Null and Alternative Hypotheses

H_0: The frequency counts agree with the claimed distribution.
H_1: The frequency counts do not agree with the claimed distribution.

Test Statistic for Goodness-of-Fit Tests

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

***P*-values:** P-values are typically provided by technology, or a range of P-values can be found from Table A-4.

Critical values:

1. Critical values are found in Table A-4 by using $k - 1$ degrees of freedom, where k is the number of categories.
2. Goodness-of-fit hypothesis tests are always *right-tailed*.

Finding Expected Frequencies

Conducting a goodness-of-fit test requires that we identify the observed frequencies, then determine the frequencies expected with the claimed distribution. Table 11-2 (on the next page) includes observed frequencies with a sum of 100, so $n = 100$. If we assume that the 100 digits were obtained from a population in which all digits are equally likely, then we *expect* that each digit should occur in $1/10$ of the 100 trials, so each of the 10 expected frequencies is given by $E = 10$. In general, if we are assuming that all of the expected frequencies are equal, each expected frequency is $E = n/k$, where n is the total number of observations and k is the number of categories. In other cases in which the expected frequencies are not all equal, we can often find the expected frequency for each category by multiplying the sum of all observed frequencies and the probability p for the category, so $E = np$. We summarize these two procedures here.

- **Expected frequencies are equal: $E = n/k$.**
- **Expected frequencies are not all equal: $E = np$ for each individual category.**

As good as these two preceding formulas for E might be, it is better to use an informal approach. Just ask, "How can the observed frequencies be split up among the different categories so that there is perfect agreement with the claimed distribution?" Also, note that the *observed* frequencies are all whole numbers because they represent actual counts, but the *expected* frequencies need not be whole numbers. If Table 11-2 had 75 observations instead of 100, each expected frequency would be 7.5.

We know that sample frequencies typically differ somewhat from the values we theoretically expect, so we now present the key question: Are the differences between the actual *observed* frequencies O and the theoretically *expected* frequencies E statistically significant? We need a measure of the discrepancy between the O and E values, so we use the test statistic given in the preceding box. (Later we will explain how this test statistic was developed, but you can see that it has differences of $O - E$ as a key component.)

The χ^2 test statistic is based on differences between the observed and expected values. If the observed and expected values are *close*, the χ^2 test statistic will be small and the P-value will be large. If the observed and expected frequencies are *far apart*, the χ^2 test statistic will be large and the P-value will be small. Figure 11-1 on the next page summarizes this relationship. The hypothesis tests of this section are always