Student Notes : Chi Square Day 1

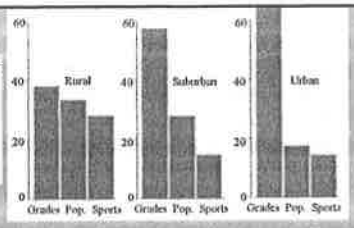
1

x^2 (Chi-Squared) Test of Independence

2

The corresponding column percentages are the following:

Goals	School Area: Rural	Suburban	Urban
Grades	35	50	65
Popular	34	25	17
Sports	20	14	14
Total	100	100	100



From the table and corresponding graphs, *it appears* that the emphasis on grades increases as the school areas become more urban, while the emphasis on popularity decreases.

Is this association significant?

3

Contingency tables are used to examine the relationship between two <u>qualitative or categorical</u> variables.

<u>For example,</u>
consider the hypothetical experiment on the effectiveness of early childhood intervention programs.

	Graduated	Failed to Graduate	Total
Experimental	73	12	85
Control	43	39	82
Total	116	51	167

4

The table shows that people in the experimental condition were more likely to graduate than were subjects in the control condition.

Thus, the <u>column</u> a person is in (graduated or failed to graduate) is ***contingent upon*** (depends on) the <u>row</u> the person is in (experimental or control)

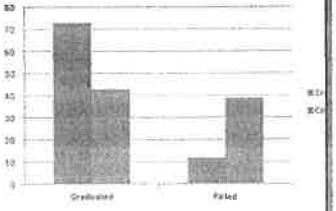
	Graduated	Failed to Graduate	Total
Experimental	73	12	85
Control	43	39	82
Total	116	51	167

5

If the columns are <u>not</u> **contingent on the rows, then the rows then the column frequencies are <u>independent</u>** from each other.

(which means no association or link between the two variables)

	Graduated	Failed to Graduate	Total
Experimental	73	12	85
Control	43	39	82
Total	116	51	167



6

The **TEST** of whether the columns are contingent on the rows is called the

Chi square test of independence.

7

The first task in computing the chi square test of independence is to compute the expected frequency for each cell with the assumption of independence.

In other words, we are assuming that the two variables (graduation and intervention) are *independent* from each other, or not linked.

8

	Graduated	Failed to Graduate	Total
Experimental	73	12	85
Control	43	39	82
Total	116	51	167

probability of being both experimental AND Graduated

•

	Graduated	Failed to Graduate	Total
Experimental			85
Control			82
Total	116	51	167

Expected frequency of people being both experimental AND Graduated =

9

Expected Frequency of Any Cell in a Contingency Table

$$f_e = \frac{r \cdot c}{n}$$

f_e = expected frequency

where, r = row total

c = column total

n = total number of observations

10

Once the expected cell frequencies are computed, it is convenient to enter them into the original table as shown below. The expected frequencies are in parentheses.

	Graduated	Failed to Graduate	Total
Experimental	73 (59.042)	12 (25.958)	85
Control	43 (56.958)	39 (25.042)	82
Total	116	51	167

OR

11

Observed frequencies

	Graduated	Failed to Graduate	Total
Experimental	73	12	85
Control	43	39	82
Total	116	51	167

Expected frequencies

	Graduated	Failed to Graduate	Total
Experimental	59.042	25.958	85
Control	56.958	25.042	82
Total	116	51	167

Warm Up Chi Sq Day 1

(1) The probability of flipping a head when a coin is tossed is 0.5 or 50% ?

What is the probability of flipping two coins and getting a head on each?

Now we'll flip one coin and roll one die. What is the probability of flipping a tail and getting a six?

(2)

	Graduated	Failed to Graduate	Total
Experimental	73	12	85
Control	43	39	82
Total	116	51	167

The probability of being in the experimental group is _____

The probability of someone graduating in the sample is _____

The probability of being BOTH in the experimental group and a graduate is:

(3) Now that we know the probability of being in both groups is _____, how many students would we expect to be a graduate from the experimental group?