Solutions to Homework 10
Statistics 302 Professor Larget

Textbook Exercises

7.14 Rock-Paper-Scissors (Graded for Accurateness) In Data 6.1 on page 367 we see a table, reproduced in the table below that shows the choices made by 119 players on the first turn of a Rock-Paper-Scissors game. Recall that rock beats scissors which beats paper which beats rock. A player gains an advantage in playing this game if there is evidence that the choices made on the first turn are not equally distributed among the three options. Use a goodness-of-fit test to see if there is evidence that any of the proportions are different from 1/3.

<table>
<thead>
<tr>
<th>Option Selected</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rock</td>
<td>66</td>
</tr>
<tr>
<td>Paper</td>
<td>39</td>
</tr>
<tr>
<td>Scissors</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>119</td>
</tr>
</tbody>
</table>

Solution
If we let $p_r$, $p_p$, and $p_s$ represent the proportions of rock, paper and scissor choices, the hypotheses are

$$H_0 : p_r = p_p = p_s = 1/3$$

$$H_a : \text{Some } p_i \neq 1/3$$

The expected count is $119(1/3) = 39.7$ for each cell. The chi-square statistic is

$$\chi^2 = \frac{(66 - 39.7)^2}{39.7} + \frac{(39 - 39.7)^2}{39.7} + \frac{(14 - 39.7)^2}{39.7}$$

$$= 17.4 + 0.01 + 16.6$$

$$= 34.01$$

The test statistic, $\chi^2 = 34.01$, lies very far in the tail of a chi-square distribution with 2 degrees of freedom, so the p-value is very close to zero. This gives strong evidence that the choices made on the first turn of a rock-paper scissors game are not all equally likely. Comparing the expected counts to the observed counts it appears that “rock” is used more often and “scissors” is less frequent than expected. Unless your opponent has also looked at this study, it might be smart to start with paper.

7.17 Birth Data and Canadian Ice Hockey (Graded for Accurateness) In his book *Outliers: The Story of Success* (2008), Malcolm Gladwell speculates that Canadian ice hockey players that are born early in the year have an advantage. This is because the birthdate cutoff for different levels of youth hockey leagues in Canada is January 1st, so youth hockey players who are born in January and February are slightly older than teammates born later in the year. Does this slight age advantage in the beginning lead to success later on? A 2010 study examined the birthdate distribution of players in the Ontario Hockey League (OHL), a high-level and selective Canadian hockey league (ages 15-20), for the 2008-2009 season. The number of OHL players born during the 1st quarter (Jan-Mar), 2nd quarter, (Apr-Jun), 3rd quarter (Jul-Sep), and 4th quarter (Oct-Dec) of the year is shown in the table below. The overall percentage of live births in Canada (year 1989) are also provided for each quarter. Is this evidence that the birthdate distribution for OHL players differs significantly from the national proportions? State the null and alternative hypotheses, calculate the chi-square statistic, find the p-value, and state the conclusion in context.
<table>
<thead>
<tr>
<th></th>
<th>Qtr 1</th>
<th>Qtr 2</th>
<th>Qtr 3</th>
<th>Qtr 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>OHL players</td>
<td>147</td>
<td>110</td>
<td>52</td>
<td>50</td>
</tr>
<tr>
<td>% of Canadian births</td>
<td>23.7%</td>
<td>25.9%</td>
<td>25.9%</td>
<td>24.5%</td>
</tr>
</tbody>
</table>

Solution
Let $p_1, p_2, p_3,$ and $p_4$ be the proportion of hockey players born in the 1st, 2nd, 3rd, and 4th quarter of the year, respectively. We are testing

$$H_0 : p_1 = 0.237, p_2 = 0.259, p_3 = 0.259 \text{ and } p_4 = 0.245$$

$$H_a : \text{Some } p_i \text{ is not specified as in } H_0$$

The total sample size is $n = 147 + 110 + 52 + 50 = 359$. The expected counts are $359(0.237) = 85$ for Qtr 1, $359(0.259) = 93$ for Qtr 2, $359(0.259) = 93$ for Qtr 3, and $359(0.245) = 88$ for Qtr 4. The chi-square statistic is

$$\chi^2 = \frac{(147 - 85)^2}{85} + \frac{(110 - 93)^2}{93} + \frac{(52 - 93)^2}{93} + \frac{(50 - 88)^2}{88} = 82.6.$$

We use the chi-square distribution with $4 - 1 = 3$ degrees of freedom, which gives a very small p-value that is essentially zero. This is strong evidence that the distribution of birth dates for OHL hockey players differs significantly from the national proportions.

7.20 Can People Delay Death? (Graded for Completeness) A new study indicates that elderly people are able to postpone death for a short time to reach an important occasion. The researchers studied deaths from natural causes among 1200 elderly people of Chinese descent in California during six months before and after the Harbor Moon Festival. Thirty-three deaths occurred in the week before the Chinese festival, compared with an estimated 50.82 deaths expected in that period. In the week following the festival, 70 deaths occurred, compared with an estimated 52. “The numbers are so significant that is would be unlikely to occur by chance,” said one of the researchers.

(a) Given the information in the problem, is the $\chi^2$ statistic likely to be relatively large or relatively small?

(b) Is the p-value likely to be relatively large or relatively small?

(c) In the week before the festival, which is higher: the observed count or the expected count? What does this tell us about the ability of elderly people to delay death?

(d) What is the contribution to the $\chi^2$-statistic for the week before the festival?

(e) In the week after the festival, which is higher: the observed count or the expected count? What does this tell us about the ability of elder people to delay death?

(f) What is the contribution to the $\chi^2$-statistic for the week after the festival?

(g) The researchers tell us that in a control group of elderly people in California who are not of Chinese descent, the same effect was not seen. Why did the researchers also include a control group?
Solution
(a) Since the results are given as statistically significant, the $\chi^2$-statistic is likely to be large.

(b) Since the results are given as statistically significant, the p-value is likely to be small.

(c) In the week before the festival, the expected count is higher than the observed count. This tells us that some elderly people may be able to delay death.

(d) For the week before the festival, the contribution to the $\chi^2$ statistic is

$$\frac{(\text{observed} - \text{expected})^2}{\text{expected}} = \frac{(33?50.82)^2}{50.82} = 6.249$$

(e) In the week after the festival, the observed count is higher than the expected count. This tells us that, although some elderly people are able to delay death, they don’t delay it for very long.

(f) For the week after the festival, the contribution to the $\chi^2$ statistic is

$$\frac{(\text{observed} - \text{expected})^2}{\text{expected}} = \frac{(70 - 52)^2}{52} = 6.231$$

(g) The control group allows us to attribute the difference specifically to the meaningful event (the Harbor Moon Festival) since the effect was only seen in the group who found this event meaningful.

7.40 Metal Tags on Penguins (Graded for Completeness) In Exercise 6.178 on page 403 we perform a test for the difference in the proportion of penguins who survive over a 10-year period, between penguins tagged with metal tags and those tagged with electronic tags. We are interested in testing whether the type of tag has an effect on penguin survival rate, this time using a chi-square test. In the study, 33 of the 167 metal-tagged penguins survived while 68 of the 189 electronic-tagged penguins survived.

(a) Create a two-way table from the information given.

(b) State the null and alternative hypotheses.

(c) Give a table with the expected counts for each of the four categories.

(d) Calculate the chi-square test statistic.

(e) Determine the p-value and state the conclusion.

Solution
(a) The two-way table of penguin survival vs type tag is shown:

<table>
<thead>
<tr>
<th></th>
<th>Metal</th>
<th>Electronic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survived</td>
<td>33</td>
<td>68</td>
<td>101</td>
</tr>
<tr>
<td>Died</td>
<td>134</td>
<td>121</td>
<td>255</td>
</tr>
<tr>
<td>Total</td>
<td>167</td>
<td>189</td>
<td>356</td>
</tr>
</tbody>
</table>
(b) The hypothesis are

\[ H_0: \text{Type of tag is not related to survival} \]
\[ H_a: \text{Type of tag is related to survival} \]

(c) The table below shows the expected counts, obtained for each cell by multiplying the row total by the column total and dividing by \( n = 356 \).

<table>
<thead>
<tr>
<th></th>
<th>Metal</th>
<th>Electronic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survived</td>
<td>47.4</td>
<td>53.6</td>
</tr>
<tr>
<td>Died</td>
<td>119.6</td>
<td>135.4</td>
</tr>
</tbody>
</table>

d) We calculate the chi-square test statistic

\[
\chi^2 = \frac{(33 - 47.4)^2}{47.4} + \frac{(68 - 53.6)^2}{53.6} + \frac{(134 - 119.6)^2}{119.6} + \frac{(121 - 135.4)^2}{135.4} \\
= 4.37 + 3.87 + 1.73 + 1.53 \\
= 11.5
\]

(e) We compare our test statistic of 11.5 from part (c) to a chi-square with 1 degree of freedom to get a p-value of 0.0007. There is strong evidence that the type of tag and survival rate of the penguins are related.

7.43 Favorite Skittles Flavor? (Graded for Accurateness) Exercise 7.13 on page 472 discusses a sample of people choosing their favorite Skittle flavor by color (green, orange, purple, red, or yellow). A separate poll sampled 91 people, again asking them their favorite skittle flavor, but rather than by color they asked by the actual flavor (lime, orange, grape, strawberry, and lemon, respectively). The table below shows the results from both polls. Does the way people choose their favorite Skittle type, by color or flavor, appear to be related to which type is chosen?

(a) State the null and alternative hypotheses.

(b) Give a table with the expected counts for each of the 10 cells.

(c) Are the expected counts large enough for a chi-square test?

(d) How many degrees of freedom do we have for this test?

(e) Calculate the chi-square test statistic.

(f) Determine the p-value. Do we find evidence that method of choice affects which is chosen?

<table>
<thead>
<tr>
<th>Color (Lime)</th>
<th>Orange</th>
<th>Purple (Grape)</th>
<th>Red (Strawberry)</th>
<th>Yellow (Lemon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>18</td>
<td>9</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Flavor</td>
<td>13</td>
<td>16</td>
<td>19</td>
<td>34</td>
</tr>
</tbody>
</table>

Solution

(a) The hypotheses are

\[ H_0: \text{Skittles choice does not depend on method of choosing (color vs flavor)} \]
(b) The expected counts under $H_0$ are

<table>
<thead>
<tr>
<th></th>
<th>Green (Lime)</th>
<th>Orange</th>
<th>Purple (Grape)</th>
<th>Red (Strawberry)</th>
<th>Yellow (Lemon)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Color</td>
<td>13.0</td>
<td>10.5</td>
<td>14.3</td>
<td>19.8</td>
<td>8.4</td>
</tr>
<tr>
<td>Flavor</td>
<td>18.0</td>
<td>14.5</td>
<td>19.7</td>
<td>27.2</td>
<td>11.6</td>
</tr>
</tbody>
</table>

(c) All of the expected counts are larger than 5, so we can use a chi-square test.

(d) We have $(5 - 1)(2 - 1) = 4$ degrees of freedom.

(e) The chi-square statistic is

$$\chi^2 = \frac{(18 - 13.0)^2}{13.0} + \frac{(9 - 10.5)^2}{10.5} + \cdots + \frac{(9 - 11.6)^2}{11.6} = 9.07$$

(f) Comparing our test statistic to a chi-square distribution with 4 degrees of freedom we get a p-value of 0.059. This is right on the border, so we see weak evidence that choosing flavor vs color might affect the choices, but not enough to reject the null hypothesis if we are using a 5% level.

7.47 Gender and Frequency of Status Updates on Facebook (Graded for Completeness)

Exercise 7.45 introduced a 2010 study about users of social networking sites such as Facebook. The table in the book shows the self-reported frequency of status updates of Facebook by gender. Are frequency of status updates and gender related? Show all details of the test.

Solution

The null hypothesis is that frequency of status updates on Facebook is not different for males and females. The alternative hypothesis is that frequency of status updates is related to gender. We compute expected counts for all 10 cells. For example, for the males who update their status every day, we have

$$\text{Expected} = \frac{130 \cdot 386}{877} = 57.2.$$ 

Computing all the expected counts in the same way, we find the expected counts shown in the following table.

| Status \ 
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Every day</td>
<td>57.2</td>
<td>72.8</td>
</tr>
<tr>
<td>3-5 days/week</td>
<td>46.2</td>
<td>58.8</td>
</tr>
<tr>
<td>1-2 days/week</td>
<td>65.6</td>
<td>83.4</td>
</tr>
<tr>
<td>Every few weeks</td>
<td>68.7</td>
<td>87.3</td>
</tr>
<tr>
<td>Less often</td>
<td>148.3</td>
<td>188.7</td>
</tr>
</tbody>
</table>

Notice that all expected counts are over 5 so we can use a chi-square distribution. We compute the contribution to the $\chi^2$-statistic, $(observed - expected)^2/expected$, for each cell and show the
results in the next table.

<table>
<thead>
<tr>
<th>Status</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Every day</td>
<td>4.05</td>
<td>3.18</td>
</tr>
<tr>
<td>3-5 days/week</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>1-2 days/week</td>
<td>0.33</td>
<td>0.23</td>
</tr>
<tr>
<td>Every few weeks</td>
<td>1.01</td>
<td>0.80</td>
</tr>
<tr>
<td>Less often</td>
<td>0.05</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Adding up all of these contributions, we obtain the $\chi^2$-statistic 9.65. Using a chi-square distribution with $df = (5 - 1) \cdot (2 - 1) = 4$, we see that the p-value is 0.047. This is very close to a 5% cutoff. At a 5% level, there is some evidence of a possible relationship between gender and frequency of status updates on Facebook. Females appear more likely to update their status at least once a day.

8.17 Laptop Computers and Sperm Count (Graded for Completeness) Studies have shown that heating the scrotum by just 1°C can reduce sperm count and sperm quality, with long-term consequences. Exercise 2.101 on page 87 introduces a study indicating that males sitting with a laptop on their laps have increased scrotal temperatures. Does a lap pad help reduce the temperature increase? Does sitting with legs apart help? The study investigated all three of these conditions: legs together and a laptop computer on the lap, legs apart and a laptop computer on the lap, and legs together with a lap pad under the laptop computer. Scrotal temperature increase over a 60-minute session was measured in °C, and the summary statistics are given in the table below.

<table>
<thead>
<tr>
<th>Condition</th>
<th>n</th>
<th>mean</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legs together</td>
<td>29</td>
<td>2.31</td>
<td>0.96</td>
</tr>
<tr>
<td>Lap pad</td>
<td>29</td>
<td>2.18</td>
<td>0.69</td>
</tr>
<tr>
<td>Legs apart</td>
<td>29</td>
<td>1.41</td>
<td>0.66</td>
</tr>
</tbody>
</table>

(a) Which conditions has the largest mean temperature increase? Which has the smallest?

(b) Do the data appear to satisfy the condition that the standard deviations are roughly the same? (The data satisfy the normality condition.)

(c) Use the fact that sum of squares is 53.2 to test whether there is a difference in mean temperature increase between the three conditions. Show all details of the test, including an analysis of variance table.

Solution
(a) Legs together with no lap pad has the largest temperature increase. Spreading the legs apart has the smallest temperature increase.

(b) Yes, the standard deviations are similar. The largest, $s_1 = 0.96$, is not more than twice the smallest, $s_3 = 0.66$. 

6
(c) The null hypothesis is that the population mean temperature increases for the three conditions are all the same and the alternative hypothesis is that at least two of the means are different. We find the mean squares by dividing the sum of squares by the respective degrees of freedom ($df = 3 - 1 = 2$ for Groups, $df = 87 - 3 = 84$ for Error). The F-statistic is the ratio of the two mean squares.

$$F = \frac{MSG}{MSE} = \frac{6.85}{0.63} = 10.9$$

These calculations are summarized in the ANOVA table below.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups</td>
<td>2</td>
<td>13.7</td>
<td>6.85</td>
<td>10.9</td>
<td>0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>84</td>
<td>53.2</td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>86</td>
<td>66.9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The p-value is the area above 10.9 in an F-distribution with numerator degrees of freedom 2 and denominator degrees of freedom 84. Using technology we see that the p-value=0.0001. We reject $H_0$ and find strong evidence that average temperature increase is not the same for these three conditions. It appears that spreading legs apart may be more effective at reducing the temperature increase.

8.22 Sandwich Ants and Bread (Graded for Accurateness) Data 8.1 on page 492 describes an experiment to study how different sandwich fillings might affect the mean number of ants attracted to pieces of a sandwich. The students running this experiment also varied the type of break for the sandwiches, randomizing between four types: Multigrain, Rye, Wholemeal, and White. The ant counts in 6 trials and summary statistics for each type of break and the 24 trials as a whole are given in the table below and stored in SandwichAnts.

<table>
<thead>
<tr>
<th>Bread</th>
<th>Ants Mean</th>
<th>Std.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multigrain</td>
<td>36.00</td>
<td>14.52</td>
</tr>
<tr>
<td>Rye</td>
<td>37.67</td>
<td>12.40</td>
</tr>
<tr>
<td>Wholemeal</td>
<td>35.83</td>
<td>13.86</td>
</tr>
<tr>
<td>White</td>
<td>42.50</td>
<td>17.41</td>
</tr>
<tr>
<td>Total</td>
<td>38.00</td>
<td>13.95</td>
</tr>
</tbody>
</table>

(a) Show how to use the summary information to compute the three sums of squares needed for using ANOVA to test for a difference in mean number of ants among these four types of bread.

(b) Use the sums of squares from part (a) to construct the ANOVA table and complete the details for this test. Be sure to give a conclusion in the context of this data situation.

Solution

(a) We find the required sum of squares using the shortcut formulas at the end of this section. We compare the group means to the overall mean:

$$SSG = \sum n_i(\bar{x}_i - \bar{x})^2 = 6(36.00 - 38)^2 + 6(37.67 - 38)^2 + 6(35.83 - 38)^2 + 6(42.50 - 38)^2 = 174.4$$
We find the variability within the groups:

\[ SSE = \sum (n_i - 1)s_i^2 = (6 - 1)14.52^2 + (6 - 1)12.40^2 + (6 - 1)13.86^2 + (6 - 1)17.41^2 = 4299.0 \]

We find the total variability:

\[ SSTotal = (n - 1)s^2 = (24 - 1)13.95^2 = 4475.9 \]

We see that \( SSG + SSE = 174.4 + 4299.0 = 4473.4 \). The difference from \( SSTotal = 4475.9 \) is due to rounding the means and standard deviations in the summary statistics.

(b) We are testing

\[ H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 \]

vs

\[ H_a : \text{Some } \mu_i \neq \mu_j \]

where the \( \mu_i \)'s represent the mean numbers of ants for the four types of bread. We have \( 4 - 1 = 3 \) degrees of freedom for the groups, \( 24 - 4 = 20 \) degrees of freedom for the error, and \( 24 - 1 = 23 \) degrees of freedom for the total. We compute mean squares by dividing sums of squares by degrees of freedom, then take the ratio of the mean squares to compute the F-statistic.

\[ MSG = \frac{174.4}{3} = 58.13 \quad MSE = \frac{4299.0}{20} = 214.95 \quad F = \frac{MSG}{MSE} = \frac{58.13}{214.95} = 0.27 \]

The ANOVA table summarizing these calculations is shown below.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bread</td>
<td>3</td>
<td>174.4</td>
<td>58.13</td>
<td>0.27</td>
<td>0.846</td>
</tr>
<tr>
<td>Error</td>
<td>20</td>
<td>4299.0</td>
<td>214.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>23</td>
<td>4473.4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Using an F-distribution with 3 and 20 degrees of freedom, we find the area beyond \( F = 0.27 \) gives a p-value of 0.846. This is a very large p-value, so we have no convincing evidence to conclude that the mean number of ants attracted to a sandwich depends on the type of bread.

Computer Exercises

For each R problem, turn in answers to questions with the written portion of the homework. Send the R code for the problem to Katherine Goode. The answers to questions in the written part should be well written, clear, and organized. The R code should be commented and well formatted.

**R problem 1 (Graded for Accurateness)** The function `chisq.test()` can be used for chi-square tests. For goodness-of-fit tests, provide a single array of counts and an optional array of the same length of probabilities for each case. (R uses uniform probabilities by default.) For example, in class for the Rock, Paper, Scissors data, we had total counts of 88, 74, and 66. Here is an example.

\[ \text{chisq.test(c}(88, 74, 66), p = \text{c}(1, 1, 1)/3) \]  

##
## Chi-squared test for given probabilities

### data: c(88, 74, 66)
### X-squared = 3.263, df = 2, p-value = 0.1956

1. Use `chisq.test()` for the data in Exercise 7.14 and compare to your previous answer.

**Solution**
Using R, we get the following output.

```
Chi-squared test for given probabilities

data: c(66, 39, 14)
X-squared = 34.1008, df = 2, p-value = 3.936e-08
```

We find that our p-value is very close to 0, which agrees with the p-value found in Exercise 7.14.

The following code in R was used to get this value.

```
chisq.test(c(66, 39, 14), p = c(1, 1, 1)/3)
```

2. Using the same data, test if the data is consistent with the null distribution $p_{\text{rock}} = 0.5, p_{\text{paper}} = 0.3, p_{\text{scissors}} = 0.2$.

**Solution**
With this new null distribution, we get the following output from R.

```
Chi-squared test for given probabilities

data: c(66, 39, 14)
X-squared = 5.0504, df = 2, p-value = 0.08004
```

We see that our p-value is not significant at the 0.05 confidence level. Thus, there is no evidence to support the fact that the data is not consistent with the null distribution of $p_{\text{rock}} = 0.5, p_{\text{paper}} = 0.3, p_{\text{scissors}} = 0.2$.

The following code in R was used to get this value.

```
chisq.test(c(66, 39, 14), p = c(0.5, 0.3, 0.2))
```

3. A randomization distribution method can also be used to calculate a p-value. Add the argument `simulate.p.value=TRUE` to the call to `chisq.test()` with `B=10000` and compare answer to that in the first part.

```
chisq.test(c(66, 39, 14), simulate.p.value = TRUE, B = 10000)
```

**Solution**
We get the following output from R.
Chi-squared test for given probabilities with simulated p-value (based on 10000 replicates)

data:  c(66, 39, 14)
X-squared = 34.1008, df = NA, p-value = 9.999e-05

The p-value is also close to 0 in this case.

The following R code was used to obtain this value.

```
chisq.test(c(66, 39, 14), simulate.p.value = TRUE, B = 10000)
```

**R Problem 2 (Graded for Accurateness)** Here you will learn to use `chisq.test()` for tests of association where the counts are in matrices. Compare the answer you get with this function with the answer from your hand calculation on Exercise 7.43.

```
x = matrix(c(18, 13, 9, 16, 15, 19, 13, 34, 11, 9), nrow = 2, ncol = 5)
chisq.test(x)
```

**Solution**

Using this code, we get the following output from R.

**Pearson’s Chi-squared test**

data:  x
X-squared = 9.0691, df = 4, p-value = 0.0594

We see that the p-value of 0.0594 matches the p-value of 0.059 found in the hand calculation of Exercise 7.43.