STATWAY® COLLEGE MODULE 7 v4.1

TABLE OF CONTENTS:

7.1: The Chi-Square Distribution

LEARNING GOALS

By the end of this collaboration, you should understand that:

- To test a claim about a single population proportion, we can use the chi-square distribution.
- With one degree of freedom, the chi-square distribution is the distribution of squares of *Z*-scores, randomly selected.
- With one degree of freedom, the chi-square distribution gives the same P-value as the normal distribution.

By the end of this collaboration, you should be able to:

- Conduct a hypothesis test for a single population proportion using a chi-square statistic.
- Use technology to determine the *P*-value for a chi-square statistic.

INTRODUCTION

In this collaboration, we consider **multinomial experiments**. In Module 2, you learned about *binomial experiments*. Multinomial experiments are like binomial experiments – there are *n* independent trials, but instead of two outcomes per trial (*success* and *failure*), we allow two *or more* outcomes per trial. The binomial experiment is a special case of the multinomial experiment.

We begin with a binomial experiment in the context of a hypothesis test for a single population proportion.

Asthma Incidence among Youth

The Centers for Disease Control and Prevention reported in a 2019 national health interview study that 8.4% of male children (aged <18) have asthma.[1] A researcher in Florida wants to compare this to the proportion of male youth in that state who have asthma.

In a recent random sample of 400 male youth in Florida, 48 responded that they had asthma. This sample is the result of a *binomial experiment* – each trial has two outcomes, asthma (success) or no asthma (fail). At the 1% level of significance, we will investigate the claim that asthma among male youth in Florida is *different* from the national figure.

1 https://www.cdc.gov/asthma/nhis/2019/data.htm

TRY THESE

1 This is a hypothesis test about a single population proportion, *p*. What are the appropriate null and alternative hypotheses?

2 The criteria for normality require that the expected number of successes, *np*, is at least 10, and the expected number of failures, *n*(1 – *p*), is at least 10.

A Compute the expected number of successes.

B Compute the expected number of failures.

C Are the criteria satisfied for approximate normality?

D What is the sample proportion? Write your answer as a decimal.

3 Compute the test statistic for the hypothesis test for a single population proportion, $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$. Because we'll be making an important point about this statistic, keep 4 places after the decimal.

4 Use technology or tables to compute the *P*-value for this two-tailed test. Round to 3 decimal places after the decimal. **Note**: If completing this problem online, follow the instructions given online to calculate the *P*-value.

5 What do you conclude about the null and alternative hypotheses?

6 Make a conclusion in the context of this problem.

NEXT STEPS

The test statistic, Z, in the previous example works well for binomial experiments, when there are only two outcomes per trial. When there are three or more outcomes per trial, things get more complicated, and the Z statistic doesn't allow for sample proportions for more outcomes.

When we want to test claims about two or more proportions In a multinomial experiment, we use the χ^2 (χ is the Greek letter *chi*), which we pronounce *chi-square*. The formula for this test statistic is:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

Each outcome corresponds to one term in the sum above. E stands for the **expected frequency** and O stands for the **observed frequency** for each possible outcome in the experiment.

For the *Asthma Incidence among Youth* example, each sample member falls into one of two categories – either the male child has asthma or not.

7 For the male children with asthma, give the observed and expected frequencies.

A What is the observed number of male children, O, with asthma?

O = _________

B What is the *expected* number of male children, E, with asthma (E is the binomial mean, np from 2A). Write to one decimal place.

E = _________

C The expected frequency, E, is based on the proportion, p, from the null hypothesis. Is it similar to the observed frequency?

D Compute the contribution of this outcome to the χ^2 test statistic. Keep 4 places after the decimal.

$\frac{(O-E)^2}{E}$ = ____________

8 For the male children *without* asthma, give the observed and expected frequencies.

A What is the observed number of male children, O, without asthma (these are *failures* in this experiment)?

O = ________

B What is the *expected* number of male children, E, without asthma (this is the mean number of failures,
$n(1 - p)$ from Question 2.B).

E = ________

C The expected frequency, E, just computed is based on the proportion, p, from the null hypothesis. Is it similar to the observed frequency?

D Compute the contribution of this outcome to the χ^2 test statistic. Keep 4 places after the decimal.

$\frac{(O-E)^2}{E}$ = ____________

E Think about the test statistic, where expected frequencies are based on p from the null hypothesis. If the null hypothesis is false, would you expect this statistic to be small or large?

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

9 The number of observed male children with asthma was observed from a random sample, and varies from sample to sample. Once this value is known, is the number of male children *without* asthma random? Explain.

10 If **degrees of freedom** represent the number of observed frequencies that vary freely (randomly), how many degrees of freedom are present among the two observed frequencies (of those with and without asthma)?

11 Complete the table below with your answers. The last cell, which is the total of the last column, is chi-square. Find the χ^2 test statistic. Round to three places after the decimal.

	O = Observed Frequency	*E* = Expected Frequency	$\frac{(O-E)^2}{E}$
Asthma			
No Asthma			
		Total:	

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \_\_\_\_\_\_\_\_$$

12 In this question we begin to understand the relationship between the *Z* and χ^2 test statistics. Compute the *square* of the *Z*-score from Question 3. Keep three places after the decimal.

13 How is the χ^2 statistic, with one degree of freedom, related to the Z statistic?

YOU NEED TO KNOW

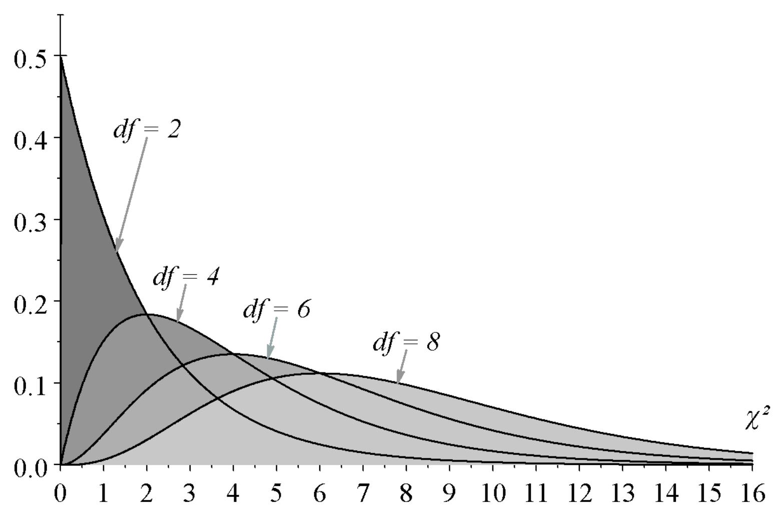
When conducting a multinomial experiment with k outcomes per trial, the statistic

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

varies (approximately) according to the χ^2 (chi-square) distribution with $k - 1$ degrees of freedom.

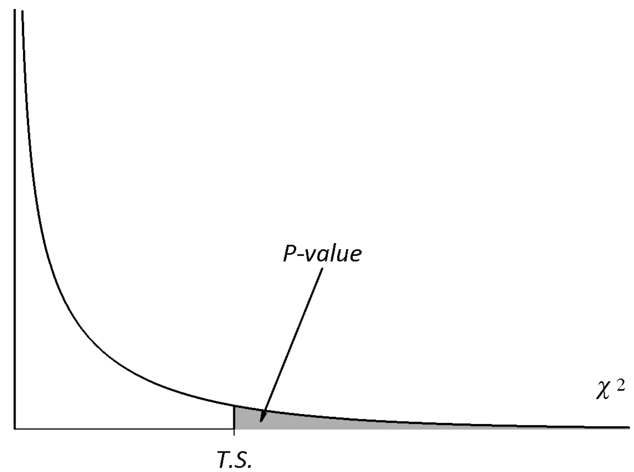
With 1 degree of freedom, the χ^2 distribution is the probability distribution of squared Z-scores from the standard normal distribution. With $k - 1$ degrees of freedom, the χ^2 distribution is the distribution of all *sums* of $k - 1$ *squared* Z-scores, randomly selected. Because of this, the distribution only allows non-negative values.

Several χ^2 curves are plotted below, with varying degrees of freedom.



The χ^2 hypothesis test which we are conducting is *always* right-tailed, so the P-value is the area to the *right* of the test statistic.

The χ^2 distribution with *one degree of freedom* is plotted below (the one degree of freedom graph is quite strange looking!). The area shaded to the right of a test statistic (*T.S.*) represents a *P*-value. Take some time with your instructor to learn how to find such *P*-values using technology.



TRY THESE

We must use the χ^2 distribution to determine the *P*-value.

14 Use technology to find the right-tailed *P*-value corresponding to your test statistic from Question 11 (such as: https://carnegiemathpathways.org/go/desmoschi or https://carnegiemathpathways.org/go/appchisq).
Round the P-value to three places after the decimal. **Note**: If completing this problem online, follow the instructions given online to calculate the *P*-value.

15 How does this *P*-value compare to the one computed in Question 4?

16 If we use a 1% level of significance, what would we conclude about our original null and alternative hypotheses?

LET'S SUMMARIZE

Because the *Z* test for a population proportion yields the same *P*-value as the corresponding χ^2 test (with one degree of freedom), the tests are equivalent.

So why use the χ^2 test? The answer is because it allows for more than two outcomes per trial. The *Z* test only works for binomial experiments, but the χ^2 test can be applied to *multinomial experiments*, where more than two outcomes per trial are possible. We will investigate examples of this in our next collaboration.

Exercise 7.1

Consider the following 2022 claim: Half of U.S. adults support stricter gun laws. In November 2021, Gallup.com surveyed by phone 206 randomly selected adults living in the West region of the United States.[2] Each person was asked if they support stricter gun laws. Of those surveyed, 107 answered yes, the rest (99 people) did not.

Perform a chi-square test on the claim that in 2022 half of Americans living in the West region of the U.S. support stricter gun laws, against the alternative that this is not the case. Use a 1% level of significance. The main difference is that, when using a chi-squared test, you must look at every possible outcome.

Step 1: Determine the Hypotheses

1 The outcome for each of the trials (a randomly selected American living in the West) above is the support or non-support of stricter gun laws. How many possible outcomes (k) are there per trial?

k = ________

2 Our claim is that, at the time, half of Americans living in the West supported stricter gun laws. What is the hypothetical proportion for each outcome?

Outcome	Hypothetical Proportion
Support stricter gun laws	
Don’t support stricter gun laws	

3 Use your previous answer to construct the null hypothesis. Define p_1 as the proportion of Americans living in the West who supported stricter gun laws in 2022 and p_2 as the proportion who did not.

H_0: p_1 = _________ (Give the hypothetical proportion of support for stricter gun laws)

p_2 = _________ (Give the hypothetical proportion of non-support for stricter gun laws)

4 If the proportions above are not correct, then at least one is wrong. Express this phrase as the alternative hypothesis below.

H_a:

2 https://news.gallup.com/poll/357317/stricter-gun-laws-less-popular.aspx

Step 2: Collect the Data

5 The expected frequencies are each equal to the sample size times the respective proportions in the null hypothesis (*E* = *np*). Fill in the table below with the appropriate expected frequencies.

Preference	*O* = Observed	*E* = Expected = *np*
Support stricter gun laws	107	
Don't support stricter gun laws	99	103

6 The χ^2 (chi-square) distribution, with 1 degree of freedom, is the distribution squared *Z*-scores, randomly selected. When $Z = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}}$ is approximately normal, $\chi^2 = \sum \frac{(O-E)^2}{E}$ will vary approximately with the χ^2 distribution. Do the expected frequencies (*np* and *n*(1 - *p*)) satisfy the conditions of an approximate χ^2 distribution? Explain.

7 What are the degrees of freedom for this test?

df = ________

Step 3: Assess the Evidence

Compute the χ^2 test statistic, $\chi^2 = \sum \frac{(O-E)^2}{E}$, by completing Questions 8-10 below.

8 Start by calculating $\frac{(O-E)^2}{E}$ for "Support stricter gun laws." Round to three decimal places.

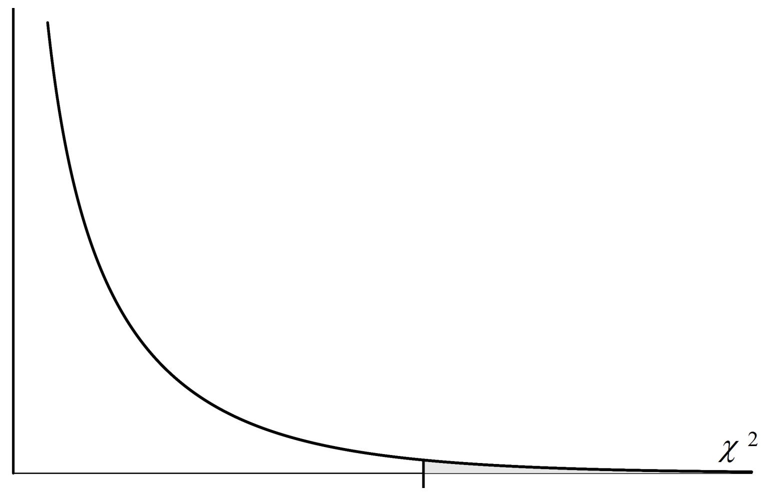
9 Calculate $\frac{(O-E)^2}{E}$ for "Don't Support." Round to three decimal places.

10 Calculate χ^2. Round the value to two decimal places.

$$\chi^2 = \sum \frac{(O-E)^2}{E} =$$

Step 3: Assess the Evidence

11 On the χ^2 distribution below (with one degree of freedom), plot the χ^2 test statistic, and label the region that represents the *P*-value.



Use technology to determine the *P*-value of this right-tailed test. (You may use the χ^2 *P*-value calculator at https://carnegiemathpathways.org/go/appchisq or https://carnegiemathpathways.org/go/desmoschi.) Round to three decimal places. **Note**: If completing this problem online, follow the instructions given online to find the *P*-value.

Step 4: Make a Conclusion

12 When comparing the *P*-value to α, what conclusion do you make regarding the null hypothesis?

7.1 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Conduct a hypothesis test for a single population proportion using a chi-square statistic.	
Use technology to determine the *P*-value for a chi-square statistic.	

7.2: Goodness of Fit

LEARNING GOALS

By the end of this collaboration, you should understand that:

- The logic, steps, and the need to check conditions in a chi-square test for goodness-of-fit are similar to hypothesis testing procedures encountered in previous modules.
- The conditions needed for a chi-square procedure are different from the conditions discussed in previous modules.

By the end of this collaboration, you should be able to:

- Given a claim about the distribution of a categorical variable, choose appropriate null and alternative hypotheses for a chi-square goodness-of-fit test.
- Given sample data, state and check the conditions needed for the chi-square goodness-of-fit test to be appropriate.
- Compute the value of the test statistic in a chi-square goodness-of-fit test.
- Use the *P*-value and the chosen significance level to reach a decision.
- Carry out a chi-square goodness-of-fit test and interpret the conclusion in context.

INTRODUCTION

Previously, we learned about the chi-square (χ^2) distribution and its relationship to the standard normal distribution. We saw that both can be used equivalently to test claims about a population proportion in binomial experiments (with two categorical outcomes per observation). Now we outline the process for χ^2 tests in multinomial experiments (with more than two outcomes). These tests are often referred to as **Goodness of Fit** tests.

Step 1: Determine the Hypotheses

The goodness of fit test makes claims about the proportions (or probabilities) for each outcome of a multinomial experiment. If there are *k* outcomes per trial, then the null hypothesis would be

$$\begin{aligned} H_0: \quad & p_1 = value_1 \\ & p_2 = value_2 \\ & \quad \dots \\ & p_k = value_k \end{aligned}$$

Because the sum of all probabilities in a distribution is always 1, the hypothetical proportions must add to 1.

For the null hypothesis to not be true, one or more of the proportions would be incorrect. Thus, the alternative hypothesis takes the form,

$$H_a: \text{ These proportions are not all correct}$$

Step 2: Collect the Data

Using a sample of *n* independent trials, each having *k* outcomes, the multinomial experiment is conducted by collecting categorical data from a random sample. For each outcome, the observed frequency, *O*, is the number of successes. The sum of the *k* observed frequencies is *always* the sample size, $n = \sum O$, so the last observed frequency is not random. There are therefore $k - 1$ degrees of freedom among the observed frequencies.

Next, we compute the expected frequencies, *E*, for each outcome. As with the binomial experiment, the expected number of successes is always *np*, where *n* is the number of trials (the sample size) and *p* is the hypothetical proportion for the corresponding outcome in the null hypothesis.

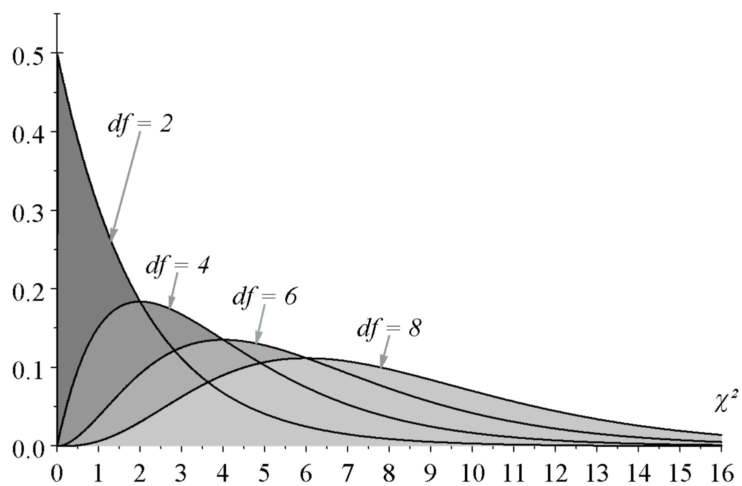
For approximate normality in a binomial experiment, we require the expected number of successes and failures to be at least 10. In the multinomial experiment, we soften this because (with potentially many categories to fill) it can require unreasonably large sample sizes.

We assume that the test statistic is approximately distributed according to the χ^2 distribution if each expected frequency, E, is at least 5.

Step 3: Assess the Evidence

For goodness of fit tests, the test statistic is: $\chi^2 = \sum \frac{(O-E)^2}{E}$

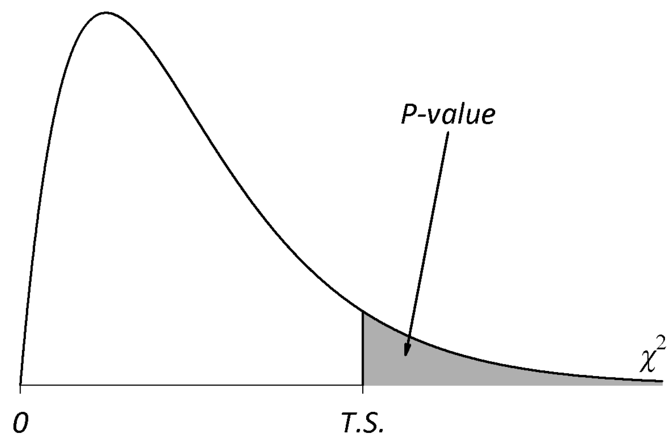
The distributions of values for this statistic has various shapes, depending on the degrees of freedom. In all cases, however, the distribution starts at 0 and is skewed to the right.



To compute the test statistic, we need to know all observed (*O*) and expected (*E*) frequencies. The null hypothesis gives probabilities for each outcome, and from these we compute the expected frequencies ($E = np$). If the null hypothesis is false, the test statistic should be large because the observed frequencies, *O*, will be quite different from the expected frequencies, *E*, so the differences, $(O - E)$ will be large.

This is why a goodness of fit test is always right tailed – when the null hypothesis is false, the test statistic should be large.

The *P*-value is determined using technology – it is the area of the right tail (starting at the test statistic) under the χ^2 distribution with $k - 1$ degrees of freedom.



Step 4: Make a Conclusion

When the null hypothesis is true, the *P*-value is the probability of obtaining a test statistic at least as extreme as the one observed. Small *P*-values (less than or equal to α) provide sufficient justification for the rejection of the null hypothesis in favor of the alternative. Large *P*-values do not justify the rejection of the null hypothesis.

TRY THESE

In the article, *Is There a Season for Homicide?* [3], the author gave the seasons in which 1361 randomly selected homicides were committed. These values are summarized below.

Season	*O* = Observed Murders
Winter	328
Spring	334
Summer	372
Fall	327

We test the claim at $\alpha = 0.05$ that the proportions of homicides in the 4 seasons are all equal (to 25%) against the alternative that they are not all equal.

[3] *Criminology*, 1988, pp. 287-296.

Step 1: Determine the Hypotheses

1 The outcome for each trial (murder) above is the season in which the murder was committed. How many possible outcomes (k) are there per trial?

k = ________

2 Construct the null hypothesis. This gives the hypothetical proportions of all murders that occur in the four seasons. The null hypothesis states that the proportion of murders that occur in each season is the same. Write the proportions as decimals.

H_0: p_1 = ________ (for winter)

p_2 = ________ (for spring)

p_3 = ________ (for summer)

p_4 = ________ (for fall)

3 What is the alternative hypothesis?

Step 2: Collect the Data

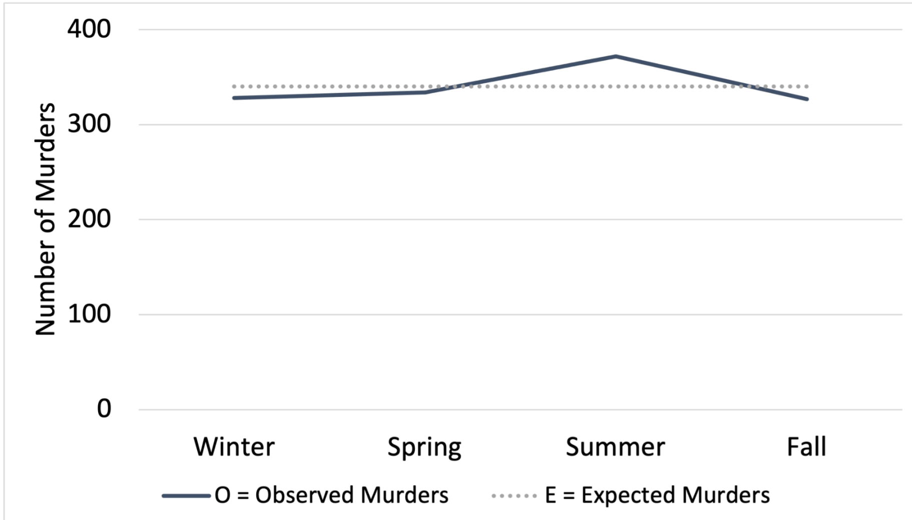
4 Compute the sample size by adding the observed frequencies.

$n = \sum O$ = ________

5 The expected frequencies are each equal to the sample size times the respective proportions in the null hypothesis ($E = np$). Because the proportions are all 0.25, they are all equal. Calculate the expected frequency. Do not round the values.

Expected # of murders = $E = np_1 = np_2 = np_3 = np_4 =$

The graph below compares the observed and expected frequencies. We can see that differences exist for each outcome (season). The Goodness of Fit Test will examine whether the differences between the actual data and hypothesized data are sufficiently large to warrant the rejection of the null hypothesis.



6 Do the expected frequencies satisfy the conditions of an approximate χ^2 distribution? Explain.

7 What are the degrees of freedom for this test? The degrees of freedom is $k - 1$ where k is the number of outcomes per trial.

df = ________

Step 3: Assess the Evidence

8 Compute the χ^2 test statistic, $\chi^2 = \sum \frac{(O-E)^2}{E}$, by completing the table below.

A Calculate $\frac{(O-E)^2}{E}$ **for winter**. Round to two decimal places.

B Calculate $\frac{(O-E)^2}{E}$ **for spring**. Round to two decimal places.

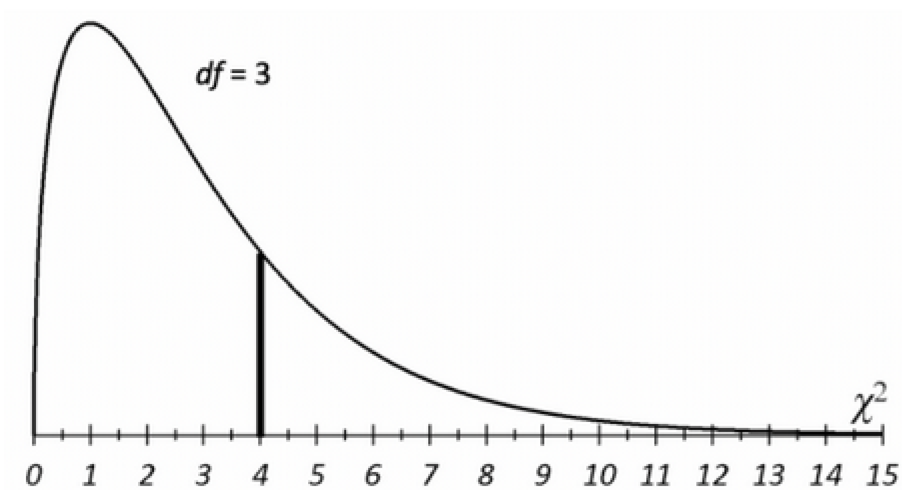
C Calculate $\frac{(O-E)^2}{E}$ **for summer**. Round to two decimal places.

D Calculate $\frac{(O-E)^2}{E}$ **for fall**. Round to two decimal places.

9 Give the test statistic by adding all values from 8A-D. Round to two places after the decimal.

χ^2 = ________

The χ^2 distribution with 3 degrees of freedom is depicted below. The *P*-value is the area to the area to the right of the χ^2 statistic. The location of this test statistic is denoted by the vertical line segment.



10 Use technology (such as https://carnegiemathpathways.org/go/appchisq or https://carnegiemathpathways.org/go/desmoschi) to determine the *P*-value of this right tailed test. Round to three places after the decimal. **Note**: If completing this problem online, follow the instructions given online to find the *P*-value.

Step 4: Make a Conclusion

11 When comparing the *P*-value to α, what conclusion do you make regarding the null hypothesis?

12 Write a brief conclusion in the context of this problem.

LET'S SUMMARIZE

- The χ^2 goodness of fit test is used to test a claim regarding the values of several proportions (more than two) corresponding to outcomes of a multinomial experiment. It is assumed that the n sample observations are gathered randomly from a population of categorical data with *k* outcomes (categories).
- The logic, steps, and need to check conditions in a chi-square goodness of fit test are similar to hypothesis testing procedures encountered in previous modules.
- For a chi-square goodness of fit test, the null hypothesis specifies the proportions of each outcome in a multinomial experiment, and the alternative hypothesis states that at least one of the specified proportions is incorrect.
- The criteria for the test is that expected frequencies are all at least 5. The test is right tailed, and the test statistic varies according to the χ^2 distribution with k - 1 degrees of freedom.
- Larger chi-square values generally correspond to smaller *P*-values and stronger evidence against the null hypothesis. Similarly, smaller chi-square values generally correspond to larger *P*-values and weaker evidence against the null hypothesis.

Exercise 7.2

Lina has reason to believe that her children are eating her Skittles® candy. Skittles come in red, orange, yellow, green, and purple. Her children all have color preferences, but Lina has no color preference. In order to prove the guilt of her children, she counts the colors in the bowl that she keeps out, with the following results.

Color	Frequency
Red	25
Orange	51
Yellow	40
Green	50
Purple	30

Skittles colors are produced in equal proportions (20% for each color) prior to being mixed together and then bagged. Is the variation in color frequencies large enough to argue that someone has been eating Lina's skittles?

Test the claim that the underlying color proportions are different from the advertised proportions of 20% each (implying that the color variations in the bowl are not due to randomness). Use a 1% level of significance.

1 The outcome for each of the trials (Skittles) above is the color. How many possible outcomes (*k*) are there per trial?

k = ________

2 What are the hypothetical proportions for each color? Write the proportion as a decimal.

3 Use your answer in the previous question to construct the null hypothesis. This gives the hypothetical proportions of Skittles for each color. Write each proportion as a decimal.

H_0: p_1 = ________ (for red)

p_2 = ________ (for orange)

p_3 = ________ (for yellow)

p_4 = ________ (for green)

p_5 = ________ (for purple)

4 What is the alternative hypothesis?

5 Compute the sample size by adding the observed frequencies.

$n = \sum O =$ ________

6 The expected frequencies are each equal to the sample size times the respective proportions in the null hypothesis ($E = np$). Because the proportions are all 0.20, the expected frequencies are all equal. Fill in the table below with the appropriate expected frequencies.

Expected frequency = $E = np_1 = np_2 = np_3 = np_4 =$

7 Do the expected frequencies satisfy the conditions of an approximate χ^2 distribution? Explain.

8 What are the degrees of freedom for this test? The degrees of freedom is $k - 1$ where k is the number of outcomes per trial.

$df =$ ________

Compute the χ^2 test statistic, $\chi^2 = \sum \frac{(O-E)^2}{E}$, by answering **Questions 9-14** below. Round answers to four decimal places.

Calculate $\frac{(O-E)^2}{E}$ for:

9 Red

10 Orange

11 Yellow

12 Green

13 Purple

14 Give the test statistic by adding all the values from the previous five questions.

$$\sum \frac{(O-E)^2}{E} =$$

15 Use technology to determine the *P*-value of this right tailed test. (You may use the χ^2 *P*-value calculator at https://carnegiemathpathways.org/go/appchisq or https://carnegiemathpathways.org/go/desmoschi.) Round to three decimal places. **Note**: If completing this problem online, follow the instructions given online to find the *P*-value.

16 When comparing the *P*-value to α, what conclusion do you make regarding the null hypothesis? We are performing this test at the 1% significance level.

17 Write a brief conclusion in the context of this problem.

7.2 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Given a claim about the distribution of a categorical variable, choose appropriate null and alternative hypotheses for a chi-square goodness-of-fit test.	
Given sample data, state and check the conditions needed for the chi-square goodness-of-fit test to be appropriate.	
Compute the value of the test statistic in a chi-square goodness-of-fit test.	
Use the *P*-value and the chosen significance level to reach a decision.	
Carry out a chi-square goodness-of-fit test and interpret the conclusion in context.	

7.3: Testing for Independence with Two-Way Tables

LEARNING GOALS

By the end of this collaboration, you should understand that:

- In a chi-square test for independence, we investigate whether two categorical variables are independent.
- The null hypothesis for a chi-square test for independence states that two categorical variables are independent.
- The alternative hypothesis for a chi-square test for independence states that two categorical variables are dependent.
- The degrees of freedom are related to the number of categories in the row and column variables.
- The logic, steps, and requirement to check conditions in a chi-square test for independence are similar to hypothesis testing procedures encountered in previous modules.

By the end of this collaboration, you should be able to:

- Given sample data, state and check the conditions needed for the use of the chi-square test for independence to be appropriate.
- Compute the value of the test statistic and find the associated P-value in a chi-square test for independence.
- Use the P-value and the chosen significance level to reach a decision, and interpret the conclusion in context.
- Evaluate whether conclusions are reasonable, given the description of a statistical study and the results of the chi-square test.

INTRODUCTION

In our previous unit, we learned to use the χ^2 (*chi-square*) distribution to conduct a goodness of fit test. This allowed us to compare theoretical expected frequencies from a null hypothesis to real observed frequencies observed in sample data. In this collaboration, we continue with the χ^2 tests in a special test for independence.

Marijuana and Politics

The data used in this problem are based on the article, *Attitudes about Marijuana and Political Views*[4]. The two-way table below summarizes data on marijuana use and political views from a random sample of 270 adults. Each frequency in an interior cell is a count of the number of adults in the sample that have two specific characteristics. The two-way table is missing several values.

[4] *Psychological Reports*, 1973, pp. 1051 to 1054

Political Views	Never Smoke	Rarely Smoke	Frequently Smoke	Totals
Liberal	96	35		155
Conservative	43	9		55
Other				60
Totals	173	53	44	270

We would like to investigate whether the variables *political views* and *smoking frequency* are independent of one-another.

1 What is the explanatory variable in this study?

2 What is the response variable?

3 Do you think these variables are independent, or dependent? Explain your answer.

4 Enter the missing values into the table above.

5 Compute the following conditional probabilities. Write the probability as a decimal rounded to 2 decimal places.

A P(*Conservative*|*Never Smoke*) =

B P(*Conservative*|*Frequently Smoke*) =

6 Given your previous answer, would you consider marijuana smoking frequency and political views as independent or dependent variables? Explain your answer.

We will now look at the number of degrees of freedom involved in this problem. Below is the two-way table from the start of this collaboration, before values in the "Frequently Smoke" column and "Other" row were entered.

Political Views	Never Smoke	Rarely Smoke	Frequently Smoke	Totals
Liberal	96	35		155
Conservative	43	9		55
Other				60
Totals	173	53	44	270

7 With the values originally given in the table, are the values you entered free, random values, or dependent on other values? Explain.

8 If the degrees of freedom in the observed frequencies (*O*) are the number of free, independent observations, how many degrees of freedom are there among the observed frequencies in the table above?

9 Suppose a two-way table has *r* rows and *c* columns (don't count the totals!). Make a rule for the degrees of freedom among the observed frequencies in the table. Express this rule as a formula for the degrees of freedom.

df = ____________

NEXT STEPS

Now that we understand the degrees of freedom in a two-way table, it is time to think about a hypothesis test. This test is similar to the goodness of fit test already discussed, but the degrees of freedom are different (as discussed) and the expected frequencies have a special formula.

We are conducting a test for independence. In the prior statistical study, the question is, "Are political views independent of marijuana smoking frequency?" The data in the two-way table are from a random sample. By examining conditional probabilities in the sample data, we saw that the variables appear to not be independent. We will now perform a chi-square goodness of fit test to examine whether the variables are independent across the entire population.

To conduct a test for independence, we construct expected frequencies based on the assumption that these variables are independent (this will be our null hypothesis).

If events A and B are independent then *P*(*A* & *B*) = *P*(*A*) · *P*(*B*). Suppose *A* = *liberal*, and *B* = *never smoke*.

Refer back to your completed two-way table.

10 Suppose we randomly pick an adult from the sample. If "being liberal" and "never smoking" are independent events, what is the probability that an adult is liberal and never smokes? Round the answer to four decimal places.

$$P(liberal \,\&\, never\ smoke) \ = \ P(liberal)\cdot P(never\ smoke) \ = \underline{\hspace{5cm}}$$

11 This probability is the proportion of people who should be in the *liberal & never smoke* category if the events are independent. If this is *p*, the population proportion of people in this category, then the expected frequency (the number of people in this sample expected to be in this category) is *E* = *np*. In this formula, *n* is the sample size (grand total of the two-way table). Find *E*. Round to two decimal places.

E = *np* = ___________

The expected frequency is actually quite close to what was observed, *O* = 96. For this event, the null hypothesis of independence seems to lead to reliable predictions of what actually occurred.

Keep in mind, this expected frequency is computed based on our null hypothesis that assumes that our variables are independent.

Look once more as we backtrack through the computation which we performed to compute the expected frequency.

$$E \ = \ np \ = \ 270\cdot 0.3678 \ = \ 270\cdot\frac{155}{270}\cdot\frac{173}{270} \ = \ \frac{155\cdot 173}{270} \ = \ \frac{row\ total\cdot column\ total}{grand\ total}$$

The key here is that the expected frequency of an event can be found directly from the row, column and grand totals.

YOU NEED TO KNOW

To compute an expected frequency for an observation in a given row and column of a two-way table, use the formula,

$$E \ = \ \frac{row\ total\cdot column\ total}{grand\ total}$$

TRY THESE

We are now ready to conduct a test for independence using a two-way table. We want to test if political views are independent of marijuana smoking frequency at the 1% significance level.

Step 1: Determine the Hypotheses

The hypotheses for this test are:

H_0: The explanatory and response variables are independent.

H_a: The explanatory and response variables are dependent.

12 Write the null and alternative hypotheses in the context of the current problem (naming the explanatory and response variables).

Step 2: Collect the Data

The test for independence between bivariate categorical variables requires that data be summarized in a two way table. Row and column totals should be computed so that expected frequencies can be computed.

The expected frequencies are computed with the formula

$$E = \frac{row\ total \cdot column\ total}{grand\ total}$$

As with the previous goodness of fit test, we require that each expected frequency is at least 5.

13 Compute the expected frequencies for the cells in the first row. Round the values to two decimal places. The expected frequencies of cells in the 2nd and 3rd rows are provided.

Political Views	Never Smoke	Rarely Smoke	Frequently Smoke	Totals
Liberal	**96** *99.31*	**35**	**24**	155
Conservative	**43** *35.24*	**9** *10.8*	**3** *8.96*	55
Other	**34** *38.44*	**9** *11.78*	**17** *9.78*	60
Totals	173	53	44	270

14 Are the criteria satisfied for the approximate χ^2 distribution for us to perform a test for independence? Explain.

NEXT STEPS

Step 3: Assess the Evidence

The test statistic for a χ^2 test for independence is:

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

This is approximately distributed according to the χ^2 distribution with degrees of freedom equal to:

$$df = (r-1) \cdot (c-1)$$

Here, r is the number of rows in the table and c is the number of columns. ***Note*** *– these do not include the total row or column!*

As before, this test for independence is always a right-tailed test.

15 Enter the expected frequencies next to the corresponding observed frequencies in the table below. For each pair, compute the contribution to the χ^2 statistic, $(O-E)^2/E$ and total these values.

O = Observed Frequency	E = Expected Frequency	$\frac{(O-E)^2}{E}$
96	*99.31*	*0.1103*
43		
34		
35		
9		
9		
24		
3		
17		
	Total:	

16 What is the χ^2 test statistic? This is the sum of the χ^2 contributions. Round your answer to 2 decimal places.

χ^2 = __________

17 What are the degrees of freedom for this test statistic? (Don't count the total row or column!)

$df = (r - 1) \cdot (c - 1)$ = __________

18 Use technology (like this applet: https://carnegiemathpathways.org/go/appchisq *or* *https://carnegiemathpathways.org/go/desmoschi*) to find the *P*-value for this right-tailed test. Round the value to 4 decimal places. **Note**: If completing this problem online, follow the instructions given online to find the *P*-value.

P-value = __________

Step 4: Make a Conclusion

19 What conclusion do you make regarding the null and alternative hypotheses? Why?

20 Write a brief conclusion in the context of this problem.

LET'S SUMMARIZE

- In a chi-square test for independence, we investigate whether two categorical variables are independent.
- The logic, steps, and need to check conditions in a chi-square test for independence are similar to hypothesis testing procedures encountered in previous modules.
- For a chi-square test for independence, the null hypothesis states that two categorical variables are independent, and the alternative hypothesis states that two categorical variables are dependent.
- As we saw for the chi-square test for goodness-of-fit, larger chi-square values generally correspond to smaller *P*-values and stronger evidence against the null hypothesis. Similarly, smaller chi-square values generally correspond to larger *P*-values and weaker evidence against the null hypothesis.

Exercise 7.3

The Pew Research Center studies many different groups in the United States. One of the center's projects is the Pew Internet and American Life Project. In this project, the research center learns how people in the United States use computers and technology.

In one study, researchers asked people, "Do you use a computer at your workplace, at school, at home, or anywhere else on at least an occasional basis?" The possible responses to this question were "Yes" and "No". Researchers also recorded information about each respondent's *urbanity*, that is whether the respondent lived in an "Urban" area (a city), a "Suburban" area (a neighborhood outside city limits), or a "Rural" area (not in a neighborhood).

Researchers obtained the following results, based on a sample of 8,296 individuals:

		Urbanity		
		Urban	**Suburban**	**Rural**
Response (“Do you use a computer?”)	**Yes**	1946	3533	943
	No	537	835	502

Do these data support the claim that there is a relationship between a person's response to the question about computer use and the person's urbanity? Execute a *complete* chi-square test for independence for this case. Use a significance level of α = 0.01.

Step 1: Determine the Hypothesis

1 What are the appropriate hypotheses for this test?

Step 2: Collect the Data

2 The table below displays the row, column and grand totals, and the expected frequencies for all but one cell. Compute and enter in the missing expected frequency. Round the value to two decimal places.

Computer Usage	Urban	Suburban	Rural	Totals
Yes		3381.30	1118.59	6422
No	560.89	986.70	326.41	1874

Step 3: Assess the Evidence

3 Each pair of observed and expected frequencies are provided in the table below. Compute the missing contribution to the χ^2 test statistic. Round the values to two decimal places.

Pairings of Values	O = Observed Frequency	E = Expected Frequency	$\frac{(O-E)^2}{E}$
Urban / Yes	1946		
Urban / No	537	560.89	1.02
Suburban / Yes	3533	3381.30	6.81
Suburban / No	835	986.70	23.32
Rural / Yes	943	1118.59	27.56
Rural / No	502	326.41	94.46

4 All expected frequencies are greater than 5, so we can proceed with the hypothesis test. What is the value of the χ^2 test statistic? Write the value to two decimal places.

5 Use technology to determine the *P*-value. For example, you can use: https://carnegiemathpathways.org/go/appchisq or https://carnegiemathpathways.org/go/desmoschi). **Note**: If completing this problem online, follow the instructions given online to find the *P*-value.

P-value = ___________

Step 4: Make a Decision

6 At the 1% significance level, write an appropriate conclusion.

7.3 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Given sample data, state and check the conditions needed for the use of the chi-square test for independence to be appropriate.	
Compute the value of the test statistic and find the associated *P*-value in a chi-square test for independence.	
Use the *P*-value and the chosen significance level to reach a decision, and interpret the conclusion in context.	
Evaluate whether conclusions are reasonable, given the description of a statistical study and the results of the chi-square test.	

7.4: The Chi-Square Test for Homogeneity in Two-Way Tables

LEARNING GOALS

By the end of this collaboration, you should understand that:

- A chi-square test for homogeneity investigates whether the distribution of a single categorical variable is similar across multiple populations (assuming distinct random sampling from each population).
- A chi-square test for independence investigates the relationship between two categorical variables in a single population (assuming one random sample is selected from the one population of interest).
- In a chi-square test for homogeneity, the null hypothesis states that the distribution of population proportions for a categorical variable is the same in all populations of interest.
- In a chi-square test for homogeneity, the alternative hypothesis states that the distribution of population proportions for a categorical variable is not the same in all populations of interest.
- The degrees of freedom in a chi-square test for homogeneity are related to the number of categories in the row and column variables.
- The procedures for a chi-square test for homogeneity are very similar to the procedures for a chi-square test for independence.

By the end of this collaboration, you should be able to:

- Determine which chi-square test is appropriate (test for independence or test for homogeneity), given data that are summarized in a two-way table and a description of how the data were collected.
- Choose appropriate null and alternative hypotheses for a chi-square test for homogeneity, given a claim about the homogeneity of population proportions.
- State and check the conditions needed for the use of the chi-square test for homogeneity to be appropriate, given sample data.
- Compute the value of the test statistic and find the associated *P*-value in a chi-square test for homogeneity.
- Use the *P*-value and the chosen significance level to reach a decision, and interpret the conclusion in context.
- Evaluate whether the conclusions are reasonable, given the description of a statistical study and the results of the chi-square test.

INTRODUCTION

Movie producers know that people who are knowledgeable about science tend to enjoy science fiction movies more. So where should they promote science fiction movies? Does a person's level of science knowledge depend on where they live? The data below represent 1,000 randomly selected individuals who were sampled from four regions of the country: Northeast, Midwest, South, and West. Each person's science knowledge is self-reported.

Table 1: Regional Survey

Knowledge Level	Northeast	Midwest	South	West
Very	105	95	145	55
Moderately	80	60	100	10
Little or None	115	95	105	35

We are interested in knowing if the distributions of science knowledge levels are *proportionally the same for each region.* In order to answer this question, we conduct a *chi-square test for homogeneity.*

The Chi-Square Test for Homogeneity

The **chi-square test for homogeneity** tests the claim that the distributions of outcomes of a categorical variable are the same, proportionally, across multiple populations. The test assumes that distinct, independent random samples have been taken from each population.

Note that the chi-square test for homogeneity is different from a chi-square test for independence. A chi-square test for independence investigates the relationship between two categorical variables in a single population using one random sample selected from the population. The test for homogeneity considers the distribution of a single categorical variable across multiple populations, using independent random samples from each population.

Even though the two tests are different, the process for conducting a test for homogeneity is the same as for a test for independence.

Step 1: Determine the Hypotheses

For a chi-square test for homogeneity, the null and alternative hypotheses state that:

H_0: The distributions of outcomes of a categorical variable are the same, proportionally, for each population.

H_a: The distributions of outcomes of a categorical variable are not the same, proportionally, for each population.

1 We will perform a chi-square test for homogeneity using the self-reported science knowledge and geographic region data in the two-way table above. Determine the hypotheses.

H_0:

H_a:

Step 2: Collect the Data

The test for homogeneity requires that the data from the independent random samples be summarized in a two way table. Row and column totals should be computed so that expected frequencies can be computed.

As before, the expected frequencies are computed with the formula

$$E = \frac{row\ total \cdot column\ total}{grand\ total}$$

As with the previous goodness of fit test, we require that each expected frequency is at least 5.

2 Enter the row totals for the two way table below.

Knowledge Level	Northeast	Midwest	South	West	Totals
Very	105	95	145	55	
Moderately	80	60	100	10	
Little or None	115	95	105	35	
Totals					

3 Now enter the column totals and grand total in the table.

4 Compute the expected frequencies for rows A, B, and C.

	Knowledge Level	Northeast	Midwest	South	West
A	Very				
B	Moderately				
C	Little or None				

5 Are the criteria satisfied for a chi-square test for homogeneity?

Step 3: Assess the Evidence

The test statistic for a χ^2 goodness of fit test is: $\chi^2 = \sum \frac{(O-E)^2}{E}$

This is approximately distributed according to the χ^2 distribution with degrees of freedom equal to:

$$df = (r-1)\cdot(c-1)$$

Here, r is the number of rows in the table and c is the number of columns. ***Note*** *– these do not include the total row or column!*

As before, this goodness of fit test is always a right-tailed test.

6 Enter the expected frequencies next to the corresponding observed frequencies in the table below. For each pair, compute the contribution to the χ^2 statistic, $(O-E)^2/E$ and total these values.

A Northeast Region

O = Observed Frequency	E = Expected Frequency	$\frac{(O-E)^2}{E}$

B Midwest Region

O = Observed Frequency	E = Expected Frequency	$\frac{(O-E)^2}{E}$

C South Region

O = Observed Frequency	E = Expected Frequency	$\frac{(O-E)^2}{E}$

D West Region

O = Observed Frequency	E = Expected Frequency	$\frac{(O-E)^2}{E}$

E Give the chi-square test statistic. This is the sum of the χ^2 contributions from all four regions.

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \_\_\_\_\_\_$$

F Give the degrees of freedom for this statistic.

df = ________

G Use technology to determine the *P*-value. (One applet for this is at: https://carnegiemathpathways.org/go/appchisq or https://carnegiemathpathways.org/go/desmoschi.) Round the value to four decimal places.
Note: If completing this problem online, follow the instructions given online to find the *P*-value.

P-value = ________

Step 4: Make a Conclusion

7 We will perform this test at the 1% level of significance. When comparing the *P*-value to α, what conclusion do you make regarding the hypotheses?

8 Write a brief conclusion in the context of this problem.

LET'S SUMMARIZE

- In a chi-square test for homogeneity, we investigate whether the distribution of a single categorical variable is similar across multiple populations.
- The logic, steps, and need to check conditions in a chi-square test for homogeneity are similar to hypothesis testing procedures encountered in previous modules (and are especially similar to the chi-square test for independence).
- For a chi-square test for homogeneity, the null hypothesis states that the distribution of the population proportions for the categorical variable is the same for all populations of interest. The alternative hypothesis states that the distribution of the population proportions for the categorical variable is not the same for all populations of interest.
- As we saw for the chi-square test for goodness-of-fit and the chi-square test for independence, larger chi-square values generally correspond to smaller *P*-values and stronger evidence against the null hypothesis. Similarly, smaller chi-square values generally correspond to larger *P*-values and weaker evidence against the null hypothesis.

Exercise 7.4

The governor of the state of Cornland, Martin Corn, ran for re-election in November, 2022. Before the election, from September 22–26, 2022, a national newspaper conducted a poll in which respondents were asked: "If the election for Cornland governor were held today, for whom would you vote?"

In an article summarizing the results of the poll, journalists reported that, "A total of 1,448 randomly selected adults in Cornland were interviewed, including… 730 voters likely to cast ballots." The sample was intended to represent the population of likely Cornland voters at that time period in September, 2022. The article summarized the responses of the likely voters using the following categories: "for Governor Martin Corn", "for Bob Maize", and "for Other/ no opinion".

The following data summarize the responses from a sample of 730 likely voters (simulated based on summary data):

for Governor Martin Corn	382 responses
for Bob Maize	302 responses
for Other/ no opinion	46 responses

A subsequent newspaper poll was conducted from October 19–22, 2022. In this poll, "2,355 randomly selected adults in Cornland were interviewed, including… 1,434 voters likely to cast ballots." Like the September sample, the October sample was intended to represent the population of likely Cornland voters for that October time period.

The following data summarize the responses from a sample of 1,434 likely voters (simulated based on summary data):

for Governor Martin Corn	774 responses
for Bob Maize	574 responses
for Other/ no opinion	86 responses

1 Suppose we are interested in investigating the similarity of the results of these two polls. Explain why the proper test is a chi-square test for homogeneity rather than a chi-square test for independence.

2 There is a claim that the distribution of responses for the population of likely Cornland voters for the October 19–22 time period is different from the distribution of responses for the population of likely Cornland voters for the September 22–26 time period. Do the data provide sufficient evidence to support that claim? We will perform a chi-square test for homogeneity at a significance level of 5%.

A table summarizing the data is provided below:

Candidate	Sept. 22-26, 2022	Oct. 19-22, 2022	Totals
Governor Martin Corn	382	774	1156
Bob Maize	302	574	876
Other/No Opinion	46	86	132
Totals:	730	1434	2164

What is the appropriate hypothesis for this test?

H_0:

H_a:

3 Compute the missing expected frequency in the table below. Round the value to two decimal places.

Candidate	Sept. 22-26, 2022	Oct. 19-22, 2022	Totals
Governor Martin Corn	389.96	766.04	1156
Bob Maize	295.51	580.49	876
Other/No Opinion	44.53		132
Totals:	730	1434	2164

4 Are the criteria satisfied for a chi-square test for homogeneity?

5 The observed frequencies from the voting results are listed next to the expected frequencies in the table below. Compute the missing contribution to the χ^2 statistic. Round the value to two decimal places.

Time Period	*O* = Observed Frequency	*E* = Expected Frequency	$\frac{(O-E)^2}{E}$
Sept. 22-26, 2022	382	389.96	0.16
Sept. 22-26, 2022	302	295.51	0.14
Sept. 22-26, 2022	46	44.53	0.05
Oct. 19-22, 2022	774	766.04	
Oct. 19-22, 2022	574	580.49	0.07
Oct. 19-22, 2022	86	87.47	0.02

6 Compute the test statistic for a χ^2 test for homogeneity. Round the test statistic to 2 decimal places.

7 Give the degrees of freedom for this statistic.

8 Use technology to determine the *P*-value. For example, you can use: https://carnegiemathpathways.org/go/appchisq or https://carnegiemathpathways.org/go/desmoschi). Round the value to three decimal places. **Note**: If completing this problem online, follow the instructions given online to find the *P*-value.

P-value = _________

9 We will perform this test at the 5% level of significance. What conclusion do you make regarding the hypotheses?

10 Write an appropriate conclusion.

7.4 Monitor (survey)

Monitor your progress on learning the objectives for this unit. If you identify any objectives you need to review, go back through the unit's activities. Your responses are not graded, but will be available for your instructor to see.

Rate how confident you are on a scale of 1–5 (1 = not confident and 5 = very confident).

Skill or Concept: I can …	**Rating from 1 to 5**
Determine which chi-square test is appropriate.	
Choose appropriate null and alternative hypotheses for a chi-square test for homogeneity.	
State and check the conditions needed for the use of the chi-square test for homogeneity to be appropriate, given sample data.	
Compute the value of the test statistic and find the associated *P*-value in a chi-square test for homogeneity.	
Use the *P*-value and the chosen significance level to reach a decision, and interpret the conclusion in context.	