Chapter 5

The Median and Mode: Alternatives to the Mean

As you learned in the previous chapter, the *mean* is the balance point in a distribution—that is, it is the point at which the deviations below the mean balance (or cancel out) the deviations above the mean. Although the mean is the most popular average, there are two conditions under which it should not be used: (1) when a distribution is clearly skewed, and (2) when the data are not equal interval. In this chapter, you will learn about two alternative averages that can be used when the mean is inappropriate.

The Median

The *median* is the average that is defined as the *middle score* (that is, the midpoint) in a distribution of ranked scores. As the middle score, it is the point that has half the scores above it and half the scores below it. Here's a simple example from the previous chapter that includes an outlier score among children's contributions to a charity expressed in cents:

4, 8, 8, 9, 11, 14, 80

We saw that the *mean* contribution is 19.1, which is not very representative of the amounts donated since none of the children gave anything like 19 cents. The *median* contribution, on the other hand, is 9, which is quite representative.

To determine the median, first put the scores in order from low to high. Then:

(1) When the number of scores is odd, the median is the middle score. (Note that there are three scores below 9 and three scores above 9 in the above example.)

or

(2) When the number of scores is even, sum the middle two scores and divide by 2. (For example, for the scores 0, 4, 5, and 12, the middle

two scores are 4 and 5. Summing 4 and 5 and dividing by 2, we get 9/2 = 4.5, which is the median.)

A minor complication arises when there are ties in the middle (that is, when two or more cases have the same score in the area where the median lies). Here's an example:

3, 6, 7, 7, 7, 20

As you can see, when you count to the middle (three scores up or three scores down), you come to 7. Two of the 7s are in the middle of the distribution, but one of them is "above" the middle. What is the median? Well, we can use the rule we used earlier to get an approximation. Since there is an even number of scores and the middle two scores are 7, we sum them and divide by 2: 7 + 7 = 14/2 = 7, which is the approximate median. For all practical purposes, this approximation is usually more than adequate. (A method for taking the ties into account is presented in Appendix A. If you apply the method in Appendix A to the scores in this example, you will get a median of 6.8, which is very close to the value of 7 we obtained using a much easier method.)

What if there is an odd number of scores with a tie in the middle? For an approximation, we can apply the rule described on page 27. Here's an example:

1, 5, 6, 7, 8, 8, 8, 11, 50

Since there are 9 scores, we count up 5 or down 5 and come to 8, which is our approximate median. (Using the more difficult method in Appendix A, we would get 7.7, which rounds to 8, illustrating again that our approximation method is quite sound.)

When should you use the median? Under two circumstances:

(1) when analyzing equal interval data for which the mean is not appropriate because the distribution is highly skewed, and

(2) when analyzing ordinal data. As you recall from Chapter 1, ordinal data put cases in rank order, such as teachers' rankings of a list of ten discipline problems from 1 (most important) to 10 (least important). For each type of problem, such as "hitting another child," we could calculate the median rank. The medians would allow us to report which problem the average teacher thought was the most important, which one was the next most important, and so on.

The Mode

The last average we will consider is the *mode*. It is defined as the *most frequently occurring score*. Here's an example we looked at earlier in this chapter, where we found that the median is 8:

1, 5, 6, 7, 8, 8, 8, 11, 50

The mode is also 8 because 8 occurs more often than any other score. Note that the mode does not always have the same value as the median and, thus, does not always have an equal number of cases on each side of it. Here's another example we looked at earlier in this chapter, where we found that the median is 9:

4, 8, 8, 9, 11, 14, 80

For this example, the mode is 8, which does not have an equal number of cases on both sides of it.

A strength of the mode is that it is easy to determine. However, the mode has several serious weaknesses. First, a distribution may have more than one mode. Here's an example, where the modes are 9 and 10:

6, 6, 8, 9, 9, 9, 10, 10, 10, 15

Since we want a single average, the mode is not a good choice here. Second, for a small population, each score may occur only once, in which case, all scores are the mode since all occur equally often. For these reasons, the mode is seldom used. Instead, almost all researchers use either the mean or median as the average.

Concluding Comment

The mean, median, and mode all belong to the "family" of statistics called "averages." They constitute a family because they all are designed to present *one type* of information. A more formal name for this family is *measures of central tendency*. In the next chapter, we will begin our consideration of another family of statistics, *measures of variability*.