7.10.09

PP

## Section 10: The Mean: An Average

The most popular average is the ***mean***.[1] It is so popular that it is sometimes simply called *the average*; however, this is an ambiguous term because there are several different types of averages used in statistics.

Computation of the mean is quite simple; just sum (i.e., add up) the scores and divide by the number of scores. Here is an example:

Scores: 5, 6, 7, 10, 12, 15
Sum of scores: 55
Number of scores: 6
Computation of mean: 55/6 = 9.166 = 9.17

Notice in the example that the answer was computed to three decimal places and rounded to two. In scientific work, the mean is usually reported to two decimal places.

There are several symbols for the mean. In scientific journals, the most commonly used symbols are ***M*** and ***m***.[2] Statisticians often use this symbol:

$\bar{X}$

It is pronounced *X-bar*. Remember that *X* without the bar stands for a score or a set of scores. Throughout this book, the symbol *M* will be used.

Although calculating the mean is simple, it is important to become familiar with the symbols in the formula because they will be used later in this book in other formulas. The formula for the mean is:

N

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

Mean = Sum of scores / # of cases

[1] Its full, formal name is the *arithmetic mean*. Other averages are described in the next section.
[2] Strictly speaking, the upper-case *M* should be used when describing an entire population and the lower-case *m* should be used when describing a sample drawn from a population. Many authors of applied research, however, ignore this convention.

The symbol $\Sigma$, which is the Greek letter sigma, is pronounced *sum of* in statistics. The $X$ stands for *score* or *scores*. $N$ is the number of cases or subjects. Thus, the formula says: *The mean equals the sum of the scores divided by the number of cases.*

An important characteristic of the *mean* is that it is the balance point of the distribution; that is, it is the *point around which the deviations sum to zero.*[3]

The example in Table 10.1 illustrates this characteristic. The sum of the scores is 60; dividing this by 5 yields a mean of 12.00. By subtracting the mean from each score, we obtain the deviations from the mean. These deviations sum to zero.[4] (Notice the negatives cancel out the positives when summing.)

Table 10.1
*Scores and Their Deviations from Their Mean*

| Score | Mean | Deviation |
|---|---|---|
| 7 | 12.00 | −5 |
| 11 | 12.00 | −1 |
| 11 | 12.00 | −1 |
| 14 | 12.00 | 2 |
| 17 | 12.00 | 5 |
| | Sum of deviations = 0 | |

A major drawback of the mean is that it is drawn in the direction of extreme scores. This is a problem if there are *either* some extremely high scores that pull it up *or* some extremely low scores that pull it down. Here's an example expressed in cents of the contributions to charity by two groups of children:

**Group A:** 1, 1, 2, 3, 3, 4, 4, 4, 5, 5, 5, 5, 6, 6, 6, 7, 8, 10, 10, 10, 11

**Mean for Group A** = 5.52

[3]This is a *defining characteristic* of the mean. There is only one value that has this characteristic for a given distribution. Any value that does *not* have this characteristic is *not* the mean.
[4]If the mean is not a whole number, the sum of the deviations may vary slightly from zero because of rounding when determining the mean. A rounded mean is not *precisely* accurate.

**Group B:** 1, 2, 2, 3, 3, 3, 4, 4, 5, 5, 5, 6, 6, 6, 6, 6, 9, 10, 10, 150, 200

**Mean for Group B** = 21.24

Notice that, overall, the two distributions are quite similar. Yet the mean for Group B is much higher than the mean for Group A because of two students who gave extremely high contributions of 150 cents and 200 cents.[5] If only the means for the two groups were reported without reporting all of the individual contributions, it would suggest that the average student in Group B gave about 21 cents when, in fact, none of the students made a contribution of about this amount. An average that provides a more accurate indication of the center for this situation is described in the next section.

Another limitation of the mean is that it is only appropriate for use with *interval* and *ratio* scales of measurement (see Section 2) because its value is dependent upon the magnitude of the scores, which create the size of their deviations from the mean, unlike the averages described in the next section.

Note that a synonym for *average* is ***measure of central tendency***. Although the latter term is seldom used in reports of scientific research, you may encounter it if you refer to other statistics texts.

## Terms to Review Before Attempting Worksheet 10

**mean, $M$, $m$, $\Sigma$, $X$, $N$, measure of central tendency**

[5] A distribution with extremes that produce a tail in one direction but not the other is called *skewed*. See Section 9.

## Worksheet 10: The Mean: An Average

**_Riddle_: What type of magic can auto mechanics do?**

DIRECTIONS: To find the answer to the riddle, write the answer to each question in the space immediately below it. The word in parentheses in the solution section next to the answer to the first question is the first word in the answer to the riddle, the word beside the answer to the second question is the second word, and so on.

1. To two decimal places, what is the mean of these scores: 10, 15, 17, 17, 20, 22?

   16.83

2. To two decimal places, what is the mean of these scores: 0, 0, 2, 5, 8, 9, 12?

   5.14

3. In scientific journals, what is the most common symbol for the mean?

   M • $m^2$

4. In statistics, how is $\Sigma$ pronounced?

   "sum of"

5. For what does *N* stand?

   number of cases or subjects

6. If the mean for a group is 15.00 and Sylvia, who is a member of the group, obtained a score of 14, what is the value of the deviation score associated with Sylvia's score? −1

7. If the mean is subtracted from each of the scores underlying it and the deviations are summed, what value is obtained? 0

   Sum of ~~X~~ deviations

## **Worksheet 10** (Continued)

8. "For a skewed distribution, the mean is pulled in the direction of the extreme scores." Is this statement true or false?

True

9. Suppose that two homes at the top of a hillside community were sold for very high prices, and twenty homes lower on the hill were sold for modest prices. Would the mean probably be good for giving an accurate indication of the average home price in the community?

~~Yes~~ NO

10. What is a synonym for the term *averages*?

Measure of central tendency
or
mean