13-7 Runs Test for Randomness

Key Concept This section describes the *runs test for randomness,* which can be used to determine whether a sequence of sample data has a random order. This test requires a criterion for categorizing each data value into one of two separate categories, and it analyzes *runs* of those categories to determine whether the runs appear to result from a random process, or whether the runs suggest that the order of the data is not random.

DEFINITIONS

After characterizing each data value as one of two separate categories, a **run** is a sequence of data having the same characteristic; the sequence is preceded and followed by data with a different characteristic or by no data at all.

The **runs test** uses the number of runs in a sequence of sample data to test for randomness in the order of the data.

Fundamental Principle of the Runs Test

The fundamental principle of the runs test can be briefly stated as follows:

Reject randomness if the number of runs is very low or very high.

- Example: The sequence of genders FFFFFMMMMM is not random because it has only 2 runs, so the number of runs is very *low*.
- Example: The sequence of genders FMFMFMFMFM is not random because there are 10 runs, which is very *high*.

CAUTION The runs test for randomness is based on the *order* in which the data occur; it is *not* based on the *frequency* of the data. For example, a sequence of 3 men and 20 women might appear to be random, but the issue of whether 3 men and 20 women constitute a *biased* sample (with disproportionately more women) is *not* addressed by the runs test.

The exact criteria for determining whether a number of runs is very high or low are found in the accompanying box, which summarizes the key elements of the runs test for randomness. The procedure for the runs test for randomness is also summarized in Figure 13-6 on page 677.

Sports Hot Streaks

It is a common belief that athletes often have "hot streaks"—that is, brief periods of extraordinary success. Stanford University psychologist Amos Tversky and other researchers used statistics to analyze the thousands of shots taken by the Philadelphia 76ers for one full season and half of another. They found that the number of "hot streaks" was no different than you would expect from random trials with the outcome of each trial independent of any preceding results. That is, the probability of a hit doesn't depend on the preceding hit or miss.

Runs Test for Randomness

Objective

Apply the runs test for randomness to a *sequence* of sample data to test for randomness in the *order* of the data. Use the following null and alternative hypotheses:

H_0: The data are in a random sequence.
H_1: The data are in a sequence that is not random.

Notation

n_1 = number of elements in the sequence that have one particular characteristic. (The characteristic chosen for n_1 is arbitrary.)
n_2 = number of elements in the sequence that have the other characteristic
G = number of runs

continued