

# Introduction To Tabulation

Contributed by John F Hall

The notes below are intended to refer to data analysis rather than statistics as such; they especially refer to the analysis of data from sample surveys. Wherever possible examples are drawn from real surveys conducted by students and/or staff at the Polytechnic of North London.

© Copyright 1990 John F Hall

The notes below are intended to refer to data analysis rather than statistics as such; they especially refer to the analysis of data from sample surveys. Wherever possible examples are drawn from real surveys conducted by students and/or staff at the Polytechnic of North London.

Bear in mind that there is a difference of approach which may at first seem strange to students in sociology and related subjects. Most surveys are conducted by a personal interview with respondents, and most analysis is of a descriptive kind, taking the people themselves as units of analysis. Another, and more rigorous approach, is what we call explanatory, in which we attempt to explain rather than describe, and in which we look at variables rather than people. Both approaches are dealt with here.

The basic idea

Social research involves many weird and wonderful methods over which debate, often bitter, rages continuously. However, at some stage even the most virulently anti-positivist and anti-empiricist will need to be able to sort and count things, or to read, understand or even act on, reports based on things which have been named, sorted and counted. Perhaps the easiest way of explaining one of the most basic skills in statistics is to try to make sense of raw data through a process of naming, sorting and counting. For instance, take the following data relating to sixth form students. Information is provided on their sex and on their intentions towards higher education. Student H.E.?

1	Male	Yes
2	Male	No
3	Female	Yes
4	Female	No
5	Female	No
6	Male	No
7	Female	No
8	Male	No
9	Female	No
10	Female	Yes
11	Male	Yes
12	Male	No
13	Male	Yes
14	Female	No
15	Male	Yes
16	Male	No
17	Female	No
18	Female	No
19	Male	No

20	Male	No
----	------	----

It is not easy to tell from these data how many males and females there are, let alone make any meaningful statement about the relationship between sex and plans for higher education. What can we do to make them easier to understand?

The first thing we need to do is to sort them into some kind of order. We can do this by arranging all the males in one group and the females in another, or we can do it by sorting all those with H.E. plans into one group and the rest into another. Thus by sex:

Female	Yes
Female	No
Female	No
Female	No



Sex:

Female 9 45%  
 Male 11 55%  
 -----  
 Total 20 100%

College:

No 14 70%  
 Yes 6 30%  
 -----  
 Total 20 100%

If we want to summarise data from both variables at the same time we need to construct a contingency table. We do this by constructing a blank table with the same number of rows as there are categories in one of the variables, and the same number of columns as there are categories in the other. Let us take "Sex" as the column variable and "College plans" as the row variable. In this case both variables have only two categories, and so the table will have 2 rows and 2 columns, and therefore 4 cells.

Sex	Male	Female
College		
No		
Yes		

These four cells form the body of the table into which we can now enter the counts from the list sorted on both variables at once. At the same time we enter outside the table the row-totals and column-totals from the original frequency distributions for each variable and the grand total for the number of cases in the whole table. Thus:

Sex

(Raw data)		Male	Female	Row Total
College				
No	14	7	7	14
Yes	6	4	2	6
Column total		11	9	20

Column total 11 9 20

This is at least a little easier to interpret than the original sorted lists, but it is still difficult to answer a question as to whether males are more likely to want to go college than are females, or vice versa. To answer this question we need to ask not, "How many?", but, "What proportion?" of each sex have college plans. One further operation necessary - to standardise the data by converting the raw counts for each sex into percentages - to enable direct comparison between sexes.

(% data)

(% data)		Male	Female	Row Total
College				
No	63.6	77.8	70.0	
Yes	36.4	22.2	30.0	
Column total	100.0	100.0	100.0	
	(Base for %)	(11)	(9)	(20)

From this table we can now state that female sixth-formers are less likely to have plans for Higher Education.

^