important parameters of mean, standard deviation, and variance for a probability distribution. Most importantly, we describe how to determine whether outcomes are *unlikely* to occur by chance. We begin with the related concepts of *random variable* and *probability distribution.*

Part 1: Basic Concepts of a Probability Distribution

> **DEFINITIONS**
>
> A **random variable** is a variable (typically represented by x) that has a single numerical value, determined by chance, for each outcome of a procedure.
>
> A **probability distribution** is a description that gives the probability for each value of the random variable. It is often expressed in the format of a table, formula, or graph.

In Section 1-3 we made a distinction between discrete and continuous data. Random variables may also be discrete or continuous, and the following two definitions are consistent with those given in Section 1-3.

> **DEFINITIONS**
>
> A **discrete random variable** has a collection of values that is finite or countable. (If there are infinitely many values, the number of values is countable if it is possible to count them individually, such as the number of tosses of a coin before getting tails.)
>
> A **continuous random variable** has infinitely many values, and the collection of values is not countable. (That is, it is impossible to count the individual items because at least some of them are on a continuous scale.)

This chapter deals exclusively with discrete random variables, but the following chapters will deal with continuous random variables.

Every probability distribution must satisfy each of the following three requirements.

Probability Distribution: Requirements

1. There is a numerical random variable x and its values are associated with corresponding probabilities.

2. $\sum P(x) = 1$ where x assumes all possible values. (The sum of all probabilities must be 1, but sums such as 0.999 or 1.001 are acceptable because they result from rounding errors.)

3. $0 \leq P(x) \leq 1$ for every individual value of the random variable x. (That is, each probability value must be between 0 and 1 inclusive.)

The second requirement comes from the simple fact that the random variable x represents all possible events in the entire sample space, so we are certain (with probability 1) that one of the events will occur. The third requirement comes from the basic principle that any probability value must be 0 or 1 or a value between 0 and 1.