Sociology 405/805
Revised January 11, 2004

Measures of Association for Hypothetical Relationship

This handout gives several possible examples of a relationship between opinions of people and their sex. The data is hypothetical. In the four tables, the relationship is analyzied using cross classifications, the chi square test of independence, and measures of association. For the last example, a test for the difference between two proportions is conducted, and the similarity of this test to the chi square test is demonstrated.

In each of the following tables, the first entry in each cell represents the observed number of respondents with each combination of characteristics. The entry in brackets represents the expected number of respondents, under the assumption of no relationship between opinion and sex. For each table, there is one degree of freedom. For each table, the null and research hypotheses are:

$$H_0 : \ O_i = E_i$$

$$H_1 : \ O_i \neq E_i$$

In words, these hypotheses are:

$$H_0 : \text{No relation between opinion and sex}$$

$$H_1 : \text{Some relation between opinion and sex}$$

$\phi = \sqrt{\frac{\chi^2}{n}}$

$C = \sqrt{\frac{\chi^2}{n+\chi^2}}$

$V = \sqrt{\frac{\chi^2}{nt}}$ where $t = \min(r-1, c-1)$, with r as the number of rows and c as the number of columns in the table. For a 2×2 table, $t = 1$ and $V = \phi$.

1. No Relationship Between Opinion and Sex

Opinion	Male	Female	Total
Agree	60 (60.0)	30 (30.0)	90
Disagree	40 (40.0)	20 (20.0)	60
Total	100	50	150

Table 1: No Relation Between Opinion and Sex

For Table 1, $\chi^2 = 0$.

$\phi = \sqrt{\frac{\chi^2}{n}} = 0.$

$C = \sqrt{\frac{\chi^2}{n+\chi^2}} = 0.$

2. Perfect Relationship between Opinion and Sex

Opinion	Male	Female	Total
Agree	100 (66.7)	0 (33.3)	100
Disagree	0 (33.3)	50 (16.7)	50
Total	100	50	150

Table 2: Perfect Relation Between Opinion and Sex

For Table 2, $\chi^2 = 150$ and this is statistically significant at less than the 0.005 level.

$\phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{150}{150}} = 1.$

$C = \sqrt{\frac{\chi^2}{n+\chi^2}} = \sqrt{\frac{150}{150+150}} = \sqrt{0.5} = 0.707.$

3. Weak Relationship between Opinion and Sex

Opinion	Male	Female	Total
Agree	65 (60.0)	25 (30.0)	90
Disagree	35 (40.0)	25 (20.0)	60
Total	100	50	150

Table 3: Weak Relation Between Opinion and Sex

For Table 3, $\chi^2 = 3.125$ and this is statistically significant at the $\alpha = 0.10$ level of significance but not at the $\alpha = 0.05$ level of significance. Exact significance is 0.0771.

$\phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{3.125}{150}} = \sqrt{0.02083} = 0.144.$

$C = \sqrt{\frac{\chi^2}{n+\chi^2}} = \sqrt{\frac{3.125}{3.125+150}} = 0.143.$

4. Strong Relationship between Opinion and Sex

Opinion	Male	Female	Total
Agree	75 (66.7)	25 (33.3)	100
Disagree	25 (33.3)	25 (16.7)	50
Total	100	50	150

Table 4: Strong Relation Between Opinion and Sex

For Table 4, $\chi^2 = 9.375$ and this is statistically significant at the $\alpha = 0.005$ level of significance. Exact significance is 0.0022.

$\phi = \sqrt{\frac{\chi^2}{n}} = \sqrt{\frac{9.375}{150}} = \sqrt{0.0625} = 0.250.$

$C = \sqrt{\frac{\chi^2}{n+\chi^2}} = \sqrt{\frac{9.375}{9.375+150}} = 0.243.$

Test for a Difference between two proportions

Let population 1 be males and population 2 be females. Also let p_1 be the true proportion of all males who agree, and let p_2 be the true proportion of all females who agree. The null hypothesis of no relationship between opinion and sex is equivalent to the null hypothesis that the proportion of males who agree is equal to the proportion of females who disagree. If one has no prior knowledge concerning the direction of the relationship, then the research hypothesis is that there is a difference in the proportion of males and females who agree. The null and research hypotheses are then

$$H_0 : p_1 = p_2$$

$$H_1 : p_1 \neq p_2$$

The test statistic is $\hat{p}_1 - \hat{p}_2$, and if the sample sizes for both samples, n_1 and n_2, are large, then

$$\hat{p_1} - \hat{p_2} \text{ is } \mathrm{Nor}\left(p_1 - p_2, \sqrt{pq\,(1/n_1 + 1/n_2)}\right).$$

For this sample, $n_1 = 100$ males and $n_2 = 50$ females, and these are large enough to ensure that $\hat{p}_1 - \hat{p}_2$ is normally distributed as above. The estimate of the values of p and q is provided by $\hat{p}$ and $\hat{q}$ where

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

and $\hat{q} = 1 - \hat{p}$. Defining success as agreeing, the number of successes among the males, population 1, is $X_1 = 65$, and the number of successes among population 2, the females, is $X_2 = 25$. For the estimate of the standard deviation of the difference in proportions,

$$\hat{p} = \frac{75 + 25}{100 + 50} = \frac{100}{150} = 0.667.$$

and $\hat{q} = 1 - 0.667 = 0.333$. The value of Z for this test is:

$$Z = \frac{\hat{p}_1 - \hat{p}_2 - (p_1 - p_2)}{\sqrt{\hat{p}\hat{q}\,(1/n_1 + 1/n_2)}}$$

Based on the sample data,

$$\hat{p}_1 = \frac{75}{100} = 0.75$$

and

$$\hat{p}_2 = \frac{25}{50} = 0.50$$

$$Z = \frac{0.75 - 0.50 - 0}{\sqrt{0.667 \times 0.333\,(1/100 + 1/50)}}$$

$$Z = \frac{0.25}{\sqrt{0.0066667}} = \frac{0.25}{0.0816497} = 3.062.$$

This Z value has an area of $0.5000 - 0.4989 = 0.0011$ to the right of it, and an equal area to the left of $Z = -3.062$. The conclusion is to reject the null hypothesis, and accept the alternative hypothesis, at all significance levels above $\alpha = 0.0022$.

This result is consistent with the earlier finding from the chi square test for independence. There, the null hypothesis was rejected at the level of 0.005 significance.

Also note that this test for a difference between two proportions is identical to the earlier test for independence between the two variables. For Table 4, $\chi^2 = 9.375$. For the test for a difference between two proportions, $Z = 3.062$, and $Z^2 = 9.375$, the same value as the chi square value. This result holds because the chi square distribution is constructed as the sum of the squares of normally distributed variables. In the case of the chi square distribution with one degree of freedom, as is the case in the above 2×2 table, $Z^2 = \chi^2$. Such a result would not hold in a table with more categories, so that there would be more than one degree of freedom. However, for a 2×2 table, it makes no difference whether one conducts a test for independence between the two variables, or whether one conducts a test for the difference between two proportions.