Name ___________________________

Chi-Square Test on Fortune Cookie Fortunes

A method of statistical analysis

Test #1: Goodness of Fit

Question: Are all fortune cookie fortunes the same? That is, do all fortunes fall under the same categories in equal numbers?

Null hypothesis: ____

____

Alternate hypothesis: ____

____

1. You will be given fortune cookies or fortunes of each of two brands. Open your cookies, read their fortunes, and classify the type of fortune as prophecy, advice, wisdom, or miscellaneous (miscellaneous will include compliments and others). Share your data with the class and fill in the frequency table below with the class totals.

	Type of Fortune				
Cookie Brand	Prophecy	Advice	Wisdom	Misc.	Totals
expected:					
expected:					

2. If we expected fortunes to be equally distributed among the four categories, calculate the expected percentages for each category and record these in the table above.
3. Using the chi-square equation, calculate the chi-square value for each of the cookie brands.
4. Use the probability table to determine if you accept or reject the null hypothesis.

Cookie Brand:	**o-e**	**$(o\text{-}e)^2$**	**$(o\text{-}e)^2/e$**
prophecy			
Advice			
Wisdom			
miscellaneous			
		Total:	

Cookie Brand:	**o-e**	**$(o\text{-}e)^2$**	**$(o\text{-}e)^2/e$**
prophecy			
Advice			
Wisdom			
miscellaneous			
		Total:	

5. What is your chi-square value for each brand? ____

6. How many degrees of freedom do you have for each brand? ___________________________________

7. Using the chart below, what is your p value for each brand? ______________________________

Degrees of Freedom	Probability										
	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.05	0.01	0.001
1	0.004	0.02	0.06	0.15	0.46	1.07	1.64	2.71	3.84	6.64	10.83
2	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	5.99	9.21	13.82
3	0.35	0.58	1.01	1.42	2.37	3.66	4.64	6.25	7.82	11.34	16.27
4	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	9.49	13.28	18.47
5	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	11.07	15.09	20.52

8. What does this mean? Do you reject or accept the null hypothesis? Do you find the fortunes in the expected ratios or is the data too different to attribute to random chance? ______________________________

____

____

____

9. Two researchers did this same research and ended up with these percentages for each of the four categories: 61.7% prophecy, 11.1% compliments (misc), 12.1% advice, and 15.1% wisdom. Using these values for your expected numbers, recalculate chi-square and see if each brand matches these percentages better. Show your work below.

Cookie Brand:	**o-e**	**$(o-e)^2$**	**$(o-e)^2/e$**
prophecy			
Advice			
Wisdom			
miscellaneous			
		Total:	

Cookie Brand:	**o-e**	**$(o-e)^2$**	**$(o-e)^2/e$**
prophecy			
Advice			
Wisdom			
miscellaneous			
		Total:	

Discussion of results: ______________________________

____

____

____

____

____

____

Extension:
Typically a two-way frequency table analysis will be extended to a chi-square hypothesis test. When analyzing data from a frequency table, there are two types of chi-square tests that could be utilized.

A test of independence answers the question, "Are the two categorical variables independent for a population under study?" It assesses whether there is a relationship between two variables for a single population. The null hypothesis for the test of independence is that the two categorical variables are not related (independent) for the population of interest.

A test of homogeneity answers the question, "Do two or more populations have the same distribution for one categorical variable?" It assesses whether a single categorical variable is distributed the same in two (or more) different populations. The null hypothesis for the test of homogeneity is that the distribution of the categorical variable is the same for the two (or more) populations.

The mechanics of tests of independence and tests of homogeneity are the same. The distinction is the way in which the data was collected. If two categorical variables are collected for each subject, then a test of independence should be performed. If a single categorical variable is collected for each of two (or more) groups, then a test of homogeneity should be performed.

Test #2: Test for ______________________________

Now, you will calculate the **conditional distribution** of the type of fortune **given** the brand of fortune cookie. That is, *for each brand*, what are the percentages of the types of fortunes. When you calculate this conditional distribution, your row totals should be approximately 100%.

	Type of Fortune			
Cookie Brand	Prophecy	Advice	Wisdom	Misc.

Based upon the percentages above, would you say that the two brands have the same type of fortunes? Explain. (Think about whether you think the differences in the percentages are significant or not)

____

____

Null hypothesis (H0): ____

Alternate hypothesis (Ha): ____

The necessary data conditions for the chi-square test of homogeneity:
1. All expected counts should be greater than 1.
2. At least 80% of the table cells should have an expected count greater than 5.

To compute the expected count for each cell: expected count = $\dfrac{\text{Row Total} \times \text{Column Total}}{\text{Total } n}$.

	Type of Fortune				
Cookie Brand	Prophecy	Advice	Wisdom	Misc.	Row Totals
Column Totals					

Calculate the expected counts in the two-way frequency table.

Are any of the expected counts less than 1? ____

What ________% of the table cells have an expected count greater than 5?

Are the necessary data conditions met for performing a chi-square test of homogeneity? Why or why not?

____

Even if you answered 'no' to the question above, calculate the chi-square test statistic:

Chi-Square Test Statistic:

	o-e	**$(o\text{-}e)^2$**	**$(o\text{-}e)^2/e$**
		TOTAL:	

Look up the *p*-value.
To determine the *p*-value we use the chi-square probability distribution with degrees of freedom where the
df = (Rows – 1)(Columns – 1)

What is the *p*-value = ___________________

Based upon the *p*-value, can you accept or reject the null hypothesis? What does that mean in regards to our extension question?

____

____

____

____