Section 11.1
Chi-Square Goodness-of-Fit Tests

State: **Carrying Out a Test**

H_0: The specified distribution of the categorical variable is correct.
H_a: The specified distribution of the categorical variable is not correct.

We can also write these hypotheses symbolically using p_i to represent the proportion of individuals that fall in category *i*:

H_0: p_1 = ___, p_2 = ___, …, p_k = ___.
H_a: At least one of the p_i's is incorrect.

Plan:

Conditions: Use the chi-square goodness-of-fit test when

- ✓ **Random** The data come from a random sample or a randomized experiment.
- ✓ **10%**: When sampling without replacement, check that $n \le 0.10N$.
- ✓ **Large Sample Size** All expected counts are at least 5.

Do:

Start by finding the expected count for each category assuming that H_0 is true. Then calculate the chi-square statistic

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

where the sum is over the k different categories. The P-value is the area to the right of χ^2 under the density curve of the chi-square distribution with $k - 1$ degrees of freedom.

Conclude:

If P-value is < α reject Ho and state your findings in the context of the problem.

Feb 14-3:24 PM

The **one-way table** below summarizes the data from a sample bag of M&M'S Milk Chocolate Candies. In general, one-way tables display the distribution of a categorical variable for the individuals in a sample.

Color	Blue	Orange	Green	Yellow	Red	Brown	Total
Count	9	8	12	15	10	6	60

The sample proportion of blue M&M's is $\hat{p} = \frac{9}{60} = 0.15$.

Since the company claims that 24% of all M&M'S Milk Chocolate Candies are blue, we might believe that something fishy is going on. We could use the one-sample *z* test for a proportion from Chapter 9 to test the hypotheses

H_0: $p = 0.24$
H_a: $p \ne 0.24$

where *p* is the true population proportion of blue M&M'S. We could then perform additional significance tests for each of the remaining colors.

However, performing a one-sample *z* test for each proportion would be pretty inefficient and would lead to the problem of multiple comparisons.

Feb 14-12:09 PM

Chi-Square Goodness-of-Fit Tests

The null hypothesis in a chi-square goodness-of-fit test should state a claim about the distribution of a single categorical variable in the population of interest. In our example, the appropriate null hypothesis is

H_0: The company's stated color distribution for M&M'S Milk Chocolate Candies is correct.

The alternative hypothesis in a chi-square goodness-of-fit test is that the categorical variable does *not* have the specified distribution. In our example, the alternative hypothesis is

H_a: The company's stated color distribution for M&M'S Milk Chocolate Candies is **not** correct.

or

We can also write the hypotheses in symbols as

H_0: $p_{\text{blue}} = 0.24$, $p_{\text{orange}} = 0.20$, $p_{\text{green}} = 0.16$, $p_{\text{yellow}} = 0.14$, $p_{\text{red}} = 0.13$, $p_{\text{brown}} = 0.13$

H_a: At least one of the p_i's is incorrect.

Feb 14-12:13 PM

The idea of the chi-square goodness-of-fit test is this: we compare the **observed counts** from our sample with the counts that would be expected if H_0 is true. The more the observed counts differ from the **expected counts**, the more evidence we have against the null hypothesis.

In general, the expected counts can be obtained by multiplying the proportion of the population distribution in each category by the sample size.

Example 1: Jerome's class collected data from a random sample of 60 M&M'S Milk Chocolate Candies. Calculate the expected counts for each color. Show your work.

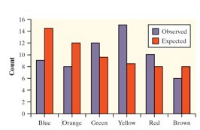
Assuming that the color distribution stated by Mars, Inc., is true, 24% of all M&M'S Milk Chocolate Candies produced are blue. For random samples of 60 candies, the average number of blue M&M'S should be (60)(0.24) = 14.40. This is our expected count of blue M&M'S® Chocolate Candies. Using this same method, we find the expected counts for the other color categories:

Orange: (0.20)(60) = 12.00
Green: (0.16)(60) = 9.60
Yellow: (0.14)(60) = 8.40
Red: (0.13)(60) = 7.80
Brown: (0.13)(60) = 7.80

Color	Observed	Expected
Blue	9	14.40
Orange	8	12.00
Green	12	9.60
Yellow	15	8.40
Red	10	7.80
Brown	6	7.80

Feb 14-12:15 PM

To see if the data give convincing evidence against the null hypothesis, we compare the observed counts from our sample with the expected counts assuming H_0 is true. If the observed counts are far from the expected counts, that's the evidence we were seeking.



We see some fairly large differences between the observed and expected counts in several color categories. How likely is it that differences this large or larger would occur just by chance in random samples of size 60 from the population distribution claimed by Mars, Inc.?

To answer this question, we calculate a statistic that measures how far apart the observed and expected counts are. The statistic we use to make the comparison is the **chi-square statistic.**

Feb 14-12:24 PM

The **chi-square statistic** is a measure of how far the observed counts are from the expected counts. The formula for the statistic is

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$

where the sum is over all possible values of the categorical variable.

Example 2: Use the formula below to compare the observed and expected counts for Jerome's sample. Calculate the chi-square statistic.

Color	Observed	Expected
Blue	9	14.40
Orange	8	12.00
Green	12	9.60
Yellow	15	8.40
Red	10	7.80
Brown	6	7.80

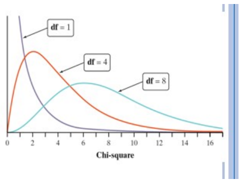
$$\chi^2 = \frac{(9-14.40)^2}{14.40} + \frac{(8-12.00)^2}{12.00} + \frac{(12-9.60)^2}{9.60} + \frac{(15-8.40)^2}{8.40} + \frac{(10-7.80)^2}{7.80} + \frac{(6-7.80)^2}{7.80}$$

$$\chi^2 = 2.025 + 1.333 + 0.600 + 5.186 + 0.621 + 0.415 = 10.180$$

Feb 14-12:21 PM

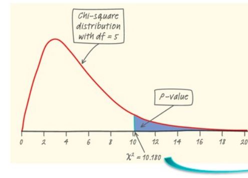
Think of χ^2 as a measure of the distance of the observed counts from the expected counts. Large values of χ^2 are stronger evidence against H_0 because they say that the observed counts are far from what we would expect if H_0 were true. Small values of χ^2 suggest that the data are consistent with the null hypothesis.

The chi-square statistic is *not* a Normal distribution. It is a right-skewed distribution that allows only positive values because χ^2 can never be negative.



Feb 14-12:23 PM

Example 3: In the last example, we computed the chi-square statistic for Jerome's random sample of 60 M&M'S Milk Chocolate Candies: $\chi^2 = 10.180$. Now let's find the *P*-value and state a conclusion.



To find the *P*-value, use Table C and look in the df = 5 row.

	P		
df	**.15**	**.10**	**.05**
4	6.74	7.78	9.49
5	8.12	9.24	11.07
6	9.45	10.64	12.59

Fail to reject H_0. Since our *P*-value is between 0.05 and 0.10, it is greater than $\alpha = 0.05$. There is not enough evidence to suggest that the company's claimed color distribution is incorrect.

In the calculator enter L_1 and L_2
Then select x^2GOF-Test

Feb 14-5:19 PM

A test for equal proportions

Christian Petersen/Getty Images

In his book *Outliers,* Malcolm Gladwell suggests that a hockey player's birth month has a big influence on his chance to make it to the highest levels of the game. Specifically, because January 1 is the cut-off date for youth leagues in Canada (where many National Hockey League (NHL) players come from), players born in January will be competing against players up to 12 months younger. The older players tend to be bigger, stronger, and more coordinated and hence get more playing time, more coaching, and have a better chance of being successful.

To see if birth date is related to success (judged by whether a player makes it into the NHL), a random sample of 80 NHL players from a recent season was selected and their birthdays were recorded. The one-way table below summarizes the data on birthdays for these 80 players:

Birthday:	Jan–Mar	Apr–Jun	Jul–Sep	Oct–Dec
Number of players:	32	20	16	12

STATE: We want to perform a test of

H_0: The birthdays of all NHL players are evenly distributed across the four quarters of the year.

H_a: The birthdays of all NHL players are not evenly distributed across the four quarters of the year. No significance level was specified, so we'll use $\alpha = 0.05$.

PLAN: If the conditions are met, we will perform a chi-square test for goodness of fit.

- ***Random:*** The data came from a random sample of NHL players.
  - 10%: Because we are sampling without replacement, there must be at least 10(80) = 800 NHL players. In the season when the data were collected, there were 879 NHL players.
- ***Large Counts:*** If birthdays are evenly distributed across the four quarters of the year, then the expected counts are all 80(1/4) = 20. These counts are all at least 5.

DO:

- *Test statistic:*

$$\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$$
$$= \frac{(32-20)^2}{20} + \frac{(20-20)^2}{20} + \frac{(16-20)^2}{20} + \frac{(12-20)^2}{20}$$
$$= 7.2 + 0 + 0.8 + 3.2 = 11.2$$

- *P*-value: **Figure 11.5** displays the *P*-value for this test as an area under the chi-square distribution with 4 − 1 = 3 degrees of freedom.

Feb 14-5:20 PM

- *P*-value: **Figure 11.5** displays the *P*-value for this test as an area under the chi-square distribution with 4 − 1 = 3 degrees of freedom.

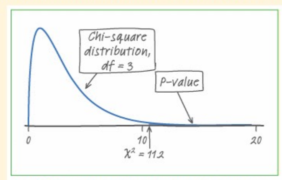


Figure 11.5 The *P*-value for the chi-square test for goodness of fit with $\chi^2 = 11.2$ and df = 3.

As the excerpt at right shows, $\chi^2 = 11.2$ corresponds to a *P*-value between 0.01 and 0.02.

Using Technology: Refer to the Technology Corner that follows the example. The calculator's χ^2 GOF-Test gives $\chi^2 = 11.2$ and *P*-value = 0.011 using df = 3.

	p		
df	**0.02**	**0.01**	**0.005**
2	7.82	9.21	10.60
3	9.84	11.34	12.84
4	11.67	13.28	14.86

CONCLUDE: Because the *P*-value, 0.011, is less than $\alpha = 0.05$, we reject H_0. We have convincing evidence that the birthdays of NHL players are not evenly distributed across the four quarters of the year.

Feb 14-5:24 PM

TI-83/84: Press STAT, arrow over to **TESTS** and choose χ^2 GOF-Test...

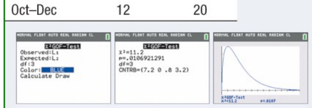
Technology Corner Videos

You can use the TI-83/84 or TI-89 to perform the calculations for a chi-square test for goodness of fit. We'll use the data from the hockey and birthdays example to illustrate the steps.

1. Enter the counts.
   - Enter the observed counts in L1/list1. Enter the expected counts in L2/list2.
2. Perform a chi-square test for goodness of fit.

Enter the inputs shown below. If you choose **Calculate,** you'll get a screen with the test statistic, *P*-value, and df. If you choose the **Draw** option, you'll get a picture of the appropriate chi-square distribution with the test statistic marked and shaded area corresponding to the *P*-value.

Birthday	Observed	Expected
Jan–Mar	32	20
Apr–Jun	20	20
Jul–Sep	16	20
Oct–Dec	12	20



Feb 14-5:26 PM

CHECK YOUR UNDERSTANDING

Biologists wish to mate pairs of fruit flies having genetic makeup RrCc, indicating that each has one dominant gene (R) and one recessive gene (r) for eye color, along with one dominant (C) and one recessive (c) gene for wing type. Each offspring will receive one gene for each of the two traits from each parent. The following table, known as a *Punnett square,* shows the possible combinations of genes received by the offspring:

	Parent 2 passes on:			
Parent 1 passes on:	**RC**	**Rc**	**rC**	**rc**
RC	RRCC (x)	RRCc (x)	RrCC (x)	RrCc (x)
Rc	RRCc (x)	RRcc (y)	RrCc (x)	Rrcc (y)
rC	RrCC (x)	RrCc (x)	rrCC (z)	rrCc (z)
rc	RrCc (x)	Rrcc (y)	rrCc (z)	rrcc (w)

Any offspring receiving an R gene will have red eyes, and any offspring receiving a C gene will have straight wings. So based on this Punnett square, the biologists predict a ratio of 9 red-eyed, straight-winged (x):3 red-eyed, curly-winged (y):3 white-eyed, straight-winged (z):1 white-eyed, curly-winged (w) offspring.

To test their hypothesis about the distribution of offspring, the biologists mate a random sample of pairs of fruit flies. Of 200 offspring, 99 had red eyes and straight wings, 42 had red eyes and curly wings, 49 had white eyes and straight wings, and 10 had white eyes and curly wings. Do these data differ significantly from what the biologists have predicted? Carry out a test at the $\alpha = 0.01$ significance level.

Correct Answer

S: H_0: The distribution of eye color and wing shape is the same as what the biologists predict versus H_a: The distribution of eye color and wing shape is not what the biologists predict. *P*: Chi-square test for goodness of fit. Random: Random sample. 10%: $n = 200 < 10\%$ of all fruit flies. Large Counts: 112.5, 37.5, 37.5, 12.5 all ≥ 5. *D*: $\chi^2 = 6.1867$, df = 3, the *P*-value is between 0.10 and 0.15 (0.1029). *C*: Because the *P*-value of 0.1029 > $\alpha = 0.01$, we fail to reject H_0. We do not have convincing evidence that the distribution of eye color and wing shape is different from what the biologists predict.

Feb 14-5:28 PM