Measures of Central Tendency*

David Lane

This work is produced by The Connexions Project and licensed under the
Creative Commons Attribution License †

In the previous section[1] we saw that there are several ways to define central tendency. This section defines the three most common measures of central tendency: the mean, the median, and the mode. The relationships between these measures of central tendency and the definitions given in the previous section will probably not be obvious to you. Rather than just tell you these relationships, we will allow you to discover them in the simulations in the sections that follow.

This section gives only the basic definitions of the mean, median and mode. A further discussion of the relative merits and proper applications of these statistics is presented in a later section[2].

1 Arithmetic Mean

The **arithmetic mean** is the most common measure of central tendency. It simply the sum of the numbers divided by the number of numbers. The symbol m is used for the mean of a population. The symbol M is used for the mean of a sample. The formula for m is shown below:

$$m = \frac{\Sigma X}{N}$$

where ΣX is the sum of all the numbers in the numbers in the sample and N is the number of numbers in the sample. As an example, the mean of the numbers $1 + 2 + 3 + 6 + 8 = \frac{20}{5} = 4$ regardless of whether the numbers constitute the entire population or just a sample from the population.

The table, Number of touchdown passes (Table 1: Number of touchdown passes), shows the number of touchdown (TD) passes thrown by each of the 31 teams in the National Football League in the 2000 season. The mean number of touchdown passes thrown is 20.4516 as shown below.

$$m = \frac{\Sigma X}{N} = \frac{634}{31} = 20.4516$$

Number of touchdown passes

37	33	33	32	29	28	28	23
22	22	22	21	21	21	20	20
19	19	18	18	18	18	16	15
14	14	14	12	12	9	6	

*Version 2.5: Oct 4, 2010 11:02 am -0500

†http://creativecommons.org/licenses/by/3.0/

[1] "Introduction to Central Tendency" <http://cnx.org/content/m10942/latest/>

[2] "Comparing Measures of Central Tendency" <http://cnx.org/content/m11011/latest/>

Table 1

Although the arithmetic mean is not the only "mean" (there is also a geometic mean), it is by far the most commonly used. Therefore, if the term "mean" is used without specifying whether it is the arithmetic mean, the geometic mean, or some other mean, it is assumed to refer to the arithmetic mean.

2 Median

The **median** is also a frequently used measure of central tendency. The median is the midpoint of a distribution: the same number of scores are above the median as below it. For the data in the table, Number of touchdown passes (Table 1: Number of touchdown passes), there are 31 scores. The 16th highest score (which equals 20) is the median because there are 15 scores below the 16th score and 15 scores above the 16th score. The median can also be thought of as the 50th percentile[3].

Let's return to the made up example of the quiz on which you made a three discussed previously in the module Introduction to Central Tendency[4] and shown in Table 2: Three possible datasets for the 5-point make-up quiz.

Three possible datasets for the 5-point make-up quiz

Student	Dataset 1	Dataset 2	Dataset 3
You	3	3	3
John's	3	4	2
Maria's	3	4	2
Shareecia's	3	4	2
Luther's	3	5	1

Table 2

For Dataset 1, the median is three, the same as your score. For Dataset 2, the median is 4. Therefore, your score is below the median. This means you are in the lower half of the class. Finally for Dataset 3, the median is 2. For this dataset, your score is above the median and therefore in the upper half of the distribution.

Computation of the Median: When there is an odd number of numbers, the median is simply the middle number. For example, the median of 2, 4, and 7 is 4. When there is an even number of numbers, the median is the mean of the two middle numbers. Thus, the median of the numbers 2, 4, 7, 12 is $\frac{4+7}{2} = 5.5$.

3 mode

The **mode** is the most frequently occuring value. For the data in the table, Number of touchdown passes (Table 1: Number of touchdown passes), the mode is 18 since more teams (4) had 18 touchdown passes than any other number of touchdown passes. With continuous data such as response time measured to many decimals, the frequency of each value is one since no two scores will be exactly the same (see discussion of continuous variables[5]). Therefore the mode of continuous data is normally computed from a grouped frequency distribution. The Grouped frequency distribution (Table 3: Grouped frequency distribution) table shows a grouped frequency distribution for the target response time data. Since the interval with the highest frequency is 600-700, the mode is the middle of that interval (650).

[3] "Percentiles" <http://cnx.org/content/m10805/latest/>
[4] "Introduction to Central Tendency" <http://cnx.org/content/m10942/latest/>
[5] "Distributions" <http://cnx.org/content/m10868/latest/>

Grouped frequency distribution

Range	**Frequency**
500-600	3
600-700	6
700-800	5
800-900	5
900-1000	0
1000-1100	1

Table 3