V. The sample variance (s^2) and standard deviation (s)

A. **Sample variance** $$S^2 = \frac{\sum (x-\bar{x})^2}{n-1}$$ **Alternate formula** $$S^2 = \frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}$$

B. Note that N has been replaced by n-1 in the denominator.

C. In chapter 11, the sample variance will be used to predict the population variance. If one is not subtracted from n when calculating the sample standard deviation, it will be bias (not representative of σ).

D. **Sample standard deviation** (Assume data on page 16 is sample data.)

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

or

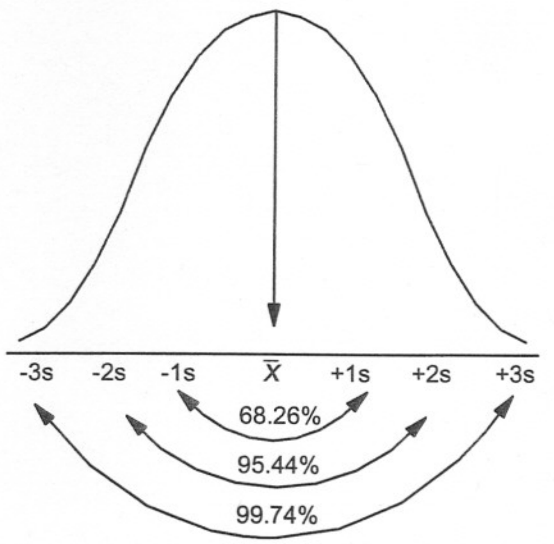
$$S = \sqrt{\frac{\sum x^2 - \frac{\left(\sum x\right)^2}{n}}{n-1}} = \sqrt{\frac{213 - \frac{(35)^2}{7}}{7-1}} = \sqrt{\frac{213-175}{6}} = 2.5$$

VI. Using the standard deviation as a measure of variability

A. The **empirical rule** is used for normal, bell-shaped data.

1. For symmetrical or bell-shaped data, 68.26% of the item will be within one standard deviation of the mean, 95.44% will be within two standard deviations of the mean, and 99.74% will be within three standard deviations of the mean. If $\mu = 500$ and $\sigma = 100$, then 95.44% of the population will be between 300 and 700.

$$500 \pm 2(100)$$
$$500 \pm 200 \rightarrow 300 \leftrightarrow 700$$



2. Students would like a small standard deviation around a test mean of 95 so everyone receives a grade of A.

B. **Chebyshev's rule** is used for nonsymmetrical distributions.

1. Russian mathematician P. Chebyshev developed a method to estimate the minimum proportion of items that are within a designated number of standard deviations from the mean for nonsymmetrical distributions with means greater than 1. As with the empirical rule, the estimate works for both samples and populations.

$$1 - \frac{1}{K^2}$$

2. The proportion of items within K standard deviations of the mean is at least 1 minus 1 over K squared provided K is a constant greater than 1.
3. The proportion of the data falling within 2 standard deviations of a mean is calculated as follows:

$$1 - \frac{1}{K^2} = 1 - \frac{1}{(2)^2} = 1 - \frac{1}{4} = \frac{3}{4}$$

VII. Coefficient of variation (CV)

A. Comparing the variability of data sets of differing magnitudes is accomplished using the coefficient of variation.

B. Department A with \$40 million in sales will have a much larger standard deviation than Department B which has only \$3 million in sales. Suppose Department A's σ was \$4 million and Department B's σ was \$400,000.

C. The **coefficient of variation,** which expresses the standard deviation as a percent of the mean, reveals which department has the largest relative sales variability. CV = $\frac{s}{\bar{x}}(100)$ for sample data.

For Department A

$$C.V. = \frac{\sigma}{\mu}(100) = \frac{\$4,000,000}{\$40,000,000}(100) = 10\%$$

For Department B

$$C.V. = \frac{\sigma}{\mu}(100) = \frac{\$400,000}{\$3,000,000}(100) = 13.3\%$$

Note: Department A had less sales dollar variability even though it had a larger standard deviation.