4 Variability

The average temperatures of San Francisco and Kansas City, Missouri, differ by only one degree Fahrenheit. The average temperature in San Francisco is 56 degrees and the average in Kansas City is 55 degrees. Does this mean that the two cities have the same weather? Not at all. In July, Kansas City tends to be 20 degrees warmer than San Francisco; in January, however, San Francisco tends to be 20 degrees warmer than Kansas City. Just because the averages are the same does not mean that the weather is the same. The weather in Kansas City is more variable than in San Francisco. The variability in the numbers is just as important as the average of the numbers.

A sample of numbers has a characteristic shape. Recall that the shape concerns the frequency of certain numbers relative to other numbers. Besides the shape, the central tendency of the numbers—or the typical value in the sample—can be determined. The mode, median, and mean can be used as measures of central tendency. A distribution of scores can also be characterized by the variability of the scores. Are the numbers bunched tightly together or do they vary over a wide range? For instance, the scores

$$3, 4, 5, 5, 5, 5, 5, 6, 6, 7$$

are bunched together near five, whereas the scores

$$1, 2, 3, 4, 4, 5, 6, 8, 11, 13$$

vary over a wider range. *Variability* refers to how much the scores differ from one another.

In everyday life, the term *consistency* is used to refer to variability. The weather in Honolulu is consistently around 75 degrees. The Boston Celtics basketball team is a consistent winner. Our favorite sports team is inconsistent. One person's weight may be very stable, whereas another's changes quite a bit from week to week. The concept of consistency implies that the person or unit has low variability, whereas inconsistency refers to high variability.

Variability is also the fundamental starting point in scientific explanation.

Because some people have higher GPAs, weigh more, live longer, and are more violent than others, there must be some reason why these people differ. Scientists observe variation and wonder what brings it about. To the extent that people vary, there must be forces that make them differ. Therefore, if it can be understood why people differ, it can be understood what causes human behavior. One goal of research is to explain what makes people vary, and one common way of stating a result from research is to give the percentage of the variability that is explained. For instance, Jencks and his colleagues (1972) claim that they are able to account for, or explain, about 30% of the variability in income earned by an individual in the United States.

The meaning of "explaining 30% of the variability" requires explanation. It does *not* mean that 3 out of 10 reasons that make persons vary are known. Nor does it mean that reasons are known why 30% of the people make money. Rather, it means that given what is known, existing variability could be reduced by 30%. If it were possible to equalize persons on education, motivation, intelligence, and other specified factors, the variability of the income distribution would be reduced by 30%.

There are three basic reasons for measuring the variability of a sample. They are as follows:

1. to determine how meaningful the measure of central tendency is,
2. to use it as a basis for determining whether a score is an outlier, and
3. to compare variability.

Consider each reason in turn.

The first reason for measuring variability is to judge how meaningful the measure of central tendency is. To the extent that the numbers vary widely, any measure of central tendency is somewhat less informative. Knowing that the average income of some group of persons is $13,432 is not very informative when some persons are earning nothing and others millions. A doctor would want to know a person's average blood pressure only if that person's blood pressure was not subject to huge swings. To the extent that the numbers are bunched together, the measure of central tendency is more descriptive of the numbers. Thus the variability tells how well the mean or any other measure of central tendency represents all the numbers.

Consider four persons who have dinner at a restaurant. The bill including tip comes to $60. One person suggests splitting the bill up evenly with everyone putting in $15. He or she has computed a measure of central tendency—the mean or arithmetic average. If one person's dinner cost $21 and the other three averaged $13, the mean would not be very representative of the cost of each person's meal. If two persons' meals were $16 and the other two $14, however, then the mean would be a reasonable approximation of the cost of each person's meal.

A measure of variability is also useful in assessing how deviant, or unusual, a given number is. To identify a genius or anyone who is exceptional some type of yardstick for assessment is needed. If it rained 15 inches more than the average this year, it is not known whether that is exceptional or typical. If it is known that the rainfall varies quite a bit from year to year, then being 15 inches above the average is not at all exceptional. But if the amount of rainfall hardly ever changes from year to year, then having 15 more inches is quite unusual.

A measure of variability is also useful for comparison purposes. Some faces are easier to read than others. That is, some persons clearly express their emotions; others are poker-faced and are not expressive. People vary in expressiveness—that is, their ability to send nonverbal messages. It also seems true that some people are better at reading or receiving other people's emotional reactions than others. Some people are sensitive and others are not. Research has shown that there is much more variability in expressiveness than in sensitivity. There is more variability among the senders of nonverbal messages than there is among the receivers (Kenny & La Voie, 1984).

Throughout this chapter, the computations will be illustrated with the same sample of numbers. During the fall of 1964, psychologists Robert Rosenthal and Lenore Jacobson entered a South San Francisco grade school and told teachers that a number of students would "bloom" during that year. The teachers were told that the students were expected to experience an unusual forward spurt in academic and intellectual performance during the year. Actually the students were no more likely to excel than their classmates. Rosenthal and Jacobson were interested in studying whether merely suggesting to teachers that certain students would improve would create an actual change. As part of this research, Rosenthal and Jacobson measured the children's intelligence, or IQ. Table 4.1 lists the scores of 22 first-grade children. An IQ score is set so that a score of 100 is normal.

TABLE 4.1 **Intelligence Tests Scores for 22 First-Grade Children**

94	95
102	98
117	92
91	86
90	92
106	120
112	97
101	116
122	130
111	117
100	108

Data were taken from Rosenthal and Jacobson (1968).

Measures of Variability

The three common measures of variability are the *range,* the *interquartile range,* and the *standard deviation.* The range looks at only the variability between the largest and smallest numbers. The interquartile range looks at the variability in the middle of the sample. The standard deviation examines the variability between all possible pairs of numbers.

The Range

The range is the difference between the largest and smallest numbers in the sample. Thus, for the sample

$$3, 5, 6, 9, 11$$

the range is 11 minus 3, which is 8. For the sample

$$101, 190, 236, 436$$

the range is 436 minus 101, or 335. *The range is the largest minus the smallest number of the sample.* For the numbers in Table 4.1 the range is $130 - 86 = 44$. The range concentrates on the extremes and ignores all the other numbers in the distribution. Because the range examines only the two most extreme numbers in the sample, it is influenced by outliers.

Although the range is a natural and simple measure of variability, it has an important limitation. The size of the range depends, in part, on how many scores there are in the sample. In general, the greater the sample size the greater the range. Consider the example of a flat distribution of numbers from 1 to 10. (Recall from Chapter 2 that a flat distribution is one in which each number is just as likely as any other number.) For the smallest possible sample size for which the range can be computed, 2, the range is, on the average, about 3.3. That is, if every combination of two numbers from 1 to 10 is taken, the range averages 3.3. As the sample size increases, the range increases. For the sample size of 100 the range is nearly 9.0, more than twice the expected range for a sample size of 2.

The reason for this situation is that extreme scores (i.e., very large or very small scores) are much more likely to appear when the sample size is large. Because extreme numbers are associated with large sample size, larger range tends to be as well. This is not to say that the range is a useless measure of variability; however, it would be misleading to compare the ranges of two samples with very different sizes.

One advantage that the range has is that it is a relatively easy measure to compute and it is simple to understand. For instance, when you receive a grade in a course, you may want to know the range of scores. The range is a natural measure of variability, but the following two measures of variability are more commonly used.

The Interquartile Range

The interquartile range is another measure of variability, and it is almost as easy to compute as the range. The range measures the variability from the largest to the smallest number. The *interquartile range* measures the variability in the middle half of the distribution and is, therefore, less influenced by outliers. Basically, to compute the interquartile range, the sample is subdivided into two groups: those above the median and those below the median. The "median" is then computed for each of these groups. The median for the group above the overall median will be called the *upper median;* for the group below the overall median, it will be called the *lower median*. The interquartile range is the difference between the median of the group above the overall median and the median of the group below the overall median. The interquartile range is then the difference between the upper and the lower medians. So, for the following distribution

1, 2, 3, 3, 3, 4, 6, 9, 11, 20

the overall median is 3.5. The values below the median are

1, 2, 3, 3, 3

and their median is 3. The values above the overall median are

4, 6, 9, 11, 20

and their median is 9. The interquartile range is 9 minus 3 which equals 6. These computations are illustrated in Table 4.2.

TABLE 4.2 **Illustration of the Interquartile Range for Two Samples**

Sample 1				Sample 2
Upper median	20 11 9 6 4	Scores of those above the median	9 7 6 6	Upper median
Overall median	3.5		6	Overall median
Lower median	3 3 3 2 1	Scores of those below the median	5 4 4 1	Lower median
Sample 1 interquartile range: 9 − 3 = 6				Sample 2 interquartile range: 6.5 − 4 = 2.5

The rule for determining which scores are above and which are below the median is quite simple. If there is an even number of scores in the sample, the first half or $n/2$ scores are below the median, and the second half or $n/2$ scores are above. So, if there are 20 scores, $n/2$ is 20/2 or 10. The median is then computed of the 10 smaller scores and the median of the 10 larger scores. If there is an odd number of scores in the sample, the score at the median is excluded. Thus, for the sample

$$1, 4, 4, 5, 6, 6, 6, 7, 9$$

exclude the 6, the overall median, and use 1, 4, 4, and 5 as the scores below the median and 6, 6, 7, and 9 as the scores above the median. Again, these computations are illustrated in Table 4.2. Once it has been determined which scores are above and which are below, the medians are computed for these groups of scores by the simple procedure described in Chapter 3. Given n scores, when n is an odd number, the median is the $(n + 1)/2$th score. If n is even, then the median is the average of the $n/2$th and $(n/2 + 1)$th scores.

For the IQ data in Table 4.1 there are 22 scores. The median is then the average of the eleventh and twelfth largest scores of the sample. It is

$$\frac{101 + 102}{2} = 101.5$$

The scores below the median are

$$86, 90, 91, 92, 92, 94, 95, 97, 98, 100, 101$$

and so the lower median is 94. The scores above the median are

$$102, 106, 108, 111, 112, 116, 117, 117, 120, 122, 130$$

and so the upper median is 116. The interquartile range is then

$$116 - 94 = 22$$

Therefore the middle half of the numbers varies over 22 IQ points.

This measure is called the interquartile range because it is based on the separation of the sample into four quartiles. The sample is separated into four groups with an equal number of scores per group: four quartiles. The first quartile contains the 25% of the scores that are the smallest. The second quartile contains the next 25%. The third and fourth quartiles are similarly defined. The boundary between the first and second quartile is the lower median. The boundary between the second and third quartile is the overall median of the sample. Finally, the boundary between the third and fourth quartile is the upper median.

The Standard Deviation

Both the range and the interquartile range are sensible ways to measure variability in the sample. But they are limited because each of them only looks at part of the data. The range uses only the two most extreme numbers, and the interquartile range throws away all the extreme numbers. It would seem sensible to have a measure of variability that looks at all the numbers. The standard deviation is just such a measure.

To measure how different the numbers are from each other, one natural measure would be to subtract them from each other. So for the numbers 9, 10, 11 the difference between all possible pairs is computed as follows:

$$10 - 9 = 1$$
$$9 - 11 = -2$$
$$10 - 11 = -1$$

Because it is not important whether the difference is positive or negative, the differences are squared.

$$(10 - 9)^2 = 1$$
$$(9 - 11)^2 = 4$$
$$(10 - 11)^2 = 1$$

The mean of these three squared differences is $(1 + 4 + 1)/3 = 6/3 = 2$. This seems to be a very sensible measure of variability. It is the average of all possible squared differences, which will be called the *average squared difference*. Although the average squared difference is sensible, it presents a computational nightmare. If there are 50 numbers in the sample there are 1225 differences! It would take hours to compute even with a calculator. Fortunately, there is a computational shortcut. It is called the standard deviation.

The most common measure of the variability is the ***standard deviation***, or as it is usually symbolized, s. Instead of first presenting the formula for the standard deviation, a rationale for it and its close relative, the variance, or s^2, is discussed. As was detailed in the previous chapter, the mean (the sum of the numbers divided by the sample size) is a measure of central tendency. One reasonable measure of variability would be to simply compute how far each score is from the mean. A deviation of each score from the mean is then computed. If the numbers are tightly bunched together, these deviations from the mean would be small and so the measure of variability would be small. Alternatively, if the numbers were spread over a wide range, the deviations from the mean would be large and so would the variability. Thus, a basic building block of a measure of variability can be the deviation of scores from the mean. For the sample

$$1, 6, 10, 12, 16$$

with a mean of 9, the deviations from the mean are

$$1 - 9 = -8$$
$$6 - 9 = -3$$
$$10 - 9 = 1$$
$$12 - 9 = 3$$
$$16 - 9 = 7$$

These deviations from the mean are each squared and then summed. Thus, the numbers −8, −3, 1, 3, and 7 are squared, to obtain 64, 9, 1, 9, 49, the sum of which equals 132. Then to adjust for sample size, the sum is divided by sample size less one. (The reasons for using sample size less one will be explained later.) So for the example above, the measure of variability is 132 divided by 4 (sample size less one) and it equals 33. This measure of variability is called the variance. The *variance*, symbolized by s^2, is the sum of squared deviations about the mean divided by sample size less one. The *standard deviation*, symbolized by s, is the positive square root of the variance.

The measure s^2, or variance, is closely related to the average squared difference. The average squared difference is determined by taking the difference between pairs of scores, squaring each of them, and then summing these squared differences and dividing that sum by the number of pairs. The variance is one-half the average squared difference. The variance and, therefore, the standard deviation are very closely related to the average squared difference. The variance is, however, much simpler to compute than the average squared difference, and hence it is generally preferred.

Consider, for instance, the sample 5, 11, 15, 17. The mean is 12 and so the deviations from the mean are −7, −1, 3, 5. To compute the variance, the deviations are squared, summed, and that sum is divided by the sample size less one. The sum of squared deviations is then 49 + 1 + 9 + 25 = 84 and the variance is 84/3 = 28. To compute the average squared difference, the difference between all possible pairs is computed:

$$
\begin{aligned}
11 - 5 &= 6 \\
15 - 5 &= 10 \\
17 - 5 &= 12 \\
15 - 11 &= 4 \\
17 - 11 &= 6 \\
17 - 15 &= 2
\end{aligned}
$$

The sum of the squares of these quantities is 36 + 100 + 144 + 16 + 36 + 4 = 336. And 336 divided by 6 (the number of pairs) equals 56. Note that one-half of 56 is 28, the variance. So, one-half the average squared difference equals the variance, as it should.

Normally, it is advisable to use the standard deviation, not the variance, as a measure of variability. By computing the square root of the variance to obtain the standard deviation, the unit of measurement becomes interpretable. For the example in Table 4.1, the variable is intelligence or as it is usually called, IQ. When variance is computed, the deviations are squared. So for intelligence, the unit of measurement for the variance is IQ points squared. It is not too clear how to interpret intelligence squared. To return the unit of measurement to that of the original metric, the square root of the variance is computed.

The standard deviation is a measure of the "average" distance from the mean. It is not a straightforward arithmetic average because of the squaring, the summing, and the taking of the square root. The effect of the squaring is to make the standard deviation more heavily reflect scores that are farther from the mean. Consider the two following samples. Sample A contains the numbers

1, 3, 7, 9

and sample B contains the numbers

1, 3, 5, 11

Both samples have a mean of 5.0 and the sum of the absolute deviations from the mean are 12.0 for both samples. However, the standard deviation for sample A is 3.65 and the standard deviation for sample B is 4.32. Why are they different? Sample B has the relatively extreme score of 11, which is 6 units from the mean while for sample A the most deviant score is only 4 units from the mean. When the deviation of 6 is squared, it dominates the variance.

The variance and, therefore, the standard deviation are quite affected by the presence of outliers. Consider the following sample.

4, 5, 5, 6, 7, 7, 8

The standard deviation of this sample is 1.41. Consider now the same sample with the addition of a single outlier.

4, 5, 5, 6, 7, 7, 8, 44

The standard deviation has now exploded to 13.50. A single outlier can dramatically affect the size of the standard deviation.

Why Squared Deviations? For the variance and the standard deviation, mean deviations are squared. An alternative is to just sum the deviations. For the sample whose numbers are 1, 6, 10, 12, and 16, the deviations are –8, –3, 1, 3, and 7. Their sum is zero. This is no accident. The sum of deviations about the mean is always zero. This is due to the definition of the mean. Recall that the mean is the balance point of a distribution; a balance point requires that the sum of deviations must be zero.

Because the sum of the deviations cannot be used as a measure of variability, an alternative to squaring them might be to compute the *absolute value* of deviations. An absolute value of a number is that number with the sign always positive. Therefore the deviations for the example above are –8, –3, 1, 3, 7, and the sum of their absolute values is 22. There is, however, no simple relationship between the absolute values of mean deviations and the absolute value of the deviation between all pairs of numbers. So, the average absolute deviation from the mean is not related to the average absolute deviation between all possible pairs.

Why Sample Size Less One? There would not be much harm in dividing by sample size instead of sample size less one.[1] However, it is just a little better to divide by sample size less one. Why? Consider what the variance would be when the sample size is one. It is, of course, impossible to measure variability when there is only a single score in the sample. However, dividing by sample size (instead of sample size less one), the variance would always be zero when sample size is one. However, if the denominator of the variance is sample size less one, the variance is undefined when sample size is one. (Division by zero is not mathematically permissible.) So one reason for "less one" is to make the variance of the sample size of one to be undefined.

A related reason for dividing by sample size less one is because the mean is computed from the numbers that are used to compute the variance. If the mean were known without having to compute it from the data, it would be correct to divide by n and not $n - 1$. In Chapter 9 the question of dividing by $n - 1$ instead of n is discussed.

Computation of the Standard Deviation. The definitional formula for the standard deviation is to

1. take each score and subtract the mean,
2. square each of the deviations from the mean,
3. sum the squared deviations,
4. divide the sum of squared deviations by sample size less one, and
5. take the square root.

In terms of a formula, the standard deviation is

$$s = \sqrt{\frac{\sum(X - \bar{X})^2}{n - 1}}$$

The numerator is the sum of squared deviations from the mean ($\bar{X}$) and the denominator is sample size less one.

The computations of $\Sigma(X - \bar{X})^2$ can be simplified by using the formula

$$\sum X^2 - \frac{(\sum X)^2}{n}$$

The first term of the formula, ΣX^2, is simply the sum of the squares of each number in the sample. The second term, $(\Sigma X)^2/n$, is the result of summing all the scores, squaring the result, and then dividing by n, the sample size. You might want to review the difference between ΣX^2 and $(\Sigma X)^2$, described in Chapter 1. The resulting complete computational formula for the standard deviation is

[1]Some texts recommend dividing by n instead of $n - 1$. However, most statistical formulas presume that the denominator is $n - 1$.

$$s = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n - 1}}$$

Table 4.3 presents the computations required for the computational formula for the standard deviation using the IQ example in Table 4.1. The first column consists of the raw data. The second column gives the squares of each score. The totals or sums of the numbers of the two columns are then written beneath each column. The total for the first column in Table 4.3 is symbolized by ΣX and the second column by ΣX^2. Taking the quantities in Table 4.3 and entering them into the computational formula yields

$$\sqrt{\frac{\sum X^2 - (\sum X)^2/n}{n - 1}} = \sqrt{\frac{242967 - (2297)^2/22}{21}} = 12.23$$

TABLE 4.3 **Computations Necessary for the Standard Deviation**

	X (score)	X^2 (score squared)
	94	8836
	102	10404
	117	13689
	91	8281
	90	8100
	106	11236
	112	12544
	101	10201
	122	14884
	111	12321
	100	10000
	95	9025
	98	9604
	92	8464
	86	7396
	92	8464
	120	14400
	97	9409
	116	13456
	130	16900
	117	13689
	108	11664
Total	2297	242967

Detection of Outliers

As discussed in Chapter 2, an outlier is a very large or small score in the sample. Now that quantitative measures of variability and central tendency have been presented, a quantitative rule for determining an outlier can be given.

Two different definitions of an outlier are presented. The first involves the interquartile range and the median. An outlier is any number that is more than two times the interquartile range away from the median. So one computes

$$\text{median} \pm 2 \times \text{interquartile range}$$

Scores larger than the median plus twice the interquartile range and smaller than the median minus twice the interquartile range are deemed outliers. For the IQ example, the median is 101.5 and the interquartile range is 22. So for a score to be an outlier it must exceed 101.5 + 44, or 145.5, or be less than 101.5 – 44, or 57.5. Certainly IQs of greater than 145.5 and less than 57.5 should be carefully examined. Using this definition of outliers, none of the scores in Table 4.1 can be considered outliers.

An outlier can be alternatively defined by using the standard deviation and the mean. The following quantity is computed:

$$\text{mean} \pm 2\tfrac{1}{2} \times \text{standard deviation}$$

The reason that it is two and one half times the standard deviation and only two times the interquartile range is that the standard deviation tends to be smaller than the interquartile range. For the IQ example the mean is 104.41 and the standard deviation is 12.23. So for a score to be an outlier it must exceed 104.41 + 30.57, or 134.98, or be less than 104.41 – 30.57, or 73.84. Using this definition of outliers, none of the scores in Table 4.1 can be considered outliers.

If one has a choice concerning the two measures of outliers, the interquartile range and the median measure is preferred to the standard deviation and the mean measure. The reason is that the mean and standard deviation are themselves influenced by the presence of an outlier. Say the sample in Table 2.1 had the three largest IQs changed to 175. This change would raise the mean and double the standard deviation, and the value of 175 would not be recognized as an outlier. However, the interquartile range and the median are less affected by the presence of outliers, and the 175 IQs would be deemed as outliers.

Computational Errors and Checks

All measures of variability must be nonnegative. Because the least amount of variability is none, the lowest value of a measure of variability is zero. When variability is zero all the values of the sample are the same.

Because the range and the interquartile range basically involve only a single subtraction, it is unlikely that they would be mistakenly computed as negative. However, the variance involves much computation, and so errors can occur. Whenever the numerator of the variance is negative, one knows with certainty that there has been an error of computation. The usual reason for such an error is that one has incorrectly summed the squares of the scores. Also, it is possible that the sum of all the scores has been incorrectly computed.

A good way to locate a computational error is to study the range, the interquartile range, and the standard deviation. First, determine whether the estimated value of variability is possible. As was just stated, zero is the lowest possible value for variability. If the numbers themselves have a lower and an upper bound, then the measure of variability can be no larger than the upper bound minus the lower bound. (This is true of the standard deviation but not the variance.) So if a group of persons rate a movie on a scale from one to ten, then the measures of variability (except the variance) computed from these numbers must be nine or less. The standard deviation must be less than or equal to the range. If it is not, then an error has been made in computing the variability or there has been an error in the recording of the numbers.

One obvious, but commonly made error is to report the variance as the standard deviation. Recall that the standard deviation is the square root of the variance. For the standard deviation, always make certain that the square root of the variance has been calculated.

Recall from the previous chapter that the mode, the median, and the mean can be equal to the same value. However, one should not expect that the range, the interquartile range, and the standard deviation to be equal or nearly equal. Although all three measure the variability in the sample, they do so in different ways. For instance, the range must always be at least as large as the interquartile range. To see this, note that the largest value of a sample must be at least as large as the median of the upper half of the distribution. Similarly, the lowest value must be at least as small as the median of the lower half of the distribution, and so the range is at least as large as the interquartile range. Generally the standard deviation is smaller than the interquartile range, as it is for the IQ example used in this chapter. For unimodal symmetric distributions a good rule of thumb is that three-quarters of the interquartile range is about equal to the size of the standard deviation. For the IQ example the interquartile range is 22 and so three-quarters of 22 is 16.5, which is still larger than but nearer to the standard deviation of 12.23.

Summary

The numbers in a sample vary. Some are larger than others and some are smaller. Measures of variability quantify how diverse the numbers are.

The simplest measure of variability is the *range*. The range is the largest number in the sample minus the smallest one. The range is not a very good measure of variability because it uses only two extreme scores and because it depends on the sample size. As the sample size increases, the range tends to increase.

The *interquartile range* is the difference between the median of the upper half of the numbers and the median of the lower half. The interquartile range measures how different the scores in the middle half of the distribution are.

The *standard deviation* is the most commonly used measure of variability. The standard deviation is the square root of the variance. The variance is one-half the average squared difference between all possible pairs of numbers. The variance is defined by the following steps.

1. Compute the deviation of each score from the mean.
2. Square each of these deviations.
3. Sum the squared deviations.
4. Divide the sum by sample size less one.

Computations can be simplified by these steps.

1. Square each score.
2. Sum the squares.
3. Subtract the square of the sum of all the scores divided by the sample size.
4. Divide by sample size less one.

Again, the square root of the variance is the standard deviation. The computational formula for the standard deviation is

$$s = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n - 1}}$$

The variability of the numbers is a fundamental part of a distribution. The range, the interquartile range, and the standard deviation provide three ways of quantifying the variability of the numbers.

Problems

1. Compute the range, interquartile range, and standard deviation for the following samples.

   a. 1, 3, 4, 6, 6, 9, 10, 11
   b. 2, 4, 6, 8, 9, 11, 13, 14, 15
   c. 3, 4, 4, 6, 7, 9, 10, 13, 15, 16

2. a. If the standard deviation of a sample is 6.5, what is the variance?
   b. What is the average squared difference?

following sample.

6, 8, 10, 11, 12, 14, 16, 22, 150

b. Compute the standard deviation, range, and interquartile range of the same sample, without the number 150.

6, 8, 10, 11, 12, 14, 16, 22

c. Which measure of variability changed the least?

4. Two automobile manufacturers, Marvel Motors and Amazing Autos, have each produced six different models. Below is the number of years each model of car typically runs before requiring major repairs.

Marvel Motors: 3 5 10 7 2 9

Amazing Auto: 6 7 7 6 5 5

The mean for both manufacturers is six years.

a. Compute the range, interquartile range, and standard deviation for each manufacturer.
b. Which manufacturer produces cars of more consistent quality?

5. The following table presents the number of milligrams of phosphorus, calcium, and vitamin C contained in one cup of various fruit juices.

Juice	*Phosphorus*	*Calcium*	*Vitamin C*
Apple	23.0	15.0	2.5
Apricot	30.0	23.0	7.5
Cranberry	7.5	13.0	40.0
Grapefruit	35.0	20.0	77.5
Orange	45.0	25.0	100.0
Pineapple	23.0	38.0	23.0

Score	*Control* (*n* = 42)	*Experimental* (*n* = 70)
65–69	1	
70–74	1	
75–79	5	
80–84	7	3
85–89	9	5
90–94	14	15
95–99	5	47

Use the class midpoint for the score for each group.

a. Compute the mean, variance, and standard deviation for each group.
b. Which group's scores are more variable?

8. Suppose two classes, each consisting of eight students, take an exam. Suppose further that the scores for class A are more variable than the scores for class B.

   a. Create a sample of scores for class A and a sample for class B, each with a mean of 70.
   b. Compute the standard deviation for each class.
   c. In which class is the mean more representative of the students' performance?

9. For the following sample determine which observations are outliers. Use the median plus or minus twice the interquartile range as the definition of an outlier.

   15, 19, 25, 22, 8, 19, 18, 15, 37

10. Below are the gender bias data discussed in the previous two chapters.

.29	.01	−.40	.00
.56	−.31	.00	.35
.00	.14	.02	.11
−.31	−.22	−.03	−.23
.56	.01	.00	−.56
−1.03	.00	−.23	.23
.00	−.10	.05	.00
.60	−.23	.24	−.36

   Compute and interpret the range, interquartile range, and the standard deviation for the data.

11. Compute the standard deviation for the following sample.

   6, 8, 12, 9, 13, 12, 10, 9

12. For the data in problem 11, add five to each score and compute the standard deviation. How does the standard deviation change?

13. The following data represent annual rainfall of two cities (A and B) over the past eight years.

    A: 24, 18, 19, 21, 28, 17, 32, 24

    B: 6, 19, 14, 7, 21, 9, 17, 4

    In which city is rainfall more variable?

14. You want to determine how variable the high temperature is in two exotic isles. You note that in San Luca the temperatures have been 87 and 93 for two days. But for Dolores you have more information. The temperatures are 84, 86, 87, 88, 93, 94, and 94 for seven days. Use the range and standard deviation to determine variability of weather of the two islands. Decide which island has more variable weather.

15. Using the median and the interquantile range, state which scores, if any, are outliers:

    14, 16, 13, 12, 16, 17, 14, 13, 17, 14, 21, 12, 14, 13, 6