Time Frame: 50 minutes

Subject Matter: Chi Square TELL ME

Objective: TSWBAT test a distribution for goodness of fit, using chi square.

Standards: DA – 3.8

 Materials: TI 83 Calculator and Worksheets

SHOW ME

Presentation of Information

Question:

 Where can one use the chi-square statistic?

* The chi-square statistic can be used to see whether a frequency distribution fits a specific pattern.

For example:

* To meet customer demands, a manufacturer of running shoes may wish to see whether buyers show a preference for a specific style.
* A traffic engineer may wish to see whether accidents occur more often on some days than on others, so that she can increase police patrols accordingly.
* An emergency service may want to see whether it receives more calls at certain times of the day than others, so it can provide adequate staffing.

Formula for the **Chi-Square** Goodness-of-Fit Test

$$χ^{2}=∑\frac{\left(O-E\right)^{2}}{E}$$

With degrees of freedom (d.f.) equals to the number of categories minus 1, and where

O = observed frequency

E = expected frequency

Example 1

A market analyst wished to see whether consumers have any preference among five flavors of a new fruit soda. Let $α=0.05$. A sample 0f 100 people provided these data:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cherry | Strawberry | Orange | Lime | Grape |
| 32 | 28 | 16 | 14 | 10 |

Solution:

* If there were no preference, one would expect each flavor to be selected with equal frequency.
* In this case, the equal frequency is $100÷5=20$.
* That is, approximately 20 people would select each flavor.

d.f. = 5 categories minus 1 = 4

Observed Frequency (O)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Frequency | Cherry | Strawberry | Orange | Lime | Grape |
| Observed | 32 | 28 | 16 | 14 | 10 |
| Expected | 20 | 20 | 20 | 20 | 20Expected Frequency(E) |

* State the hypotheses:

$H\_{0}:$ Consumers show no preferences for flavors.

$H\_{1}:$ Consumers show a preference.

* Find the critical value:

Use table G

* Compute the test value:

Observed Frequency (O)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Frequency | Cherry | Strawberry | Orange | Lime | Grape |
| Observed | 32 | 28 | 16 | 14 | 10 |
| Expected | 20 | 20 | 20 | 20 | 20Expected Frequency(E) |

$$χ^{2}=∑\frac{\left(O-E\right)^{2}}{E}$$

$χ^{2}=\frac{\left(32-20\right)^{2}}{20}$ + $\frac{\left(28-20\right)^{2}}{20}+\frac{\left(16-20\right)^{2}}{20}+\frac{\left(14-20\right)^{2}}{20}+\frac{\left(10-20\right)^{2}}{20}$

$χ^{2}=\frac{\left(12\right)^{2}}{20}$ + $\frac{\left(8\right)^{2}}{20}+\frac{\left(-4\right)^{2}}{20}+\frac{\left(-6\right)^{2}}{20}+\frac{\left(-10\right)^{2}}{20}$

$χ^{2}=7.2$ + $3.2+0.8+1.8+5$

Compare this to the critical value

$$χ^{2}=18.0$$

* Make the decision:

Accept $H\_{1}$ since the computed value is greater than the critical value.

* Summarize the result:

Consumers show a preference.

Example 2

A staff member of an emergency medical service wishes to determine whether the number of accidents is equally distributed during the week. A week was selected at random and the following data were obtained. Is there evidence to reject the hypothesis that the number of accidents is equally distributed throughout the week at $α=0.10$?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Day | Mon. | Tue. | Wed. | Thurs. | Fri. | Sat. | Sun. |
| No. of accidents | 28 | 32 | 15 | 14 | 38 | 43 | 19 |

Solution:

* The total number of accidents is 189.
* In this case, the equal frequency is $189÷7=27$.
* That is, approximately 27 accidents each day.

d.f. = 7 days minus 1 = 6

Observed Frequency (O)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Frequency | Mon. | Tue. | Wed. | Thurs. | Fri. | Sat. | Sun. |
| Observed | 28 | 32 | 15 | 14 | 38 | 43 | 19 |
| Expected | 27 | 27 | 27 | 27 | 27 | 27 | 27Expected Frequency(E) |

* State the hypotheses:

$H\_{0}:$ The number of accidents is equally distributed.

$H\_{1}:$ The number of accidents is not equally distributed.

* Find the critical value:

Use table G

* Compute the test value:

Observed Frequency (O)

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Frequency | Mon. | Tue. | Wed. | Thurs. | Fri. | Sat. | Sun. |
| Observed | 28 | 32 | 15 | 14 | 38 | 43 | 19 |
| Expected | 27 | 27 | 27 | 27 | 27 | 27 | 27Expected Frequency(E) |

$$χ^{2}=∑\frac{\left(O-E\right)^{2}}{E}$$

$χ^{2}=\frac{\left(28-27\right)^{2}}{27}$ + $\frac{\left(32-27\right)^{2}}{27}+\frac{\left(15-27\right)^{2}}{27}+\frac{\left(14-27\right)^{2}}{27}+\frac{\left(38-27\right)^{2}}{27}+\frac{\left(43-27\right)^{2}}{27}+\frac{\left(19-27\right)^{2}}{27}$

$χ^{2}=\frac{\left(1\right)^{2}}{27}$ + $\frac{\left(5\right)^{2}}{27}+\frac{\left(-12\right)^{2}}{27}+\frac{\left(-13\right)^{2}}{27}+\frac{\left(11\right)^{2}}{27}+\frac{\left(16\right)^{2}}{27}+\frac{\left(-8\right)^{2}}{27}$

$χ^{2}=0.04$ + $0.93+5.33+6.26+4.48+9.48+2.37$

Compare this to the critical value

$$χ^{2}=28.89$$

* Make the decision:

Accept $H\_{1}$ since the computed value is greater than the critical value.

* Summarize the result:

The number of accidents is not equally distributed.

Classwork

Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Period: \_\_\_\_\_\_\_\_\_\_\_\_\_\_ Date: April 12, 2011

A children’s raincoat manufacturer wants to know whether costumers prefer any specific color over other colors in children’s raincoats. He selected a random sample of 50 raincoats sold and notes the colors. The data are shown here. At $α=0.10$, is there a color preference for the raincoats?

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Color | Yellow | Red | Green | Blue |
| No. Sold | 17 | 13 | 8 | 12 |

Solution:

* The total number of raincoats is 50.
* In this case, the equal frequency is $50÷4=12.5$.
* That is, approximately 12.5 raincoats are sold each color.

d.f. = 4 colors minus 1 = 3

Observed Frequency (O)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Frequency | Yellow | Red | Green | Blue |
| Observed | 17 | 13 | 8 | 12 |
| Expected | 12.5 | 12.5 | 12.5 | 12.5Expected Frequency(E) |

* State the hypotheses:

$H\_{0}:$ Consumers show no preferences on colors

$H\_{1}:$ Consumers show preferences on colors.

* Find the critical value:

Use table G

Observed Frequency (O)

* Compute the test value:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Frequency | Yellow | Red | Green | Blue |
| Observed | 17 | 13 | 8 | 12 |
| Expected | 12.5 | 12.5 | 12.5 | 12.5Expected Frequency(E) |

$$χ^{2}=∑\frac{\left(O-E\right)^{2}}{E}$$

$χ^{2}=\frac{\left(17-12.5\right)^{2}}{12.5}$ + $\frac{\left(13-12.5\right)^{2}}{12.5}+\frac{\left(8-12.5\right)^{2}}{12.5}+\frac{\left(12 -12.5\right)^{2}}{12.5}$

$χ^{2}=\frac{\left(4.5\right)^{2}}{12.5}$ + $\frac{\left(0.5\right)^{2}}{12.5}+\frac{\left(-4.5\right)^{2}}{12.5}+\frac{\left(-0.5\right)^{2}}{12.5}$

$$χ^{2}=1.62+0.02+1.62+0.02$$

Compare this to the critical value

$$χ^{2}=3.28$$

* Make the decision:

Accept $H\_{0}$ since the computed value is less than the critical value.

* Summarize the result:

Consumers show no preferences on colors.