CHAPTER 11: INFERENCE FOR DISTRIBUTIONS OF CATEGORICAL DATA

11.2 INFERENCE FOR TWO-WAY TABLES

OUTCOME: I WILL COMPARE CONDITIONAL DISTRIBUTIONS FOR DATA IN A TWO-WAY TABLE, STATE APPROPRIATE EXPECTED COUNTS FOR A CHI-SQUARE TEST BASED OFF OF THE TWO-WAY TABLES, CALCULATE THE CHI-SQUARE TEST STATISTIC, DEGREES OF FREEDOM, AND P-VALUE FOR A CHI SQUARE TEST BASED OFF OF A TWO-WAY TABLE, AND PERFORM (AND DISTINGUISH BETWEEN) A CHI-SQUARE TEST FOR HOMOGENEITY AND INDEPENDENCE.

Confession: I don't know what I'm doing.

RELATIONSHIPS BETWEEN TWO CATEGORICAL VARIABLES

■ Another common situation that leads to a two-way table is when a single random sample of individuals is chosen from a *single* population and then classified based on two categorical variables.

■ In that case, our goal is to analyze the relationship between the variables.

■ Our **null hypothesis** is that there is no association between the two categorical variables in the population of interest.

■ The **alternative hypothesis** is that there is an association between the variables.

BIGGEST DIFFERENCE BETWEEN THE TWO TESTS

- Homogeneity tests compare the distribution of **one** variable in **two or more** populations/groups

- Independence tests we are investigating the relationship between **two** variables in **one** population.

DO

- Find the expected counts

- Find the chi-square test statistic and p-value

EXAMPLE

Are men and women equally likely to suffer lingering fear from watching scary movies as children? Researchers asked a random 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?	Gender		Total
	Male	Female	
Yes	7	29	36
No	31	50	81
Total	38	79	117

Chi-Square Test: Male, Female

Expected counts are printed below observed counts ✳

Chi-Square contributions are printed below expected counts ✳

	Male	Female	Total
1	7	29	36
	11.69	24.31	
	1.883	0.906	
2	31	50	81
	26.31	54.69	
	0.837	0.403	
Total	38	79	117

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

Chi-Sq = 4.028, DF = 1, P-Value = 0.045

WHY WOULD WE USE A CHI-SQUARE TEST FOR INDEPENDENCE AND NOT A CHI-SQUARE TEST FOR HOMOGENEITY?

STATE AN APPROPRIATE PAIR OF HYPOTHESES FOR RESEARCHERS TO TEST IN THIS SETTING.

H_0: There is no association between gender and fright symptoms.

H_a: There is an association between...

WHICH CELL CONTRIBUTES MOST TO THE CHI-SQUARE STATISTIC? IN WHAT WAY DOES THIS CELL DIFFER FROM WHAT THE NULL HYPOTHESIS SUGGESTS?

DONE WITH CH. 11 NOTES!

- Woo!

- Classwork!

- Homework!