

## DESCRIBING THE DATA

In the preceding chapter we looked at ways of organizing data that would give us an overall picture of its general shape. In most research this overall picture is helpful but does not give us figures which we can use to justify our answers to our research questions. If you want to be able to defend your answers, it is not enough to say that the frequency distribution was such and such.

Clearly, when the director of courses walks up to you after the midterm and asks how your foreign students did on the exam, you cannot tell her: 47, 56, 39, 77, 88, 23, etc. And it won't work to tell her that 77% are in the 59th percentile either unless you have something more to say about 77% of what by whom or the 59th percentile of what measured against whom. What you need is a figure which will let her know the most typical score the students received and just how typical it was.

The purpose of this chapter is to show how you can arrive at these most typical scores and also the reservations that you must keep in mind when you interpret them. These reservations are important if we are to be as reasonable and logical as everybody expects researchers to be. The typical score is also important, for it allows us to compare different groups.

Suppose you conducted an experiment using a pretest posttest control group design. (If you don't remember what that is, review Chapter 3.) You have four different distributions: two for the pretest (experimental and control) and two for the posttest. Suppose the representative typical scores of each distribution on a scale of 50 were the following:

Control	25	27
Experimental	26	40
	Pretest	Posttest

This is valuable information because it tells you:

1. At the time of the pretest there appears to be no difference between the groups.
2. At the time of the posttest (after the treatment), there appears to be a large difference between the groups (your treatment worked!).

3. The difference is in favor of your experimental group, not the control. Thus, the most typical score is both useful and crucial to your research.

The term *central tendency* is used to talk about the central point in the distribution of scores in the data. There are three measures of central tendency: the *mode*, the *median*, and the *mean*.

### MODE

The *mode* is the *most frequently obtained score* in the data. For example, in the following data the mode is 25:

20 22 23 23 25 25 25 25 27 29 30

If you draw a frequency polygon, you can just drop a line from the peak to the baseline, and the number on the baseline will be the mode. In bimodal distributions there are two values which are obtained most often, e.g.:

2 3 4 4 4 4 5 7 7 9  
10 10 10 10 12 12 13 15

This distribution has two modes, 4 and 10.

The mode is the easiest measure of central tendency to identify. It does not need any calculations. However, it has some limitations which we should consider. The most serious limitation in using the mode as a measure of central tendency is that it is easily affected by chance scores. Suppose we had the following data:

35 35 40 40 45 45 45 46 46  
48 52 55 60 60 60 60 65 70

The mode of this distribution is 60. However, suppose that one of the *Ss* who scored 46 just by chance had scored 45 and one of the *Ss* who had scored 60 had scored 61 instead. Then the mode would shift from 60 to 45. Quite a difference. There are problems in using the mode when you have only a few *Ss*; if the number of *Ss* is very large, these limitations will be diminished.

### MEDIAN

The median score is also easy to find. Arrange your scores in rank order. The *median* is the score which is at the center of the distribution. Half of the scores are above the median and half are below. If the number of scores is odd, the median is the middle score: 4 4 5 7 9 10 11. If the number of scores is even, use the midpoint between the two middle scores as the median: 4 5 7 9 10 12 ( $7 + 9 \div 2 = 8$ ). The median score is often used as the measure of central tendency when the number of scores is small and/or when the data are obtained by ordinal measurement.

## MEAN

The *mean* is the most commonly used measure of central tendency because it takes all scores into account. It is an arithmetic concept. The arithmetic concept is that if the scores are distributed along a scale, the mean will fall exactly at the balance point. Think of the scale as a playground seesaw and the scores as the weights on the seesaw. To make it balance you have to move the plank back and forth until you get the exact balance point. That balance point is the mean.

You already know how to compute the mean because it is the same as the "average": add up all the scores and divide by the number of scores. Suppose you gave a vocabulary test in your class. There were 20 items and the scores you obtained were: 16 10 5 6 8 15 20 14 16 10. You add the scores and get 120. There are 10 scores; so you divide 120 by 10. The mean is 12. Although this computation is obvious, at this point let's start to learn some of the shorthand that will help you to read statistical studies.  $\bar{X}$  ("X-bar") is the symbol for the mean of a sample. The formula for obtaining the mean is

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

Doesn't that look scientific! All it means is that to get the mean ( $\bar{X}$ ), you have to add up ( $\Sigma$  means to add up or to sum) all the individual observations of  $X$  and divide by  $N$  (the total number of observations). Did you remember the symbols  $\Sigma$  and  $N$  from previous chapters?

Although the mean is the most frequently used measure of central tendency, it too has a limitation. It is seriously sensitive to extreme scores. As an example, the following sets of data are identical except for one score. Imagine you gave a test to your two ESL classes, all of whom are beginners. By mistake, a native speaker of English showed up in your class and thought she was supposed to take the test too.

Native-speaker score	80	20
	18	18
	16	16
	12	12
	10	10
	4	4
	<hr/>	<hr/>
	$\Sigma X = 144$	$\Sigma X = 84$
	$\bar{X} = 21$	$\bar{X} = 12$

Her score has changed from the mean for her group drastically. Such mistakes don't always happen fortunately, and if the  $N$  is large, the effect of extreme scores is very small. (There are also ways of adjusting for extreme scores.) In any case, the mean is the best, most practical, most useful measure of central tendency.

Look at Table 6.1, which shows the reading speed of students in a remedial reading program. First scan the data for the mode. The largest number of *S*s fell into the 125–149 group, right? Are there fast or slow readers in the group that seem to be extreme? Note that two *S*s scored over 300 and that the next *S* is in the 225–249 range. If you use the mean as the measure of central tendency, which way would it be slanted (higher or lower than the mode)?

If you do not have a huge number of cases or a huge number of observations, you can easily look at your frequency distribution first for the mode as a quick eyeball-look at the data. Then look at the median and note whether there are some really extreme observations that might distort the picture. Then calculate the mean. If you have a normal distribution, you will have a bell-shaped (symmetric) curve, and all three measures (mode, median, and mean) will be the same.

Table 6.1. Reading speed of students in a remedial group

Words per minute	<i>f</i>
300+	2
275 – 299	0
250 – 274	0
225 – 249	1
200 – 224	0
175 – 199	2
150 – 174	5
125 – 149	8
100 – 124	4
75 – 99	0
50 – 74	1

### MEASURES OF VARIABILITY

Once you have decided on your measure of central tendency and found your most typical score, there are still some reservations to keep in mind about the typical score. Suppose we gave a test to measure reading speed to two different classes and they both turned out to have the same mean score. Does this imply

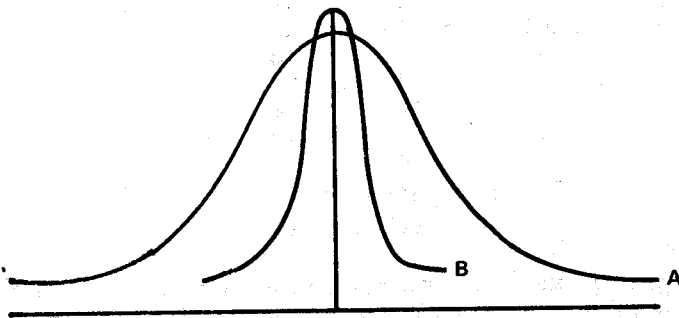


Figure 6-1

that the two classes are really the same? No, of course it doesn't. The variability among the scores, how they spread out from the central point, may be quite different in the two groups. Compare the polygons of Figure 6-1 for two distributions both of which have the same mean score. Notice that in Class A the

spread of scores is larger than in Class B. Therefore, to be able to talk about data more accurately, we have to measure the degree of variability of the data from our measure of central tendency.

Just as there are three ways of talking about the most typical score in your data, there are three major ways, too, to show how the data are spread out from that point: the range, the standard deviation, and variance.

### Range

The easiest, most informal way to talk about the spread of the distribution of scores is the range. Again, if you told your director of courses that the mean score of your class on the midterm was 70, you could give her an idea of the spread of the scores by also telling her the range. If the scores ranged from a high of 87 to a low of 61, the range would be 26. To calculate the range:

1. Arrange the scores from the highest to the lowest.
2. Subtract the lowest score from the highest score.

$$\text{Range} = X_{\text{highest}} - X_{\text{lowest}}$$

The problem with using range as an index of variability is that it changes drastically with the magnitude of the extreme scores. Imagine if your class had one person who scored a zero on the test, the range would be dramatically changed just by that one score. Since it is an unstable measure, it is rarely used for statistical analyses. But it is a useful, informal measure.

### Standard Deviation

The most frequently used measure of variability is the standard deviation. It is "standard" in the sense that it looks at the average variability of all the scores around the mean; all the scores are taken into account. The larger the standard deviation, the more variability from the central point in the distribution. The smaller the standard deviation, the closer the distribution is to the central point.

When you give back test papers in your Methods of Teaching ESL class and tell the teacher-trainees that the mean score on the exam was 76, you can be sure that every person in that class immediately checks to see how many points above or below the mean his or her score was. If my score was 82, I know I scored 6 points above the mean. That is the deviation of my score from the mean. While students may find this information useful, for research we want to know more than just individual deviation from the mean. We want to know the average deviation of all scores from the mean. To do that we start with all the individual deviation scores. *Little x* symbolizes the deviation of the individual score from the mean.

Consider, for example, the data on scores of ten Ss on a short cloze passage: 2, 3, 3, 4, 5, 5, 5, 6, 6, 8. The mean ( $\bar{X} = \Sigma X \div N$ ) is  $47 \div 10 = 4.7$ . Right? The individual deviation of each score is:

Sample cloze data

$X$ (scores)	$X - \bar{X}$	$x$
2	2 - 4.7	-2.7
3	3 - 4.7	-1.7
3	3 - 4.7	-1.7
4	4 - 4.7	-0.7
5	5 - 4.7	.3
5	5 - 4.7	.3
5	5 - 4.7	.3
6	6 - 4.7	1.3
6	6 - 4.7	1.3
8	8 - 4.7	3.3

At this point, if you add up all the individual deviation scores, you may be surprised because the  $\Sigma x$  is always zero. But remember the mean is the balance point on the seesaw. If you add up the minus weights on one side of that seesaw and the plus weights on the other side, you will get zero because they balance each other out. Obviously, adding them up and dividing by the number of cases is not going to work to get us the standard deviation. So, what we do is square each of the individual deviation scores and then add them up.

Deviation scores for cloze data

$X$ (scores)	$x$ (individual deviations)	$x^2$
2	-2.7	7.29
3	-1.7	2.89
3	-1.7	2.89
4	-0.7	.49
5	.3	.09
5	.3	.09
5	.3	.09
6	1.3	1.69
6	1.3	1.69
8	3.3	10.89
$\Sigma X = 47$	$\Sigma x = 0$	$\Sigma x^2 = 28.10$
$N = 10$		
$\bar{X} = 4.7$		

You might reason now that since we want to know the average amount of variation from the score, we could simply divide our total by  $N$  (in this case 10). This would be a perfectly valid procedure if you have a large  $N$  (over 100). However, with a small sample, mathematicians have determined that it is more accurate to divide by  $N - 1$ .

$$\frac{\Sigma x^2}{N - 1} = \frac{28.10}{9}$$

Now you will remember that in order to get rid of the minus scores and the problem of its balancing out to zero, we squared the individual variation scores. So, now we need to change it back. We take the square root to get the standard deviation.

$$\begin{aligned} s &= \sqrt{\frac{\Sigma x^2}{N-1}} \\ &= \sqrt{\frac{28.10}{9}} \\ &= \sqrt{3.12} \\ &= 1.77 \end{aligned}$$

To reiterate, the steps are:

1. Calculate the mean:  $\bar{X}$ .
2. Subtract the mean from each score to get the individual deviation scores:  $x = X - \bar{X}$ .
3. Square each individual deviation and then add them up:  $\Sigma x^2$ .
4. Divide by  $N - 1$ :  $\Sigma x^2 / (N - 1)$ .
5. Take the square root of the result:  $\sqrt{\Sigma x^2 / (N - 1)}$ .

Can you see that the two directions below say the same thing? If you can't, review the directions for getting individual deviation scores.

$$s = \sqrt{\frac{\Sigma(X - \bar{X})^2}{N - 1}} \quad \text{or} \quad s = \sqrt{\frac{\Sigma x^2}{N - 1}}$$

To clarify the concept of standard deviation we have used this method. However, there is an easier way to find it—one which allows you to use raw scores rather than finding each individual deviation score. The formula is

$$s = \sqrt{\frac{\Sigma X^2 - [(\Sigma X)^2 / N]}{N - 1}}$$

All it says is first to square each of your scores. Using our cloze scores again, this would give us:

$X$	$X^2$	$X$	$X^2$
2	4	5	25
3	9	5	25
3	9	6	36
4	16	6	36
5	25	8	64

The  $\Sigma X$  is 47 and the  $\Sigma X^2$  is 249. Plugging these totals into the formula, we get

$$\begin{aligned}
 s &= \sqrt{\frac{\Sigma X^2 - [(\Sigma X)^2 / N]}{N - 1}} &&= \sqrt{\frac{249 - 220.9}{9}} \\
 &= \sqrt{\frac{249 - [(47)^2 / 10]}{9}} &&= \sqrt{3.12} \\
 & &&= 1.77
 \end{aligned}$$

If you are confused by the two symbols  $\Sigma X^2$  and  $(\Sigma X)^2$ , remember that  $\Sigma X^2$  means to first square every score and then add them up and  $(\Sigma X)^2$  means to add up all the scores and then square the sum.

What can the standard deviation tell us? We said that the figure tells us the standard of how far out from the point of central tendency the individual scores are distributed. If you were a teacher new to the English Language Institute and were told that you could pick one class to teach from four sections of Intermediate English, which of the classes in Table 6.2 would you pick? You look at the mean scores and decide there isn't much difference between sections 1 and 2 or between 3 and 4

Table 6.2. Language aptitude scores of intermediate group

Section	$\bar{X}$	$s$
1	55.6	23.4
2	56.4	4.8
3	39.1	5.3
4	38.1	18.7

Then you look at the standard deviations. You know the scores for sections 1 and 4 are much more widely spread than those for class sections 2 and 3. How different do you think section 1 is from 2? Do you want a homogeneous class or would you rather pick one where you can use your latest techniques for mixed groups, peer teaching, and small-group work? Or do you like to work with classes where everyone keeps together and you can use whole-group instruction? Which would you select? When we make such decisions on the basis of data, the standard deviation gives us information which the mean alone cannot give us. It can be as important (or even more important) than the mean.

*Variance*

In most statistical analyses the variance is used as the measure of variability. Variance is the sum of the squared deviation scores divided by  $N - 1$ . To find it, then, you simply stop short of the last step in calculating the standard deviation. You do not need to bother with finding the square root.

$$s = \sqrt{\frac{\Sigma(X - \bar{X})^2}{N - 1}} \qquad \text{Variance} = \frac{\Sigma(X - \bar{X})^2}{N - 1}$$

With the raw score the formulas would be

$$s = \sqrt{\frac{\Sigma X^2 - (\Sigma X)^2 / N}{N - 1}} \qquad \text{Variance} = \frac{\Sigma X^2 - (\Sigma X)^2 / N}{N - 1}$$



You will frequently find variance annotated as  $s^2$ .

$$\text{Variance} = s^2 \text{ (standard deviation squared)}$$

or

$$\text{Standard deviation} = \sqrt{\text{variance}} \text{ (square root of variance)}$$

In reporting our research results, the basic information that we have covered in this chapter, mean scores (occasionally median or mode scores) *and* the standard deviations must be presented. Not only is the standard deviation important for decisions such as those above, but the amount of deviation from the central score will be important in our statistical analyses.

### ACTIVITIES

1. Interview as many ESL teachers as you can regarding their hourly rate of pay. Find the mean, mode, and median for their pay. Give two reasons to explain why teachers are paid by the hour rather than by yearly contract.
2. The ages of Ss enrolled in one Detroit Community Adult School ESL class are: 21, 28, 42, 31, 24, 26, 24, 23, 29, 32, 33, 41, 37, 22, 24, 21, 21, 26, 22, 33, 48, 21, 28, 26, 21, 24, 25, 21, 22, 36, 31, 23. Find the mean, mode, and median for age.
3. Manfred Evans School in Los Angeles has a championship soccer team. Here is information about some of the players. Which data are appropriate for discovering the mean? Which are not and why? Which information would their opponents be most interested in? Does the range appear to be extreme for any of the data?

Player No.	ESL Level	Age	Weight	Height	Country
54	4	21	183	5'8"	Egypt
23	3	18	192	5'9"	Brazil
17	5	17	141	5'5"	Cuba
14	7	23	165	5'6"	Algeria
21	7	19	155	5'10"	Germany
48	5	23	145	5'5"	France
42	2	28	190	5'7"	Bulgaria
29	4	25	175	5'7"	Egypt
16	2	16	145	5'6"	Costa Rica
20	3	17	190	5'8"	Mexico
44	6	17	175	5'8"	Mexico
52	6	21	147	5'11"	Colombia
60	2	19	152	5'10"	Peru

4. The practice typing paper in typewriters on display at Fedco Department Store provides the following data. Find the mean of any factor you wish. What else could you do if you had the real raw data to work with? What would your research question(s) be?

Message 1: length 23 words, 12 English, 9 Spanish, 2 uncertain. Misspelled words, 2. Content—description of typewriter.

Message 2: length 48 words, English. Misspelled words 14 (typos on 8). Content—political message.

Message 3: length 52 words, English with 5 Japanese names. Misspelled words, 0. Content—names, numbers, brand names.

Message 4: length 29, English. Misspelled words, 4. Content—description of person's family.

Message 5: Length 170, English. Misspelled words, 15. Content—obscene.

Message 6: Length 138, English. Misspelled words, 17. Tirade against obscenity.

Message 7: Length 16, English. 0 misspelled. Content—"Now is the time for all good men. . ."

Message 8: Length, 43, English 21, Spanish 22. 5 typos (corrected by striking over). Content—slogans and names.

Message 9: Length ?. Content—the alphabet and numbers.

Message 10: Length 6 words, English. Content: "This machine stinks don't buy it."

5. The Spanish version of the Peabody Picture Vocabulary Test was given to Spanish-speaking children in your kindergarten. The Unified School District asks each school to submit means and standard deviations for each kindergarten. Their scores are: 24, 33, 41, 18, 27, 26, 19, 18, 14, 32, 29, 27, 17. What is your report?
6. Your language lab assistant found that during the past quarter the language lab was used as follows: 1 class (elementary) 40 hours; 2 classes (lower intermediate) 8 hours, 10 hours; 3 classes (upper intermediate) 8 hours, 0 hours, 6 hours; 5 classes (advanced), 0 hours, 10 hours, 4 hours, 2 hours, 0 hours. Because he just got a new calculator, he decided to report the mean and standard deviation for lab use. What were his figures? If he actually did give you the mean and standard deviation for these data, what would you say to him? Why?
7. The children in your ESL class have a great deal of difficulty spelling words with long vowels. You spot-check their written work and then take a 500-word sample of each student's writing to count the long vowel errors. Write out your findings as raw data for 10 children. Find the mean and standard deviation.
8. You have collected a number of observations on a second grade child who has already been diagnosed as dyslexic. He has copied a list of 10 words, each approximately 4 or 5 letters in length, at three-day intervals. The number of errors (including omission) are as follows: 3, 9, 23, 8, 14, 28, 2, 30, 5, 32. What is the mean and standard deviation for his errors? Would the median or mode be a more appropriate measure of central tendency? What do you imagine the scores might be for any other second grade child? What particular errors would you expect either child to make?
9. Figure the *range* of age in the Detroit class in problem 2. What are the variance and standard deviation for age?
10. In problem 3, what are the mean and standard deviation for weight? What are the mean, variance, and standard deviation for age?
11. Here are the scores of the first exam you gave in an introductory linguistics class. What are the mean, mode, and median? What is the standard deviation? 88, 87, 66, 54, 97, 34, 48, 56, 99, 87, 73, 86, 74, 69, 88, 87, 86, 87.

Suggested further reading for this chapter: Shavelson.