Example 7.6

A lot contains 12 items of which four are defective. Three items are drawn at random from the lot one after the other. Find the probability p that all three are nondefective.

The probability that the first item is nondefective is $\frac{8}{12}$ since eight of 12 items are nondefective. If the first item is nondefective, then the probability that the next item is nondefective is $\frac{7}{11}$ since only seven of the remaining 11 items are nondefective. If the first two items are nondefective, then the probability that the last item is nondefective is $\frac{6}{10}$ since only 6 of the remaining 10 items are now nondefective. Thus by the multiplication theorem,

$$p = \frac{8}{12} \cdot \frac{7}{11} \cdot \frac{6}{10} = \frac{14}{55} \approx 0.25$$

7.5 INDEPENDENT EVENTS

Events A and B in a probability space S are said to be *independent* if the occurrence of one of them does not influence the occurrence of the other. More specifically, B is independent of A if $P(B)$ is the same as $P(B|A)$. Now substituting $P(B)$ for $P(B|A)$ in the Multiplication Theorem $P(A \cap B) = P(A)\ P(B|A)$ yields

$$P(A \cap B) = P(A)\ P(B).$$

We formally use the above equation as our definition of independence.

Definition: Events A and B are *independent* if $P(A \cap B) = P(A)\ P(B)$; otherwise they are *dependent*.

We emphasize that independence is a symmetric relation. In particular, the equation

$$P(A \cap B) = P(A)\ P(B) \quad \text{implies both} \quad P(B|A) = P(B) \quad \text{and} \quad P(A|B) = P(A)$$

Example 7.7

A fair coin is tossed three times yielding the equiprobable space

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

Consider the events:

$$A = \{\text{first toss is heads}\} = \{HHH, HHT, HTH, HTT\}$$
$$B = \{\text{second toss is heads}\} = \{HHH, HHT, THH, THT\}$$
$$C = \{\text{exactly two heads in a row}\} = \{HHT, THH\}$$

Clearly A and B are independent events; this fact is verified below. On the other hand, the relationship between A and C or B and C is not obvious. We claim that A and C are independent, but that B and C are dependent. We have

$$P(A) = \frac{4}{8} = \frac{1}{2}, \qquad P(B) = \frac{4}{8} = \frac{1}{2}, \qquad P(C) = \frac{2}{8} = \frac{1}{4}$$

Also,

$$P(A \cap B) = P(\{\text{HHH, HHT}\}) = \frac{1}{4},\ P(A \cap C) = P(\{\text{HHT}\}) = \frac{1}{8},\ P(B \cap C) = P(\{\text{HHT, THH}\}) = \frac{1}{4}$$

Accordingly,

$$P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} = P(A \cap B), \quad \text{and so } A \text{ and } B \text{ are independent}$$

$$P(A)P(C) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} = P(A \cap C), \quad \text{and so } A \text{ and } C \text{ are independent}$$

$$P(B)P(C) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8} \neq P(B \cap C), \quad \text{and so } B \text{ and } C \text{ are dependent}$$

Frequently, we will postulate that two events are independent, or the experiment itself will imply that two events are independent.

Example 7.8

The probability that A hits a target is $\frac{1}{4}$, and the probability that B hits the target is $\frac{2}{5}$. Both shoot at the target. Find the probability that at least one of them hits the target, i.e., that A or B (or both) hit the target.

We are given that $P(A) = \frac{1}{2}$ and $P(B) = \frac{2}{5}$, and we seek $P(A \cup B)$. Furthermore, the probability that A or B hits the target is not influenced by what the other does; that is, the event that A hits the target is independent of the event that B hits the target, that is, $P(A \cap B) = P(A)\,P(B)$. Thus

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - P(A)\,P(B) = \frac{1}{4} + \frac{2}{5} - \left(\frac{1}{4}\right)\left(\frac{2}{5}\right) = \frac{11}{20}$$

7.7 RANDOM VARIABLES

Let S be a sample of an experiment. As noted previously, the outcome of the experiment, or the points in S, need not be numbers. For example, in tossing a coin the outcomes are H (heads) or T (tails), and in tossing a pair of dice the outcomes are pairs of integers. However, we frequently wish to assign a specific number to each outcome of the experiment. For example, in coin tossing, it may be convenient to assign 1 to H and 0 to T; or, in the tossing of a pair of dice, we may want to assign the sum of the two integers to the outcome. Such an assignment of numerical values is called a *random variable*. More generally, we have the following definition.

Definition: A *random variable* X is a rule that assigns a numerical value to each outcome in a sample space S.

We shall let R_X denote the set of numbers assigned by a random variable X, and we shall refer to R_X as the *range space*.

Remark: In more formal terminology, X is a function from S to the real numbers $\mathbf{R}$, and R_X is the range of X. Also, for some infinite sample spaces S, not all functions from S to $\mathbf{R}$ are considered to be random variables. However, the sample spaces here are finite, and every real-valued function defined on a finite sample space is a random variable.

Example 7.11

A pair of fair dice is tossed. (See Problem 7.3.) The sample space S consists of the 36 ordered pairs (a, b) where a and b can be any integers between 1 and 6; that is,

$$S = \{(1, 1), (1, 2), \ldots, (6, 6)\}$$

Let X assign to each point in S the sum of the numbers; then X is a random variable with range space

$$R_X = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

Let Y assign to each point the maximum of the two numbers; then Y is a random variable with range space

$$R_Y = \{1, 2, 3, 4, 5, 6\}$$

Example 7.12

A box contains 12 items of which three are defective. A sample of three items is selected from the box. The sample space S consists of the $\binom{12}{3} = 220$ different samples of size 3. Let X denote the number of defective items in the sample; then X is a random variable with range space $R_X = \{0, 1, 2, 3\}$.

Probability Distribution of a Random Variable

Let $R_X = \{x_1, x_2, \ldots, x_t\}$ be the range space of a random variable X defined on a finite sample space S. Then X induces an assignment of probabilities on the range space R_X as follows:

$$p_i = P(x_i) = P(X = x_i) = \text{sum of probabilities of points in } S \text{ whose image is } x_i$$

The set of ordered pairs $(x_1, p_1), \ldots, (x_t, p_t)$, usually given by a table

x_1	x_2	$\cdots$	x_t
p_1	p_2	$\cdots$	p_t

is called the *distribution* of the random variable X.

In the case that S is an equiprobable space, we can easily obtain the distribution of a random variable from the following result.

Theorem 7.8: Let S be an equiprobable space, and let X be a random variable on S with range space

$$R_X = \{x_1, x_2, ..., x_t\}.$$

Then

$$p_i = P(x_i) = \frac{\text{number of point in } S \text{ whose image is } x_i}{\text{number of points in } S}$$

Example 7.13

Consider the random variable X in Example 7.11 which assigns the sum to the toss of a pair of dice. We use Theorem 7.8 to obtain the distribution of X.

There is only one outcome (1, 1) whose sum is 2; hence $P(2) = \frac{1}{36}$. There are two outcomes, (1, 2) and (2, 1), whose sum is 3; hence $P(3) = \frac{2}{36}$. There are three outcomes, (1, 3), (2, 2), (3, 1), whose sum is 4; hence $P(4) = \frac{3}{36}$. Similarly $P(5) = \frac{4}{36}$, $P(6) = \frac{5}{36}$, ..., $P(12) = \frac{1}{36}$. The distribution of X consists of the points in R_X with their respective probabilities; that is,

x_i	2	3	4	5	6	7	8	9	10	11	12
p_i	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Example 7.14

Let X be the random variable in Example 7.12. We use Theorem 7.8 to obtain the distribution of X.

There are $\binom{9}{3} = 84$ samples of size 3 with no defective items; hence $P(0) = \frac{84}{220}$. There are $3\binom{9}{2} = 108$ samples of size 3 containing one defective item; hence $P(1) = \frac{108}{220}$. There are $\binom{3}{2} \cdot 9 = 27$ samples of size 3 containing two defective items; hence $P(2) = \frac{27}{220}$. There is only one sample of size 3 containing the three defective items; hence $P(3) = \frac{1}{220}$. The distribution of X follows:

x_i	0	1	2	3
p_i	$\frac{84}{220}$	$\frac{108}{220}$	$\frac{27}{220}$	$\frac{1}{220}$

Remark: Let X be a random variable on a probability space $S = \{a_1, a_2, ..., a_m\}$, and let $f(x)$ be any polynomial. Then $f(X)$ is the random variable which assigns $f(X(a_i))$ to the point a_i or, in other words, $f(X)(a_i) = f(X(a_i))$. Accordingly, if X takes on the values $x_1, x_2, ..., x_n$ with respective probabilities $p_1, p_2, ..., p_n$, then $f(X)$ takes on the values $f(x_1), f(x_2), ..., f(x_n)$ with the same corresponding probabilities.

Expectation of a Random Variable

Let X be a random variable. There are two important measurements (or parameters) associated with X, the *mean* of X, denoted by μ or μ_X, and the *standard deviation* of X denoted by σ or σ_X. The mean μ is also called the *expectation* of X, written $E(X)$. In a certain sense, the mean μ measures the "central tendency" of X, and the standard deviation σ measures the "spread" or "dispersion" of X. (The mean is sometimes called the *average* value since it corresponds to the average of a set of numbers where the probability of a number is defined by the relative frequency of the number in the set.) This subsection discusses the expectation $\mu = E(X)$ of X, and the next subsection discusses the standard deviation σ of X.

Let X be a random variable on a probability space $S = \{a_1, a_2, ..., a_m\}$. The *mean* or *expectation* of X is defined by

$$m = E(X) = X(a_1)\ P(a_1) + X(a_2)\ P(a_2) + \cdots + X(a_m)\ P(a_m) = \Sigma X(a_i)\ P(a_i)$$

In particular, if X is given by the distribution

x_1	x_2	$\cdots$	x_n
p_1	p_2	$\cdots$	p_n

then the expectation of X is

$$\mu = E(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n = \Sigma x_ip_i$$

(For notational convenience, we have omitted the limits in the summation symbol Σ.)

Example 7.15

(a) Suppose a fair coin is tossed six times. The number of heads which can occur with their respective probabilities are as follows:

x_i	0	1	2	3	4	5	6
p_i	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

Then the mean or expectation or expected number of heads is

$$m = E(X) = 0\left(\frac{1}{64}\right) + 1\left(\frac{6}{64}\right) + 2\left(\frac{15}{64}\right) + 3\left(\frac{20}{64}\right) + 4\left(\frac{15}{64}\right) + 5\left(\frac{6}{64}\right) + 6\left(\frac{1}{64}\right) = 3$$

(b) Consider the random variable X in Example 7.12 whose distribution appears in Example 7.14. It gives the possible numbers of defective items in a sample of size 3 with their

Remark: Let X be a random variable on a probability space $S = \{a_1, a_2, ..., a_m\}$, and let $f(x)$ be any polynomial. Then $f(X)$ is the random variable which assigns $f(X(a_i))$ to the point a_i or, in other words, $f(X)(a_i) = f(X(a_i))$. Accordingly, if X takes on the values $x_1, x_2, ..., x_n$ with respective probabilities $p_1, p_2, ..., p_n$, then $f(X)$ takes on the values $f(x_1), f(x_2), ..., f(x_n)$ with the same corresponding probabilities.

Expectation of a Random Variable

Let X be a random variable. There are two important measurements (or parameters) associated with X, the *mean* of X, denoted by μ or μ_X, and the *standard deviation* of X denoted by σ or σ_X. The mean μ is also called the *expectation* of X, written $E(X)$. In a certain sense, the mean μ measures the "central tendency" of X, and the standard deviation σ measures the "spread" or "dispersion" of X. (The mean is sometimes called the *average* value since it corresponds to the average of a set of numbers where the probability of a number is defined by the relative frequency of the number in the set.) This subsection discusses the expectation $\mu = E(X)$ of X, and the next subsection discusses the standard deviation σ of X.

Let X be a random variable on a probability space $S = \{a_1, a_2, ..., a_m\}$. The *mean* or *expectation* of X is defined by

$$m = E(X) = X(a_1)\ P(a_1) + X(a_2)\ P(a_2) + \cdots + X(a_m)\ P(a_m) = \Sigma X(a_i)\ P(a_i)$$

In particular, if X is given by the distribution

x_1	x_2	$\cdots$	x_n
p_1	p_2	$\cdots$	p_n

then the expectation of X is

$$\mu = E(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n = \Sigma x_ip_i$$

(For notational convenience, we have omitted the limits in the summation symbol Σ.)

Example 7.15

(a) Suppose a fair coin is tossed six times. The number of heads which can occur with their respective probabilities are as follows:

x_i	0	1	2	3	4	5	6
p_i	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$

Then the mean or expectation or expected number of heads is

$$m = E(X) = 0\left(\frac{1}{64}\right) + 1\left(\frac{6}{64}\right) + 2\left(\frac{15}{64}\right) + 3\left(\frac{20}{64}\right) + 4\left(\frac{15}{64}\right) + 5\left(\frac{6}{64}\right) + 6\left(\frac{1}{64}\right) = 3$$

(b) Consider the random variable X in Example 7.12 whose distribution appears in Example 7.14. It gives the possible numbers of defective items in a sample of size 3 with their respective probabilities. Then the expectation of X or, in other words, the expected number of defective items in a sample of size 3, is

$$m = E(X) = 0\left(\frac{84}{220}\right) + 1\left(\frac{108}{220}\right) + 2\left(\frac{27}{220}\right) + 3\left(\frac{1}{220}\right) = 0.75$$

(c) Three horses a, b, and c are in a race; suppose their respective probabilities of winning are $\frac{1}{2}, \frac{1}{3}$, and $\frac{1}{6}$. Let X denote the payoff function for the winning horse, and suppose X pays \$2, \$6, or \$9 according as a, b, or c wins the race. The expected payoff for the race is

$$E(X) = X(a)\ P(a) + X(b)\ P(b) + X(c)\ P(c)$$

$$= 2\left(\frac{1}{2}\right) + 6\left(\frac{1}{3}\right) + 9\left(\frac{1}{6}\right) = 4.5$$

Variance and Standard Deviation of a Random Variable

Consider a random variable X with mean μ and probability distribution

x_1	x_2	x_3	$\cdots$	x_n
p_1	p_2	p_3	$\cdots$	p_n

The *variance* $\text{Var}(X)$ and *standard deviation* σ of X are defined by

$$\text{Var}(X) = (x_1 - \mu)^2\, p_1 + (x_2 - \mu)^2\, p_2 + \cdots + (x_n - \mu)^2 p_n = \Sigma(x_i - \mu)^2\, p_i = E((X - \mu)^2)$$

$$\sigma = \sqrt{\text{Var}(X)}$$

The following formula is usually more convenient for computing $\text{Var}(X)$ than the above:

$$\text{Var}(X) = x_1^2\, p_1 + x_2^2 p_2 + \cdots + x_n^2 p_n - \mu^2 = \Sigma x_i^2 p_i - \mu^2 = E(X^2) - \mu^2$$

Remark: According to the above formula, $\text{Var}(X) = \sigma^2$. Both σ^2 and σ measure the weighted spread of the values x_i about the mean μ; however, σ has the same units as μ.

Example 7.16

(a) Let X denote the number of times heads occurs when a fair coin is tossed six times. The distribution of X appears in Example 7.15(a), where its mean $\mu = 3$ is computed. The variance of X is computed as follows:

$$\text{Var}(X) = (0 - 3)^2\ \frac{1}{64} + (1 - 3)^2\ \frac{6}{64} + (2 - 3)^2\ \frac{15}{64}\ (6 - 3)\ \frac{1}{64} = 1.5$$

Alternatively:

$$\text{Var}(X) = 0^2\,\frac{1}{64} + 1^2\,\frac{6}{64} + 2^2\,\frac{15}{64} + 3^2\,\frac{20}{64} + 4^2\,\frac{15}{64} + 5^2\,\frac{6}{64} + 6^2\,\frac{1}{64} - 3^2 = 1.5$$

Thus the standard deviation is $\sigma = \sqrt{1.5} \approx 1.225$ (heads).