References:

**Degree of Freedom (DF)**

The concept of degree of freedom derives from the idea that once you know that values of a certain number of cells, the values in the rest of the cells are determined. Let’s illustrate this point with a simple example. Suppose we construct a contingency table with only four cells, like the one in table 4, to examine the association between age and happiness. We select 25 respondents at random from the 1996 GSS and classify them according to how old they are and how happy they are. Of the 25 respondents, we notice right away that 15 are ages 18 to 39 and 10 are 40 and up. We enter those values I our contingency table as the column marginal totals for the two categories of the dependent variable. Next we start to fill in the cells of the table. We begin by finding out how many belong in the first cell, young and not to happy. We learn that 13 belong in the first cell, so we enter the number 13.

Table 4.1 Understanding Degrees of Freedom Using a Hypothetical Set of Data

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 Age 18 through 39 40 and up Row totals

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Happiness Not to happy 9 22

 Happy 3

 Column Totals 15 10 N= 25

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Once we have done that, notice that all of the values in the rest of the cells are determined by the row and column marginal totals- we don’t have to count the respondents cell by cell. We know that there are 2 respondents in the 18 through 39 and happy cell. (How? Because 13 + 2 = the column marginal total, 15.) We know there are 9 respondents in the 40 and up and not too happy cell (because 13 + 9 = the row marginal total 22). There is one respondent who is 40 and up and happy.

When we know the row and column marginal totals, we only have to know the contents of one of the cells in the table to complete the whole contingency table. Only one cell, therefore, is free to vary. Consequently, this table is one in which there is only one degree of freedom. The concept degrees of freedom tells us how many cells in a table are free to vary, once the row and column marginal totals are known.

Table 4.2 Contingency Table with row and column totals fro a random sample

 Of 48 1996 GSS respondents for the variables *agecat* and *happy*

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Happy General Happiness \* AGECAT recoded age- four categories Crosstabulation

 AGECAT recoded age – Four categories

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 1 18 thru 25 | 2 26 thru 39 | 3 40 thru 60 | 4 65 thru 89 | total |
|  |  |  |  | 19 |
|  |  |  |  | 26 |
|  |  |  |  | 3 |
| 7 | 13 | 15 | 13 | 48 |
|  |  |  |  |  |

Happy General 1 not to

Happiness 2 pretty

 3 very happy

Total
Skills Practice 5 see whether you can figure out how many degrees of freedom there are in Table 4.2, a table of 48 respondents selected at random from the 1996 GSS. How many cell frequencies do you have to fill in before the values in the rest of the cells are determined? You will see the answer when you get to Example 2 in the next section. (Hint: Try filling in the cells with hypothetical frequencies to see how many you have to fill in before the rest are determined.)

Although it’s important to understand how degrees of freedom are determined, the degrees of freedom in a table can be found more easily using a formula.

**Formula:**  Degrees of freedom for chi-square

 *df* = (r-1) (c-1)

where r represents the number of rows in a contingency table, and c represents the number of columns.

**Example:** For table 4.2, we can compute the degrees of freedom with the formula by counting the number of rows and number of columns and entering them in the formula.

 *df =* (r – 1) (c -1 ) = (2 -1) (2 – 1) = (1)(1) = 1

There is only one degree of freedom in table 4.1. The frequency for only one cell is free to vary. Once that frequency has been determined, the row and column marginal totals dictate the contents of the rest of the cells.

**Example:** For table 4.2, we can compute the degrees of freedom as follows:

 *df=* (r -1)(c -1) = (4 -1)(3 -1) = (3)(2) = 6

J. Richard Kendrick, Jr. 2000. “Degree of Freedom” in Social Statistics. Mayfield Publishing Company. Pp.476-477