Class Time:

Names:

The Chi Square Distribution: Chi-Square Goodness of Fit Lab

Student Learning Outcome:

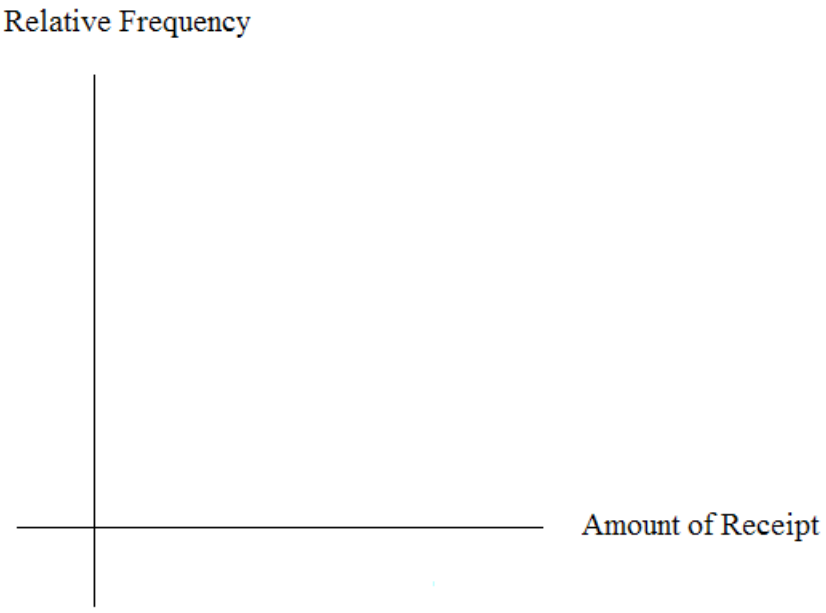
The student will evaluate data collected to determine if they fit either the uniform or exponential distributions.

Collect the data:

Go to your local supermarket. Ask 30 people as they leave for the total amount on their grocery receipts. (Or, ask 3 cashiers for the last 10 amounts. Be sure to include the express lane, if it is open.)

1. Record the values.

2. Construct a histogram of the data. Make 5 - 6 intervals. Sketch the graph using a ruler and pencil. Scale the axes.



3. Calculate the following:

   a. $\bar{x}$ = _____
   b. s = _____
   c. s^2 = _____

Uniform Distribution

Test to see if grocery receipts follow the uniform distribution.

1. Using your lowest and highest values, X ~ U(_____, _____)
2. Divide the distribution above into fifths.

3. Calculate the following

   a. lowest value = _____
   b. 20th percentile = _____
   c. 40th percentile = _____
   d. 60th percentile = _____
   e. 80th percentile = _____
   f. highest value = _____

4. For each fifth, count the observed number of receipts and record it. Then determine the expected number of receipts and record that.

Fifth	Observed	Expected
1st		
2nd		
3rd		
4th		
5th		

5. H_o:

6. H_a:

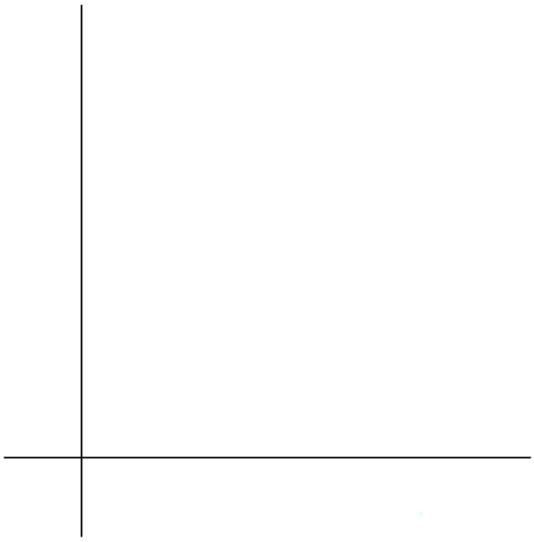
7. What distribution should you use for a hypothesis test?

8. Why did you choose this distribution?

9. Calculate the test statistic.

10. Find the p-value.

11. Sketch a graph of the situation. Label and scale the x-axis. Shade the area corresponding to the p-value.



12. State your decision.

13. State your conclusion in a complete sentence.

Exponential Distribution

Test to see if grocery receipts follow the exponential distribution with decay parameter $1/\bar{x}$.

1. Using $1/\bar{x}$ as the decay parameter, X ~ Exp(_____)
2. Calculate the following.

    a. lowest value = _____
    b. first quartile = _____
    c. 37th percentile = _____
    d. median = _____
    e. 63rd percentile = _____
    f. 3rd quartile = _____
    g. highest value = _____

3. For each cell, count the observed number of receipts and record it. Then determine the expected number of receipts and record that.

Cell	**Observed**	**Expected**
1st		
2nd		
3rd		
4th		
5th		
6th		

4. H_o

5. H_a

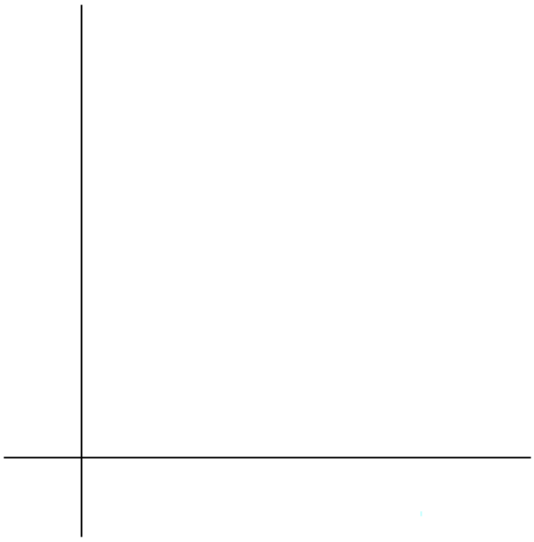
6. What distribution should you use for a hypothesis test?

7. Why did you choose this distribution?

8. Calculate the test statistic.

9. Find the p-value.

10. Sketch a graph of the situation. Label and scale the x-axis. Shade the area corresponding to the p-value.



11. Decision

12. State your conclusion in a complete sentence.

Discussion Questions

1. Did your data fit either distribution? If so, which one?

2. In general, do you think it's likely that data could fit more than one distribution? In complete sentences, explain why or why not.