Variance *between* Samples Find the variance *between* samples by evaluating $ns_{\bar{x}}^2$, where $s_{\bar{x}}^2$ is the variance of the sample means and n is the size of each of the samples. That is, consider the sample means to be an ordinary set of values and calculate the variance. (From the central limit theorem, $\sigma_{\bar{x}} = \sigma/\sqrt{n}$ can be solved for σ to get $\sigma = \sqrt{n} \cdot \sigma_{\bar{x}}$, so that we can estimate σ^2 with $ns_{\bar{x}}^2$.) For example, the sample means for Data Set A in Table 12-2 are 5.5, 6.0, and 6.0; these three values have a variance of $s_{\bar{x}}^2 = 0.0833$, so that

$$\text{variance between samples} = ns_{\bar{x}}^2 = 4(0.0833) = 0.3332$$

Variance *within* Samples Estimate the variance *within* samples by calculating s_p^2, which is the pooled variance obtained by finding the mean of the sample variances. The sample variances in Table 12-2 are 3.0, 2.0, and 2.0, so that

$$\text{variance within samples} = s_p^2 = \frac{3.0 + 2.0 + 2.0}{3} = 2.3333$$

Calculate the Test Statistic Evaluate the F test statistic as follows:

$$F = \frac{\text{variance between samples}}{\text{variance within samples}} = \frac{ns_{\bar{x}}^2}{s_p^2} = \frac{0.3332}{2.3333} = 0.1428$$

The critical value of F is found by assuming a right-tailed test, because large values of F correspond to significant differences among means. With k samples each having n values, the numbers of degrees of freedom are as follows.

Degrees of Freedom: k = number of samples and n = sample size

$$\text{numerator degrees of freedom} = k - 1$$

$$\text{denominator degrees of freedom} = k(n - 1)$$

For Data Set A in Table 12-2, $k = 3$ and $n = 4$, so the degrees of freedom are 2 for the numerator and $3(4 - 1) = 9$ for the denominator. With $\alpha = 0.05$, 2 degrees of freedom for the numerator, and 9 degrees of freedom for the denominator, the critical F value from Table A-5 is 4.2565. If we were to use the critical value method of hypothesis testing with Data Set A in Table 12-2, we would see that this right-tailed test has a test statistic of $F = 0.1428$ and a critical value of $F = 4.2565$, so the test statistic is not in the critical region. We therefore fail to reject the null hypothesis of equal means.

To really see how the method of analysis of variance works, consider both collections of sample data in Table 12-2. Note that the three samples in Data Set A are identical to the three samples in Data Set B, except that each value in Sample 1 of Data Set B is 10 more than the corresponding value in Data Set A. The three sample means in A are very close, but there are substantial differences in B. However, the three sample variances in A are identical to those in B.

Adding 10 to each data value in the first sample of Table 12-2 has a dramatic effect on the test statistic, with F changing from 0.1428 to 51.5721. Adding 10 to each data value in the first sample also has a dramatic effect on the P-value, which changes from 0.8688 (not significant) to 0.0000118 (significant). Note that the variance between samples in A is 0.3332, but for B it is 120.3332 (indicating that the sample means in B are farther apart). Note also that the variance within samples is 2.3333 in both parts, because the variance *within* a sample isn't affected when we add a constant to every sample value. *The change in the F test statistic and the P-value is attributable only to the change in* $\bar{x}_1$. This illustrates the key point underlying the method of one-way analysis of variance: **The F test statistic is very sensitive to sample *means,* even**