AP Statistics
Ch. 12
Chi-Square Goodness of Fit Tests

Chi-Square GOF Test

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

H_o: p_1 = _____, p_2 = _____, p_3 = _____, …, p_i = _____.
H_a: At least one of the pi's is incorrect

$\propto$ = _____

Plan: Chi-Square goodness of fit test

Random: The data come from a random sample or a randomized experiment

10%: When sampling without replacement

Large Sample Size – All expected counts are at least 5

Do

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

Find the P-value and list degrees of freedom.

(The P-value is the area to the right of χ^2 under the density curve of the chi-square distribution with n – 1 degrees of freedom. (n is the number of categories))

Conclude

P-value < $\propto$, we reject H_o, we have convincing evidence for H_a

P-value > $\propto$, we fail to reject H_o, we Do not have convincing evidence for H_a

When were you born?

Are births evenly distributed across the days of the week? The one-way table below shows the distribution of births across the days of the week in a random sample of 140 births from local records in a large city:

Day	**Sun.**	**Mon.**	**Tue.**	**Wed.**	**Thu.**	**Fri.**	**Sat.**
Births	13	23	24	20	27	18	15

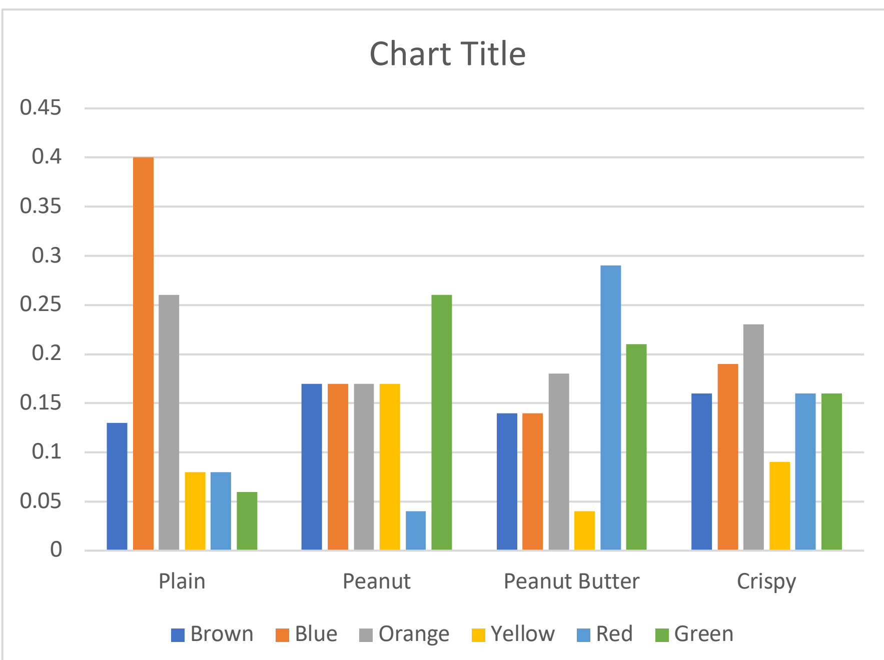
Do these data give significant evidence that local births are not equally likely on all days of the week?
Use $\alpha = .05$.

Chi-Square Tests for Homogeneity

Are the distributions of the colors the same for the different varieties of M&Ms?

	Plain	Peanut	Peanut Butter	Crispy	Total
Brown	7	4	4	7	22
Blue	21	4	4	8	37
Orange	14	4	5	10	33
Yellow	4	4	1	4	13
Red	4	1	8	7	20
Green	3	6	6	7	22
Total	53	23	28	43	147

A graph might help.



Are these differences significant? Let's do a significance test to see!!!

Hypotheses:

H_o: There is no difference in the distribution of colors for the different types of M&Ms.

H_a: There is a difference in the distribution of colors for the different types of M&Ms.

Now let's see how our observed counts compare to what we would expect to get if Ho is true. First we need our expected counts.

Expected Cell Counts

$$\text{expected count} = \frac{\text{row total} \times \text{column total}}{n}$$

n is the total number of data.

Fill in the expected counts for Brown.

	Plain	**Peanut**	**Peanut Butter**	**Crispy**	
Brown					
Blue	13.3	5.8	7.0	10.8	
Orange	12.0	5.2	6.3	9.7	
Yellow	4.7	2.0	2.5	3.8	
Red	7.2	3.1	3.8	5.9	
Green	7.9	3.4	4.2	6.4	

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

Now we need our Chi-Square Statistic to see how far our observed counts were from the expected counts. Fill in the contributions of Plain M&Ms.

________+ ________+ ________+ ________+ ________+ ________+

$$\frac{(4-3.4)^2}{3.4} + \frac{(4-5.8)^2}{5.8} + \frac{(4-5.2)^2}{5.2} + \frac{(4-2.0)^2}{2.0} + \frac{(1-3.1)^2}{3.1} + \frac{(6-3.4)^2}{3.4} +$$

$$\frac{(4-4.2)^2}{4.2} + \frac{(4-7.0)^2}{7.0} + \frac{(5-6.3)^2}{6.3} + \frac{(1-2.5)^2}{2.5} + \frac{(8-3.8)^2}{3.8} + \frac{(6-4.2)^2}{4.2} +$$

$$\frac{(7-6.4)^2}{6.4} + \frac{(8-10.8)^2}{10.8} + \frac{(10-9.7)^2}{9.7} + \frac{(4-3.8)^2}{3.8} + \frac{(7-5.9)^2}{5.9} + \frac{(7-6.4)^2}{6.4} = 24.58$$

Let's find our P-value. But first we need degrees of freedom.

Degrees of Freedom = (#columns – 1)(#rows – 1)

DF = _______

Similar to GOF test, the p-value is always the right tail of the Chi-Square distribution.

Pvalue = _______

Follow-up Analysis

If the test allows us to reject the null hypothesis of no difference, you then want to do a follow-up analysis that examines the differences in detail.

Look at the individual components to see which terms contribute most to the chi-square statistic

Chi-Square Tests for Homogeneity

State

Ho: There is not difference in the distribution of (the categorical variable) for the different samples.
Ha: There is a difference in the distribution of (the categorical variable) for the different samples.

$\propto =$ _____

Plan: Chi-square test for homogeneity

Random: The data come from separate random samples from each population of interest or from groups in a randomized experiment

10%: When sampling without replacement, if experiment, not needed.

Large Sample Size – All expected counts are at least 5

Do

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

Find the P-value and list degrees of freedom {(#columns – 1)(#rows – 1)}

Conclude

P-value < $\propto$, we reject H_o, we have convincing evidence for H_a

P-value > $\propto$, we fail to reject H_o, we Do not have convincing evidence for H_a

Canada has universal health care. The United States does not but often offers more elaborate treatment to patients with access. How do the two systems compare in treating heart attacks? Researchers compared random samples of 2600 US and 400 Canadian heart attack patients. One key outcome was the patient's own assessment of their quality of life relative to what it had been before the heart attack

Quality of Life	Canada	USA
Much Better	75	541
Somewhat Better	71	498
About the same	96	779
Somewhat worse	50	282
Much worse	19	65
Total	311	2165

Is there a significant difference between the two distributions of quality-of-life ratings? Carry out an appropriate test at the $\alpha = 0.01$ level.

Chi-Square Tests for Independence/ Association

State

H_o: There is no association between two categorical variables in the population of interest.
H_a: There is an association between two categorical variables in the population of interest.
OR
H_o: Two categorical variables are independent in the population of interest
H_a: Two categorical variables are not independent in the population of interest

$\propto =$ _____

Plan: Chi-square test for association/independence

Random: The data come from a random sample

10%: When sampling without replacement

Large Sample Size – All expected counts are at least 5

Do

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

Find the P-value and list degrees of freedom {(#columns – 1)(#rows – 1)}

Conclude

P-value < $\propto$, we reject H_o, we have convincing evidence for H_a

P-value > $\propto$, we fail to reject H_o, we Do not have convincing evidence for H_a

Are men and women equally likely to suffer lingering fear from watching scary movies as children? Researchers asked a sample of 117 college students to write narrative accounts of their exposure to scary movies before the age of 13. More than one-fourth of the students said that some of the fright symptoms are still present when they are awake. The following table breaks down these results by gender.

Fright Symptoms	**Male**	**Female**	**Total**
Yes	7	29	36
No	31	50	81
Total	38	79	117

a. Explain why a chi-square test for association/independence and not a chi-square test for homogeneity should be used in this setting.

b. Do these data provide convincing evidence that there is an association between gender and lingering fright symptoms? $\alpha = 0.$

c. Which cell contributes most to the chi-statistic? In what way does this cell differ from what the null hypothesis suggests?

Which Chi-Square?

1. Focus on how the data were produced. If the data come from two or more independent random samples or treatment groups in a randomized experiment, do the chi-square test for homogeneity.

2. If the data come from a single random sample, with the individuals classified according to two variables, use a chi-square test for association/independence.

3. If the data comes from a single random sample and classified according to one variable, use chi-square goodness of fit test.

For each of the following situations decide what type of chi square test is appropriate. Explain.

a. Shopping at secondhand stores is becoming more popular and has even attracted the attention of business schools. A study of customers' attitudes toward secondhand stores interviewed separate random samples of shoppers at two secondhand stores of the same chain in different cities. The two-way table shows the breakdown of respondents by gender.

		Store		
		A	B	Total
Gender	Male	38	68	106
	Female	203	150	353
	Total	241	218	459

b. The General Social Survey (GSS) asked a random sample of adults their opinion about whether astrology is very scientific, sort of scientific, or not at all scientific. Here is a two-way table of counts for people in the sample who had three levels of higher education:

		Degree held			
		Associate's	Bachelor's	Master's	Total
Opinion about astrology	Not at all scientific	169	256	114	539
	Very or sort of scientific	65	65	18	148
	Total	234	321	132	687

c. Casinos are required to verify that their games operate as advertised. American roulette wheels have 38 slots—18 red, 18 black, and 2 green. In one casino, managers record data from a random sample of 200 spins of one of their American roulette wheels. The table displays the results.

Color	Red	Black	Green
Count	85	99	16