

## 2.1: Frequency Tables

Most material in this text is introduced first at an abstract level, then generally a step-by-step recipe is given and finally example problems are solved. This general to specific approach to learning statistics is the opposite of how many introductory statistics tests for the social sciences teach. For our first topic of frequency tables, the abstract concept is counting so let's dive into the recipe with the expectation that you won't get the complete picture until an example or two is worked.

The construction of a frequency table proceeds in two steps :

**Step 1 :** Determine the classes. There are two possibilities here, either the classes are given to you (pre-defined) or you have to define the classes based on the number of groups you want. So either

- i. Classes are given – nothing to do.
- ii. Define classes based on the number of groups you want. There are a number of different ways to group data into classes. We will cover a method here, different from Bluman's, that works for whole number data only. Here are the steps for that method :

(a) determine high data limit,  $H$  and the low data limit,  $L$ .

(b) compute the range  $R = H - L$

(c) compute the class width :

$$W = \frac{R + 1}{G} \quad (2.1.1)$$

where  $G$  is the number of groups (or classes) you want.

(d) Begin the frequency table's first two columns :

Class	Class Boundaries
$L$ to $(L + W - 1)$	$(L - 0.5)$ to $(L - 0.5 + W)$
$(L + W)$ to $(L + 2W - 1)$	$(L - 0.5 + W)$ to $(L - 0.5 + 2W)$
$\vdots$	$\vdots$
	$(H + 0.5 - W)$ to $(H + 0.5)$

**Note :** If the classes are given, you won't have, or need, the second column.

In the class column above a specific way of labelling classes is given. (We will see how this works exactly in the upcoming example.) This is to make the class names useful for seeing that the classes are uniquely defined — there will be no data points on the boundaries of the classes. The numbers in the labels will be whole numbers, since we are assuming that the data are whole numbers<sup>[1]</sup>. In general we can label the classes any way we like.

Also we need to note that this procedure of defining classes using the formula given in step (2)(c) will only work for whole number data. In general the process of defining classes is a lot looser; there are few rules beyond thinking about what kind of information you hope to capture by defining the classes. Since I want to keep you focused on learning the basic ideas and not worry about stuff that is not really statistics all assignment and exam questions that ask for the construction of classes from quantitative data will be for whole number data only. The procedure given here does work in general but some data points may end up on class boundaries and will have to make up an arbitrary rule about which class the data point should go in.

**Step 2 :** Construct the frequency table and fill it in :

Class	Class Boundaries	Tally	Frequency	Cumulative Freq.	Relative Freq.
			$a$	$a$	$a/n$
			$b$	$a + b$	$b/n$
			$c$	$a + b + c$	$c/n$

				⋮	⋮
				$n$	

The last number in the cumulative frequency column,  $n$ , should equal number of data points as a check since it is the sum of the frequencies. And the sum of the relative frequencies will be 1 — we will see that this is an essential feature of probabilities. The tally column is optional.

**Example 2.1 :** 25 army inductees were tested for blood type. The data are :

A	B	B	AB	O
O	O	B	AB	B
B	B	O	A	O
A	O	O	O	AB
AB	A	O	B	A

Construct a frequency table.

**Solution :**

**Step 1 :** Classes are given : A B O AB

**Step 2 :** Construct frequency table :

Class	Tally	Frequency	Cumulative Freq.	Relative Freq.
A		5	5	$5/25 = 0.20$
B		7	12	$7/25 = 0.28$
O		9	21	$9/25 = 0.36$
AB		4	25	$4/25 = 0.16$

The tally is actually silly in this case because you count<sup>[2]</sup> all the instances of A for the class A, etc., and you're done. The tally column will be more useful for the next example.

**Example 2.2 :** Given the high temperature data for each of 50 states for the month of July :

112	100	127	120	134	118	105	110	109	112
110	118	117	116	118	122	114	114	105	109
107	112	114	115	118	117	118	122	106	110
116	108	110	121	113	120	119	111	104	111
120	113	120	117	105	110	118	112	114	114

Construct a frequency table using 7 classes.

**Solution :**

**Step 1 :**

(a) High limit,  $H = 134$

Low limit,  $L = 100$

(b) Range:  $R = H - L = 134 - 100 = 34$

(c) Class width:  $W = \frac{R+1}{G} = \frac{34+1}{7} = 5$

(d) (and continue to Step 2) :

Step 2 :

Class	Class Boundaries	Tally	Frequency	Cumulative Freq.	Relative Freq.
100 — 104	99.5 to 104.5		2	2	0.04
105 — 109	104.5 to 109.5		8	10	0.16
110 — 114	109.5 to 114.5	etc.	18	28	0.36
115 — 119	114.5 to 119.5		13	41	0.26
120 — 124	119.5 to 124.5		7	48	0.14
125 — 129	124.5 to 129.5		1	49	0.02
130 — 134	129.5 to 134.5		1	50	0.02
					= 1

Note how we can now use the tally column to keep track of our counting. For example, for the class 100 — 104, we first count all the instances of 100 (there is 1), then 101 (none), 102 (none), 103 (none) and 104 (one). The sum of the frequencies is  $n = 50$  and the sum of the relative frequencies is 1. Imagine that this data set represented the whole population and not just a sample. Then if you picked a random state there would be a 0.16 probability that the temperature would be between 105 and 109 inclusive. On other words relative frequency = probability for a population. Hence the term *frequentist* definition of probability.  $\square$

You can also compute cumulative relative frequency in a frequency table. When you use SPSS to make a frequency table you will run up against the limitations of using black box canned software. SPSS produces only one style of frequency table and it doesn't match what we've been doing. In fact SPSS won't compute relative frequency; instead it computes "percentage". You need to convert percentage to relative frequency in your brain by dividing by 100.

1. Whole numbers are 0 and the positive integers. [↩](#)
2. The frequency of A is the number of times A is in the dataset, etc. [← the take-home concept here. ↩](#)

This page titled [2.1: Frequency Tables](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Gordon E. Sarty](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.