1 Experiments and Probability

1.1 DEFINITION OF AN EXPERIMENT

To fully appreciate the meaning of probability and acquire a strong mathematical foundation for analytical work, it is necessary to define precisely the concept of an experiment and sample space mathematically. These definitions provide consistent methods for the assignment of elementary probabilities in paradoxical situations, and thus allow for meaningful calculation of probabilities of events other than the elementary events. Although at the beginning this approach may seem stilted, it will lead to a concrete concept of probability and an interpretation of derived probabilities.

An **experiment** $\mathscr{E}$ is specified by the three tuple $(S, \mathscr{F}, \mathscr{P}(.))$, where S is a finite, countable, or noncountable set called the **sample space**, $\mathscr{F}$ is a Borel field specifying a set of events, and $\mathscr{P}(.)$ is a probability measure allowing calculation of probabilities of all events.

1.1.1 The Sample Space

The sample space S is a set of elements called **outcomes** of the experiment $\mathscr{E}$ and the number of elements could be finite, countable, or noncountable infinite. For example, S could be the set containing the six faces of a die, $S = \{f_1, f_2, f_3, f_4, f_5, f_6\}$, or the positive integers, $S = \{i : i = 1, 2, \ldots\}$, or the real values between zero and one, $S = \{x : 0 < x < 1\}$, respectively.

An **event** is defined as any subset of S. On a single trial of the experiment an outcome is obtained. If that outcome is a member of an event, it is said that the event has occurred. In this way many different events occur at each trial of the experiment. For example, if f_1 is the outcome of a single trial of the experiment then the events $\{f_1\}, \{f_1, f_2\}, \{f_1, f_3\}, \ldots, \{f_1, f_6\}, \ldots, \{f_1, f_3, f_5\}, \ldots,$ all occur. Events consisting of single elements, like f_i, are called **elementary events**. The **impossible event** corresponds to the empty set ϕ and never occurs, while the **certain event**, S, contains all outcomes and thus always occurs no matter what the outcome of the trial is. Events A and B are called **mutually exclusive** or disjoint if $A \cap B = \phi$, where ϕ is the **null set**.

Two events A and B are called **independent** if $P(A \cap B) = P(A) \cdot P(B)$. The events $A_1, A_2, \ldots, A_n$ are defined to be independent if the probabilities of all intersections two, three, ..., and n events can be written as products. This implies for all $i, j, k, \ldots$, that the following conditions must be satisfied for independence

$$\begin{aligned} P(A_i \cap A_j) &= P(A_i)P(A_j) \\ P(A_i \cap A_j \cap A_k) &= P(A_i)P(A_j)P(A_k) \\ &\vdots \\ P(A_1 \cap A_2 \cap \ldots \cap A_n) &= P(A_1)P(A_2)\ldots P(A_n) \end{aligned} \tag{1.1}$$

1.1.2 The Borel Field

A **field** can be defined as a nonempty class of sets such that (1) if $a \in \mathscr{F}$, then the complement of $a \in \mathscr{F}$ and (2) if $a \in \mathscr{F}$ and $b \in \mathscr{F}$ then $a \cup b \in \mathscr{F}$. Thus a field contains all finite unions, and by virtue of compliments and DeMorgan's theorem, all intersections of the collection. If we further require that all infinite unions and intersections are present in the collection, a **Borel field** is defined.

The set of all events of our experiment that will have probabilities assigned to them (measurable events) must be a Borel field to have mathematical consistency. If A, a collection of events, has a finite number of elements, a Borel field can be formed as the set of those events plus all possible subsets obtained by unions and intersections of those events including the null set ϕ and entire set S.

If a set is noncountable, it is little harder to describe a Borel field. The most common Borel field, containing the real numbers, is the smallest Borel field containing the following intervals: $\{x : x \leq x_1\}$ for all $x_1 \in$ real numbers. This will contain all finite and infinite closed and open intervals of the form $[a, b], [a, b), (a, b]$, and (a, b), where a and b are real numbers and the intersections and unions of those intervals thereof.

1.1.3 The Probability Measure

The **probability measure**, $\mathscr{P}(.)$, must be a consistent assignment of probabilities such that the following conditions are satisfied

(1) For any event $A \in \mathscr{F}$, the probability of the event $A, P(A)$, is such that $P(A) \geq 0$.

(2) For the certain event, S, $P(S) = 1$.

(3) If A and B are any two events such that $A \cap B = \phi$ then $P(A \cup B) = P(A) + P(B)$.

(3a) If $A_i \in \mathscr{F}$ for $i = 1, 2, \ldots$, and $A_i \cap A_j = \phi$ for all $i \neq j$, then
$$P(A_1 \cup A_2 \cup \ldots \cup A_i \cup \ldots) = P(A_1) + P(A_2) + \cdots + P(A_i) + \cdots.$$

With these conditions satisfied, the probability of any event $A \in \mathscr{F}$ can be calculated. How does one go about assigning probabilities such that we satisfy the conditions above?

For the case where S is a set with a finite number of elements the conditions above can be satisfied by assigning probabilities to all the events with only single outcomes, $\{\zeta_i\}$, where $\zeta_i \in S$, such that conditions (1) and (2) above are satisfied. This mapping from the sample space S to the positive reals is called the **distribution function** for the probability measure and equivalently specifies the probability measure.

When S is a set with a noncountably infinite number of elements the assignment above is not useful as the probabilities of most elementary events will be zero. In this case the assignment of probabilities is consistent if the probabilities of the events $\{x : x \leq x_1\}$ for all $x_1 \in$ real numbers are assigned such that

(1) $0 \leq P\{x : x \leq x_1\} \leq 1$ (**bounded**).
(2) $P\{x : x \leq x_2\} \geq P\{x : x \leq x_1\}$ for all $x_2 > x_1$ (**nondecreasing function** of x) and
(3) lim as $\varepsilon \to 0$ of $P\{x : x \leq x_1 + \varepsilon\}$ equals $P\{x : x \leq x_1\}$ (**continuous from the right side**).

This mapping: $x \to P\{x : x \leq x_1\}$, defined for all x, is called the **cumulative distribution function** and equivalently describes the probability measure for the noncountable case. From this distribution function we are able to calculate all probabilities of events that are members of the Borel field $\mathscr{F}$. The cumulative distribution function could have also been used to specify the probability measure for the case where S has a countable or finite number of elements.

A number of examples of experiments will now be presented. They will include a couple of coin-tossing experiments and a die-rolling experiment. The experiments will be described by specifying their sample space, Borel field, and probability measures.

EXAMPLE 1.1

This experiment consists of a single flipping of a coin that results in either a head or a tail showing. Give its description by specifying as $(S, \mathscr{F}, \mathscr{P}(.))$.

SOLUTION

The possible outcomes of the experiment are either a head or a tail. Thus the sample space can be described as the set $S = \{\text{head, tail}\}$.

The Borel field $\mathscr{F}$ consists of the elementary events $\{\text{head}\}$ and $\{\text{tail}\}$, the null set $\varnothing$, and S.

To complete the description of the experiment, a probability measure must be assigned. This particular assignment could be based on previous experience, careful experimentation, use of favorable to total alternatives, or any other interpretation of the concept of probability. There is no right or wrong assignment, but certain assignments (models) may be more appropriate in explaining the results of corresponding physical experiments. For the purpose of this example, we assume that this coin has been tampered with for more often a head comes up than a tail, which we specify by $P\{\text{head}\} = p$ and $P\{\text{tail}\} = 1 - p$. These two assignments comprise the distribution function and thus the probability measure $\mathscr{P}(.)$ for the experiment. ■

EXAMPLE 1.2

A more realistic assignment for the experiment of flipping a coin could be the experiment $(S, \mathscr{F}, \mathscr{P}(.)))$ defined as follows: $S = \{\text{head, tail, edge}\}$ with $\mathscr{P}(.)$ described by $P\{h\} = 0.49, P\{t\} = 0.49, P\{e\} = 0.02$. The Borel field $\mathscr{F}$ is defined as the power set of S.

(a) Identify all the events (elements) of the Borel field. (b) Calculate the probabilities of those events.

SOLUTION

(a) The Borel field $\mathscr{F}$ specifying the measurable events is the power set of S (all possible subsets of S) given by

$$\mathscr{F} = \{\varnothing, \{h\}, \{t\}, \{e\}, \{h, t\}, \{h, e\}, \{t, e\}, \{h, t, e\}\}$$

where $\varnothing$ is the impossible event, and $\{h, t, e\}$ is the certain event. The other events consist of all possible proper subsets of S, single elements, and combinations of two elements.

(b) By definition $P(\varnothing) = 0$ and the probabilities of the elementary events are given in the specification of the probability measure of the experiment as $P\{h\} = 0.49$, $P\{t\} = 0.49$, and $P\{e\} = 0.02$. The probabilities of the other events can be determined by repeated application of property (3) for the probability measure. For example the events $\{h\}$ and $\{e\}$ are mutually exclusive therefore $P\{h, e\}$ can be found as follows:

$$\begin{aligned} P\{h, e\} &= P(\{h\} \cup \{e\}) = P\{h\} + P\{e\} \\ &= 0.49 + 0.02 = 0.51 \end{aligned}$$

Similarly $P\{h, t\} = 0.98$, $P\{t, e\} = 0.51$, and $P\{h, t, e\} = 1$. ■

EXAMPLE 1.3

An experiment consists of the "random" selection of a point somewhere in the closed interval [0,1]. Let the sample space S be described as $S = \{x : 0 \leq x \leq 1\}$. Define the Borel field $\mathscr{F}$ to be the smallest field containing the events $\{x : 0 \leq x \leq x_1\}$ for all $x_1 \in$ real numbers. Define the probability measure $\mathscr{P}(.)$, for the experiment by the following: $P\{x : x \leq x_1\} = x_1$ for all $0 \leq x_1 \leq 1$.

(a) Give three or four outcomes of this experiment. (b) Give a couple examples of regions in the event space (Borel field). (c) Determine the probabilities of the following events:

$$\{0 \leq x \leq 0.5\}, \{0.25 < x \leq 0.75\}, \{x > 0.75\}, \{x < 0.5\}, \text{and } \{x = 0.5\}.$$

SOLUTION

(a) The outcomes of the experiment are any real numbers in the closed interval [0,1]. Some examples are 0.333, $\pi/4$, 0.00001, and 0.2346.

(b) Possible regions in the event space are (0.5, 0.75], [0.6235, 0.8976), [0.2, 0.3], (0.627, 0.8), and any union of these types of regions.

(c) The probabilities of events can be determined by using set operations, the probability measure, and property (3). By direct application of the probability measure given, $P\{0 \leq x \leq 0.5\}$ is seen as

$$P\{0 \leq x \leq 0.5\} = 0.5$$

$P\{0.25 < x \leq 0.75\}$ can be determined indirectly as follows: We know that

$$\{0 \leq x \leq 0.75\} = \{0 \leq x \leq 0.25\} \cup \{0.25 < x \leq 0.75\}$$

Since $\{0 \leq x \leq 0.25\} \cap \{0.25 < x \leq 0.75\} = \varnothing$, we have by property (3) that

$$P\{0 < x \leq 0.75\} = P\{0 \leq x \leq 0.25\} + P\{0.25 < x \leq 0.75\}$$

Then rearranging and using the definition for the probability measure, we have

$$\begin{aligned} P\{0.25 < x \leq 0.75\} &= P\{0 \leq x \leq 0.75\} - P\{0 \leq x \leq 0.25\} \\ &= 0.75 - 0.25 = 0.5 \end{aligned}$$

$P\{x > 0.75\}$ is found indirectly using $\{0 \leq x \leq 1\} = \{0 \leq x \leq 0.75\} \cup \{x > 0.75\}$. Since $\{0 \leq x \leq 0.75\}$ and $\{x > 0.75\}$ are mutually exclusive, we have

$$P\{0 \leq x \leq 1\} = P\{0 \leq x \leq 0.75\} + P\{x > 0.75\}$$

After rearranging, $P\{x > 0.75\}$ is easily seen to be

$$P\{x > 0.75\} = P\{0 \leq x \leq 1\} - P\{0 \leq x \leq 0.75\}$$
$$= 1 - 0.75 = 0.25$$

Using similar concepts, it can be shown that $P\{x < 0.5\} = 0.5$ and $P\{x = 0.5\} = 0$. ■

1.2 COMBINED EXPERIMENTS

Combined experiments play an important role in probability theory applications. There are many ways we can combine experiments, including cartesian products in which independent trials of the same or different experiments can be described. In some cases the probabilities of events will depend on the results of previous trials of experiments or random selection of different experiments. A number examples of combined experiments are now explored beginning with the classical case of sampling with replacement.

1.2.1 Cartesian Product of Two Experiments

Consider the case of having two separate experiments specified by the following: $\mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.))$ and $\mathscr{E}_2 : (S_2, \mathscr{F}_2, \mathscr{P}_2(.))$. The sample spaces S_1 and S_2 are usually different sets, for example, results of a coin toss and results of a die roll, but they could be the same sets representing separate trials of the same experiment as in repeated coin-tossing experiments. We can define a new combined experiment by using the cartesian product concept as $\mathscr{E} = \mathscr{E}_1 \circledast \mathscr{E}_2$, where the new sample space $S = S_1 \circledast S_2$ is the **cartesian product** of the two sample spaces expressible by the ordered pair of elements where the first element is from S_1 and the second is from S_2.

EXAMPLE 1.4

Let experiment $\mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.))$ and $\mathscr{E}_2 : (S_2, \mathscr{F}_2, \mathscr{P}_2(.))$ be defined as follows: $\mathscr{E}_1$ is the experiment of flipping a coin with outcomes head (h) and tail (t) with equal probability of occurrence. $\mathscr{E}_2$ is the experiment of random selection of a colored ball from a box with outcomes red (r), white (w), and blue (b) with replacement. Define the new experiment $\mathscr{E}$ as $\mathscr{E} = \mathscr{E}_1 \circledast \mathscr{E}_2$ with experiments $\mathscr{E}_1$ and $\mathscr{E}_2$ being performed independently of each other; that is, the outcome of experiment $\mathscr{E}_1$ in no way effects the outcome of $\mathscr{E}_2$, and vice versa. Set up a reasonable model for this new experiment.

SOLUTION

To specify the model it suffices to give $S, \mathscr{F}, \mathscr{P}(.)$ of the new experiment $\mathscr{E}$.

The new sample space S is the cartesian product of the two experiments and given by $S = S_1 \circledast S_2$. Elements of S are ordered pairs with the first element coming from $S_1 = \{h, t\}$ and the second from $S_2 = \{r, w, b\}$; therefore

$$S = \{(h, r), (h, w), (h, b), (t, r), (t, w), (t, b)\}$$

The new Borel field $\mathscr{F}$ is selected as the power set of S, that is all possible unions and intersections of S. This includes the null set, the entire set, and all possible pairs, triples, and so on, as shown below:

$$\mathscr{F} = \left\{ \begin{array}{l} \varnothing, \{(h, r)\}, \{(h, w)\}, \{(h, b)\}, \{(t, r)\}, \{(t, w)\}, \{(t, b)\} \\ \{(h, r), (h, w)\}, \{(h, r), (h, b)\}, \{(h, r), (t, r)\}, \{(h, r), (t, w)\}, \ldots \\ \{(h, r), (h, w), (h, b)\}, \{(h, r), (h, w), (t, r)\}, \{(h, r), (h, w), (t, w)\}, \ldots \\ \qquad\qquad\qquad\qquad \vdots \\ \{(h, r), (h, w), (h, b), (t, r), (t, w), (t, b)\} \end{array} \right\}$$

The probability measure $\mathscr{P}(.)$ can be described by specifying the probability of the elementary events. Once these are known, the probability of any event can be found by writing the event as a union of those events and using property (3). Since the experiments are independent, it is reasonable to assign probabilities of the elementary events as a product of the probabilities from each experiment. For example, $P\{(h, r)\} = P\{h\} \cdot P\{r\}$. If head and tail are equally probable in $\mathscr{E}_1$ then it is reasonable for $\mathscr{P}_1(.)$ to be described by $P\{h\} = P\{t\} = \frac{1}{2}$, and if we have reason to believe that red, white, and blue are not equally probable in $\mathscr{E}_2$, then $\mathscr{P}_2(.)$ could be given by $P\{r\} = 0.5, P\{w\} = 0.3, P\{b\} = 0.2$. Thus, for this example, the probability measure $\mathscr{P}(.)$ can be described by specifying the probabilities of the elementary events (the distribution function):

$$\begin{array}{ll} P\{(h, r)\} = P\{(h)\} \cdot P\{(r)\} = 0.25 & P\{(t, r)\} = P\{(t)\} \cdot P\{(r)\} = 0.25 \\ P\{(h, w)\} = P\{(h)\} \cdot P\{(w)\} = 0.15 & P\{(t, w)\} = P\{(t)\} \cdot P\{(w)\} = 0.15 \\ P\{(h, b)\} = P\{(h)\} \cdot P\{(b)\} = 0.1 & P\{(t, b)\} = P\{(t)\} \cdot P\{(b)\} = 0.1 \end{array}$$

The event of a head in the new experiment is $H = \{(h, r), (h, w), (h, b)\}$, and its probability of occurrence can be determined using property (3) as

$$\begin{aligned} P(H) &= P\{(h, r), (h, w), (h, b) = P(\{(h, r)\} \cup \{(h, w)\} \cup \{(h, b)\}\} \\ &= 0.25 + 0.15 + 0.1 = 0.5 \end{aligned}$$

■

1.2.2 Cartesian Product of *n* Experiments

Consider the case of having n separate experiments specified by $\mathscr{E}_k : (S_k, \mathscr{F}_k, \mathscr{P}_k(.))$ for $k = 1, 2, \ldots, n$. Define a new combined experiment $\mathscr{E} : (S, \mathscr{F}, \mathscr{P}(.))$ as a cartesian product: $\mathscr{E} = \mathscr{E}_1 \circledast \mathscr{E}_2 \circledast \cdots \circledast \mathscr{E}_n$ where the new sample space

$S = S_1 \circledast S_2 \circledast \cdots \circledast S_n$ is the cartesian product of the n spaces and expressible by the ordered n-tuples of elements whose first element is from S_1 and the second from $S_2 \ldots$, the nth from S_n. The $\mathscr{E}_k$ are, in general, different, but in many cases the experiment could be formed from independent trials of the same experiment. Also there is an important class of problems where the experiments are the same, yet they cannot be thought of as independent. A good example of this is the random selection of outcomes without replacement. Examples of each type are now presented.

Binomial Distribution. Consider the experiment $\mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.))$ where the outcomes of the experiment are either failure indicated by a 0 or a success indicated by a 1; therefore $S_1 = \{0, 1\}$. Assume that the probability measure $\mathscr{P}_1(.)$ for the experiment is given by $P\{\text{success}\} \triangleq P\{1\} = p$ and $P\{\text{failure}\} \triangleq P\{0\} = 1 - p$ and that the $\mathscr{F}_1$ is the set $\{\varnothing, \{0\}, \{1\}, \{0, 1\}\}$. Define a new experiment by $\mathscr{E} = \mathscr{E}_1 \circledast \mathscr{E}_1 \circledast \cdots \circledast \mathscr{E}_1$ where $\mathscr{E} : (S, \mathscr{F}, \mathscr{P}(.))$ describes the new experiment. Assume that this represents independent trials of the same experiment $\mathscr{E}_1$ where the probability of success or failure is the same for each trial.

The $S, \mathscr{F}, \mathscr{P}(.)$ are now described for this new experiment. The new S is the cartesian product $S = \mathrm{S}_1 \circledast \mathrm{S}_1 \circledast \ldots \circledast \mathrm{S}_1$ and consists of all possible n-tuples where the elements are either 0 or 1 as shown below:

$$S = \begin{Bmatrix} (0, 0, \ldots, 0, 0) \\ (0, 0, \ldots, 0, 1) \\ (0, 0, \ldots, 1, 0) \\ (0, 0, \ldots, 1, 1) \\ \vdots \\ (1, 1, \ldots, 1, 1) \end{Bmatrix} \tag{1.2}$$

The new probability measure is specified once the distribution function or probabilities of the elements of S are determined. By virtue of the independent experiment assumption, the probability of each elementary event of the new experiment is the product of the probabilities for the elementary events in the single trial. For example, the $P\{(0, 1, 1, 0, \ldots, 0, 1, 1)\}$ is given by

$$\begin{aligned} P\{(0, 1, 1, 0, \ldots, 0, 1, 1)\} &= P\{0\} \cdot P\{1\} \cdot P\{1\} \cdot P(0) \cdot \ldots \cdot P\{0\} \cdot P\{1\} \cdot P\{1\} \\ &= p^4(1 - p)^{n-4} \end{aligned} \tag{1.3}$$

As a matter of fact every sequence that has only four ones (successes) will have this same probability. The total number of these sequences is the combination of n things taken four at a time, since we have n locations and we want only four of them with ones. The probability distribution function for the new experiment can then be determined as follows, where the first entry is the elementary sequence and after the arrow is the corresponding probability:

Outcome **Probability**

$$
\begin{aligned}
&(0,0,\ldots,0,0) \rightarrow P(0,0,\ldots,0,0) = P(0)\cdot P(0)\cdots P(0)\cdot P(0) = p^0(1-p)^n \\
&(0,0,\ldots,0,1) \rightarrow P(0,0,\ldots,0,1) = P(0)\cdot P(0)\cdots P(0)\cdot P(1) = p^1(1-p)^{n-1} \\
&(0,0,\ldots,1,0) \rightarrow P(0,0,\ldots,1,0) = P(0)\cdot P(0)\cdots P(1)\cdot P(0) = p^1(1-p)^{n-1} \\
&(0,0,\ldots,1,1) \rightarrow P(0,0,\ldots,1,1) = P(0)\cdot P(0)\cdots P(1)\cdot P(1) = p^2(1-p)^{n-2} \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
&(1,1,\ldots,1,1) \rightarrow P(1,1,\ldots,1,1) = P(1)\cdot P(1)\cdots P(1)\cdot P(1) = p^n(1-p)^0
\end{aligned}
\tag{1.4}
$$

The Borel field will be the power set associated with S and specifies all events for which probabilities are assigned.

Probabilities of different type of events for the above example can be determined by using the distribution function described. There are a wide number of applications as a success can mean all kinds of things. For example, a success could be obtaining an ace in drawing a card from a standard deck of cards with replacement, or obtaining successful reception of a binary symbol from a random communication channel.

The event of **exactly k successes out of n independent trials** appears frequently in physical situations and its probability will now be derived using the results above. Define the event A_1 as the event of exactly one success out of n trials. From the results above we see that the probability of exactly one success out of n trials is

$$
\begin{aligned}
P(A_1) &= P(\text{exactly one success out of } n \text{ trials}) \\
&= P(\{(0,0,\ldots,0,1)\} \cup \{(0,0,\ldots,1,0)\} \cup \ldots \cup \{(1,0,\ldots,0,0)\} \\
&= P\{(0,0,\ldots,0,1)\} + P\{(0,0,\ldots,1,0)\} + \cdots + P\{(1,0,\ldots,0,0)\} \\
&= \binom{n}{1} p^1(1-p)^{n-1}
\end{aligned}
\tag{1.5}
$$

Similarly the probability of exactly k successes out of n trials can be found to be

$$
\begin{aligned}
P(A_k) &= P(\text{exactly } k \text{ success out of } n \text{ trials}) \\
&= \binom{n}{k} p^k(1-p)^{n-k}
\end{aligned}
\tag{1.6}
$$

In some problems we may wish to know the probability that **the number of successes out of n trials is within a range of values**. The probability that the number of successes k is in the range $m \le k \le n$ can be obtained by adding the probabilities of exactly k successes for that range, since the events A_k and A_j are mutually exclusive events for all k and j such that $k \neq j$. Therefore

$$P\left(J \leq \begin{matrix}\text{number of}\\ \text{successes}\end{matrix} \leq K\right) = P\left(\begin{pmatrix}\text{exactly } J\\ \text{successes}\end{pmatrix} + \begin{pmatrix}\text{exactly } J+1\\ \text{successes}\end{pmatrix} + \cdots + \begin{pmatrix}\text{exactly } K\\ \text{successes}\end{pmatrix}\right)$$

$$= P\left(\bigcup_{k=J}^{K} A_k\right) = \sum_{k=J}^{K} \binom{n}{k} p^k (1-p)^{n-k} \tag{1.7}$$

EXAMPLE 1.5

Consider the experiment of tossing a fair coin with the sample space $S = \{h, t\}$ and probability distribution $P\{h\} = 0.6$ and $P\{t\} = 0.4$. Suppose the experiment is performed 10 times independently to define a new cartesian product experiment.

(a) Determine the probability that we get exactly 5 heads in the 10 trials.
(b) Determine the probability that we get greater than 7 heads in the 10 trials.
(c) Determine the probability that we get less than or equal 9 heads in the 10 trials.
(d) Determine the probability that the number of tails is greater than or equal to 4 and less than or equal 5.

SOLUTION

The desired probabilities are determined for Eqs. (1.6) and (1.7) as follows:

(a) $P(A_1) = P(\text{exactly 5 heads out of 10 trials})$

$$= \binom{10}{5}(0.6)^5(0.4)^{10-5} = 0.20066$$

(b) $$P\left(8 \leq \begin{matrix}\text{number}\\ \text{of heads}\end{matrix} \leq 10\right) = P\left(\begin{pmatrix}\text{exactly}\\ \text{8 heads}\end{pmatrix} + \begin{pmatrix}\text{exactly}\\ \text{9 heads}\end{pmatrix} + \begin{pmatrix}\text{exactly}\\ \text{10 heads}\end{pmatrix}\right)$$

$$= \binom{10}{8}(0.6)^8(0.4)^2 + \binom{10}{9}(0.6)^9(0.4)^1 + \binom{10}{10}(0.6)^{10}(0.4)^0$$

$$= 0.1269790$$

(c) $P(\text{less than or equal 9 heads out of 10 trials})$

$$= 1 - P(\text{exacly 10 heads out of 10 trials})$$

$$= 1 - \binom{10}{10}(0.6)^{10}(0.4)^0 = 0.99395$$

(d) $$P\left(4 \leq \begin{matrix}\text{number}\\ \text{of heads}\end{matrix} \leq 5\right) = P\left(\begin{pmatrix}\text{exactly}\\ \text{4 heads}\end{pmatrix} + \begin{pmatrix}\text{exactly}\\ \text{5 heads}\end{pmatrix}\right)$$

$$= \binom{10}{4}(0.6)^4(0.4)^6 + \binom{10}{5}(0.5)^5(0.4)^5 = 0.1315425$$

■

Approximations for Binomial Probabilities. The probabilities of k successes in n trials of a Bernoulli experiment was found to be as in Eq. (1.6) and evaluating probabilities of ranges of successes is given in Eq. (1.7). In Figure 1.1 the $P(k$ successes out of n trials) is plotted for values of n equal to 5, 10, and 50 for all values of k, and a $P(\text{success}) = 0.6$. The graphs show a tendency toward a hill-shaped curve similar to a Gaussian function. These plots would be symmetrical around the point 0.5 n if $p = 0.5$. However, when p does not equal 0.5, as for the cases given, the plot is not quite symmetrical. If the number of trials is large, the calculation load due to the factorials is considerable, and certain approximations to these probabilities become useful. Many of these approximations are good provided that the number of trials n, number of successes k, and probability of success p satisfy a given set of conditions. Of these approximations we will present the DeMoivre and Poisson approximations and the regions where the approximations are reliable.

DeMoivre-Laplace Approximation. For $np(1-p) \gg 1$ and $|k - np|$ of the order of $\sqrt{np(1-p)}$,

$$\binom{n}{k} p^k (1-p)^{n-k} \approx \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{-\frac{(k-np)^2}{2np(1-p)}\right\} \tag{1.8}$$

Poisson Approximation. For $n \gg 1, p \ll 1$, and np of order 1,

$$\binom{n}{k} p^k (1-p)^{n-k} \approx e^{(-np)} \frac{(np)^k}{k!} \tag{1.9}$$

Approximation of Regions of Successes in Bernoulli Trials. Say that we are interested in approximating the probability that the number of successes k in n repeated trials of a Bernoulli experiment is in the range $k_1 \le k \le k_2$. If this range of successes

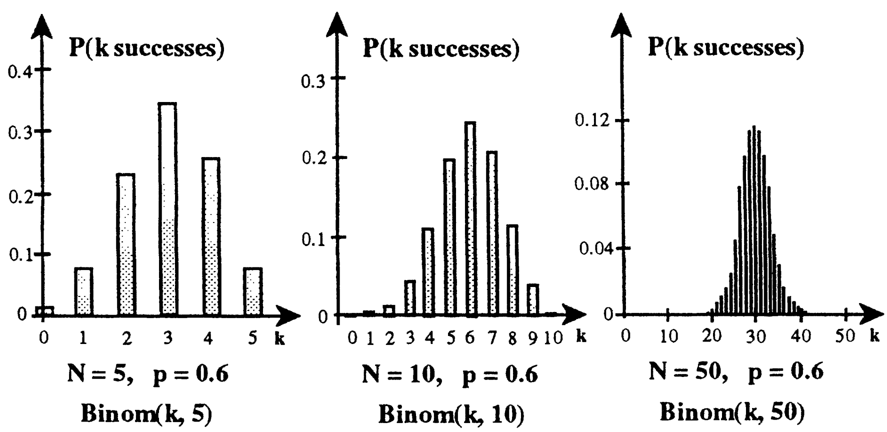

Figure 1.1 Plots of P (k successes out of n trials) for $p = 0.6$ and $n = 5, 10, 50$.

contains values that satisfy the DeMoivre-Laplace approximation, then the summation can be approximated by using the error function or the $\Phi(x)$ function as follows:

$$\sum_{k=k_1}^{k_2} \binom{n}{k} p^k (1-p)^{n-k} \approx \Phi\left(\frac{k_2 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{np(1-p)}}\right) \tag{1.10}$$

$$\text{where } \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

The $\Phi(x)$ can only be determined by numerical integration as there is no antiderivative and it is convenient to use the table given in Appendix E.

EXAMPLE 1.6 (APPROXIMATIONS)

Consider the experiment of transmitting a single binary symbol across a communication channel that has an error probability $P(\text{error}) = 0.01$ (a pretty poor channel). Suppose that 1000 symbols are transmitted independently.

(a) Determine exactly the probability that exactly 10 errors occur in the 1000 trials.
(b) Are the conditions satisfied to use the DeMoivre-Laplace approximation? If so, evaluate the probability using the approximation and compare to the results in (a).
(c) Are the conditions satisfied to use the Poisson approximation? If not, calculate the probability using the approximation and compare to exact answer.
(d) Determine exactly the probability that the number of errors is greater than 7 but less than 16. Also use the DeMoivre approximation to determine the probability and compare it to the exact answer.
(e) Assuming 10,000 independent transmissions across a channel that has an error probability of 0.0001, determine the probability of exactly one error in the total transmission. Are the conditions satisfied to use the Poisson approximation? If so, determine the approximate results and compare to the exact results.

SOLUTION

(a) Using Eq. (1.6), the Bernoulli trials result, gives

$$P(\text{exactly 10 errors in 1000 trials}) = \binom{1000}{10}(0.01)^{10}(0.99)^{990} = 0.12574$$

(b) To see if the conditions are satisfied, it is necessary to calculate the following:

$$np(1-p) = 1000(0.01)(0.99) = 9.9$$
$$|k - np| = |10 - 1000(0.01)| = 0$$
$$\sqrt{np(1-p)} = \sqrt{1000(0.01)(0.99)} = 3.1464$$

Since $np(1-p) > 1$ and $|k - np|$ is of the same order of as $\sqrt{np(1-p)}$, the probability of exactly 10 errors can be approximated by the Gaussian function from Eq. (1.8) as follows:

$$\binom{1000}{10}0.01^{10}(0.99)^{990} \approx \frac{1}{\sqrt{2\pi}3.1464}\exp\left\{-\frac{(10-10))^2}{2(9.9)}\right\} = 0.126809$$

The exact result and approximate result differ in the third decimal. Although close, since $np(1-p)$ is only an order of magnitude rather than $\gg 1$, there is a loss in accuracy.

(c) To use the Poisson approximation the conditions $n \gg 1, p \ll 1$, and np of order 1 must be satisfied. At 1000, n is big enough, at 0.01 p is is small enough, but $np = 10$ is only marginally of order 1. Therefore we expect that the Poisson approximation will be marginal. Using the approximation given in Eq. (1.9) anyway gives

$$\binom{1000}{10}(0.01)^{10}(0.99)^{990} \approx e^{(-10)}\frac{(10)^{10}}{10!} = 0.12511$$

It is seen that this result is still pretty close to the exact value of 0.12574 determined in part (a).

(d) Since $|8 - 10|$ and $|15 - 10|$ are of order 3.1464 and $1000(0.01)(0.09) = 9.9$ is marginally $\gg 1$, then the approximation given in Eq. (1.10) can be used as follows:

$$\begin{aligned}\sum_{k=8}^{15}\binom{1000}{k}(0.01)^k(0.99)^{1000-k} &\approx \Phi\left(\frac{15-10}{3.1464}\right) - \Phi\left(\frac{8-10}{3.1464}\right)\\ &= \Phi(1.589) - \Phi(-0.635647)\\ &= 0.94408 - (1 - 0.73748) = 0.68156\end{aligned}$$

(e) The probability of exactly one error is given by the formula (1.6) as

$$\begin{aligned}P\begin{pmatrix}\text{exactly 1 error}\\ \text{in 10,000 trials}\end{pmatrix} &= \binom{10{,}000}{1}(0.0001)^1(0.9999)^{9999}\\ &= 0.367898\end{aligned}$$

To use the Poisson approximation the conditions $n \gg 1, p \ll 1$, and np of order 1 must be satisfied. At 10,000, n is big enough, at 0.0001, p is is small enough, and

$np = 1$ is definitely of order 1, so all conditions are satisfied and the approximation should be very good. The approximation from Eq. (1.9) is as follows:

$$\binom{10{,}000}{1} 0.0001^1 (0.9999)^{9999} \approx e^{(-1)} \frac{(1)^1}{1!} = 0.36788$$

It is seen that the approximation is accurate to the fourth decimal place. ■

Multinomial Distribution. A combined experiment that is an extension of the Bernoulli trial experiment just described, which resulted in the binomial distribution, is the case of multiple occurrences of several different events on multiple trials of the same experiment. Consider an experiment $\mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.))$. Define a set of events $A_i, i = 1$ to k, consisting of elements of S_1 such that their union is the certain event, they are pairwise disjoint, and their probabilities are given by $P(A_i) = p_i$ for $i = 1$ to k.

Define a new experiment $\mathscr{E}$ by $\mathscr{E} = \mathscr{E}_1 \circledast \mathscr{E}_1 \circledast \ldots \circledast \mathscr{E}_1$ described by $\mathscr{E} : (S, \mathscr{F}, \mathscr{P}(.))$. Assume that this represents n independent trials of the same experiment $\mathscr{E}_1$. Let $n_i, i = 1$ to k, be the number of times event A_i occurs in the new experiment, which is composed of n repeated trials. Also assume that $n_i \geq 0$ and $n_1 + n_2 + \cdots + n_k = n$.

It can be shown that the probability that A_i occurs exactly n_i times in the n trials is

$$P\begin{pmatrix} A_1 \text{ occurs exactly } n_1 \text{ times,} \\ A_2 \text{ occurs exactly } n_2 \text{ times,} \\ \vdots \\ A_k \text{ occurs exactly } n_k \text{ times} \end{pmatrix} = \frac{n!}{n_1! n_2! \ldots n_k!} p_1^{n_1} p_2^{n_2} \ldots p_k^{n_k} \tag{1.15}$$

EXAMPLE 1.7

An archer shoots at a target on a single trial of an experiment. Assume that the following events are defined: A_1 the arrow hits the bull's eye, A_2 the arrow hits the first ring of the target not on the bulls eye, and A_3 the arrow misses the target entirely. The probabilities of these events are given by

$$P(A_1) = 0.1, \quad P(A_2) = 0.7, \quad \text{and} \quad P(A_3) = 0.2$$

Assume that the archer shoots a number of arrows and that the results of each arrow shot are independent from the others. Find:

(a) The probability that on six arrows shot, the archer hit the bulls eye twice, hits the ring three times, and misses the target once.

(b) The probability that on five arrows shot, the archer hits the bulls eye three times and the ring twice.

SOLUTION

The desired probabilities are determined from Eq. (1-15) as follows

$$\text{(a)}\quad P\begin{pmatrix} A_1 \text{ occurs exactly 2 times,} \\ A_2 \text{ occurs exactly 3 times,} \\ A_3 \text{ occurs exactly 1 times} \end{pmatrix} = \frac{6!}{2!3!1!}(0.1)^2(0.7)^3(0.2)^1 = 0.04116$$

$$\text{(b)}\quad P\begin{pmatrix} A_1 \text{ occurs exactly 3 times,} \\ A_2 \text{ occurs exactly 2 times,} \\ A_3 \text{ occurs exactly 0 times} \end{pmatrix} = \frac{5!}{3!2!0!}(0.1)^3(0.7)^2(0.2)^0 = 0.049$$

Hypergeometric Distribution. In the previous compound experiments the trials of the experiment were considered to be independent. A very important class of experimentation problems is sampling done without replacement, for a finite number of elements, and without independent trials. For example, if a is the number of successful (s) elements and b is the number of failure (f) elements and n elements are drawn at random but not replaced, the corresponding outcomes of the experiment do not contain all possible n-tuples of s and f. So the n-tuple $\{s, s, \ldots, s\}$ and sequences containing more than a successes are not possible.

It can be shown for this experiment that the probability of k successes out of n trials can be determined as

$$P(k \text{ successes in } n \text{ trials}) = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}, \quad k = 0, 1, \ldots \tag{1.16}$$

The next example is typical of the type of problems for which the hypergeometric formula above can be used.

EXAMPLE 1.8

Consider a standard deck of 52 playing cards with 4 aces, 4 kings, . . . , and 4 twos. Say we draw 5 cards at random without replacement. What is the probability that our draw consists of (a) exactly 3 aces? (b) no aces? (c) more than two aces?

SOLUTION

(a) Using the formula (1.16) above, we find the probability of drawing exactly three aces as

$$P(3 \text{ successes in 5 trials}) = \frac{\binom{4}{3}\binom{48}{5-3}}{\binom{4+48}{5}} = 0.001736$$

(b) Using the formula (1.16) above, we find the probability of drawing exactly zero aces as

$$P(0 \text{ aces in } 5 \text{ trials}) = \frac{\binom{4}{0}\binom{48}{5-0}}{\binom{4+48}{5}} = 0.6588$$

(c) Using the formula (1.16) for two independent events, we find the probability of drawing more than two aces as

$$\begin{aligned} &P(\text{more than 2 aces in 5 draws}) \\ &= P((\text{exactly 3 aces in 5 draws}) \cup (\text{exactly 4 aces in 5 draws}) \\ &= \frac{\binom{4}{3}\binom{48}{5-3}}{\binom{48+4}{5}} + \frac{\binom{4}{4}\binom{48}{5-4}}{\binom{48+4}{5}} = 0.001754469 \end{aligned}$$

■

1.2.3 Counting Experiments

A special class of compound experiments deal with successive application of experiments where the total experiment may be stopped at any point depending on the results from the current experiment. Two common distributions result from these compound experiments.

Geometric Distribution. The geometric distribution is a result of another form of compound experiment. Define a countable number of identical experiments $\mathscr{E}_k : (S_k, \mathscr{F}_k, \mathscr{P}_k(.))$ for $k = 1, 2, \ldots$. Each of these experiments will be of the Bernoulli type (two possible outcomes) previously discussed, with identical probabilities of success; the result of the experiment, if performed, is independent of previous or future experiments. The new experiment is described as follows: Perform experiment $\mathscr{E}_1$ if a success is obtained, then stop; if a success does not happen, then perform experiment $\mathscr{E}_2$; if a success occurs, then stop otherwise. Continue this process until a success occurs. Thus the number of trials is not the same each time the compound experiment is performed, and thus the sample space is not a cartesian product.

If p is the probability of success on each experiment, then the sample space, Borel field and probability measure can be described as follows: The elementary events are sequences of 1's (for success) and 0's for failures but with the property that they end on a 1 and only have 0's preceding the 1. Thus S can be described by

$$S = \{(1), (0, 1), (0, 0, 1), \ldots, (0, 0, \ldots, 0, 1), \ldots\} \tag{1.17}$$

Notice that (1, 0), (0, 1, 0), (1, 1), etc., are not elements of the sample space since the new experiment would stop at a 1 and not proceed to the next one. Events that are collections of outcomes are not of the same sequence length. For example, $\{(1), (0, 1)\}$ is an event where a success is obtained in less than or equal to 2 trials.

The probability measure $\mathscr{P}(.)$ can be specified by giving the distribution function for the elementary events. The $P\{(1)\}$ is the probability of getting a success on the very first trial. Since the probability of a success on the first experiment is p, the probability of getting $\{(1)\}$ is also p,

$$P\{(1)\} = p \tag{1.18}$$

The probability of getting $\{(0, 1)\}$ equals the probability of getting a failure on performing the first experiment and getting a success on the performance of the second experiment. Since results of each experiment are independent if performed, the $P\{(0, 1)\}$ can be written as a product:

$$P\{(0, 1)\} = (1-p)p \tag{1.19}$$

Similarly the probability of the elementary event consisting of a sequence of $k-1$ failures (0's) followed by a single success (1) is

$$\underset{(k-1)\text{zeros}}{P\{(0, 0, \ldots, 0, 1)\}} = (1-p)^{k-1}p \tag{1.20}$$

These probabilities as $k \to \infty$ determine the distribution function, or equivalently the probability measure $\mathscr{P}(.)$. It can be shown that

$$\sum_{k=1}^{\infty}(1-p)^{k-1}p = 1 \tag{1.21}$$

The Borel field is again defined as the power set of S, that is, all possible subsets of S. The probabilities of arbitrary events that are the subsets of S can then be found by adding up the probabilities of the elementary elements in the event. For example, define the event A to be the set of outcomes such that a success occurs on the performance of an odd number of experiments. Thus A can be written as

$$A = \{(1), (0, 0, 1), (0, 0, 0, 0, 1), \ldots, (0^{(2k)}, 1), \ldots\} \tag{1.22}$$

The probability of A can be determined by adding up the probabilities of the elementary events and using the properties of a geometric sequence as follows.

$$\begin{aligned} P(A) &= P(\{(1), (0, 0, 1), (0, 0, 0, 0, 1), \ldots, (0^{(2k)}, 1), \ldots\}) \\ &= \sum_{k=0}^{\infty}(1-p)^{2k}p = p\left(\frac{1}{1-(1-p)^2}\right) = \frac{1}{2-p} \end{aligned} \tag{1.23}$$

Negative Binomial Distribution. An extension of the geometric experiment is a compound experiment in which the number of trials necessary to obtain k successes is desired rather than the number of trials needed for a single success. The basic underlying experiment is to define a countable number of identical experiments $\mathscr{E}_k : (S_k, \mathscr{F}_k, \mathscr{P}_k(.))$ for $k = 1, 2, \ldots$. Each of these experiments will be of the Bernoulli type with identical probabilities of success, and the result of the experiment, if performed, is independent of previous or future experiments. The new experiment is described as follows: Perform experiment $\mathscr{E}_i$ if a success is obtained and it is the kth, then stop; if not, continue to the next experiment $\mathscr{E}_{i+1}$. Continue this process until the kth success occurs.

Let x be the number of trials in which the kth success occurs; then the probability of that event can be shown to be

$$P(\text{event that the kth success occurs on the } x\text{th trial}) = \binom{x-1}{k-1} p^k (1-p)^{x-k} \qquad x = k, k+1, k+2, \ldots \quad (1.24)$$

The following example illustrates the type of problems that will use the negative Binomial distribution given above.

EXAMPLE 1.9

Suppose the probability of getting a f_2 on the single toss of a die is 0.2. Find (a) the probability that the fourth f_2 occurs on the 6th trial; (b) that the fifth f_2 occurs before the 7th trial.

SOLUTION

(a) $P(\text{event that the 4th success occurs on the 6th trial})$

$$= \binom{6-1}{4-1} 0.2^4 (1-0.2)^{6-4} = 0.02048 \quad (1.25)$$

(b) $P(\text{event that the 5th success occurs before the 7th trial})$

$= P((\text{event that the 5th success occurs on the 5th trial})$

$\cup$ event that the 5th success occurs on the 6th trial))

$$= \binom{5-1}{5-1} 0.2^5 (1-0.2)^{5-5} + \binom{6-1}{5-1} 0.2^5 (1-0.2)^{6-5} \quad (1.26)$$

■

1.2.4 Selection Combined Experiment

A different type of combined experiment will now be considered that is a combination of three or more experiments. For purpose of illustration only three component experiments are presented. Let the three experiments be defined as

$$\mathscr{E}_0 : (S_0, \mathscr{F}_0, \mathscr{P}_0(.)), \quad \mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.)), \quad \mathscr{E}_2 : (S_2, \mathscr{F}_2, \mathscr{P}_2(.))$$

The sample space for $\mathscr{E}_0$ consists of only two outcomes (e_1 and e_2). These outcomes serve to help us select which one of the other experiments will be performed. If the outcome is e_1, then $\mathscr{E}_1$ is performed with outcome α_i; if the outcome is e_2 then $\mathscr{E}_2$ is performed with outcome β_i. Thus the combination experiment $\mathscr{E} : (\mathrm{S}, \mathscr{F}, \mathscr{P}(.))$ can be described as follows: The sample space S will consist of ordered pairs of elements the first either e_1 or e_2 and the second either $\alpha_i \in S_1$ or $\beta_i \in S_2$. Thus a trial of the combined experiment results in an outcome $\zeta = (e, \delta)$, where $e \in \{e_1, e_2)$ and $\delta \in S_1 \cup S_2$.

The Borel field $\mathscr{F}$ will be defined to be the set of all possible subsets of S, and the associated probability measure can be described as

$$\begin{aligned} P(\{(e_i, \delta_j)\}) &= P_0(e_i) \cdot P(\delta_j) \quad \text{where} \\ P(\delta_j) &= P_1(\alpha_j) \quad \text{if } e_i = e_1 \quad \text{or} \\ P(\delta_j) &= P_2(\beta_j) \quad \text{if } e_i = e_2 \end{aligned} \tag{1.27}$$

The product of the probabilities is because of the assumption that the experiment $\mathscr{E}_0$ is independent of the experiments $\mathscr{E}_1$ and $\mathscr{E}_2$.

EXAMPLE 1.10

Define three experiments by $\mathscr{E}_0 : (S_0, \mathscr{F}_0, \mathscr{P}_0(.))$, $\mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.))$, and $\mathscr{E}_2 : (S_2, \mathscr{F}_2, \mathscr{P}_2(.))$. Experiment $\mathscr{E}_0$ is performed with $S_0 = \{e_1, e_2\}$. If the outcome is e_1 then box 1 is chosen and experiment $\mathscr{E}_1$ performed, while if e_2 is the outcome, $\mathscr{E}_2$ is performed. Experiment $\mathscr{E}_1$ is the random selection of a colored ball from a box with outcomes $S_1 = \{r, w, b\}$ where r means red, w means white, and b means blue ball chosen. The probability distribution for this experiment is given by

$$P_1(r) = 0.5, \quad P_1(w) = 0.3, \quad P_1(b) = 0.2$$

Similarly $\mathscr{E}_2$ is the random selection of a colored ball from box 2 having sample space with outcomes $S_2 = \{r, w, b\}$ and probability distribution given by

$$P_2(r) = 0.4, \quad P_2(w) = 0.1, \quad P_2(b) = 0.5$$

(a) Describe the sample space S for the this new compound experiment.
(b) Assume $P_0(e_1) = 0.6$ and describe the probability measure $\mathscr{P}(.)$ for this new compound experiment.
(c) Show the event that the ball drawn is red.
(d) Show the event that box 1 is selected.

SOLUTION

(a) The new sample space S consists of ordered pairs

$$S = \{(e_1, r), (e_1, w), (e_1, b), (e_2, r), (e_2, w), (e_2, b)\}$$

This sample space is illustrated in Figure 1.2(a).

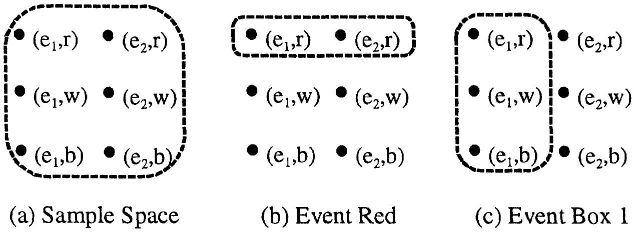

Figure 1.2 Set diagram for (a) Sample Space S, (b) Event red, (c) Event Box 1.

(b) The $\mathscr{P}(.)$ is specified by giving the distribution function. Using the results shown in the previous section because the experiments were given as independent, the distribution function is determined as

$$\begin{aligned}
&\text{Outcome} \rightarrow \text{Probability}\\
&(e_1, r) \rightarrow P\{(e_1, r)\} = P_0\{e_1\} \cdot P_1\{r\} = 0.6 \cdot 0.5 = 0.3\\
&(e_1, w) \rightarrow P\{(e_1, w)\} = P_0\{e_1\} \cdot P_1\{w\} = 0.6 \cdot 0.3 = 0.18\\
&(e_1, b) \rightarrow P\{(e_1, b)\} = P_0\{e_1\} \cdot P_1\{b\} = 0.6 \cdot 0.2 = 0.12\\
&(e_2, r) \rightarrow P\{(e_2, r)\} = P_0\{e_2\} \cdot P_2\{r\} = 0.4 \cdot 0.5 = 0.2\\
&(e_2, w) \rightarrow P\{(e_2, w)\} = P_0\{e_2\} \cdot P_2\{w\} = 0.4 \cdot 0.3 = 0.12\\
&(e_2, b) \rightarrow P\{(e_2, b)\} = P_0\{e_2\} \cdot P_2\{b\} = 0.4 \cdot 0.2 = 0.08
\end{aligned}$$

(c) The event the ball is red is shown in the Figure 1.2 and given by

$$\text{Ball is red} = \{(e_1, r), (e_2, r)\}$$

(d) The event box 1 is chosen, also shown in Figure 1.2, is as follows:

$$\text{Box 1 is chosen} = \{(e_1, r), (e_1, w), (e_1, b)\}$$

■

1.3 CONDITIONAL PROBABILITY

The probability measure allows us to calculate the probabilities of all events that are members of the Borel field for the defined experiment. It will become useful to define the concept of conditional probability of events. Given a conditioning event C such that $P(C) > 0$, the **conditional probability of any event A assuming C** is defined as

$$P(A \mid C) \triangleq \frac{P(A \cap C)}{P(C)} \tag{1.28}$$

The following examples consider the calculation of conditional probabilities for discrete and continuous sample spaces. It is seen that the fundamental set operations and the probability measure for the underlying experiment are used to obtain conditional probabilities.

EXAMPLE 1.11

Consider an experiment defined as a single toss of a "crooked" die with sample space $S = \{f_1, f_2, f_3, f_4, f_5, f_6\}$. The probability measure, maybe based on past history, is known to be

$$P\{f_1\} = \tfrac{1}{2} \quad P\{f_2\} = \tfrac{1}{4} \quad P\{f_3\} = \tfrac{1}{8}$$
$$P\{f_4\} = \tfrac{1}{16} \quad P\{f_5\} = \tfrac{1}{32} \quad P\{f_6\} = \tfrac{1}{32}$$

Let the conditioning event C be given as $C = \{f_1, f_3, f_5\}$, the face of the die is odd, and calculate the $P(A \mid C)$ where $A = \{f_1, f_2\}$, the face is less than or equal 2.

SOLUTION

By the definition of conditional probability, Eq. (1.28), $P(A \mid C)$ is

$$P(A \mid C) = \frac{P(A \cap C)}{P(C)} = \frac{P(\{f_1, f_2\} \cap \{f_1, f_3, f_5\})}{P(\{f_1, f_3, f_5\})}$$
$$= \frac{P(\{f_1\})}{P(\{f_1, f_3, f_5\})} = \frac{1/2}{1/2 + 1/8 + 1/32} = \frac{16}{21} \qquad \blacksquare$$

EXAMPLE 1.12

Consider the experiment of tossing two altered dice independently where the individual probabilities are as given in Example 1.11. Determine the probability of the event that one die is a six conditioned on the event that the sum of the two dice is 10.

SOLUTION

The sample space for the new experiment is the cartesian product of the single toss experiment, and thus the sample space S consists of the set of all ordered pairs with f_i as first element and f_j as second example:

$$S = \{(f_i, f_j) : i, j = 1, 2, \ldots, 6\}$$

The probability distribution for the sample space is the mapping from S to the reals, and by the independence condition, it is determined as the product

$$(f_i, f_j) \to P\{(f_i, f_j)\} = P\{f_i\} \cdot P\{f_j\} \qquad \text{for } i, j = 1, 2, \ldots, 6$$

The events under consideration are as follows:

$$\{\text{one die is } 6\} = \{(f_6,f_j) : j = 1, 2, \ldots, 5\} \cup \{(f_i,f_6) : i = 1, 2, \ldots, 5\} \cup \{(f_6,f_6)\}$$
$$\{\text{sum of dice equals } 10\} = \{(f_4,f_6), (f_6,f_4), (f_5,f_5)\}$$
$$\{\text{sum of dice equals } 10\} \cap \{\text{one die is } 6\} = \{(f_4,f_6), (f_6,f_4)\}$$

From the sets above the conditional probability desired is seen to be

$$\begin{aligned}
P(\{\text{one die is a } 6\} \mid \{\text{sum is } 10\}) &= \frac{P(\{\text{one die is a } 6\} \cap \{\text{sum is } 10\})}{P(\{\text{sum is } 10\})} \\
&= \frac{P\{(f_4,f_6)\} + P\{(f_6,f_4)\}}{P\{(f_4,f_6)\} + P\{(f_4,f_6)\} + P\{(f_5,f_5)\}} \\
&= \frac{(1/16)(1/32) + (1/32)(1/16)}{(1/16)(1/32) + (1/32)(1/16) + (1/32)(1/32)} \\
&= \frac{4}{5}
\end{aligned}$$

■

EXAMPLE 1.13

Define an experiment that has as outcomes, t, the set of real numbers greater than or equal to zero and that the outcome represents the time until a certain device fails. The probability measure for the experiment is described as

$$P(\text{failure time } t \leq x) = 1 - e^{-x} \qquad \text{for } x \geq 0$$

The Borel field is given by the smallest field containing the intervals $\{t : 0 \leq t \leq x\}$ for all $x \geq 0$. Find the $P(\text{failure time}\, t \leq 2 \mid \text{ failure time}\, t > 1)$.

SOLUTION

From the definition of conditional probability the probability that the time to failure is less than or equal 2 for the device given that the failure time is greater than 1 is

$$P(\{\text{failure } t \leq 2\} \mid \{\text{failure } t > 1\}) = \frac{P(\{\text{failure } t \leq 2\} \cap \{\text{failure } t > 1\})}{P(\{\text{failure } t > 1\})}$$

Using set operations gives

$$\{\text{failure } t \leq 2\} \cap \{\text{failure } t > 1\} = \{\text{failure } 1 < t \leq 2\}$$

and the denominator is seen to be

$$\begin{aligned}
P(\{\text{failure } t > 1\}) &= 1 - P(\{\text{failure } t \leq 1\}) \\
&= 1 - (1 - e^{-1}) = e^{-1}
\end{aligned}$$

Substituting these two results into the first equation gives us the following result

$$P(\{\text{failure } t \le 2\} \,|\, \{\text{failure } t > 1\}) = \frac{P(\{\text{failure} 1 < t \le 2\})}{P(\{\text{failure } t > 1\})} = \frac{(1-e^{-2}) - (1-e^{-1})}{e^{-1}} = 1 - e^{-1} \quad \blacksquare$$

There are a number of special cases that deserve mentioning with regard to conditional probability. Usually A and C are overlapping but not necessarily included. The following special cases are for inclusion and nonoverlapping events:

$$\begin{aligned} &\text{If } A \subset C, \text{ then } A \cap C = A; \text{ thus } P(A \,|\, C) = \frac{P(A)}{P(C)} \\ &\text{If } A \supset C, \text{ then } A \cap C = C; \text{ thus } P(A \,|\, C) = \frac{P(C)}{P(C)} = 1 \\ &\text{If } A \cap C = \varnothing, \text{ then } P(A \,|\, C) = \frac{P(\varnothing)}{P(C)} = 0 \end{aligned} \tag{1.29}$$

If events A and C are independent, $P(A \cap C) = P(A)P(C)$, then conditional probabilities for the events can be written as

$$\begin{aligned} P(A \,|\, C) &= \frac{P(A \cap C)}{P(C)} = \frac{P(A)P(C)}{P(C)} = P(A) \\ P(C \,|\, A) &= \frac{P(C \cap A)}{P(A)} = \frac{P(C)P(A)}{P(A)} = P(C) \end{aligned} \tag{1.30}$$

Therefore, if events are independent, the conditional probabilities of the events are the same as the unconditional probabilities

1.3.1 Total Probability Theorem

Given n events $A_1, A_2, \ldots, A_n$ such that

$$\begin{aligned} &A_i \cap A_j = \varnothing \quad \text{for all } i \ne j \quad \text{(mutually exclusive)} \\ &\bigcup_{i=1}^{n} A_i = S \quad \text{(exhaustive)} \end{aligned} \tag{1.31}$$

Then it can be shown that the total probability of an arbitrary event B can be written in terms of the following conditional probabilities as

$$\begin{aligned} P(B) &= P(B \,|\, A_1)P(A_1) + P(B \,|\, A_2)P(A_2) + \cdots + P(B \,|\, A_n)P(A_n) \\ &= \sum_{i=1}^{n} P(B \,|\, A_i)P(A_i) \end{aligned} \tag{1.32}$$

1.3.2 Bayes's Theorem

A very important theorem that has many applications is Bayes's theorem which involves the determination of conditional probabilities $P(A_k \mid B)$ under the same framework as above. It is easily shown that

$$\begin{aligned} P(A_k \mid B) &= \frac{P(B \mid A_k)P(A_k)}{P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \cdots + P(B \mid A_n)P(A_n)} \\ &= \frac{P(B \mid A_k)P(A_k)}{\sum_{i=1}^{n} P(B \mid A_i)P(A_i)} \end{aligned} \tag{1.33}$$

An example is now presented that is representative of the type of problem that can be solved using Bayes's theorem and the total probability theorem.

EXAMPLE 1.14

Consider an experiment involving a random selection of one of three boxes. The random selection is of a single ball from the box chosen. The boxes contain red, white, and blue balls with specified probabilities of selection. Assume that

$$\begin{array}{rrr} P(\text{box1}) = 0.5 & P(\text{box2}) = 0.3 & P(\text{box3}) = 0.2 \\ P(\text{red} \mid \text{box1}) = 0.4 & P(\text{red} \mid \text{box2}) = 0.5 & P(\text{red} \mid \text{box3}) = 0.1 \\ P(\text{white} \mid \text{box1}) = 0.3 & P(\text{white} \mid \text{box2}) = 0.3 & P(\text{white} \mid \text{box3}) = 0.2 \\ P(\text{blue} \mid \text{box1}) = 0.3 & P(\text{blue} \mid \text{box2}) = 0.2 & P(\text{blue} \mid \text{box3}) = 0.7 \end{array}$$

(a) Find the probability of getting a red ball.
(b) Find the conditional probability, $P(A \mid B)$, where B is the event box 2 selected and A is the event a red ball is selected.

SOLUTION

(a) The total probability theorem can be used to obtain the probability of getting a red ball as follows:

$$\begin{aligned} P(\text{red}) &= P(\text{red} \mid \text{box1})P(\text{box1}) + P(\text{red} \mid \text{box2})P(\text{box2}) + P(\text{red} \mid \text{box3})P(\text{box3}) \\ &= 0.4 \cdot 0.5 + 0.3 \cdot 0.5 + 0.1 \cdot 0.2 = 0.37 \end{aligned}$$

(b) The probability of box 2 given that the ball is red can be determined by using Bayes's theorem and the results of part (a) as

$$\begin{aligned} P(\text{box } 2 \mid \text{red}) &= \frac{P(\text{red} \mid \text{box2})P(\text{box2})}{p(\text{red})} \\ &= \frac{0.5 \cdot 0.3}{0.37} = 0.4054 \end{aligned}$$

■

1.4 RANDOM POINTS

The random placement of points in an interval is an important problem. Conceptually it is analogous to random arrival times used in basic inventory problems, and it is used as a basis for shot noise in communication theory. The random placement of the points can be uniformly or nonuniformly distributed on an interval as seen in the following sections.

1.4.1 Uniform Random Points in an Interval

Define an experiment $\mathscr{E} : (S, \mathscr{F}, \mathscr{P}(.))$ as the random placement of a point t somewhere in the closed interval $[0, T]$ as shown in Figure 1.3(a). The sample space is $S = \{t : 0 \le t \le T\}$, $\mathscr{F}$ is the smallest field containing the sets $\{t : t \le t_1\}$ for all $t_1 \in S$, and $\mathscr{P}(.)$ is defined by $P\{t : t \le t_1\} = t_1/T$, for all $t_1 \in S$. Thus it can be seen that the probability that the point selected will be in any given interval is the ratio of that interval's length to the total length T. Probabilities of other events that are unions of nonoverlapping intervals can be obtained by adding up the probabilities for each of the intervals.

The purpose of this section is to talk about the random placement of n points, not just one point, in an interval $[0, T]$ see Figure 1.3(b). A convenient way to design such an experiment is to form a new compound experiment $\mathscr{E}_n : (S_n, \mathscr{F}_n, \mathscr{P}_n(.))$ composed of ordered n-tuples of times, $(t_1, t_2, \ldots, t_n)$, obtained from repeating the experiment $\mathscr{E}$ defined above independently. Assume an independence of the trials so that the probability measure can be described as the product

$$P\{t_1 \le x_1, t_2 \le x_2, \ldots, t_n \le x_n\} = P\{t_1 \le x_1\}P\{t_2 \le x_1\} \ldots P\{t_n \le x_n\} \qquad (1.34)$$

The Borel field is defined to be the smallest field containing the events $\{t_1 \le x_1, t_2 \le x_2, \ldots, t_n \le x_n\}$ for all $x_1, x_2, \ldots, x_n \in [0, T]$. We are interested in answering questions relating to calculating the probabilities that a certain number of points fall in a given interval or intervals.

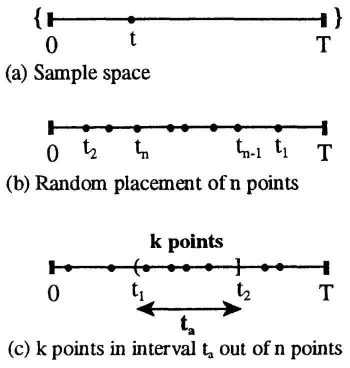

Figure 1.3 Random times in interval [0, T].

Say that the probability that exactly k of the points fall in a given interval $(t_1, t_2]$ of length t_a as shown in Figure 1.3(c) is desired. Let A be the event that on a single trial the point selected at random falls in the given interval. Using the uniform probability measure $\mathscr{P}(.)$, we give the probability of A by $P(A) = (t_2 - t_1) = t_a$. Thus, on n independent trials, the probability of getting k points in the interval is binomially distributed as

$$P(k \text{ points in } (t_1, t_2) \text{ for } n \text{ trials}) = \binom{n}{k} p^k (1-p)^{n-k}$$
$$\text{where } p = \frac{t_a}{T} \tag{1.35}$$

Further assume that n, the number of trials, is very large, $n \gg 1$, and that the relative width of the interval is very small, $t_a/T \ll 1$, and k is of the order nt_a/T. The resulting Poisson approximation is written as

$$P\begin{pmatrix} \text{exactly } k \text{ points in } (t_1, t_2) \\ \text{out of } n \text{ trials} \end{pmatrix} \approx \exp\left(\frac{-nt_a}{T}\right) \frac{(nt_a/T)^k}{k!} \tag{1.36}$$
$$\text{where } t_a = t_2 - t_1$$

In the limiting case this result will give an interpretation in terms of an average number of points per unit interval. If $n \to \infty$, $T \to \infty$, and $n/T \to \lambda$, then it can be shown that

$$\lim_{t_a \to 0} \frac{P(\text{exactly 1point in } t_a)}{t_a} = \lambda \tag{1.37}$$

Another probability of interest is that of getting exactly k_a points in interval t_a and exactly k_b in interval t_b as indicated in Figure 1.4.

Let A, B, C equal the events that on a single trial of the experiment exactly one point falls in t_a, t_b, and not in t_a or t_b, respectively. Then the probability of getting exactly k_a points in interval t_a, exactly k_b points in interval t_b, and exactly $n - k_a - k_b$ points not in t_a or t_b out of n trials can be obtained from the multinomial result as

$$P\begin{pmatrix} k_a \text{ in } t_a \\ k_b \text{ in } t_b \\ n - k_a - k_b \\ \notin t_a \text{ or } t_b \end{pmatrix} = \frac{n!}{k_a! k_b! (n - k_a - k_b)!} \left(\frac{t_a}{T}\right)^{k_a} \left(\frac{t_b}{T}\right)^{k_b} \left(1 - \frac{t_a}{T} - \frac{t_b}{T}\right)^{n-k_a-k_b} \tag{1.38}$$

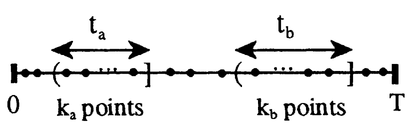

Figure 1.4 Exactly k_a points in interval t_a and exactly k_b in interval t_b.

The individual events exactly k_a points in t_a and exactly k_b points in t_b out of n trials have probabilities as follows:

$$P(k_a \text{ in } t_a) = \binom{n}{k_a}\left(\frac{t_a}{T}\right)^{k_a}\left(1-\frac{t_a}{T}\right)^{n-k_a}$$
$$P(k_a \text{ in } t_a) = \binom{n}{k_a}\left(\frac{t_b}{T}\right)^{k_b}\left(1-\frac{t_b}{T}\right)^{n-k_b} \tag{1.39}$$

Since the product of the above two probabilities does not equal that of Eq. (1.38), the events exactly k_a points in t_a and exactly k_b points in t_b out of n trials are not independent events.

1.4.2 Nonuniform Random Points in an Interval

In certain problems the random points are not placed uniformly in the interval. A common way to describe a nonuniform rate is to assign the probability measure by using a weighting function $\alpha(t)$ that satisfies the following properties:

$$\alpha(t) \geq 0 \quad \text{for all } t \in [0, T]$$
$$\int_0^T \alpha(t)\,dt = 1 \tag{1.40}$$

On a single trial of the experiment, the random placement of a single point in the interval $[0, T]$, the probability that the point selected is in $\{t_1, t_2\}$ is given by

$$P\{t \in (t_1, t_2]\} = \int_{t_1}^{t_2} \alpha(t)\,dt \tag{1.41}$$

Thus, if $\alpha(t)$ has a peak at $t = t_1$, it means that the point selected at random has a higher probability of being close to t_1 than other values of t. If n independent trials of this experiment are performed, the probability of exactly k points out of n trials being in the interval $\{t_1, t_2\}$ can be determined from Eq. (1.6) as

$$P\begin{pmatrix} k \text{ points in } (t_1, t_2) \\ \text{out of } n \text{ trials} \end{pmatrix} = \binom{n}{k} p^k (1-p)^{n-k}$$
$$\text{where } p = P\{t \in (t_1, t_2]\} = \int_{t_1}^{t_2} \alpha(t)\,dt \tag{1.42}$$

For the case of a nonuniform rate with the assumption that n, the number of trials, is very large, $n \gg 1$, and that the relative width of the interval $t_a = (t_1, t_2)$ is very small, $t_a/T \ll 1$, and k of the order nt_a/T, the Poisson approximation results in the following probability where p is given as above:

$$P\begin{pmatrix} \text{exactly } k \text{ points in } (t_1, t_2) \\ \text{out of } n \text{ trials} \end{pmatrix} \approx \exp(-np)\frac{(np)^k}{k!} \tag{1.43}$$

1.5 SUMMARY

In this chapter the mathematical definitions of an experiment in terms of a sample space, a Borel field, and a probability measure were given. The concept of an experiment is basic to the understanding of random variables and random processes to be discussed in later chapters. The assignment of the probability measure for several experiments obtained by combining other experiments led to the binomial, multinormal, and hypergeometric distributions.

The important concept of conditional probability was defined, and several examples were presented to illustrate the relationship to the sample space and the probability measure. Using this definition a discussion of the total probability theorem and the Bayes theorem for obtaining the a posteriori probability followed. These concepts have a fundamental role in the detection and estimation of random variables and random processes as will be seen in Chapters 2, 8, and 9.

The chapter concluded with a short discussion on the random placement of points in a given interval. These experiments are important in analyzing problems involving random times of arrival and other related problems.

It was not the intent of this chapter to give an exhaustive presentation on experiments and their use but to provide background material that would be used in the remainder of the text. For those wishing a more thorough presentation there are many excellent texts, including Papoulis [1], Freund [3], Drake [4], and Parzen [7].

PROBLEMS

1.1 A regular 6 sided die is tossed twice. **(a)** What is the probability that the sum of the tosses is 6? **(b)** What is the probability that the die is a 1 on the first throw given that the sum is 6? **(c)** What assumptions did you make in obtaining your answers?

1.2 For a standard deck of 52 playing cards, what is

(a) the probability of being dealt at least 3 aces in a 5 card poker hand?

(b) the probability that a card drawn at random is a red card or an ace?

1.3 A small deck of cards contains four kings and four queens. Two cards are drawn without replacement. For this problem the suit is of no concern. Define an experiment, assuming that at the first draw each card is equally likely and that the second card drawn is equally likely from the remaining cards.

1.4 If the events A and B are mutually exclusive events and if A^c and B^c represent the compliments of A and B, respectively, prove or disprove that

(a) A^c and B are mutually exclusive events.

(b) A^c and B^c are mutually exclusive events.

(c) A and B are independent events.

(d) A^c and S are mutually exclusive events.

(e) A and $\varnothing$, the null event, are mutually exclusive events.

(f) A^c and B^c are independent events.

(g) A and B^c are independent events.

(h) A and S (the certain event) are independent events.

(i) A^c and $\varnothing$, the null event, are independent events.

1.5 A fisherman caught 10 fish, 3 of which were under six inches. Two of them are selected at random and measured. What is the probability that both fish selected are under six inches?

1.6 For three tosses of a fair coin, determine the probability of the following sequences:

(a) HHH.

(b) HTH.

(c) Two heads and one tail in any order.

(d) More heads than tails.

(e) More heads than tails given at least one tail.

(f) More heads than tails given less than two heads.

1.7 We have a box full of 50 balls. A trial of an experiment is the selection of a ball at random from a box. There are 5 balls of each of 10 different colors—call them $C_1, C_2, \ldots, C_{10}$.

(a) What is the probability of getting exactly 3 balls of color C_1 in 10 trials of the experiment *with* replacement?

(b) What is the probability of getting exactly 3 balls of color C_1 in 10 trials of the experiment *without* replacement?

(c) What is the probability of getting exactly 3 balls of C_1, 2 balls of C_2, and 3 balls of C_4 in 10 trials of the experiment with replacement.

1.8 We have a "crooked" coin, call it c, with the probability of getting a head being $\frac{3}{4}$ and a "fair" coin, call it f with probability of getting a head of $\frac{1}{2}$. An experiment is defined as selecting one of these coins at random, with probability of selecting the fair coin of 0.6, and flipping the coin selected. A single trial results in a head (h) or a tail (t).

(a) Determine the probability of getting a head on a single trial of the experiment.

(b) Given that on a single trial a head has occurred, determine the probability that the crooked coin was selected.

(c) Given two trials and that both outcomes were heads determine the probability that the crooked coin was selected.

1.9 Consider an urn with 3 red, 2 white, and 1 blue ball. An experiment consists in the drawing of an ordered pair of balls without replacement.

(a) Give the sample space for this experiment.

(b) What are the probabilities of all the elementary events. Explain any assumptions you made in assigning these probabilities.

(c) Calculate the probability of the event that a blue ball is drawn in a single trial of the experiment above.

1.10 Consider an urn with 3 red, 2 white, and 1 blue ball. An experiment is the drawing of an ordered triple of balls without replacement.

(a) Give the sample space for this experiment.

(b) What are the probabilities of all the elementary events? Explain any assumptions you made in assigning these probabilities.

(c) Calculate the probability of the event that none of the three balls drawn is blue in a single trial of the experiment above.

1.11 An urn contains two white balls and six red ones.

(a) What is the probability that one ball drawn at random is white?

(b) What assumptions did you make in arriving at your answer?

1.12 The probability that a rifleman hits the bull's-eye is 0.3, the probability that he hits the target but not the bull's-eye is 0.6 and the probability that he misses the target all together is 0.1. Determine the probability that in five shots at the target he will hit the bull's eye twice, the target but not the bull's eye twice and miss the target once.

1.13 In playing an opponent of equal ability, which is more probable: to win exactly three games out of four or to win exactly five games out of eight? Substantiate your answer.

1.14 A weighted coin is tossed ten times. On each trial the probability of a head is 0.8 and the probability a tail is 0.2. Determine the probabilities of the following events. **(a)** Exactly 5 heads. **(b)** At least 5 heads. **(c)** More than 5 heads. **(d)** Number of heads is greater than or equal to 5 and less than or equal to 7.

1.15 Given that a team is leading the world series 3 games to 2. What is the probability they will win the series if each team has an equal probability of winning at each game? (Remember the first team to win 4 games wins the series.)

1.16 We have 10 people each on a bus to Las Cruces which has four stops. Each person, acting independently from the others, has a probability of $\frac{1}{4}$ of getting off at each stop.

(a) What is the probability that on one trip to Las Cruces that exactly three people get off on stop 1, exactly two people get off at stop 2, exactly two people get off at stop 3, and exactly three people get off at stop 4?

(b) What is the probability that none get off at stop 1?

(c) What is the probability that less than 8 get off at stop 1?

1.17 The probability of getting a head on a single flip of a coin is p. The coin is tossed 10 times independently.

(a) Define the sample space for this experiment.

(b) Describe the probability assignment for the experiment.

(c) Give the event that an odd number of heads occur.

(d) Determine the probability of the event described in (c).

1.18 Given an honest standard deck of 52 playing cards a dealer draws one card at random from the deck and records it. He then slips the card back into the deck, shuffles, and is ready to draw another card.

(a) What is the probability that in 10 independent draws of the specified procedure above he will draw two or more aces?

(b) If the card was not replaced after each draw but kept out of the deck, what is the probability that of two cards drawn both are aces?

1.19 Say we have a hat with 10 letters, 3 a's, 5 b's, and 2 c's and a letter is drawn at random from the hat recorded and replaced. This experiment is repeated independently a number of times.

(a) What is the probability that on 10 draws that we get exactly k a's? Evaluate for $k = 3$.

(b) What is the probability that on 10 draws we get exactly 2 a's, 6 b's, and 2 c's?

(c) If we draw 10 times, what is the probability that we get more than 2 a's?

(d) If we draw 10 times, what is the probability that we get at least 2 a's?

(e) Determine the probability that we have 3 or 4 a's given that we have at least two a's.

(f) Suppose that on the first 10 trials we get 5 a's, 2 b's, and 3 c's. What is the probability that on the 11th trial that we get an a?

(g) If the letters are not replaced, find the probabilities for the following sequences: (a, a, a), (a, b, c), (c, c, c), and (a, a, b).

1.20 Given experiments $\mathscr{E}_1$ and $\mathscr{E}_2$ which are defined as follows:

$$\mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.)), S_1 = \{1, 2\}, \mathscr{F}_1 = \text{power set of } S_1,$$
$$\mathscr{P}_1(.) : P\{1\} = 2/3, P\{2\} = 1/3$$
$$\mathscr{E}_2 : (S_2, \mathscr{F}_2, \mathscr{P}_2(.)), S_2 = \{a, b\}, \mathscr{F}_2 = \text{power set of } S_2,$$
$$\mathscr{P}_2(.) : P\{a\} = 1/4, P\{b\} = 3/4$$

(a) Give all elements of $\mathscr{F}_2$.

(b) Considering $\mathscr{E}_1$ and $\mathscr{E}_2$ to be independent experiments, define a new experiment $\mathscr{E} = \mathscr{E}_1 \otimes \mathscr{E}_2$, the cartesian product, and specify its sample space, S, Borel field, $\mathscr{F}$, and probability measure, $\mathscr{P}(.)$. Be specific.

(c) Specify, for the combined experiment $\mathscr{E}$, the event that a 1 is obtained, and find the probability for this event.

1.21 Let $\mathscr{E}_1 : (S_1, \mathscr{F}_1, \mathscr{P}_1(.))$ and $\mathscr{E}_2 : (S_2, \mathscr{F}_2, \mathscr{P}_2(.))$ be independent experiments with

$S_1 = \{a, b, c\}$, $\mathscr{F}_1$ all subsets of S_1, and

$$\mathscr{P}_1 : (P\{a\} = 1/2, P\{b\} = 1/4, P\{c\} = 1/4)$$

$S_2 = \{1, 2, 3\}$, $\mathscr{F}_2$ all subsets of S_2, and

$$\mathscr{P}_1 : (P\{1\} = 1/3, P\{2\} = 1/3, P\{3\} = 1/3)$$

(a) List the elements of the cartesian product space $S = S_1 \otimes S_2$.

(b) What is the new $\mathscr{P}(.)$ for this combined experiment?

(c) For this new sample space, list the elements of the event that the number drawn is odd.

1.22 Say we have a big bag of red, blue, black, and white marbles. It is known that the probabilities of taking a red, blue, black, or white marble from the bag at random are, respectively, 0.3, 0.4, 0.2, and 0.1. Say we draw five marbles (with replacement) from the bag and the results at each trial are independent.

(a) Define the sample space for this new combined experiment.

(b) How many elements are there in this sample space?

(c) Specify the event R that at least one of the marbles is red.

(d) Specify the event Q that exactly two of the marbles are red.

(e) Give any elementary event for the new experiment and find its probability.

(f) Determine the probability that of the five marbles drawn exactly two of them are white and exactly one of them is red.

(g) What is the probability that the number of red marbles is greater than 1? Less than 2?

1.23 Evaluate the following integrals:

(a) $$\int_{-10}^{20} \frac{1}{\sqrt{2\pi}10} \exp\left\{-\frac{(x-10)^2}{200}\right\} dx$$

(b) $$\int_{-425}^{525} \frac{1}{\sqrt{2\pi}100} \exp\left\{-\frac{(x-500)^2}{20{,}000}\right\} dx$$

1.24 A fair coin, $P(\text{head}) = 0.5$, is tossed five times. Determine the probability that the number of heads occuring is greater than 2. Less than 2. Exactly equal to 2. Either 1 or 2.

1.25 An honest die is rolled 1000 times.

(a) Determine the probability that the number of 3's is between 120 and 150.

(b) Find the probability that the number of 3's is less than 20.

1.26 Consider the number of 3's that result from 600 tosses of a fair six-sided die. Find the probability that the number of 3's is between 60 and 140 inclusive.

1.27 An experiment of tossing a coin has the following outcomes: h, t, and e with given probabilities $P\{h\} = 0.5, P\{t\} = 0.495$, and $P\{e\} = 0.005$ (e means landed on edge).

(a) Give the exact expression for finding the probability that in 1000 independent trials that the number of heads is between and including 485 and 500.

(b) Give an approximate answer for the results in (a), if possible, and justify your approximation.

(c) In 1000 independent trials, what is the exact expression for the probability that the coin lands on edge less than two times.

(d) Give an approximate answer for your results in (c), if possible, and justify your approximation if you make one.

(e) Determine the probability that on 5 trials of the experiment we get exactly two h's, two t's and one e.

1.28 In a binary communication channel, the probability of making an error in the transmission of a single binary symbol is p. Assume that digits are sent along the channel independently.

(a) It is desired that the probability of 2 or fewer errors in the transmission of 100 symbols is 0.99. What is the largest value of p that satisfies this condition?

(b) Let $p = 0.001$. If less than 2 errors occur in 100 trials, determine the probability that there were no errors.

(c) If $p = 1$, "a worthless channel," determine the probability that the number of errors N out of 1000 trials is $400 \leq N \leq 600$.

1.29 In a binary communication channel the probability of making an error on a given transmission of a binary symbol is 0.01.

(a) In sending of 10,000 independent symbols, determine the probability that the number of errors is between and including 95 to 100 errors.

(b) What is the probability that the number of errors is 2 or less in the transmission of 100 symbols?

1.30 A player sits on the origin of a graph and flips a crooked coin where bias to heads is given by $P(H) = \frac{2}{3}$. If the result of the toss is a head, he moves one unit to the right, and if it is a tail, he moves one unit to the left. The coin is flipped again, and he moves from his new position in the same way. Repeating this procedure a total of 1200 times, what is the probability that the player ends up sitting on a number N where $350 \leq N \leq 450$?

1.31 If the probability that a person will catch a fish in 0.01 minutes is 0.02,

(a) Determine the probability that a person will not catch a single fish in one hour.

(b) Find the probability that in 2 hours he will catch at least 4 fish.

1.32 Suppose that there is a diagnostic test for measles. Define C as the event that a person tested has measles and A as the event that the test states that the person tested has the measles. Assume that $P(A \mid C) = 0.95$ and $P(A^c \mid C^c) = 0.95$. Say the probability that a person taking the test actually has measles is 0.005.

(a) Compute the probability that a person who tests positive actually does have measles.

(b) Compute the probability of A.

1.33 Consider a source that sends a 0 or a 1 with known probabilities

$$P(0 \text{ sent}) = p \text{ and } P(1 \text{ sent}) = 1 - p$$

The digit is sent over a noisy channel with the following transition probabilities:

$$P(0 \text{ received} \mid 1 \text{ sent}) = P(1 \text{ received} \mid 0 \text{ sent}) = 1/4$$
$$P(0 \text{ received} \mid 0 \text{ sent}) = P(1 \text{ received} \mid 1 \text{ sent}) = 3/4$$

(a) Find the probability that a 0 is received.

(b) What does p need to be for this probability to be $\frac{3}{8}$?

(c) What is the probability that a 1 was sent given that a 1 was received?

1.34 Given two boxes B_1 and B_2 where B_1 contains 8 red and 8 white balls and B_2 contains 5 red and 10 white balls. An experiment consists of selecting one of the boxes with $p(B_1) = 0.4$ and $P(B_2) = 0.6$, and then selecting a ball at random from the selected box.

(a) Find the probability that the ball drawn is red.

(b) If a red ball is chosen, determine the probability that it is from box B_1.

1.35 In a year's time if a tree is not watered, it will die with a probability of $\frac{1}{2}$, and if watered, it will die with probability of $\frac{1}{4}$. After a year, the tree is dead. Determine the probability that the tree was watered.

1.36 A crooked coin ($P\{\text{head}\} = 0.25$) is flipped twice with experimental outcome that at least one toss resulted in a head. Given this partial information, what is the conditional probability that both tosses resulted in heads.

1.37 From a master handwritten manuscript composed of the letters $a, b,$ and c (the author has a limited vocabulary) with $P(a) = 0.5, P(b) = 0.2,$ and $P(c) = 0.3$. A typist copies the manuscript making random errors when typing the letters $a, b,$ and c with probabilities as follows $P(b|a) = 0.05$, $P(c|a) = 0.1$, $P(a|b) = 0.8$, $P(c|b) = 0.7$, $P(a|c) = 0.2$, $P(b|c) = 0.2$.

(a) What is the probability that a letter a typed was really a letter b in the original document?

(b) Given that a letter c was typed, what is the probability that there was a letter c in that position in the original manuscript?

REFERENCES

1. Papoulis, Athanasios, *Probability, Random Variables, and Stochastic Processes*, McGraw Hill, 1965.
2. Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes, *Introduction to the Theory of Statistics*, 3rd ed. McGraw-Hill, 1974.
3. Freund, John E., *Mathematical Statistics*, Prentice-Hall, 1964.
4. Drake, Alvin W., *Fundamentals of Applied Probability Theory*, McGraw-Hill, 1967.
5. Davenport, Jr., William B., *Probability and Random Processes*, 1970.
6. Stark, Henry and John W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*, Prentice-Hall, 1986.
7. Parzen, Emanuel, *Modern Probability Theory and Its Applications*, Wiley, 1960.