Connexions module: m17049

THE CHI-SQUARE DISTRIBUTION: LAB I*

Susan Dean Barbara Illowsky, Ph.D.

This work is produced by The Connexions Project and licensed under the Creative Commons Attribution License †

Abstract

This module provides a lab on Chi-Square Distribution as a part of Collaborative Statistics collection (col10522) by Barbara Illowsky and Susan Dean.

Class Time:

Names:

1 Student Learning Outcome:

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

2 Collect the Data

NOTE: You may need to combine two categories so that each cell has an expected value of at least 5.

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.

Table 1

^{*}Version 1.9: Jun 19, 2012 8:58 pm -0500

[†]http://creativecommons.org/licenses/by/3.0/

Connexions module: m17049

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

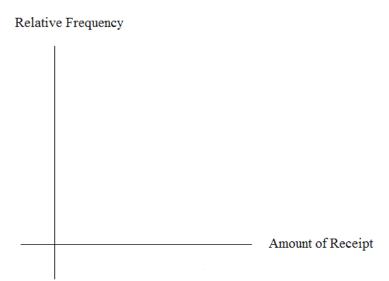


Figure 1

3. Calculate the following:

a. $\overline{x} =$

b. s =

 $c. s^2 =$

3 Uniform Distribution

Test to see if grocery receipts follow the uniform distribution.

- 1. Using your lowest and highest values, $X \sim U(\underline{}, \underline{}, \underline{},$
- 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 =

 $\textbf{f.} \ \ Highest \ value =$

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

Connexions module: m17049 3

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

Table 2

- $5 H_o$
- $6 H_a$:
- 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.

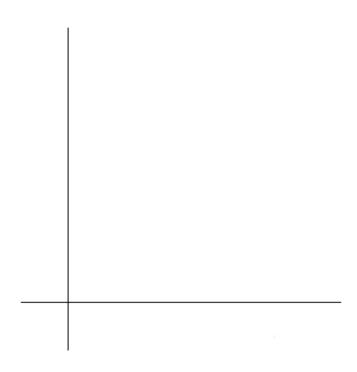


Figure 2

Connexions module: m17049

- 12. State your decision.
- 13. State your conclusion in a complete sentence.

4 Exponential Distribution

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

- 1. Using $\frac{1}{\overline{x}}$ as the decay parameter, $X \sim \mathrm{Exp}\,(_____).$
- 2. Calculate the following:
 - **a.** Lowest value =
 - **b.** First quartile =
 - **c.** 37th percentile =
 - \mathbf{d} . Median =
 - e. 63rd percentile =
 - $\mathbf{f.}$ 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		
$3\mathrm{rd}$		
$4 \mathrm{th}$		
$5 \mathrm{th}$		
$6 \mathrm{th}$		

Table 3

- 4. H_o
- 5. H.
- 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.

Connexions module: m17049 5

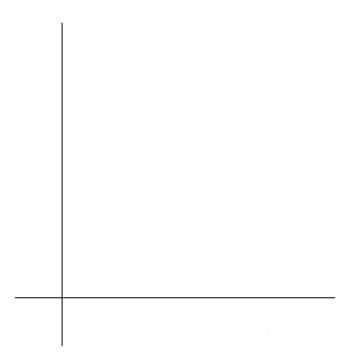


Figure 3

- 11. State your decision.
- 12. State your conclusion in a complete sentence.

5 Discussion Questions

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