STATWAY™ INSTRUCTOR NOTES

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

INSTRUCTOR SPECIFIC MATERIAL IS INDENTED AND APPEARS IN GREY

ESTIMATED TIME

50 minutes

MATERIALS REQUIRED

- Students need calculators for this lesson

BRIEF DESCRIPTION

This lesson introduces the concept of a chi-square test statistic as a means of measuring the extent to which the distribution of proportions observed in a sample differs from the expected distribution of proportions based on hypothesized population parameters.

LEARNING GOALS

Students will understand that:

- Larger chi-square values give greater cause to question an initial claim about a variable's distribution.
- Categories with the greatest "observed–expected" differences relative to the expected count contribute the most to the chi-square statistic's value.
- The default claim (null hypothesis) is that the variable is distributed as expected; the alternative is that there are departures from the expectation.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

Students will be able to:

- Recognize hypothesis testing situations where a chi-square goodness-of-fit test can be used to answer a question of interest.
- Compute the value of the test statistic in a chi-square goodness-of-fit test.

INTRODUCTION

In this series of tasks, you are not formally developing the full framework for executing a chi-square goodness-of-fit test. Rather, one categorical variable's population distribution is stated (as percentages), and students are asked to compute expected counts for each category of a hypothetical sample of 1,000 items based on the variable's stated population distribution. Using other hypothetical samples, students then consider the following:

- How different an observed count must be from an expected count in order to be judged a cause for concern, and
- If it matters how these discrepancies between observed and expected counts are spread out (i.e., Does it matter if just two categories are off by quite a bit? Is it worse if all categories are off but only by a small amount?).

Remind students that in a previous module, they were tasked with designing a children's cereal that struck an effective balance of taste and nutrition in order to receive a favorable rating from *Consumer Reports*. Consider showing a box of a multicolored cereal where the cereal pieces are fairly identical except for their colors. Pour out some of the contents if you wish. Have students envision what the cereal would look like if it were entirely composed of just one of the colors or if one color were considerably more (or less) dominant. Explain that presentation of the cereal is potentially very important in terms of its marketability (i.e., even if the cereal is healthy or tastes good, children may not be interested in an unappealing looking cereal).

Using the introductory material that is part of the student handout (also given below), have students imagine what a bowl of cereal with the desired characteristics may look like. As in previous modules, have students work on the task alone for a few minutes and then in small groups for slightly longer.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

STUDENT MATERIAL IS NOT INDENTED AND APPEARS IN BLACK

INTRODUCTION

Manufacturing Cereal

In a previous lesson, you were asked to design a children's cereal that struck an effective balance of taste and nutrition in order to receive a favorable rating from *Consumer Reports*. In this lesson, you will discuss some aspects of effectively marketing the cereal to consumers. Effective marketing is important because the goal is to keep costs low while increasing sales and profits.

It costs too much to make many different kinds of shapes for cereal. Because of manufacturing costs and concerns, all the cereal pieces, in the cereal that you develop, will be the same in terms of size, shape, and weight. However, suppose that the research shows that certain colors and certain distributions of colors are more appealing to consumers and are relatively inexpensive to produce.

Assume that the desired color distribution for the cereal is **20% red, 35% white,** and **45% blue.** To check these proportions, you take a simple random sample of 1,000 cereal pieces and count how many of each color are present in the sample. When there is strong evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue this implies that something may be wrong with the production process.

TRY THESE

1 How many pieces of each color would you expect in your sample of 1,000 pieces if the true population proportions were in fact 20% red, 35% white, and 45% blue? (**Note:** These three expected counts should add up to 1,000 since you have 1,000 pieces in your sample.)

> **Answer:** The expected count of red should be $0.20 \cdot 1{,}000 = 200$; white: $0.35 \cdot 1{,}000 = 350$; blue: $0.45 \cdot 1{,}000 = 450$. Expected counts should add up to 1,000 since there are 1,000 pieces in the sample.

2 If you obtained a sample of 1,000 pieces that exactly matched the three counts you computed in Question 1, it would *not* suggest that there are any problems in the manufacturing process since your actual counts were the same as your expected counts.

Recall that you have discussed sampling variability in a previous module. You know that it is unusual to obtain a sample with the exact counts you just computed in Question 1.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

A For a sample of 1,000 pieces, give another example of counts for red, white, and blue pieces that *would not* suggest that there are any problems in the manufacturing process.

Red	White	Blue

> **Answer:** approximately red = 200, white = 350, and blue = 450 (i.e., answers vary but differences between observed and expected counts are small).

B For a sample of 1000 pieces, give an *example* of counts for red, white, and blue pieces that would suggest that there are some problems in the manufacturing process.

Red	White	Blue

> **Answer:** answers should be very different from red = 200, white = 350, blue = 450 (i.e., differences between observed and expected counts are large).

3 The tables that follow show the distributions of four random samples. Each sample has a size of 1,000 and is taken from four different production days.

Do the following for each sample:

- First, examine the distribution of the actual counts obtained from the sample in the table.
- Decide whether you think the sample provides strong evidence, moderate evidence, or weak evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue and record your response in the space provided.
- Finally, list the characteristics of the sample that led you to your decision.

Sample: Day 1

Red	White	Blue
210	360	430

It appears that the Day 1 sample provides *(circle one)*

strong **moderate** **weak**

evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue because …

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

Sample: Day 2

Red	White	Blue
200	425	375

It appears that the Day 2 sample provides *(circle one)*

strong **moderate** **weak**

evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue because …

Sample: Day 3

Red	White	Blue
180	360	460

It appears that the Day 3 sample provides *(circle one)*

strong **moderate** **weak**

evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue because …

Sample: Day 4

Red	White	Blue
235	355	410

It appears that the Day 4 sample provides *(circle one)*

strong **moderate** **weak**

evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue because …

> **Answer:** The first and third samples have small differences between observed and expected counts and thus provide weak or no evidence *against* the claim that the

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

> population proportions are 20% red, 35% white, and 45% blue. On the other hand, the second and fourth samples have large differences between observed and expected counts and thus provide strong evidence against the claim that the population proportions are 20% red, 35% white, and 45% blue.

4 Based on your work in Question 3, which sample gives the strongest evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue? What characteristics of the sample were most important in your decision?

> **Answer:** Sample 2 gives the strongest evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue due to the fact that the white and blue categories represent such large discrepancies between the expected and observed. Sample 4 is also a possible choice.

5 Sometimes, in the production process, you may not get the distribution of 20% red, 35% white, and 45% blue that you had hoped to achieve. When might you become concerned about a large difference? Would you be concerned if an actual sample count differed from its expected count by about 10 pieces? How about by about 50 pieces or 100 pieces? Why? Explain your reasoning.

> **Answer:** See the Part 1 Wrap-Up for some suggestions for discussion and recording responses. Encourage students to share their reasoning requested in Question 5 and some creative ideas developed in addressing Question 6. Keep in mind that Part II is where you address any errors in student reasoning and computation.

6 It would be useful to have a statistical measure of deviation. The statistical measure of deviation helps you decide how much the distribution of a sample (such as those shown previously) deviates from what is expected.
Create a method (and/or a statistic) to *measure* which sample deviates the most from the ideal expected distribution.

> **Answer:** See the Part 1 Wrap-Up for some suggestions for discussion and recording responses. Encourage students to share their reasoning requested in Question 5 and some creative ideas developed in addressing Question 6. Keep in mind that Part II is where you address any errors in student reasoning and computation.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

WRAP-UP

As has been the case in previous lessons, the previous task should provide students the opportunity to struggle with the important ideas—in this case, what are the expected counts for a sample of size 1,000 based on the original claim and how much deviation from those expected counts is reasonable. Students may have considered formal sampling distribution logic from previous modules in determining how much of a deviation is reasonable from an expected count, but that is not required. In Part II, you will address any errors in student reasoning and computation.

Consider polling the class on its responses to Question 4 (consider even using a categorical variable graph covered earlier in the course to chart the responses). Also consider allowing students to discuss as a group the sample characteristics that influenced their decisions in Question 4 and their reasoning in Question 5. Encourage comparisons of the deviations *and* examination of the deviations relative to the corresponding expected counts, particularly in light of Question 6.

Have students share some of their inventions and ideas regarding Question 6, mindful that some methods may not work well (e.g., just adding up the deviations will not work; they always add to 0). Also, bring up the fact that in Samples 1 and 3 the sum of the squares of the deviations (also the sum of the absolute value of the deviations) are the same, but the sample on the right deviates more from the ideal because a deviation of 20 from an expected count of 200 (red) is a bigger deal than a deviation of 20 from a larger expected count 450 (blue). This also serves as preparation for Question 11.
If appropriate, consider reminding students that when similar deviation measures were discussed earlier in the course, *squaring* was often employed (e.g., deviations are squared as a step in computing standard deviation, least-squares regression).

INTRODUCTION

Tell students that they are now going to develop a way of more carefully quantifying some of the previous concepts. For the following problems where the final chi-square value is provided, encourage students to seek assistance (if needed) until they obtain calculations that are consistent with the presented values. Answer any questions and address any issues that come up regarding the calculation of differences, squares, quotients, and sums and/or general order or operations.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

Quantifying the Strength of the Evidence

Earlier in this lesson, we talked about strong evidence against a claim about population proportions. Strong evidence against the claim that the population proportions are 20% red, 35% white, and 45% blue implies that something may be wrong with the production process.

For each sample, you can compute a chi-square statistic (*chi* is pronounced *ki* as in *kite*). The chi-square statistic is useful in assessing the strength of your evidence that something might be wrong with the production process. The chi-square statistic relies on the comparison of *how different an observed count in a given category is from the count that was expected for that given category*.
Keep this information in mind. In later lessons, you will use this statistic in a way that is similar to how you used *Z*-scores for testing hypotheses in previous modules.

Important Information: The steps for computing a chi-square value (written as χ^2) are as follows:

(1) For each category, compute the difference between the observed, or actual, count for that category (obtained from the sample) and the expected count for that category:

Observed Count – Expected Count

(2) For each category, compute the square of this difference obtained in Step 1:

$(\textit{Observed Count} - \textit{Expected Count})^2$

(3) For each category, divide the squared difference obtained in Step 2 by the expected count for the category:

$(\textit{Observed Count} - \textit{Expected Count})^2/\textit{Expected Count}$

(4) Add up the Step 3 calculation results from each category; this is the chi-square value.

Example

Using the Day 1 sample data set as an example:

	Red	**White**	**Blue**
Observed Count (from sample)	210	360	430
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count - Expected Count	10	10	-20
Step 2: (Observed Count - Expected Count)^2	100	100	400
Step 3: (Observed Count - Expected Count)^2/Expected Count	0.5	0.286	0.889

Step 4: 0.5 + 0.286 + 0.889 = 1.675

For the Day 1 sample, the chi-square value generated is 1.675.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

Use the steps above to answer the following questions:

7 Compute the chi-square values for the other samples (Days 2–4) by filling in the tables below. Compute the Step 3 calculations to three decimal places as shown in the previous example.

Day 2

	Red	White	Blue
Observed Count (from sample)	200	425	375
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count – Expected Count	0	75	-75
Step 2: (Observed Count –Expected Count)^2	0	5625	5625
Step 3: (Observed Count – Expected Count)^2/Expected Count	0		

Step 4: 0 + + = 28.571

For the Day 2 sample, the chi-square value generated is 28.571.

Day 3

	Red	White	Blue
Observed Count (from sample)	180	360	460
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count - Expected Count			10
Step 2: (Observed Count - Expected Count)^2	400		100
Step 3: (Observed Count - Expected Count)^2/Expected Count			

Step 4: + + =

For the Day 3 sample, the chi-square value generated is _____________.

Day 4

	Red	White	Blue
Observed Count (from sample)	235	355	410
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count – Expected Count			
Step 2: (Observed Count –Expected Count)^2			
Step 3: (Observed Count – Expected Count)^2/Expected Count			

Step 4: + + = 9.752

For the Day 4 sample, the chi-square value generated is 9.752.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

Answer:

Day 2

	Red	White	Blue
Observed Count (from sample)	200	425	375
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count - Expected Count	0	75	-75
Step 2: (Observed Count - Expected Count)^2	0	5625	5625
Step 3: (Observed Count - Expected Count)^2/Expected Count	0	16.071	12.500

chi-square value = 28.571

Day 3

	Red	White	Blue
Observed Count (from sample)	180	360	460
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count - Expected Count	-20	10	10
Step 2: (Observed Count - Expected Count)^2	400	100	100
Step 3: (Observed Count - Expected Count)^2/Expected Count	2	0.286	0.222

chi-square value = 2.508

Day 4

	Red	White	Blue
Observed Count (from sample)	235	355	410
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count - Expected Count	35	5	-40
Step 2: (Observed Count - Expected Count)^2	1225	25	1600
Step 3: (Observed Count - Expected Count)^2/Expected Count	6.125	0.071	3.556

chi-square value = 9.752

8 Which of the daily samples had the highest chi-square value?

First, think about the size of the chi-square values for these four samples. Do they appear to correspond in any way to the initial strength-of-evidence statements that you made in Question 3 about the cereal population proportions?

Answer: Sample 2 had the highest chi-square value. Students should begin to see that higher chi-square values are stronger evidence against the claim that the population proportions of the cereal pieces are 20% red, 35% white, and 45% blue.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

---

9 What happens if one category has an observed count and an expected count that are exactly the same? Is it possible to still have a large chi-square value generated by that sample? If so, make up a sample that shows an observed count and an expected count that are exactly the same. Did any of the samples in Question 7 exhibit this behavior?

> **Answer**: Yes, it is possible that if one category has an observed count and expected count that are exactly the same, a large chi-square value is still possible for that sample. Sample 2 is an example.

10 Now suppose that <u>no</u> category in a sample has a case where the observed count and expected count are exactly the same. Is it still possible to have a small chi-square value generated by that sample? If so, make up a sample that shows how this could happen. Did any of the samples in Question 7 exhibit this behavior?

> **Answer**: Yes, it is possible. See Samples 1 and 3.

11 In the Day 2 data set, the white and blue category counts had an actual count that differed from the expected count by 75 pieces.

Which difference made a larger contribution to the size of the chi-square value:

- the difference of size 75 relative to an expected count of 350 (white)
- or the difference of size 75 relative to an expected count of 450 (blue)?

In other words, would a difference of 75 make a larger contribution to the size of the chi-square value for a category with an expected count of 350 than it would be for a category with an expected count of 450? Explain your thinking.

> **Answer**: The difference of size 75 relative to the expected count of 350 (white) had a bigger effect on the chi-square statistic (its contribution was 16.071) than the difference of size 75 relative to the expected count of 450 (blue—its contribution was 12.5). The difference of 75 was a bigger deal for a category with an expected count of 350 (a smaller expected count) than it was for a category with an expected count of 450 (a larger expected count).
>
> **WRAP-UP**
>
> Although expected counts are provided in the fill-in-the-table cases, make sure that students see that the expected counts are the hypothesized proportions multiplied by the sample size. Monitor for any difficulty with the calculation of differences, squares, quotients, and sums and/or general order or operations. Encourage students to examine how the size of a chi-square statistic might be related to the strength of evidence against the claim that the population is distributed as expected.

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

STUDENT NAME ______________________________ **DATE** ______________

TAKE IT HOME

1 Based on the samples you worked on today, answer the following questions:

A Which seems to provide greater evidence *against* the claim that the population proportions are 20% red, 35% white, and 45% blue:

- a high chi-square value (such as you computed for Days 2 and 4)
- or a low chi-square value (such as you computed for Days 1 and 3)?

> **Answer**: A higher chi-square value (such as those computed for Days 2 and 4) provides greater evidence against the claim that the population proportions are 20% red, 35% white, and 45% blue.

B Earlier in this lesson we learned that strong evidence against the claim that the population proportions are 20% red, 35% white, and 45% blue implies that something may be wrong with the production process.
Based on your answer above, would a high chi-square value or a low chi-square value give stronger evidence that something may be wrong with the production process?

> **Answer**: A high chi-square value gives stronger evidence that something may be wrong with the production process.

C Is it possible to obtain a *negative* chi-square statistic? If so, explain how. If not, explain why not.

> **Answer:** It is not possible to obtain a negative chi-square statistic. Expected counts cannot be negative, and the square of an observed–expected difference cannot be negative either.

2 On Day 5, the following sample was collected:

Red	White	Blue
178	364	458

Compute the chi-square value. Determine if this value provides strong evidence against the claim that the population proportions for the cereal pieces are 20% red, 35% white, and 45% blue.

(**Note**: For reasons that will be explained in a future lesson, in a case where the categorical variable contains three categories:

Lesson 11.1.1
Introduction to Chi-Square Tests for One-Way Tables

- consider a chi-square value of 5.99 or greater to be statistically significant evidence against the claim that the population proportions of cereal pieces are 20% red, 35% white, and 45% blue.)

Answer:

Day 5

	Red	White	Blue
Observed Count (from sample)	178	364	458
Expected Count (based on desired population distribution)	200	350	450
Step 1: Observed Count - Expected Count	-22	14	8
Step 2: (Observed Count - Expected Count)^2	484	196	64
Step 3: (Observed Count - Expected Count)^2/Expected Count	2.42	0.560	0.142

chi-square value = 3.122

This value does not provide strong evidence against the claim that the population proportions for the cereal pieces are 20% red, 35% white, and 45% blue. It is a small chi-square value and it is less than the critical chi-square value of 5.99 mentioned in the comment that accompanies the question. (The P-value here is about 0.21.)