Notes 7: Chi-Square tests (χ^2 tests)

The Z-test, t-test, and Pearson's r all assume that at least one of the variables (usually the dependent variable) is measured on the interval scale. When variables of interest are nominal or categorical, these statistical tests oftentimes are inappropriate, but the chi-square may prove useful. Two types of chi-square tests exist: goodness-of-fit and the contingency table chi-square (or test of association).

1. Goodness-of-Fit (one nominal or categorical variable present)

(a) Example: Sex Distribution Within a Class

With the goodness-of-fit test one is usually interested in determining whether a given distribution of data follows an expected pattern. For example, suppose one wishes to know whether the sex distribution of students in a class is random according to the distribution found within the school. Students in the school are about 50% female and 50% male. If students are assigned at random to classes one should therefore expect about a 50:50 split in females and males within a given classroom. One classroom, however, appears to have a student sex distribution that may not be the result of random assignment, or so thought the classroom teacher who seems to have a disproportionate number of troublesome boys.

Hypotheses

The null hypothesis states that the distribution of students by sex within a class fits the larger distribution found within the school, i.e.

H_0: males and females are equally distributed,

or

H_0: $f_1 = f_2$

or

H_0: $\text{distribution}_{pop} = \text{distribution}_{theory}$

and the alternative hypothesis is:

H_1: not H_0; student sex not equally distributed,

or

H_1: the frequencies f_1 and f_2 are not equal

or

H_1: $\text{distribution}_{pop} \neq \text{distribution}_{theory}$

Observed Data for Classroom Sex Distribution

The observed frequencies of students by sex in this class are:

	Observed
Female	6
Male	16

Total = 22

Does this distribution seem likely if indeed students are randomly assigned to classes? Given the sex distribution of students within the school, one should expect half to be females and half to be males.

	Observed	Proportion Expected	Expected Frequency (Total Frequency × Proportion)
Female	6	.5	22 × .5 = 11
Male	16	.5	22 × .5 = 11

Total = 22

Calculating χ^2 (chi-square)

The chi-square goodness-of-fit statistic to test H_0 can be calculated using the following formula:

$$\chi^2 = \sum \frac{(O_j - E_j)^2}{E_j} = n.\sum \frac{(p_j - \pi_j)^2}{\pi_j}$$

The formula on the far right is presented by some authors and is based upon proportions. The formula on the left is presented in most other textbooks and is based upon frequencies. The formula on the left is the one to be used for the remainder of this section.

The chi-square goodness-of-fit formula can be explained as follows:

(1) j = the unique cells or categories in the table of frequencies;
(2) O = the observed frequency in cell j;
(3) E = the expected frequency in cell j;
(4) $\sum$ = a summation sign—add up all squared terms once division has occurred;

The expected frequencies, E_j, are determined by theory. In the example above, one may expect students to be distributed equally by sex, so 22/2 = 11 for each of the two possible cells.

The value of χ^2 is obtained as follows:

$$\chi^2 = \frac{(6-11)^2}{11} + \frac{(16-11)^2}{11}$$

$$= \frac{25}{11} + \frac{25}{11}$$

= 2.273 + 2.273

= 4.546

The χ^2 distributions are (a) positively skewed, (b) have a minimum of zero, and (c) have just one parameter which is their degree of freedom (df).

Supplemental Note: Yates' Correction

Some software may present a corrected χ^2 value that adjusts for small samples or small cells sizes [count < 5 per cell]. This adjustment is known as the correction for continuity, or Yates' correction and the formula is simply to subtract -.5 from each O-E deviation before squaring:

$$\chi^2 = \sum \frac{(|O_j - E_j| - .5)^2}{E_j}$$

Note that one must use the absolute value of $|O_j\text{-}E_j|$ for the formula to work properly. In most cases this correction is not needed; sometimes this correction may over-correct and present values that are too small. Also, this correction is applied usually to χ^2 tests with df = 1 (2 cells for goodness-of-fit and 2x2 tables for tests of association [discussed elsewhere]).

Degrees of freedom

The df for goodness-of-fit chi-squares is defined as:

df (or ν) = J - 1

where J is the number of categories present. Since there were two categories in the example data (i.e., males and females), there is

df = 2 - 1 = 1

one df.

Testing H_0

To statistically test the tenability of the null hypothesis, one must determine whether the calculated value of χ^2 exceeds what would be expected by chance given that H_0 is true, i.e., does the calculated χ^2 exceed the critical value of χ^2?

The critical χ^2 or ${}_{crit}\chi^2$, can be found in Table 1 below.

Table 1: Upper Critical Values for Chi-square ($_{crit}\chi^2$)

	α				α		
df	0.10	0.05	0.01	df	0.10	0.05	0.01
1	2.706	3.841	6.635	51	64.295	68.669	77.386
2	4.605	5.991	9.210	52	65.422	69.832	78.616
3	6.251	7.815	11.345	53	66.548	70.993	79.843
4	7.779	9.488	13.277	54	67.673	72.153	81.069
5	9.236	11.070	15.086	55	68.796	73.311	82.292
6	10.645	12.592	16.812	56	69.919	74.468	83.513
7	12.017	14.067	18.475	57	71.040	75.624	84.733
8	13.362	15.507	20.090	58	72.160	76.778	85.950
9	14.684	16.919	21.666	59	73.279	77.931	87.166
10	15.987	18.307	23.209	60	74.397	79.082	88.379
11	17.275	19.675	24.725	61	75.514	80.232	89.591
12	18.549	21.026	26.217	62	76.630	81.381	90.802
13	19.812	22.362	27.688	63	77.745	82.529	92.010
14	21.064	23.685	29.141	64	78.860	83.675	93.217
15	22.307	24.996	30.578	65	79.973	84.821	94.422
16	23.542	26.296	32.000	66	81.085	85.965	95.626
17	24.769	27.587	33.409	67	82.197	87.108	96.828
18	25.989	28.869	34.805	68	83.308	88.250	98.028
19	27.204	30.144	36.191	69	84.418	89.391	99.228
20	28.412	31.410	37.566	70	85.527	90.531	100.425
21	29.615	32.671	38.932	71	86.635	91.670	101.621
22	30.813	33.924	40.289	72	87.743	92.808	102.816
23	32.007	35.172	41.638	73	88.850	93.945	104.010
24	33.196	36.415	42.980	74	89.956	95.081	105.202
25	34.382	37.652	44.314	75	91.061	96.217	106.393
26	35.563	38.885	45.642	76	92.166	97.351	107.583
27	36.741	40.113	46.963	77	93.270	98.484	108.771
28	37.916	41.337	48.278	78	94.374	99.617	109.958
29	39.087	42.557	49.588	79	95.476	100.749	111.144
30	40.256	43.773	50.892	80	96.578	101.879	112.329
31	41.422	44.985	52.191	81	97.680	103.010	113.512
32	42.585	46.194	53.486	82	98.780	104.139	114.695
33	43.745	47.400	54.776	83	99.880	105.267	115.876
34	44.903	48.602	56.061	84	100.980	106.395	117.057
35	46.059	49.802	57.342	85	102.079	107.522	118.236
36	47.212	50.998	58.619	86	103.177	108.648	119.414
37	48.363	52.192	59.893	87	104.275	109.773	120.591
38	49.513	53.384	61.162	88	105.372	110.898	121.767
39	50.660	54.572	62.428	89	106.469	112.022	122.942
40	51.805	55.758	63.691	90	107.565	113.145	124.116
41	52.949	56.942	64.950	91	108.661	114.268	125.289
42	54.090	58.124	66.206	92	109.756	115.390	126.462
43	55.230	59.304	67.459	93	110.850	116.511	127.633
44	56.369	60.481	68.710	94	111.944	117.632	128.803
45	57.505	61.656	69.957	95	113.038	118.752	129.973
46	58.641	62.830	71.201	96	114.131	119.871	131.141
47	59.774	64.001	72.443	97	115.223	120.990	132.309
48	60.907	65.171	73.683	98	116.315	122.108	133.476
49	62.038	66.339	74.919	99	117.407	123.225	134.642
50	63.167	67.505	76.154	100	118.498	124.342	135.807

If α = .05, the critical value for the example data is

${}_{crit}\chi^2 = 3.841$.

To test H_0, simply compare the obtained χ^2 against the critical, and if the obtained is larger, then reject H_0.

Decision Rule

If $\chi^2 \geq {}_{crit}\chi^2$, then reject H_0, otherwise FTR H_0.

With the current example, the decision rule is:

If 4.546 ≥ 3.841, then reject H_0, otherwise FTR H_0.

So reject the null (at alpha equal to .05) and conclude that the distribution of students by sex does not appear to be randomly distributed; i.e., there appears to be more males than females in this class and the numbers are larger than one would expect by chance alone.

If one were writing this result for a paper, it would be written as follows:

> The statistical results, χ^2 (df = 1) = 3.841, $p < .05$, indicate that the frequencies of students by sex are not equally distributed within this class. Males are disproportionately represented in this class and females are under-represented.

(b) Example: Sex Distribution with Unequal Proportions

Instead of the school population consisting of about 50% female and 50% male, assume the school population has more males than females. Specifically, assume that females represent 33% of students and males represent 67%. If students are assigned at random to classes one should therefore expect about 2/3 of students to be male.

What impact will this have on the expected frequencies? The revised table of expected frequencies appears below.

	Observed	Proportion Expected	Expected Frequency (Total Frequency × Proportion)
Female	6	.33	22 × .33 = 7.26
Male	16	.67	22 × .67 = 14.74

Total = 22

Calculating χ^2 (chi-square)

Using these revised proportions and expected frequencies result in the following χ^2 value:

$$\chi^2 = \frac{(6-7.26)^2}{7.26} + \frac{(16-14.74)^2}{14.74}$$

$$= \frac{1.588}{7.26} + \frac{1.588}{14.74}$$

$$= 0.219 + 0.108$$

$$= 0.327$$

Degrees of freedom and Critical Value

The df remain the same at

df = 2 - 1 = 1

so the critical value will also remain the same for $\alpha = .05$:
one df.

${}_{crit}\chi^2 = 3.841$.

The decision rule now states the following:

If 0.327 ≥ 3.841, then reject H_0, otherwise FTR H_0.

Since 0.327 is less than the critical value of 3.841 one would fail to reject the null and therefore conclude that the sex distribution of students within the class are similar to the distribution found within the school.

(c) Example: Birth Distribution

As a second example suppose one wishes to know whether the distribution of births throughout the year is random, with equal probabilities or frequencies during the year.

Hypotheses

The null hypothesis states that the distribution of births throughout the year is random and has equal frequencies, i.e.

H_0: frequency of births equally distributed throughout year,

or

H_0: $f_1 = f_2 = f_3 = ... = f_6$

or

H_0: $\text{distribution}_{pop} = \text{distribution}_{theory}$

and the alternative hypothesis is:

H_1: not H_0; births not equally distributed throughout year,

or

H_1: the frequencies $f_1, f_2, ..., f_6$ are not all equal

or

H_1: $\text{distribution}_{pop} \neq \text{distribution}_{theory}$

Observed Data for Birth by Month Distribution

Assume the researcher obtained information about frequencies of birth from the local hospital. The following frequencies were observed:

	Jan.-Feb.	March-April	May-June	July-August	Sept.-October	Nov.-Dec.
Observed	71	78	83	94	112	114

Total births = 552

Does this distribution seem random if indeed births are equally likely throughout the year; that is, does the null hypothesis of equality of births throughout the year seem tenable?

	Jan.-Feb.	March-April	May-June	July-August	Sept.-October	Nov.-Dec.
Observed	71	78	83	94	112	114
Proportion Expected	1/6 (.16667)	1/6 (.16667)	1/6 (.16667)	1/6 (.16667)	1/6 (.16667)	1/6 (.16667)
Expected Frequency (Total Frequency × Proportion)	552 × .16667 = 92	552 × .16667 = 92	552 × .16667 = 92	552 × .16667 = 92	552 × .16667 = 92	552 × .16667 = 92

Total births = 552; expected = 552/6 = 92

Calculating χ^2 (chi-square)

As noted above, the chi-square goodness-of-fit statistic to test H_0 can be calculated using the following formula:

$$\chi^2 = \sum \frac{(O_j - E_j)^2}{E_j}$$

The expected frequencies, E_j, are determined by theory. In the example above, it was expected that the frequencies of births would be equally distributed across the year. Since 552 births were observed during the year, this means that 552/6 = 92 births are expected every two months. If, for example, one wanted to find the expected frequency for every month, the expected frequency would be 552/12 = 46.

The value of χ^2 is obtained as follows:

$$\chi^2 = \frac{(71-92)^2}{92} + \frac{(78-92)^2}{92} + \frac{(83-92)^2}{92} + \frac{(94-92)^2}{92} + \frac{(112-92)^2}{92} + \frac{(114-92)^2}{92}$$

$$= \frac{441}{92} + \frac{196}{92} + \frac{81}{92} + \frac{4}{92} + \frac{400}{92} + \frac{484}{92}$$

= 4.793 + 2.130 + .880 + .043 + 4.348 + 5.261

= 17.455

Degrees of freedom

Since there were six categories in the example data, there are

df = 6 - 1 = 5

five df.

Testing H_0

The critical χ^2 or $_{crit}\chi^2$, can be found in Table 1 above. If α = .05, the critical value for the example data is

$_{crit}\chi^2$ = 11.07.

To test H_0, compare the obtained χ^2 against the critical, and if the obtained is larger, then reject H_0.

Decision Rule

If $\chi^2 \geq {}_{crit}\chi^2$, then reject H_0, otherwise FTR H_0.

With the current example, the decision rule is:

If 17.455 ≥ 11.07, then reject H_0, otherwise FTR H_0.

So reject the null (at alpha equal to .05) and conclude that the distribution of births does not appear to be randomly distributed; i.e., the distribution of births seems to be higher during the fall and earlier winter months than during other times of the year.

If one were writing this result for a paper, it would be written as follows:

> The statistical results, χ^2 (df = 5) = 17.455, $p < .05$, indicate that the distribution of births are not equally distributed throughout the year. Based upon the observed frequencies it appears that the birth rate is highest for the months of September to December, and lowest for the spring and summer months.

Exercises

1. Horse-racing fans often maintain that in a race around a circular track significant advantages accrue to the horses in certain post positions. Any horse's post position is his assigned post in the starting line-up. Position 1 is closest to the rail on the inside of the track; and position 8 is on the outside, farthest from the rail in an 8-horse race. Test whether post position is related to race results. (This example taken from S. Siegel, 1956, Nonparametric statistics for the behavioral sciences, McGraw-Hill.)

Listed below are the observed frequencies of 1st place finishers for each post position during a regular month at the tracks.

Post Position

	1	2	3	4	5	6	7	8
No. of wins	29	19	18	25	17	10	15	11

Total wins = 144.

(a) What are the null and alternative hypotheses?
(b) What are the expected frequencies if post position is unrelated to winning?
(c) What is the obtained and critical chi-square statistics and df if alpha is set at the .05 level?
(d) What are the results of this test?

2. The director of athletics at the local high school wonders if the sports program is getting a proportional amount of support from each of the four classes represented in the high school. If there are roughly equal numbers of students in each of the classes, what does the following breakdown of attendance figures from a random sample of students in attendance at a recent basketball game suggest? (This example taken from J. F. Healey, 1993, Statistics: A tool for social research, 3rd ed., Wadsworth.)

Class	Frequency
Freshmen	200
Sophomores	150
Juniors	120
Seniors	110

Total attendance = 580.

(a) What are the null and alternative hypotheses?
(b) What are the expected frequencies for attendance by class?
(c) What is the obtained and critical chi-square statistics and df if alpha is set at the .01 level?
(d) What are the results of this test?

3. Using the same data as above, assume that class rank is not equally distributed within the high school. Instead, about 30% of students are freshmen, 28% are sophomores, 22% are juniors, and 20% are seniors. Using these percentages to estimate expected values for attendance at the basketball game, is there any evidence that attendance differs from what one would expect given the class rank distribution of students enrolled in the school?

4. Suppose someone has a hypothesis that the "Transylvania effect" of the full moon is related to incidence of drug overdose. A search of medical files at a hospital yielded 1182 drug overdose cases which included the date. The full-moon phase was based on the actual dates of the full moon plus or minus 2 days, yielding 75 full-moon days and 381 non-full-moon days (total number of days was 456 for the period investigated). The observed frequency of drug overdose was 196 during the full-moon days and 986 during the non-full-moon days.

(a) What are the null and alternative hypotheses?
(b) What are the expected frequencies; how does one calculate their values for this example?
(c) What is the obtained and critical chi-square statistics and df if alpha is set at the .05 level?
(d) What are the results of this test?

2. Chi-Square Test of Association or Chi-Square Test of Contingency Tables (two nominal or categorical variable present)

With the goodness-of-fit test one is interested in determining whether a given distribution of data follows an expected pattern. For the test of association, one is interested in learning whether two (or more) categorical variables are related. It is most typical to find two categorical variables depicted in a contingency table (a cross-tabulation of the frequencies for various combinations of the variables). Note that contingency tables are referred to as 2-by-3, 3-by-3, etc. where the numerals are determined by the number of rows (R) and columns (C) in the table. If, for example, there is a table with two rows and two columns, the table is a R x C or 2 x 2 (2-by-2) table.

At issue in the following research question is whether the policy of allowing college faculty to take-on outside consultation for a fee is supported uniformly between tenured and untenured faculty. The data are as follows (example taken from D. E. Hinkle et al., 1979, Applied statistics for he behavioral sciences, Rand McNally):

	Support Policy	Do not Support Policy
Tenured	88	17
Nontenured	84	11

Total = 200.

Hypotheses

The null hypothesis states that there is no relationship between the two variables, i.e., that support for the consulting policy is independent of the tenure status of the faculty; or, that there is no difference between tenured and nontenured faculty regarding their support of the consulting policy.

H_0: distribution$_{tenured}$ = distribution$_{nontenured}$ (or the distributions are equal)

or

H_0: variable A (tenure status) is independent of variable B (policy support)

and the alternative hypothesis is:

H_1: some difference in the distributions

or

H_1: variables A and B are associated, not independent

Determining Expected Values

Expected values are determined by the column and row marginal frequencies. Marginal frequencies are pointed out below.

	Support Policy	Do not Support Policy	Marginal Row Frequencies
Tenured	88	17	88 + 17 = **105**
Nontenured	84	11	84 + 11 = **95**
Marginal Column Frequencies	88 + 84 = **172**	17 + 11 = **28**	Grand Total = 172 + 28 = **200**

Total = 200.

The following formula can be used to calculate expected frequencies for a given row r and column c, e.g., r = 1 and c = 1, which corresponds to cell "Tenured" and "Support Policy."

$$E_{rc} = \frac{(row_r total)(column_c total)}{N}$$

where **E_{rc}** is the expected value for row r and column c, **row_r total** is the marginal frequency for row r, **$column_c$ total** is the marginal frequency for column c, and **N** is the total sample size.

For the current example, the expected values are:

1. r = 1, c = 1 (tenured and support policy):

$$E_{11} = \frac{(105)(172)}{200} = \frac{18060}{200} = 90.3$$

2. r = 1, c = 2 (tenured and do not support policy):

$$E_{12} = \frac{(105)(28)}{200} = \frac{2940}{200} = 14.7$$

3. r = 2, c = 1 (nontenured and support policy):

$$E_{21} = \frac{(95)(172)}{200} = \frac{16340}{200} = 81.7$$

4. r = 2, c = 2 (nontenured and do not support policy):

$$E_{22} = \frac{(95)(28)}{200} = \frac{2660}{200} = 13.3$$

	Support Policy	Do not Support Policy	Marginal Row Frequencies
Tenured	88 (**90.3**)	17 (**14.7**)	105
Nontenured	84 (**81.7**)	11 (**13.3**)	95
Marginal Column Frequencies	172	28	200

Note: Expected values in parentheses.

Calculating χ^2 (chi-square)

The chi-square test of association statistic used to test H_0 can be calculated using the following formula:

$$\chi^2 = \sum \frac{(O_{rc} - E_{rc})^2}{E_{rc}}$$

The chi-square test of association formula can be explained as follows:

(1) rc = the unique cells or categories in the table of frequencies;
(2) O = the observed frequency in cell rc;
(3) E = the expected frequency in cell rc;
(4) $\sum$ = a summation sign—add up all squared terms once division has occurred;

The expected frequencies, E_{rc}, are determined in the manner demonstrated above in part (b).

The value of χ^2 is obtained as follows:

$$\chi^2 = = \frac{(88-90.3)^2}{90.3} + \frac{(84-81.7)^2}{81.7} + \frac{(17-14.7)^2}{14.7} + \frac{(11-13.3)^2}{13.3}$$

$$= \frac{5.29}{90.3} + \frac{5.29}{81.7} + \frac{5.29}{14.7} + \frac{5.29}{13.3}$$

$$= 0.06 + 0.06 + 0.36 + 0.40$$

$$= 0.88$$

The χ^2 distributions are (a) positively skewed, (b) have a minimum of zero, and (c) have just one parameter which is their degree of freedom (df).

Degrees of freedom

The df for association chi-squares is defined as:

df (or ν) = (R - 1)(C - 1)

where R is the number of rows present and C is the number of columns present.

Since there were two rows and two columns in the example data, there is

df = (2 - 1)(2 - 1) = 1.

Testing H_0

To statistically test the tenability of the null hypothesis, one must determine whether the calculated value of χ^2 exceeds what would be expected by chance given that H_0 is true, i.e., does the calculated χ^2 exceed the critical value of χ^2?

The critical χ^2 or ${}_{crit}\chi^2$, can be found in Table 1 above. If $\alpha = .05$, the critical value for the example data is

${}_{crit}\chi^2 = 3.84$.

To test H_0, simply compare the obtained χ^2 against the critical, and if the obtained is larger, then reject H_0.

Decision Rule

If $\chi^2 \geq {}_{crit}\chi^2$, then reject H_0, otherwise FTR H_0.

With the current example, the decision rule is:

If $0.88 \geq 3.84$, then reject H_0, otherwise FTR H_0.

So fail to reject the null (at alpha equal to .05) and conclude that policy support does not depend upon tenure status.

If one were writing this result for a paper, it would be written like follows:

> The statistical results, χ^2 (df = 1) = 0.88, $p > .05$, indicate that one's decision to support the policy of consultations does not appear to be associated with one's tenure status. In short, support of the consulting policy is independent of the tenure status of the faculty.

Computer output for the previous example

```
           | col
        row|         1          2 |     Total
-----------+----------------------+----------
         1 |        88         17 |       105
         2 |        84         11 |        95
-----------+----------------------+----------
      Total|       172         28 |       200

          Pearson chi2(1) =   0.8809   Pr = 0.348
```

Exercises

1. A researcher wishes to determine whether an experimental treatment (RPT) enhances achievement and academic self-efficacy. The researcher must use two intact classes for the experiment since random assignment is not possible. Good research requires that the experimental and control groups be as equivalent as possible at the start of the experiment to ensure adequate internal validity. To help establish that the two classes are equivalent, the researcher plans to collect IQ and ITBS scores to determine whether a statistical difference exists between the two groups on these measures. In addition, the researcher will try to show that the two groups also have similar racial distributions. The following data are collected for the two classes:

	Black	Hispanic	White
Class 1	15	7	10
Class 2	12	6	17

(a) What are the null and alternative hypotheses?
(b) What are the expected frequencies?
(c) What is the obtained and critical chi-square statistics and df if alpha is set at the .05 level?
(d) What is the decision rule?
(e) What are the results of this test?

2. Is there a relationship between high school program of study and whether the student eventually dropped out of college? Some educators argue that students who study under college preparatory programs are much better prepared for college than are students who studied under general education programs or vocational education programs. Listed below are dropout figures for students enrolled in a medium sized, midwestern university. Determine whether dropping out is related to program of study in high school.

High School Program of Study	Dropped Out of College	Graduated from College
Vocational	289	323
General	334	456
College Prep.	230	698

(a) What are the null and alternative hypotheses?
(b) What are the expected frequencies?
(c) What is the obtained and critical chi-square statistics and df if alpha is set at the .01 level?
(d) What is the decision rule?
(e) What are the results of this test?

3. A research psychologist wants to investigate the impact of instructor feedback upon mastery of a complex learning task. Four groups of ten subjects each are selected to participate. One group receives only positive feedback, another only negative feedback. The third receives both positive and negative feedback, the fourth receives no feedback at all. The following are the results:

Type of Feedback	Successful	Nonsuccessful
Positive	6	4
Negative	4	6
Both	8	2
None	3	7

(a) What are the null and alternative hypotheses?
(b) What are the expected frequencies?
(c) What is the obtained and critical chi-square statistics and df if alpha is set at the .01 level?
(d) What is the decision rule?
(e) What are the results of this test?

<u>Compute answers for extra examples</u>

(1) Class by Race

```
           | col
       row|         1          2          3 |     Total
-----------+---------------------------------+----------
         1 |        15          7         10 |        32
         2 |        12          6         17 |        35
-----------+---------------------------------+----------
     Total|        27         13         27 |        67

          Pearson chi2(2) =   2.0949   Pr = 0.351
```

(2) Dropout by Program of Study

```
           | col
       row|         1          2 |     Total
-----------+----------------------+----------
         1 |       289        323 |       612
         2 |       334        456 |       790
         3 |       230        698 |       928
-----------+----------------------+----------
     Total|       853       1477 |      2330

          Pearson chi2(2) =  96.5579   Pr = 0.000
```

(3) Success by Type of Feedback

```
           | col
       row|         1          2 |     Total
-----------+----------------------+----------
         1 |         6          4 |        10
         2 |         4          6 |        10
         3 |         8          2 |        10
         4 |         3          7 |        10
-----------+----------------------+----------
     Total|        21         19 |        40

          Pearson chi2(3) =   5.9148   Pr = 0.116
```