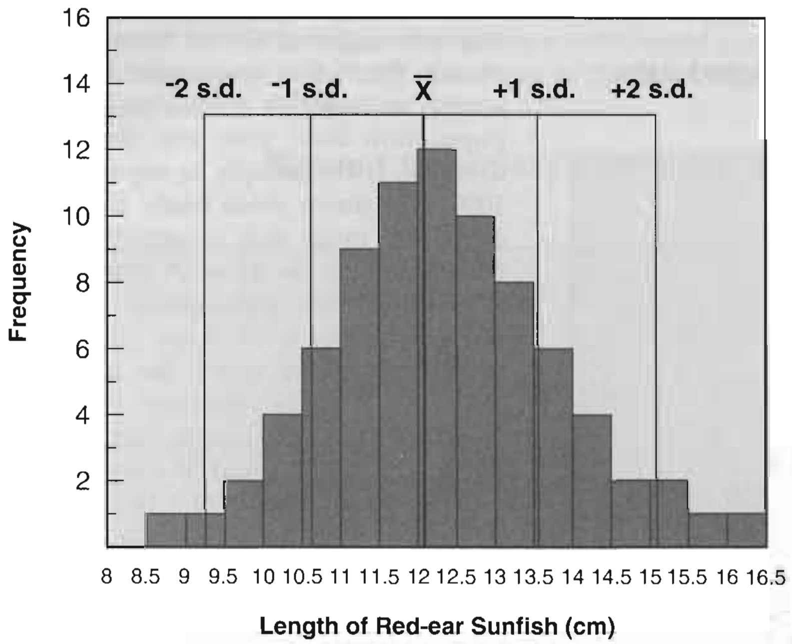
Figure 1.10 Standard deviations.



To calculate the size of a standard deviation, it is actually easier first to calculate a related statistic called the **variance**. The variance is the square of the standard deviation, so we use s^2 = **the sample variance**, and σ^2 = **the population variance.** Calculation of the variance is based on the difference between each observation and the mean. If all these differences are squared, and we calculate an average of the squared values, we have the variance. (See Appendix 1.) It is good to remember that the units on variance are the original measurement units squared. If we measure length in cm, then the sample variance is reported in cm^2. In calculating standard deviations, we take the square root of the variance, which returns us to our original measurement units, which is length in cm.

$$s = \sqrt{s^2} \quad \textbf{and} \quad \sigma = \sqrt{\sigma^2}$$

s = sample s.d. — **σ = population s.d.**

s^2 = sample variance — **σ^2 = population variance**

Check your progress:

If the standard deviation of a population is 9.5, what is the population variance?

Answer: 90.25

Confidence Intervals

There is one more statistic that you will find useful when characterizing the typical sunfish in your population with a mean. Incorporating sample size and variation, you can develop a measure of reliability of the mean called the **standard error (S.E.)**.

Assume there are 25 people in your ecology class. Each of you goes to the same pond sometime this week, nets a sample of sunfish in the same way, and measures 100 randomly selected individuals. Releasing the fish unharmed, you return to the lab and calculate a mean and standard deviation from your data.

Everyone else does the same. Would your 25 sample means be identical? No, but the variation in means would be considerably smaller than the total variation among the fish in the pond. Repeating a sampling program 25 times is usually impractical. Fortunately, statistics gives us a way to measure reliability when we have only one mean developed from one sample. The variation among all possible sample means can be predicted from the sample size and the variation in the pond's sunfish with the following formula:

$$\text{S.E.} = s/\sqrt{n}$$

S.E. = standard error of the mean
s = standard deviation of sample
n = sample size

Looking at the formula, you can see the relationship between error in our estimate, the variability of sunfish, and the sample size. The smaller the S.E. is, the more trustworthy your calculated mean. Note that the sample standard deviation is in the numerator of the calculation, so the more variable the size of the fish, the less accurate the estimate you made from a random sample. Sample size, on the other hand, is in the denominator. This implies that a large sample makes your mean more reliable. The formula shows that the more variable the population, the larger our sample must be to hold S.E. to an acceptably small margin of error.

Since standard errors tend to be normally distributed, it is a safe assumption that 95% of the variation in all possible means will fall within 1.96 S.E. of the actual mean. This fact can be used to calculate a 95% confidence interval as follows:

95% Confidence interval = $\bar{x} \pm 1.96$ S.E.

$\bar{x}$ = sample mean

S.E. = standard error

Check your progress:

Calculate the 95% confidence interval for a mean of 14.3, derived from a sample of 25, where the standard deviation is 4.2. What are the upper and lower limits?

Answer: 12.65 to 15.95

To go back to our number line, the confidence intervals can be represented by brackets around the sample mean. A 95% confidence interval implies that the actual population mean (μ) will fall within the brackets you have placed around your estimate ($\bar{x}$) 95% of the time under these experimental conditions (Figure 1.11).