{10.5 - 7.5}{\sqrt{15}/2} \approx 1.54919.
\]
Using an online calculator, I got
\[
\Pr(z_1 < Z < z_2) \approx 0.939 - 0.5 = 0.439.
\]
Using the binomial distribution, we get
\[
\left( \frac{1}{2} \right)^{15} \left[ \binom{15}{8} + \binom{15}{9} + \binom{15}{10} \right] \approx 0.440765.
\]
The normal approximation is thus quite close.

\[\square\]

**Problem 8** (Problem 8.36 in the book). Breakfast food packages are listed as containing 12 ounces of cereal. The filling machine is subject to errors that follow a normal curve with a standard deviation of 0.1 ounces.

1. If the machine is set to fill a package with 12.1 ounces of cereal, what percentage of packages will be at least 0.1 ounces short of the listed 12 ounce weight?

2. If the producer wishes at most 2% of the packages to have a shortage of 0.1 ounces or more (that is, weight under 11.9 ounces), what mean filling weight should be used?

**Solution.** Solution to (1): The \(z\)-score is
\[
\frac{11.9 - 12.1}{0.1} = -2.
\]
Let \(Z\) be a standard normal random variable. The table inside the front cover of the book gives \(\Pr(Z < -2) \approx 0.02275\), so about 2.275% of packages will be less than 11.9 ounces.

Solution to (2): Let \(Z\) be a standard normal random variable. We need to find \(z\) such that \(\Pr(Z < z) = 0.02\). The mean filling weight \(\mu\) should then satisfy
\[
\frac{11.9 - \mu}{0.1} = z,
\]
that is, \( \mu = 11.9 - (0.1)z \). An online calculator gave \( z \approx -2.054 \), so the filling weight should be about \( 11.9 + (0.1)(2.054) = 12.1054 \) ounces.  

**Problem 9** (Problem 10.2 in the book). Consider discrete random variables \( X \) and \( Y \) whose joint probability density function is 

\[
 f_{X,Y}(x,y) = \begin{cases} 
 \frac{x+y}{30} & x = 0, 1, 2, 3 \text{ and } y = 0, 1, 2 \\
 0 & \text{otherwise} 
\end{cases}
\]

(1) Find \( \Pr(X = 1 \text{ and } Y = 2) \).
(2) Find \( \Pr(X + Y < 2) \).
(3) Find \( \Pr(XY \leq 2) \).

**Solution.** Solution to (1):

\[
 \Pr(X = 1 \text{ and } Y = 2) = f_{X,Y}(1, 2) = \frac{1+2}{30} = \frac{1}{10}.
\]

Solution to (2): The pairs \((x, y)\) for which \( f_{X,Y}(x, y) \) might be nonzero and \( x + y < 2 \) are \((0, 0), (0, 1), \) and \((1, 0)\). Therefore

\[
 \Pr(X + Y < 2) = f_{X,Y}(0, 0) + f_{X,Y}(0, 1) + f_{X,Y}(1, 0) = 0 + \frac{1}{30} + \frac{1}{30} = \frac{1}{15}.
\]

(This is approximately 0.0666667.)

Solution to (3): We find \( 1 - \Pr(XY > 2) \). The pairs \((x, y)\) for which \( f_{X,Y}(x, y) \) might be nonzero and with \( xy > 2 \) are \((2, 2), (3, 1), \) and \((3, 2)\). Therefore

\[
 \Pr(XY > 2) = \frac{4}{30} + \frac{4}{30} + \frac{5}{30} = \frac{13}{30},
\]

so

\[
 \Pr(XY \leq 2) = 1 - \Pr(XY > 2) = 1 - \frac{13}{30} = \frac{17}{30}.
\]

(This is approximately 0.5666667.)

Alternate solution to (3) (not recommended): The pairs \((x, y)\) for which \( f_{X,Y}(x, y) \) might be nonzero and with \( xy \leq 2 \) are \((0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), \) and \((3, 0)\). Therefore

\[
 \Pr(XY \leq 2) = f_{X,Y}(0, 0) + f_{X,Y}(0, 1) + f_{X,Y}(0, 2) + f_{X,Y}(1, 0) + f_{X,Y}(1, 1) + f_{X,Y}(1, 2) + f_{X,Y}(2, 0) + f_{X,Y}(2, 1) + f_{X,Y}(3, 0)
\]

\[
 = \frac{0}{30} + \frac{1}{30} + \frac{2}{30} + \frac{1}{30} + \frac{2}{30} + \frac{3}{30} + \frac{2}{30} + \frac{3}{30} + \frac{3}{30} = \frac{17}{30}.
\]

\(\square\)

**Problem 10** (Problem 10.3 in the book). The joint probability mass function of two discrete random variables \( X \) and \( Y \) is given by the following table:

<table>
<thead>
<tr>
<th>( f_{X,Y}(x, y) )</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1/24</td>
<td>3/24</td>
<td>1/24</td>
<td>1/24</td>
</tr>
<tr>
<td>1</td>
<td>2/24</td>
<td>2/24</td>
<td>6/24</td>
<td>2/24</td>
</tr>
<tr>
<td>2</td>
<td>2/24</td>
<td>1/24</td>
<td>2/24</td>
<td>1/24</td>
</tr>
</tbody>
</table>

\(\square\)
(1) Find $\Pr(X = 1 \mid Y = 3)$.

(2) Are $X$ and $Y$ independent? Why or why not?

**Solution.** For convenience, we reproduce the table with the marginal probabilities added. (Most of the extra information is not actually needed.)

<table>
<thead>
<tr>
<th>$f_{X,Y}(x,y)$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>$f_X(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1/24</td>
<td>3/24</td>
<td>1/24</td>
<td>1/24</td>
<td>1/4</td>
</tr>
<tr>
<td>1</td>
<td>2/24</td>
<td>2/24</td>
<td>6/24</td>
<td>2/24</td>
<td>1/2</td>
</tr>
<tr>
<td>2</td>
<td>2/24</td>
<td>1/24</td>
<td>2/24</td>
<td>1/24</td>
<td>1/4</td>
</tr>
</tbody>
</table>

| $f_Y(y)$      | 5/24| 6/24| 9/24| 4/24 |

Solution to (1):

$$\Pr(X = 1 \mid Y = 3) = \frac{\Pr(X = 1 \text{ and } Y = 3)}{\Pr(Y = 3)} = \frac{f_{X,Y}(1,3)}{f_Y(3)} = \frac{\frac{1}{24}}{\frac{4}{24} + \frac{1}{24} + \frac{1}{24}} = \frac{1}{3}.$$  

Solution to (2): The random variables $X$ and $Y$ are not independent. Given Part (1), the shortest proof to write is the computation

$$\Pr(X = 1) = f_X(1) = \frac{2}{24} + \frac{6}{24} + \frac{2}{24} = \frac{1}{2},$$

from which we see that $\Pr(X = 1 \mid Y = 3) \neq \Pr(X = 1)$. If $X$ and $Y$ are independent, these are supposed to be equal.

As alternate solutions, one can observe many pairs $(x,y)$ such that $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$. For example,

$$f_X(0)f_Y(2) = \left(\frac{1}{4}\right)\left(\frac{5}{24}\right) = \frac{5}{96} \neq \frac{1}{24} = f_{X,Y}(0,2).$$

In fact, $f_{X,Y}(x,y) \neq f_X(x)f_Y(y)$ for most pairs $(x,y)$ occurring in the table (although not for all of them).  

**Problem 11** (Problem 10.8 in the book). Consider discrete random variables $X$ and $Y$ whose joint probability mass function is

$$f_{X,Y}(x,y) = \begin{cases} 
\frac{x+y}{30} & x = 0,1,2,3 \text{ and } y = 0,1,2 \\
0 & \text{otherwise.}
\end{cases}$$

(1) Compute Cov$(X,Y)$.

(2) Compute the mean of $X - Y$.

(3) Compute the variance of $X - Y$.

**Solution.** It is most convenient to do this problem from a table of the joint probability mass function, including the marginal probabilities. Here it is.
We now find $E(X)$, $E(X^2)$, $E(Y)$, $E(Y^2)$, and $E(XY)$. For all but the last, we only need the marginal probabilities, which shortens the work. We get:

$$E(X) = \sum_{x=0}^{3} x f_X(x) = 2$$

$$E(X^2) = \sum_{x=0}^{3} x^2 f_X(x) = 5$$

$$E(Y) = \sum_{y=0}^{4} y f_Y(y) = \frac{19}{15}$$

$$E(Y^2) = \sum_{y=0}^{4} y^2 f_Y(y) = \frac{5 + 4 \cdot 7}{15} = \frac{11}{5}$$

$$E(XY) = \sum_{x=0}^{3} \sum_{y=0}^{4} xy f_{X,Y}(x,y) = \frac{12}{5}.$$

(This gives $E(Y) \approx 1.266667$.)

Solution to (1): We calculate

$$\text{Cov}(X,Y) = E(XY) - E(X)E(Y) = \frac{12}{5} - 2 \left( \frac{19}{15} \right) = -\frac{2}{15}.$$

(This is approximately $-0.133333$.)

Solution to (2): We calculate

$$E(X - Y) = E(X) - E(Y) = 2 - \frac{19}{15} = \frac{11}{15}.$$

(This is approximately $-0.733333$.)
Solution to (3): We calculate
\[
\text{Var}(X - Y) = E((X - Y)^2) - (EX)^2 - E(X^2) - 2E(XY) + E(Y^2) - E(X - Y)^2
\]
\[
= 5 - 2 \left( \frac{12}{5} + \frac{11}{5} - \left( \frac{11}{15} \right)^2 \right)
\]
\[
= \frac{12}{5} - \frac{121}{225} = \frac{419}{225}.
\]
(This is approximately 1.862222.)

There are other ways to shorten the computation of \(\text{Var}(X - Y)\). For example, one can compute \(\text{Var}(X)\) and \(\text{Var}(Y)\) from the results already obtained, and then use the formula (on page 286 of the book)
\[
\text{Var}(X - Y) = \text{Var}(X) - 2\text{Cov}(X,Y) + \text{Var}(Y).
\]
\[
\square
\]

Problem 12 (Problem A). One day, an exotic pet store had in stock 16 spiders, of which 11 were Carolina wolf spiders and 5 were tarantulas. Customers that day bought 4 tarantulas and 2 Carolina wolf spiders. It is suspected that customers are more likely to buy tarantulas than Carolina wolf spiders. Test this hypothesis at the \(\alpha = 0.05\) level of significance, using both the binomial exact test and Fisher’s exact test. Assume that the customers on the day in question can be treated as a simple random sample of all customers.

Be sure to do the following:

(1) Define every variable etc. used.
(2) State the hypotheses, both in ordinary language in terms of the problem situation and mathematically (for example, from a different context, \(H_0: \mu = 28.5\)). In the ordinary language version, make sure that it is clear that the hypotheses are about a population, not about a sample.
(3) For both tests, compute the \(P\)-value exactly as a fraction with integer numerator and denominator, showing work for the computation, before evaluating it numerically. In particular, if it is a sum of several terms, show those terms. You need not simplify the numerator and denominator of your fraction.
(4) Include a statement of the conclusion in ordinary language in terms of the problem situation.

This information is included in solutions to problems in a previous homework set.

Solution. The hypotheses are:

\(H_0\): Customers are equally likely to buy tarantulas and Carolina wolf spiders.
\(H_a\): Customers are more likely to buy tarantulas than Carolina wolf spiders.

(It is not correct for the hypotheses to say anything about the relative numbers of the two kinds of spiders bought on the day in question. That is a statement about the sample.)

We did not start with a pool of 16 spiders and randomly assign 5 of them to be tarantulas and 11 of them to be Carolina wolf spiders. So there is no justification for a priori taking the probability of being in the first group to be \(\frac{1}{2}\). Instead,
we take the probability of being in the first group to be determined by the actual numbers in the groups, which gives

$$\frac{5}{11 + 5} = \frac{5}{16}.$$ 

Assuming $H_0$, the probability $\pi$ that each of the $6 = 2 + 4$ purchased spiders was a tarantula should have been $\frac{5}{16}$. Therefore the hypotheses can be reformulated as:

- $H_0: \pi = \frac{5}{16}$.
- $H_a: \pi > \frac{5}{16}$.

The test statistic is 4. At least a larger proportion than expected, namely $\frac{4}{6} > \frac{5}{16}$, of the purchased spiders were tarantulas. (Otherwise, the evidence would point towards Carolina wolf spiders being more likely to be purchased, and there would be no point in further work.)

For the binomial exact test, the $P$-value is therefore the probability that, in 6 tosses of a coin biased to give tails $\frac{5}{16}$ of the time, one sees tails 4 or more times. (This is the probability of seeing an outcome as extreme as or more extreme than what was actually observed.) This probability is

$$\begin{align*}
\binom{6}{4} \left( \frac{5}{16} \right)^4 \left( \frac{11}{16} \right)^2 + \binom{6}{5} \left( \frac{5}{16} \right)^5 \left( \frac{11}{16} \right)^1 + \left( \frac{5}{16} \right)^6 \\
= \frac{5^4}{16^6} \left[ 15 \cdot 11^2 + 6 \cdot 11 \cdot 5 + 5^2 \right] \\
= \frac{5^4 \cdot 2170}{16^6} = \frac{5^5 \cdot 434}{16^6} = \frac{1356250}{16777216} \approx 0.0808388.
\end{align*}$$

(You need not multiply out the numerator and denominator to get $\frac{1356250}{16777216}$; this expression is provided so those who did can check their work.) This number is greater than $\alpha = 0.05$. We therefore fail to reject the null hypothesis. We do not find convincing evidence (at the significance level $\alpha = 0.05$) that customers are more likely to buy tarantulas than Carolina wolf spiders.

A solution need not contain the fair coin analogy or the comment on being “as extreme as or more extreme than” the observed outcome; these are included for explanation. However, the null and alternate hypotheses must be explicitly stated, and the test statistic and $P$-value must be explicitly calculated and labelled. It must be clear that the null and alternate hypotheses are about the population, not about the sample. Moreover, the plain language version of the conclusion is an essential part of a full solution.

For Fisher’s exact test, the hypotheses

- $H_0: \pi = \frac{5}{16}$.
- $H_a: \pi > \frac{11}{16}$.

and the the test statistic 4 are the same as for the binomial exact test, as is the justification for proceeding with the test.

The $P$-value is the probability that, if 6 cards are drawn without replacement from a deck containing 5 red cards and 11 black cards, 4, 5, or 6 of them are red. (This is the probability of seeing an outcome as extreme as or more extreme than
what was actually observed.) It isn’t possible to draw 6 red cards, so we need only consider 4 or 5 red cards. Also,
\[
\binom{16}{6} = \frac{16!}{6! \cdot 10!} = \frac{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6} = 2 \cdot 14 \cdot 13 \cdot 2 \cdot 11.
\]
The \(P\)-value is thus
\[
\binom{16}{6}^{-1} \left[ \left( \binom{5}{5} \right) \left( \binom{11}{1} \right) + \left( \binom{5}{4} \right) \left( \binom{11}{2} \right) \right] = \frac{11 + 55 \cdot 5}{2 \cdot 14 \cdot 13 \cdot 2 \cdot 11} = \frac{1}{28} \approx 0.0357143.
\]
We therefore reject the null hypothesis. We find convincing evidence (at the significance level \(\alpha = 0.05\)) that customers are more likely to buy tarantulas than Carolina wolf spiders.

A solution need not contain the card drawing analogy or the comment on being “as extreme as or more extreme than” the observed outcome; these are included for explanation. However, the null and alternate hypotheses must be explicitly stated, and the test statistic and \(P\)-value must be explicitly calculated and labelled. It must be clear that the null and alternate hypotheses are about a population, not about the sample. Moreover, the plain language version of the conclusion is an essential part of a full solution. \(\square\)