Name ______________________________

Chapter 11 Learning Objectives	**Section**	**Related Example on Page(s)**	**Relevant Chapter Review Exercise(s)**	**Can I do this?**
State appropriate hypotheses and compute expected counts for a chi-square test for goodness of fit.	11.1	681	R11.1	
Calculate the chi-square statistic, degrees of freedom, and *P*-value for a chi-square test for goodness of fit.	11.1	683, 685	R11.1	
Perform a chi-square test for goodness of fit.	11.1	688	R11.1	
Conduct a follow-up analysis when the results of a chi-square test are statistically significant.	11.1, 11.2	Discussion on 690–691, 716	R11.4	
Compare conditional distributions for data in a two-way table.	11.2	697, 711	R11.3, R11.5	
State appropriate hypotheses and compute expected counts for a chi-square test based on data in a two-way table.	11.2	701, 713	R11.2, R11.3, R11.4, R11.5	
Calculate the chi-square statistic, degrees of freedom, and *P*-value for a chi-square test based on data in a two-way table.	11.2	704	R11.3, R11.5	
Perform a chi-square test for homogeneity.	11.2	708	R11.3	
Perform a chi-square test for independence.	11.2	715	R11.5	
Choose the appropriate chi-square test.	11.2	718	R11.4	

11.1 Chi-square tests
Read 678–687
What is a one-way table?

A **one-way table** is often used to display the distribution of a single categorical variable for a sample of individuals.

Example:

Grade on Test	A	B	C	D	F
Count	20	25	22	8	5

The **chi-square test for goodness of fit** tests the null hypothesis that a categorical variable has a specified distribution in the population of interest.

Examples of null and alternative hypotheses for some chi-square goodness-of-fit tests:

H_0: ACT scores are normally distributed.
H_a: The distribution of ACT scores is not a normal distribution.

H_0: Letter grades for Prof. Cohen's class are uniformly distributed.
H_a: Letter grades for Prof. Cohen's class are not uniformly distributed.

H_0: Flower colors for that species are distributed according to: 25% red, 50% pink, & 25% white.
H_a: Flower colors for that species are distributed differently.

This test compares the **observed count** in each category with the counts that would be expected if H_0 were true. The **expected count** for any category is found by multiplying the sample size by the proportion in each category according to the null hypothesis.

Even though observed counts will be whole numbers, **<u>don't round the expected counts</u>** to whole numbers.

This is the chi-square test statistic as it appears on the formula sheet:

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

where the sum is over all possible categories.

It measures (I'm stating this very informally) the total squared relative weirdness. That is, how weird is it to expect certain counts for each category and observe these particular counts, relative to the expected numbers.

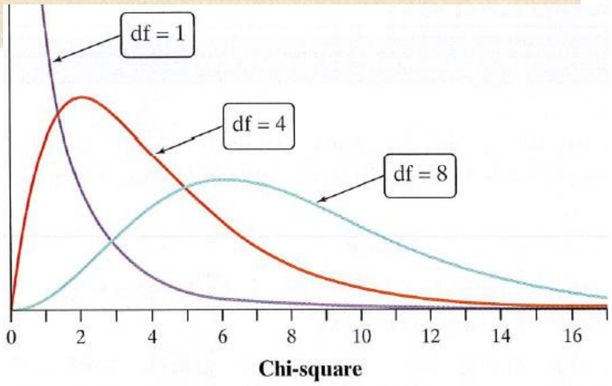
The conditions for performing a chi-square test for goodness of fit are:

- **Random:** The data were produced by a well-designed random sample or randomized experiment.
  - **10%:** When sampling without replacement, check that the population is at least 10 times as large as the sample.
- **Large Counts:** All expected counts must be at least 5.

When the conditions are met, the sampling distribution of the statistic χ^2 can be modeled by a **chi-square distribution**.

Information about the chi-square distributions:

The chi-square distributions are a family of density curves that take only non-negative values and are skewed to the right. A particular chi-square distribution is specified by giving its degrees of freedom. The chi-square test for goodness of fit uses the chi-square distribution with degrees of freedom = the number of categories − 1.



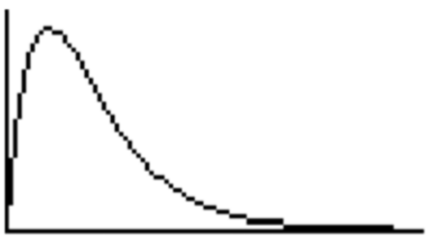
Other facts (just for the curious; you don't need to know these):

The mean for a Chi-square distribution = its df.

If df>2, the peak or mode for a Chi-square distribution is at (df-2).

To calculate *p*-values for chi-square distributions, use the chi-square table or X^2cdf on the calculator.

Example: If there are 5 categories, how much area (our p-value) is under the X^2 distribution curve to the right of our X^2 statistic, when $X^2 = 5.2$?



Tom made a tetrahedron out of heavy cardstock in his geometry class and is using it as a 4-sided die. He rolled it 60 times to test if it was equally likely to land on each side.

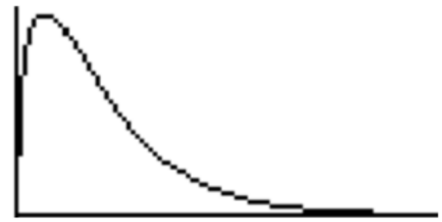
(a) State the hypotheses Tom is interested in testing.

(b) Assuming that his die is fair, calculate the expected counts for each possible outcome.

(c) Here are the results of 60 rolls of Tom's die. Calculate the chi-square statistic.

Outcome	**Observed**
1	13
2	12
3	17
4	18
Total	60

(d) Find the *P*-value for Tom's chi-square test.



(e) Make an appropriate conclusion about Tom's 4-sided die.

HW #31: page 693 (1–6, 23–25)

11.1 Chi-square Tests for Goodness of Fit

Read 687–691

According to a July 2015 study, 44.2% of U.S. smartphones in use are made by Apple, 27.3% are Samsung phones, and 8.7% are made by LG, and 19.8% are made by other manufacturers like Motorola and HTC. The table shows the distribution of phone manufacturers for a random sample of smartphone users at our school. Do these data provide convincing evidence that the manufacturer distribution in our school is not the same as the manufacturer distribution for all U.S. smartphones?

Category	Count
Apple	105
Samsung	10
LG	3
Other	2
Total	120

You can certainly use your calculator to conduct a chi-square goodness-of-fit test, but it is a good idea to show a few of the "contributing fractions" of the form (observed-expected)/expected.

When we reject H_0, we sometimes do a follow-up analysis in which we look for the largest contributing fraction(s) <u>and</u> identify whether there were more than or fewer than expected. In our smartphone test on the previous page, which manufacturer contributed most to the X^2 statistic?

HW #32 page 693 (9, 11, 14, 17)

11.2 Chi-Square Tests for Homogeneity

Read 697–705

How is section 11.2 different than section 11.1?

Ch. 11 Section 1	Ch. 11 Section 2	
One-way tables	Two-way tables	
Is this 1 variable distributed in a certain way in this 1 population?	Is this 1 variable distributed the same way in 2+ populations?	Is there an association between these 2 variables in this 1 population?
Chi-square Goodness of Fit Test	Chi-square Test of Homogeneity	Chi-Square Test of Independence
Ask 1 sample 1 question	Ask 2+ samples 1 question	Ask 1 sample 2 questions
Uses Lists on the calculator	Uses Matrices on the calculator	Uses Matrices on the calculator

Meg bought a bag of her favorite candy-coated chocolates at the store and Tim bought a bag of his favorite gummy candy, as well. Each candy comes in 5 colors and Meg and Tim counted the number of each color they had in their bags. They were surprised by the differences in the distributions of colors in their bags. The table shows the distributions in their respective bags:

	Red	Orange	Yellow	Green	Blue
Meg-chocolate	12	16	15	17	15
Tim-gummies	32	18	17	16	17

What are the two explanations for the differences in the distributions of colors in their bags?

Hypotheses for a test of homogeneity:

H_0: There is no difference in the distribution of a categorical variable for several populations or treatments.

H_a: There is a difference in the distribution of a categorical variable for several populations or treatments.

We could do several 2-Proportion Z Tests, but that's a lot more work than just doing one Chi-square Test of Homogeneity.

How do you calculate the expected counts for a test that compares the distribution of a categorical variable in multiple groups or populations?

$$\text{expected count} = \frac{\text{row total} \cdot \text{column total}}{\text{table total}}$$

Example:

These are the observed counts for a random sample of vehicles from some state:

	Cars	Trucks	Totals
Manual (stick) Transmission	8	14	22
Automatic Transmission	51	10	61
Totals	59	24	83

Find the expected counts for this table.

	Cars	Trucks
Manual (stick) Transmission		
Automatic Transmission		

These are the conditions for a chi-square test for homogeneity:

- **Random:** The data come from independent random samples or from the groups in a randomized experiment.
  - **10%:** When sampling without replacement, check that $n \le \frac{1}{10}N$.
- **Large Counts:** All *expected* counts are at least 5.

This is the chi-square test statistic as it appears on the formula sheet: $\chi^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$

To calculate the degrees of freedom for a chi-square test for homogeneity, use df=(# rows - 1)(# columns - 1)

Examples: Find the **degrees of freedom** for each of these situations:

A 2 x 3 table:

9	12	18
21	17	25

df=

A 5 x 3 table:

7	11	8
12	16	21
14	15	11
20	30	17
13	19	16

df=

A 4 x 8 table:

df=

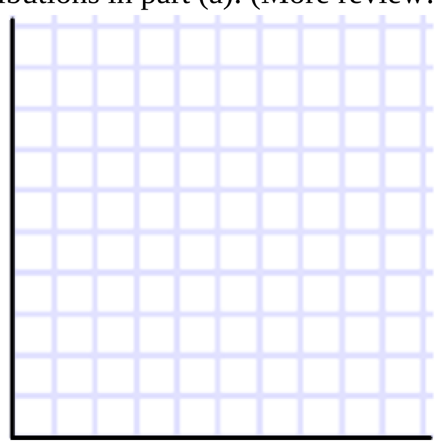
One way that linguistic scholars can try to determine whether a work that is *suspected* of being written by someone really *was* written by the person is by analyzing the use of certain words or sentence structures. Seven works attributed to Aristotle were analyzed for their inclusion of the word "gar" (in Greek).

(a) Calculate the conditional distribution (in proportions) of sentences with gar for each work. This is a review of how we analyzed 2-way table data last semester.

	Sentences with gar	Sentences without gar	
1st work	63	137	**200**
2nd work	58	142	**200**
3rd work	84	216	**300**
4th work	103	197	**300**
5th work	25	75	**100**
6th work	72	228	**300**
7th work	45	101	**146**
	450	**1096**	**1546**

	Sentences with gar	Sentences without gar
1st work		
2nd work		
3rd work		
4th work		
5th work		
6th work		
7th work		

(b) Make an appropriate graph for comparing the conditional distributions in part (a). (More review!)



(c) Write a few sentences comparing the distributions of gar for each work.

Do these data provide evidence that the distribution of “gar” in these 7 literary works is not the same?
(d) **State** the hypotheses.

(e) **Plan**: Verify that the conditions are met & name the procedure.

(f) **Do**: Calculate the expected counts, chi-square statistic, and *p*-value.

(g) **Conclude**: Make an appropriate conclusion.

HW #33 page 695 (19–22, 27–31 odd, 57)
Read 706–710
You can use the calculator to conduct a chi-square test of homogeneity, but it is a good idea to show a few of the "contributing fractions".

A graduate student decided to investigate whether the percent of knee surgery patients who return with a need for additional physical therapy 6 months after surgery differs based on the manner in which patients are instructed to do their post-operative therapy. Working with a physician, she obtained 60 volunteer knee surgery patients, who all had the same surgical procedure done. She randomly assigned 20 patients to get a video demonstrating how to do their post-operative therapy, 20 patients to get an instructional brochure, & 20 volunteers to get both a brochure & videos. The exercises in the video and brochure were the same & a physical therapist modeled the exercises with each patient. Do the data in the table provide convincing evidence at the $\alpha=0.5$ level that there is a difference in the rate of referral for additional rounds of physical therapy for the three types of post-operative care instruction?

	Video only	**Brochure & video**	**Brochure only**	**Total**
Referred for more Therapy after 6 months	6	5	7	**18**
Released from therapy After 6 months	14	15	13	**42**
Total	**20**	**20**	**20**	**60**

Remember that when we reject H_0, we might be asked do a follow-up analysis. That means we will look for the largest contributing fraction(s) and identify whether some category had more than or fewer than expected.

HW #34 page 725 (33–39 odd)

11.2 Chi-Square Test for Independence

Read pages 711–717

What does it mean if two variables have an association?

What does it mean if two variables are independent?

For a test of independence & a test of homogeneity, the math is the same, but the wording is different.

This part of the table, copied from page 5 of the notes, will help.

Ch. 11 Section 2	
Two-way tables	
Is this 1 variable distributed the same way in 2+ populations?	Is there an association between these 2 variables in this 1 population?
Chi-square Test of Homogeneity	Chi-Square Test of Independence
Ask 2+ samples 1 question	Ask 1 sample 2 questions
Uses Matrices on the calculator	Uses Matrices on the calculator

Hypotheses for a test of independence:

H_0: 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, alternatively,

H_0: Two categorical variables are independent in the population of interest.

H_a: Two categorical variables are not independent in the population of interest.

For a test of independence, you calculate the expected counts, $\text{expected count} = \dfrac{\text{row total} \cdot \text{column total}}{\text{table total}}$

the test statistic, $\chi^2 = \sum \dfrac{(\text{Observed} - \text{Expected})^2}{\text{Expected}}$

& the degrees of freedom df=(# rows - 1)(# columns - 1)

just like you did for the test of homogeneity.

Conditions for a test of association/independence:

- **Random:** The data come from a well-designed random sample or randomized experiment.
  - **10%:** When sampling without replacement, check that $n \le \frac{1}{10}N$.
- **Large Counts:** All expected counts are at least 5.

Horseshoe Crabs revisited: Two members of the University of Florida at Gainesville Department of Zoology collected data on Horseshoe Crabs on a Delaware beach during 4 days in the late spring of 1992. Based on the color of the shells, they classified each crab as Young, Intermediate, or Old and whether the crabs could right themselves when flipped on their backs or whether they were stranded for at least a certain period of time. Here are the results.

	Young	Intermediate	Old	Total
Stranded	214	384	295	893
Not Stranded	1668	1204	216	3088
Total	1882	1588	511	3981

(a) Do the data provide convincing evidence at the α = 0.05 level of an association between age and strandedness for Horseshoe Crabs?

(b) If your conclusion in part (a) was in error, which type of error did you commit? Explain.

HW #35 page 726 (41, 43, 45, 47, 51–56)
11.2 Using Chi-square Tests Wisely / FRAPPY!
Read 717–721
Based on a 2012 Pew Internet Tracking Survey, the following table was constructed for a random sample of adults, age 18 or older. Suppose that you decide to analyze these data using a chi-square test. However, without any additional information about how the data were collected, it isn't possible to know which chi-square test is appropriate.

	Annual Household Income ($)				
	<30,000	30,000-49,999	50,000-74,999	75,000+	Total
Use social networking on mobile phones:	170	126	131	242	669
Do not use social networking on mobile phones:	277	190	141	296	904
Total:	447	316	272	538	1573

(a) Explain why it is OK to use annual household income as a categorical variable rather than a quantitative variable.

(b) Explain how you know that a goodness-of-fit test is not appropriate for analyzing these data.

(c) Describe how these data could have been collected so that a test for homogeneity is appropriate.

(d) Describe how these data could have been collected so that a test for independence is appropriate.

In a study reported in the *Journal of Orthodontics* (2003), researchers conducted a randomized clinical trial to compare the adhesive failure rates of identical pre-coated brackets and non-coated brackets; all bonded using a light-cure system. (The orthodontists who conducted the study also recorded the time required to affix the brackets to each patient's teeth—a t-test showed no significant difference in the time required to bond each type.) Each patient had both adhesive systems used with one type of adhesive being randomly assigned to the upper left and lower right quadrants of the mouth and the other type being assigned to the remaining quadrants. Here are the results:

	Pre-coated brackets	Non-coated brackets	Total
Bracket adhesive failed within 6 months	30	25	55
No bracket adhesive failure within 6 months	342	349	691
Total	372	374	746

(a) Which type of chi-square test is appropriate here? Explain.

(b) Use the calculator to find the chi-square statistic and p-value.

(c) Find the p-value 2-proportion z test on the calculator and compare your result to the p-value in part b above.

For H_a:$p_1 \neq p_2$, the chi-square test & 2-proportion z test are equivalent.
For H_a:$p_1 < p_2$ or Ha:$p_1 > p_2$, use the 2-proportion z test.

What can we do if some of the expected counts are < 5 and we still want to run a chi-square test?

Frappy: 2008 #5 – Moose!
HW #36: page 732 Chapter 11 Review Exercises
Review Chapter 11
HW #37: page 733 Chapter 11 AP Statistics Practice Test
Chapter 11 Test