Agenda

Look at the Distribution of the Colors of M & M's

An e-mail was sent to the Company that makes M & M's asking about the color distribution

The company replied with:

Brown	13%
Yellow	14%
Orange	20%
Green	16%
Blue	24%
Red	13%

The company replied with:

Brown	13%
Yellow	14%
Orange	20%
Green	16%
Blue	24%
Red	13%

We're going to take a sample to try and find evidence against this claim

Bag of M&M's

think of this bag as being a random sample of the entire population of M&M's.

Bag of M&M's

think of this bag as being a random sample of the entire population of M&M's.

You'll each be given a small handful to count.

We'll combine all as our sample data.

your counts ↓ %

6	2 5 / 5 7	3 6 / 6 8	10	8			= 66	Brown	
1	5 4 / 9 5	5 8 / 9 3	4	6			= 61	Yellow	
14	5 10 / 9 5	15 7 / 11 11	8	7			= 102	Orange	
9	2 1 / 6 8	4 7 / 4 7	3	4			= 55	Green	
3	9 9 / 8 4	10 12 / 11 9	7	11			= 93	Blue	
6	3 1 / 4 2	5 2 / 6 1	9	3			= 42	Read	

419

	Claimed	Sample $\hat{p}$
Brown	13%	
Yellow	14%	
Orange	20%	
Green	16%	
Blue	24%	
Red	13%	

Sometimes the observed is greater than the expected; sometimes it is less.

How will we compare?

How will we compare?

With the

Chi-Square Test for
Goodness of Fit
(informally)

Which color M&M is the most common?


The company that makes milk chocolate M&Ms claims the following distribution:
13% Brown, 14% Yellow, 20% Orange, 16% Green, 24% Blue, and 13% Red. Is this true?

1. Observed values: Brown:_____ Yellow:_____ Orange:_____ Green:_____ Blue:_____ Red:_____

Total number of M&Ms:_______

The company that makes milk chocolate M&Ms claims the following distribution:
13% Brown, 14% Yellow, 20% Orange, 16% Green, 24% Blue, and 13% Red. Is this true?

1. Observed values: Brown: 66 Yellow: 61 Orange: 102 Green: 55 Blue: 93 Red: 42

Total number of M&Ms: 419

2. As a class, write down hypotheses for a significance test.

H_0: The company's claimed distribution is true

H_a: The claimed color distrib. is not true

3. Let's suppose that M&Ms claimed distribution is correct. If they are correct, how many of each color would we expect to get in our sample.

Expected values: Brown: 54.47 Yellow: 58.66 Orange: 83.8 Green: 67.04 Blue: 100.56 Red: 54.47

13% 14% 20% 16% 24% 13%

13% of 419

$$\frac{(L_1 - L_2)^2}{L_2}$$

Use the table to calculate the test statistic.

	Observed	Expected	(Observed - Expected)	(Observed - Expected)^2	$\frac{(Observed - Expected)^2}{Expected}$
Brown	66	54.47	11.53	132.9409	2.44
Yellow	61	58.66	2.34	5.4756	0.093
Orange	102	83.8	18.2	331.24	
Green	55	67.04	-12.04	144.9616	
Blue	93	100.56	-7.596	57.699	
Red	42	54.47	-12.47	155.500	

Add up all the numbers in the last column. This is our test statistic: 12.07

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

4. What value would we get for the test statistic if our sample was very close to what is expected? Explain.

It would be close to 0
χ^2

5. What value would we get for the test statistic if our sample was very far from what is expected? Explain.

It would be large

ESSENTIAL QUESTION *How do we perform significance tests for distributions of categorical variables and for relationships between categorical variables?*

The Chi-Square Test for Goodness of Fit

is a significance test that can help determine if a distribution of a claimed population [Company claimed 13% Brown, 14% Yellow, 20% orange, etc] differs from a sample distribution.

The Big Picture:
Where Chapter 11 Fits

Chapter 11: Inference for Distributions of Categorical Data

11.1 Chi-Square Tests for Goodness of Fit	2 Days
11.2 Inference for Two-way Tables	2 Days
Review, FRAPPY, and Test	2 Days

which puts the Ch. 11 test on
next Friday, Feb 21st

Chapter 11: The Big Ideas

- Chi-square tests are about categorical variables
- Three types of chi-square tests
  - Goodness of Fit: Distribution of 1 categorical variable in 1 population
  - Homogeneity: Distribution of 1 categorical variable for 2 or more populations/treatments
  - Independence: Relationship between 2 categorical variables in 1 population

No Confidence intervals this chapter

How To Pronounce ?

Walking into a room full of statisticians and referring to the "Chai" square test statistic will result in immediate loss of credibility.

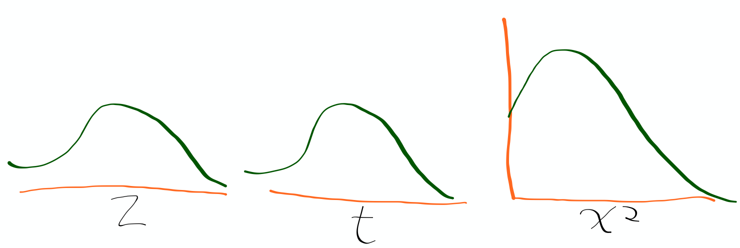
["Chai" should only be used when ordering a drink at Starbucks.]

Please help your biology teachers with this one!

Chi-Square tests are similar to other tests.

- Four Step process
- Need to Check Conditions

but there are differences



today:

✓STATE appropriate hypotheses and COMPUTE the expected counts and chi-square test statistic for a chi-square test for goodness of fit.

✓STATE and CHECK the Random, 10%, and Large Counts conditions for performing a chi-square test for goodness of fit.

✓CALCULATE the degrees of freedom and *P*-value for a chi-square test for goodness of fit.

Stating Hypotheses

***Ho* :** The distribution of color in the large bag of M&M'S Milk Chocolate Candies **is the same as** the claimed distribution.

Ha: The distribution of color in the large bag of M&M'S Milk Chocolate Candies **is *not* the same as** the claimed distribution.

In this chapter, there will be no two-sided tests. :)

The alternative hypothesis will always be "the null hypothesis is not correct"

We can also write the hypotheses in symbols.

H_0: $p_{brown} = 0.125$, $p_{red} = 0.125$, $p_{yellow} = 0.125$, $p_{green} = 0.125$, $p_{orange} = 0.25$, $p_{blue} = 0.25$
H_a: At least two of the p_i's are incorrect

where p_{color} = the true proportion of M&M'S Milk Chocolate Candies in the large bag of that color

A third way to write hypotheses:

H_0: The company's claimed distribution of color is correct for this bag of M&M's.

H_a: The company's claimed distribution of color is incorrect for this bag of M&M's.

CAUTION:

Don't state the alternative hypothesis in a way that suggests that all the proportions in the hypothesized distribution are wrong.

it's a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

it's a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

H_0: The claimed distribution is true

H_a: The claimed distribution is not true

it's a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

H_0: The claimed distribution is true

context

H_a: The claimed distribution is not true.

~~En~~

Expected values $= P_i \cdot n$
for each category

it's a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

H_0: The claimed distribution is true

context

H_a: The claimed distribution is not true.

Expected values $= P_i \cdot n$

claimed prob. of each category

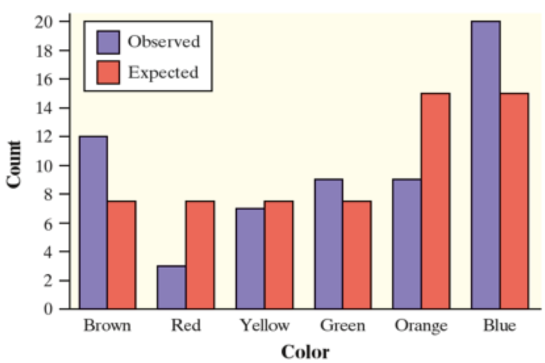
sample size

AP Exam Tip

Don't round the expected counts!
(just because you think they should be integers)

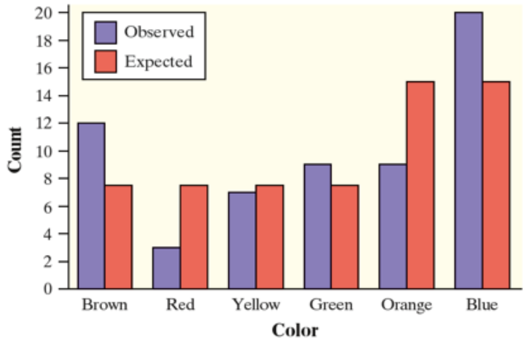
- it's the average number of observations in a given category in many many random samples.

Comparing Observed and Expected Counts: The Chi-Square Test Statistic



Color	Observed	Expected
Brown	12	7.5
Red	3	7.5
Yellow	7	7.5
Green	9	7.5
Orange	9	15.0
Blue	20	15.0

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.?



Color	Observed	Expected
Brown	12	7.5
Red	3	7.5
Yellow	7	7.5
Green	9	7.5
Orange	9	15.0
Blue	20	15.0

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.?

The statistic we use to make the comparison is the **chi-square test statistic** χ^2.

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

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

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

its a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

context

H_0: The claimed distribution is true

H_a: The claimed distribution is not true

Expected values $= P_i \cdot n$

prob. of each category

sample size

Chi-Square Statistic

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

$$\cancel{y} = \bar{x} = \frac{\sum x_i}{n}$$

Why divide by the expected count?

We are interested in how far away the observed count is from the expected

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

always include summation symbol.

Observed counts, not observed proportions

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

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

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

A typical calculation, by hand, look like this:

$$\chi^2 = \frac{(12-7.5)^2}{7.5} + \frac{(3-7.5)^2}{7.5} + \frac{(7-7.5)^2}{7.5} + \frac{(9-7.5)^2}{7.5} + \frac{(9-15)^2}{15} + \frac{(20-15)^2}{15}$$

category 1 2 3 3 4 . . .

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

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

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

For Jerome's data, we add six terms—one for each color category:

$$\chi^2 = \frac{(12-7.5)^2}{7.5} + \frac{(3-7.5)^2}{7.5} + \frac{(7-7.5)^2}{7.5} + \frac{(9-7.5)^2}{7.5} + \frac{(9-15)^2}{15} + \frac{(20-15)^2}{15}$$

$$= 2.7 + 2.7 + 0.03 + 0.30 + 2.4 + 1.67$$

$$= 9.8$$

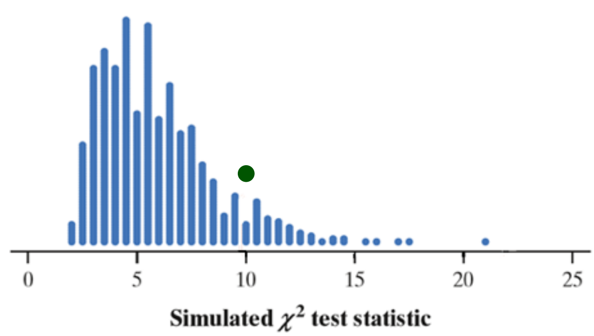
always positive!

The Chi-Square Distributions and P-Values

If we used software to simulate taking 1000 random samples of size 60 from the population distribution of M&M'S Milk Chocolate Candies given by Mars, Inc. Here are the values of the chi-square test statistic for these 1000 samples.

The Chi-Square Distributions and P-Values

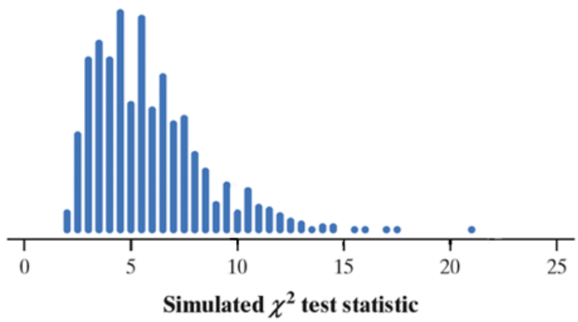
If we used software to simulate taking 1000 random samples of size 60 from the population distribution of M&M'S Milk Chocolate Candies given by Mars, Inc. Here are the values of the chi-square test statistic for these 1000 samples.



The Chi-Square Distributions and *P*-Values

If we used software to simulate taking 1000 random samples of size 60 from the population distribution of M&M'S Milk Chocolate Candies given by Mars, Inc. Here are the values of the chi-square test statistic for these 1000 samples.

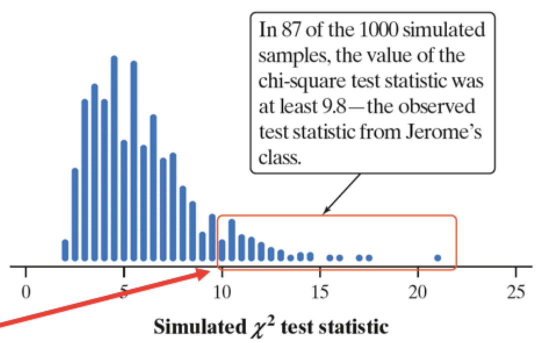
Larger values of χ^2 give more convincing evidence against H_0 and in favor of H_a.



We used software to simulate taking 1000 random samples of size 60 from the population distribution of M&M'S Milk Chocolate Candies given by Mars, Inc. Here are the values of the chi-square test statistic for these 1000 samples.

Larger values of χ^2 give more convincing evidence against H_0 and in favor of H_a.

Our estimated *P*-value is 87/1000 = 0.087



it's a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

H_0: The claimed distribution is true

context

H_a: The claimed distribution is not true.

Expected values $= P_i \cdot n$

prob. of each category

sample size

Chi-Square Statistic

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

df = # of categories − 1

it's a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

H_0: The claimed distribution is true

context

H_a: The claimed distribution is not true.

Expected values $= P_i \cdot n$

prob. of each category

sample size

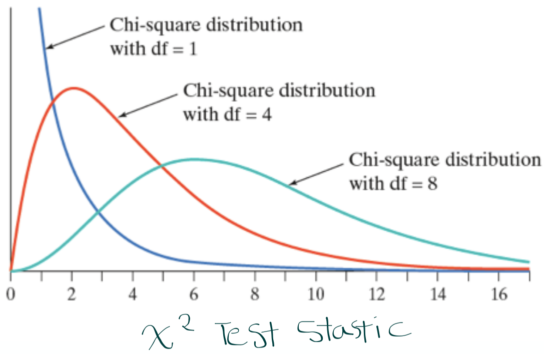
Chi-Square Statistic

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

df = # of categories − 1

P-value → Table C

A **chi-square distribution** is defined by a density curve that takes only nonnegative values and is skewed to the right. A particular chi-square distribution is specified by its degrees of freedom.



- As the degrees of freedom (df) increase, the density curves become less skewed.
- The mean of a particular chi-square distribution is equal to its degrees of freedom.
- For df > 2, the mode (peak) of the chi-square density curve is at df – 2.

The χ^2 distributions are the third and final type of distribution that you will use regularly.

[Normal distrib., t-distributions, χ^2 distrib.]

The χ^2 distributions are the third and final type of distribution that you will use regularly.

[Normal distrib., t-distributions, χ^2 distrib.]

normalcdf t_{cdf} χ^2_{cdf}

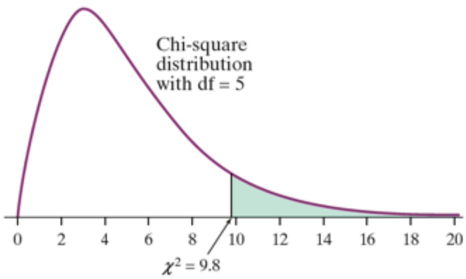
not on all calculators

find table C

There are 6 color categories for M&M's® Milk Chocolate Candies, so df = 6 – 1 = 5.

The *P*-value is the probability of getting a χ^2 value of as large as or larger than 9.8 when H_0 is true.

TEST STATISTIC



	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

The *P*-value for a test based on Jerome's data is between 0.05 and 0.10.

Conditions for Performing a Chi-Square Test for Goodness of Fit

Random: The data come from a random sample from the population of interest.

10%: When sampling without replacement, $n < 0.10N$.

Large Counts: All *expected* counts are at least 5.

Add to your notes

Why Check

Random: So.... we can generalize to the appropriate population

10%: So.... sampling without replacement is ok.

Large Counts: So.... so the sampling distribution is approximately a chi-square distribution and we can use χ^2 to find a P-Value

A chi-square density curve is a good model for the sampling distribution of the chi-square statistic but only when the sample size is large enough that the expected counts are all at least 5

[similar to Large Counts Condition for $\hat{p}$]

ie

Outcome of roll	1	2	3	4	5	6	Total
Observed count	12	28	12	13	10	15	90
Expected count	15	15	15	15	15	15	90

its a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

context

H_0: The claimed distribution is true

H_a: The claimed distribution is not true

Expected values = $P_i \cdot n$

prob. of each category

sample size

Chi-Square Statistic

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

df = # of categories − 1

P-value → Table C

Test is valid only if expected counts are ≥ 5

Large Counts Condition

without rounding

Check

Your

Understanding

Carrie made a 6-sided die in her ceramics class and rolled it 90 times to test if each side was equally likely to show up. The table summarizes the outcomes of her 90 rolls.

Outcome of roll	1	2	3	4	5	6	Total
Frequency	12	28	12	13	10	15	90

(a) State the hypotheses that Carrie should test.

(b) Calculate the expected count for each of the possible outcomes.

(c) Calculate the value of the chi-square test statistic.

Carrie made a 6-sided die in her ceramics class and rolled it 90 times to test if each side was equally likely to show up. The table summarizes the outcomes of her 90 rolls.

Outcome of roll	1	2	3	4	5	6	Total
Frequency	12	28	12	13	10	15	90

(a) State the hypotheses that Carrie should test.

H_0: The claimed distribution of $\frac{1}{6}$ prob for each roll is true

(b) Calculate the expected count for each of the possible outcomes.

(c) Calculate the value of the chi-square test statistic.

Carrie made a 6-sided die in her ceramics class and rolled it 90 times to test if each side was equally likely to show up. The table summarizes the outcomes of her 90 rolls.

Outcome of roll	1	2	3	4	5	6	Total
Frequency	12	28	12	13	10	15	90

(a) State the hypotheses that Carrie should test.

H_0: The claimed distribution of 1/6 prob for each roll is true

H_a: The claimed distribution is not true (die isn't fair)

(b) Calculate the expected count for each of the possible outcomes.

(c) Calculate the value of the chi-square test statistic.

Carrie made a 6-sided die in her ceramics class and rolled it 90 times to test if each side was equally likely to show up. The table summarizes the outcomes of her 90 rolls.

Outcome of roll	1	2	3	4	5	6	Total
Frequency	12	28	12	13	10	15	90

(a) State the hypotheses that Carrie should test.

H_0: The claimed distribution of 1/6 prob for each roll is true

H_a: The claimed distribution is not true (die isn't fair)

(b) Calculate the expected count for each of the possible outcomes.

$\frac{1}{6} \times 90 = 15$ for each

(c) Calculate the value of the chi-square test statistic.

$$\chi^2 = \frac{(12-15)^2}{15} + \dots \frac{(15-15)^2}{15} =$$

Carrie made a 6-sided die in her ceramics class and rolled it 90 times to test if each side was equally likely to show up. The table summarizes the outcomes of her 90 rolls.

Outcome of roll	1	2	3	4	5	6	Total
Frequency	12	28	12	13	10	15	90

(a) State the hypotheses that Carrie should test.

H_0: The claimed distribution of $\frac{1}{6}$ prob for each roll is true

H_a: The claimed distribution is not true (die isn't fair)

(b) Calculate the expected count for each of the possible outcomes.

$\frac{1}{6} \times 90 = 15$ for each

(c) Calculate the value of the chi-square test statistic.

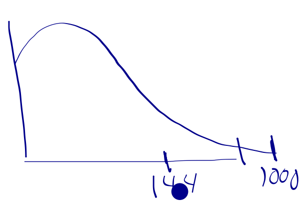
$$\chi^2 = \frac{(12-15)^2}{15} + \ldots\ldots \frac{(15-15)^2}{15} = 14.4$$

(d) Which degrees of freedom should you use?

6 groups $\quad df = 6 - 1 = 5$

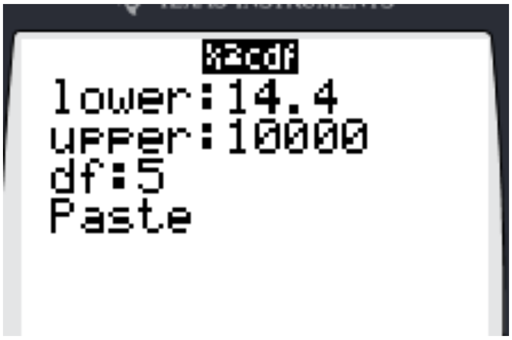
(e) Use table C to find the p-value. What conclusion would you make?

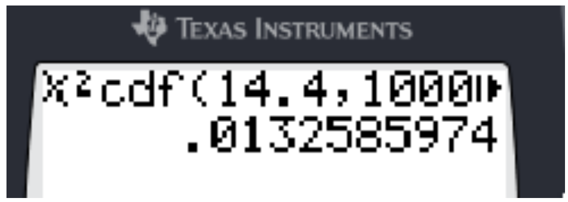
from Table C → P-Value is in between 0.01 and 0.02



This is convincing evidence.

Calculator → $\chi^2\text{cdf}(14.4 \text{ lower}, 1000 \text{ upper}, 5 \text{ df}) = .013$





it's a significance test

Lesson 11.1: Day 1: Chi-Square Test: Goodness of Fit

H_0: The claimed distribution is true

context

H_a: The claimed distribution is not true.

Expected values $= P_i \cdot n$

prob. of each category

sample size

Chi-Square Statistic

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

df = # of categories − 1

P-value → Table C

or

$\chi^2\text{cdf}(\text{Lower}, \text{Upper}, df)$

Test is valid only if expected counts are ≥ 5

without rounding

See your
Ch. 10 Test

11.1: 1, 3, 5, 7, 26

Ch. 3

Study pp.709-716