PART I

Probability

This page intentionally left blank

CHAPTER 1

Basic Probability

Random Experiments

We are all familiar with the importance of experiments in science and engineering. Experimentation is useful to us because we can assume that if we perform certain experiments under very nearly identical conditions, we will arrive at results that are essentially the same. In these circumstances, we are able to control the value of the variables that affect the outcome of the experiment.

However, in some experiments, we are not able to ascertain or control the value of certain variables so that the results will vary from one performance of the experiment to the next even though most of the conditions are the same. These experiments are described as *random*. The following are some examples.

EXAMPLE 1.1 If we toss a coin, the result of the experiment is that it will either come up "tails," symbolized by T (or 0), or "heads," symbolized by H (or 1), i.e., one of the elements of the set $\{H, T\}$ (or $\{0, 1\}$).

EXAMPLE 1.2 If we toss a die, the result of the experiment is that it will come up with one of the numbers in the set $\{1, 2, 3, 4, 5, 6\}$.

EXAMPLE 1.3 If we toss a coin twice, there are four results possible, as indicated by $\{HH, HT, TH, TT\}$, i.e., both heads, heads on first and tails on second, etc.

EXAMPLE 1.4 If we are making bolts with a machine, the result of the experiment is that some may be defective. Thus when a bolt is made, it will be a member of the set {defective, nondefective}.

EXAMPLE 1.5 If an experiment consists of measuring "lifetimes" of electric light bulbs produced by a company, then the result of the experiment is a time t in hours that lies in some interval—say, $0 \leq t \leq 4000$—where we assume that no bulb lasts more than 4000 hours.

Sample Spaces

A set S that consists of all possible outcomes of a random experiment is called a *sample space*, and each outcome is called a *sample point*. Often there will be more than one sample space that can describe outcomes of an experiment, but there is usually only one that will provide the most information.

EXAMPLE 1.6 If we toss a die, one sample space, or set of all possible outcomes, is given by $\{1, 2, 3, 4, 5, 6\}$ while another is {odd, even}. It is clear, however, that the latter would not be adequate to determine, for example, whether an outcome is divisible by 3.

It is often useful to portray a sample space graphically. In such cases it is desirable to use numbers in place of letters whenever possible.

EXAMPLE 1.7 If we toss a coin twice and use 0 to represent tails and 1 to represent heads, the sample space (see Example 1.3) can be portrayed by points as in Fig. 1-1 where, for example, (0, 1) represents tails on first toss and heads on second toss, i.e., TH.

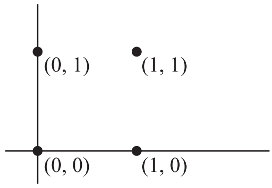


Fig. 1-1

If a sample space has a finite number of points, as in Example 1.7, it is called a *finite sample space*. If it has as many points as there are natural numbers 1, 2, 3, . . . , it is called a *countably infinite sample space*. If it has as many points as there are in some interval on the x axis, such as $0 \leq x \leq 1$, it is called a *noncountably infinite sample space*. A sample space that is finite or countably infinite is often called a *discrete sample space*, while one that is noncountably infinite is called a *nondiscrete sample space*.

Events

An *event* is a subset A of the sample space S, i.e., it is a set of possible outcomes. If the outcome of an experiment is an element of A, we say that the event A *has occurred.* An event consisting of a single point of S is often called a *simple* or *elementary event*.

EXAMPLE 1.8 If we toss a coin twice, the event that only one head comes up is the subset of the sample space that consists of points (0, 1) and (1, 0), as indicated in Fig. 1-2.

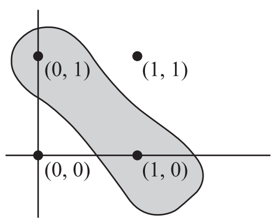


Fig. 1-2

As particular events, we have S itself, which is the *sure* or *certain event* since an element of S must occur, and the empty set $\varnothing$, which is called the *impossible event* because an element of $\varnothing$ cannot occur.

By using set operations on events in S, we can obtain other events in S. For example, if A and B are events, then

1. $A \cup B$ is the event "either A or B or both." $A \cup B$ is called the *union* of A and B.
2. $A \cap B$ is the event "both A and B." $A \cap B$ is called the *intersection* of A and B.
3. A' is the event "not A." A' is called the *complement* of A.
4. $A - B = A \cap B'$ is the event "A but not B." In particular, $A' = S - A$.

If the sets corresponding to events A and B are disjoint, i.e., $A \cap B = \varnothing$, we often say that the events are *mutually exclusive*. This means that they cannot both occur. We say that a collection of events $A_1, A_2, \ldots, A_n$ is mutually exclusive if every pair in the collection is mutually exclusive.

EXAMPLE 1.9 Referring to the experiment of tossing a coin twice, let A be the event "at least one head occurs" and B the event "the second toss results in a tail." Then $A = \{HT, TH, HH\}$, $B = \{HT, TT\}$, and so we have

$$A \cup B = \{HT, TH, HH, TT\} = S \qquad A \cap B = \{HT\}$$

$$A' = \{TT\} \qquad A - B = \{TH, HH\}$$

The Concept of Probability

In any random experiment there is always uncertainty as to whether a particular event will or will not occur. As a measure of the *chance*, or *probability*, with which we can expect the event to occur, it is convenient to assign a number between 0 and 1. If we are sure or certain that the event will occur, we say that its probability is 100% or 1, but if we are sure that the event will not occur, we say that its probability is zero. If, for example, the probability is $\frac{1}{4}$, we would say that there is a 25% chance it will occur and a 75% chance that it will not occur. Equivalently, we can say that the *odds* against its occurrence are 75% to 25%, or 3 to 1.

There are two important procedures by means of which we can estimate the probability of an event.

1. **CLASSICAL APPROACH.** If an event can occur in h different ways out of a total number of n possible ways, all of which are equally likely, then the probability of the event is h/n.

EXAMPLE 1.10 Suppose we want to know the probability that a head will turn up in a single toss of a coin. Since there are two equally likely ways in which the coin can come up—namely, heads and tails (assuming it does not roll away or stand on its edge)—and of these two ways a head can arise in only one way, we reason that the required probability is $1/2$. In arriving at this, we assume that the coin is *fair*, i.e., not *loaded* in any way.

2. **FREQUENCY APPROACH.** If after n repetitions of an experiment, where n is very large, an event is observed to occur in h of these, then the probability of the event is h/n. This is also called the *empirical probability* of the event.

EXAMPLE 1.11 If we toss a coin 1000 times and find that it comes up heads 532 times, we estimate the probability of a head coming up to be $532/1000 = 0.532$.

Both the classical and frequency approaches have serious drawbacks, the first because the words "equally likely" are vague and the second because the "large number" involved is vague. Because of these difficulties, mathematicians have been led to an *axiomatic approach* to probability.

The Axioms of Probability

Suppose we have a sample space S. If S is discrete, all subsets correspond to events and conversely, but if S is nondiscrete, only special subsets (called *measurable*) correspond to events. To each event A in the class C of events, we associate a real number $P(A)$. Then P is called a *probability function*, and $P(A)$ the *probability* of the event A, if the following axioms are satisfied.

Axiom 1 For every event A in the class C,

$$P(A) \geq 0 \tag{1}$$

Axiom 2 For the sure or certain event S in the class C,

$$P(S) = 1 \tag{2}$$

Axiom 3 For any number of mutually exclusive events $A_1, A_2, \ldots,$ in the class C,

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots \tag{3}$$

In particular, for two mutually exclusive events A_1, A_2,

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \tag{4}$$

Some Important Theorems on Probability

From the above axioms we can now prove various theorems on probability that are important in further work.

Theorem 1-1 If $A_1 \subset A_2$, then $P(A_1) \leq P(A_2)$ and $P(A_2 - A_1) = P(A_2) - P(A_1)$.

Theorem 1-2 For every event A,

$$0 \leq P(A) \leq 1, \tag{5}$$

i.e., a probability is between 0 and 1.

Theorem 1-3

$$P(\varnothing) = 0 \tag{6}$$

i.e., the impossible event has probability zero.

Theorem 1-4 If A' is the complement of A, then

$$P(A') = 1 - P(A) \tag{7}$$

Theorem 1-5 If $A = A_1 \cup A_2 \cup \cdots \cup A_n$, where $A_1, A_2, \ldots, A_n$ are mutually exclusive events, then

$$P(A) = P(A_1) + P(A_2) + \cdots + P(A_n) \tag{8}$$

In particular, if $A = S$, the sample space, then

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1 \tag{9}$$

Theorem 1-6 If A and B are any two events, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{10}$$

More generally, if A_1, A_2, A_3 are any three events, then

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3) = {} & P(A_1) + P(A_2) + P(A_3) \\ & - P(A_1 \cap A_2) - P(A_2 \cap A_3) - P(A_3 \cap A_1) \\ & + P(A_1 \cap A_2 \cap A_3) \end{aligned} \tag{11}$$

Generalizations to n events can also be made.

Theorem 1-7 For any events A and B,

$$P(A) = P(A \cap B) + P(A \cap B') \tag{12}$$

Theorem 1-8 If an event A must result in the occurrence of one of the mutually exclusive events $A_1, A_2, \ldots, A_n$, then

$$P(A) = P(A \cap A_1) + P(A \cap A_2) + \cdots + P(A \cap A_n) \tag{13}$$

Assignment of Probabilities

If a sample space S consists of a finite number of outcomes $a_1, a_2, \ldots, a_n$, then by Theorem 1-5,

$$P(A_1) + P(A_2) + \cdots + P(A_n) = 1 \tag{14}$$

where $A_1, A_2, \ldots, A_n$ are elementary events given by $A_i = \{a_i\}$.

It follows that we can arbitrarily choose any nonnegative numbers for the probabilities of these simple events as long as (14) is satisfied. In particular, if we assume *equal probabilities* for all simple events, then

$$P(A_k) = \frac{1}{n}, \quad k = 1, 2, \ldots, n \tag{15}$$

and if A is any event made up of h such simple events, we have

$$P(A) = \frac{h}{n} \tag{16}$$

This is equivalent to the classical approach to probability given on page 5. We could of course use other procedures for assigning probabilities, such as the frequency approach of page 5.

Assigning probabilities provides a *mathematical model*, the success of which must be tested by experiment in much the same manner that theories in physics or other sciences must be tested by experiment.

EXAMPLE 1.12 A single die is tossed once. Find the probability of a 2 or 5 turning up.

The sample space is $S = \{1, 2, 3, 4, 5, 6\}$. If we assign equal probabilities to the sample points, i.e., if we assume that the die is fair, then

$$P(1) = P(2) = \cdots = P(6) = \frac{1}{6}$$

The event that either 2 or 5 turns up is indicated by $2 \cup 5$. Therefore,

$$P(2 \cup 5) = P(2) + P(5) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

Conditional Probability

Let A and B be two events (Fig. 1-3) such that $P(A) > 0$. Denote by $P(B \mid A)$ the probability of B *given that* A has occurred. Since A is known to have occurred, it becomes the new sample space replacing the original S. From this we are led to the definition

$$P(B \mid A) \equiv \frac{P(A \cap B)}{P(A)} \tag{17}$$

or

$$P(A \cap B) \equiv P(A)\, P(B \mid A) \tag{18}$$

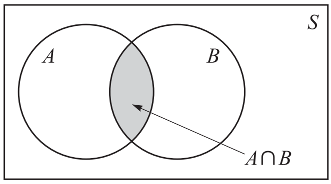


Fig. 1-3

In words, (18) says that the probability that both A and B occur is equal to the probability that A occurs times the probability that B occurs given that A has occurred. We call $P(B \mid A)$ the *conditional probability* of B given A, i.e., the probability that B will occur given that A has occurred. It is easy to show that conditional probability satisfies the axioms on page 5.

EXAMPLE 1.13 Find the probability that a single toss of a die will result in a number less than 4 if (a) no other information is given and (b) it is given that the toss resulted in an odd number.

(a) Let B denote the event {less than 4}. Since B is the union of the events 1, 2, or 3 turning up, we see by Theorem 1-5 that

$$P(B) = P(1) + P(2) + P(3) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$$

assuming equal probabilities for the sample points.

(b) Letting A be the event {odd number}, we see that $P(A) = \frac{3}{6} = \frac{1}{2}$. Also $P(A \cap B) = \frac{2}{6} = \frac{1}{3}$. Then

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{1/3}{1/2} = \frac{2}{3}$$

Hence, the added knowledge that the toss results in an odd number raises the probability from $1/2$ to $2/3$.

Theorems on Conditional Probability

Theorem 1-9 For any three events A_1, A_2, A_3, we have

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \tag{19}$$

In words, the probability that A_1 and A_2 and A_3 all occur is equal to the probability that A_1 occurs times the probability that A_2 occurs given that A_1 has occurred times the probability that A_3 occurs given that both A_1 and A_2 have occurred. The result is easily generalized to n events.

Theorem 1-10 If an event A must result in one of the mutually exclusive events $A_1, A_2, \ldots, A_n$, then

$$P(A) = P(A_1)\, P(A \mid A_1) + P(A_2)\, P(A \mid A_2) + \cdots + P(A_n)\, P(A \mid A_n) \tag{20}$$

Independent Events

If $P(B \mid A) = P(B)$, i.e., the probability of B occurring is not affected by the occurrence or non-occurrence of A, then we say that A and B are *independent events*. This is equivalent to

$$P(A \cap B) = P(A)P(B) \tag{21}$$

as seen from (18). Conversely, if (21) holds, then A and B are independent.

We say that three events A_1, A_2, A_3 are *independent* if they are pairwise independent:

$$P(A_j \cap A_k) = P(A_j)P(A_k) \qquad j \neq k \qquad \text{where} \quad j, k = 1, 2, 3 \tag{22}$$

and

$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3) \tag{23}$$

Note that neither (22) nor (23) is by itself sufficient. Independence of more than three events is easily defined.

Bayes' Theorem or Rule

Suppose that $A_1, A_2, \ldots, A_n$ are mutually exclusive events whose union is the sample space S, i.e., one of the events must occur. Then if A is any event, we have the following important theorem:

Theorem 1-11 **(Bayes' Rule):**

$$P(A_k \mid A) = \frac{P(A_k)\,P(A \mid A_k)}{\sum_{j=1}^{n} P(A_j)\,P(A \mid A_j)} \tag{24}$$

This enables us to find the probabilities of the various events $A_1, A_2, \ldots, A_n$ that can *cause* A to occur. For this reason Bayes' theorem is often referred to as a *theorem on the probability of causes*.

Combinatorial Analysis

In many cases the number of sample points in a sample space is not very large, and so direct enumeration or counting of sample points needed to obtain probabilities is not difficult. However, problems arise where direct counting becomes a practical impossibility. In such cases use is made of *combinatorial analysis*, which could also be called a *sophisticated way of counting*.

Fundamental Principle of Counting: Tree Diagrams

If one thing can be accomplished in n_1 different ways and after this a second thing can be accomplished in n_2 different ways, . . . , and finally a kth thing can be accomplished in n_k different ways, then all k things can be accomplished in the specified order in $n_1 n_2 \ldots n_k$ different ways.

EXAMPLE 1.14 If a man has 2 shirts and 4 ties, then he has $2 \cdot 4 = 8$ ways of choosing a shirt and then a tie.

A diagram, called a *tree diagram* because of its appearance (Fig. 1-4), is often used in connection with the above principle.

EXAMPLE 1.15 Letting the shirts be represented by S_1, S_2 and the ties by T_1, T_2, T_3, T_4, the various ways of choosing a shirt and then a tie are indicated in the tree diagram of Fig. 1-4.

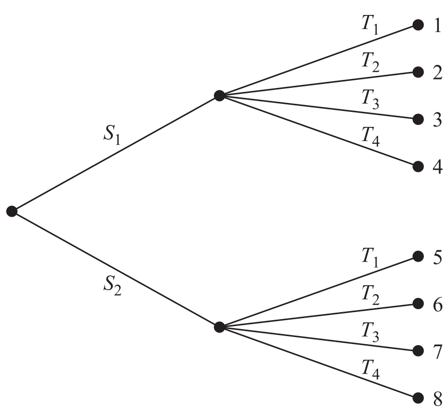


Fig. 1-4

Permutations

Suppose that we are given n distinct objects and wish to *arrange* r of these objects in a line. Since there are n ways of choosing the 1st object, and after this is done, $n - 1$ ways of choosing the 2nd object, . . . , and finally $n - r + 1$ ways of choosing the rth object, it follows by the fundamental principle of counting that the number of different *arrangements*, or *permutations* as they are often called, is given by

$$_nP_r = n(n-1)(n-2) \cdots (n-r+1) \tag{25}$$

where it is noted that the product has r factors. We call $_nP_r$ the *number of permutations of n objects taken r at a time*.

In the particular case where $r = n$, (25) becomes

$$_nP_n = n(n-1)(n-2) \cdots 1 = n! \tag{26}$$

which is called *n factorial*. We can write (25) in terms of factorials as

$$_nP_r = \frac{n!}{(n-r)!} \tag{27}$$

If $r = n$, we see that (27) and (26) agree only if we have $0! = 1$, and we shall actually take this as the definition of 0!.

EXAMPLE 1.16 The number of different arrangements, or permutations, consisting of 3 letters each that can be formed from the 7 letters A, B, C, D, E, F, G is

$$_7P_3 = \frac{7!}{4!} = 7 \cdot 6 \cdot 5 = 210$$

Suppose that a set consists of n objects of which n_1 are of one type (i.e., indistinguishable from each other), n_2 are of a second type, . . . , n_k are of a kth type. Here, of course, $n = n_1 + n_2 + \cdots + n_k$. Then the number of different permutations of the objects is

$$_nP_{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! n_2! \cdots n_k!} \tag{28}$$

See Problem 1.25.

EXAMPLE 1.17 The number of different permutations of the 11 letters of the word *M I S S I S S I P P I*, which consists of 1 *M*, 4 *I*'s, 4 *S*'s, and 2 *P*'s, is

$$\frac{11!}{1!4!4!2!} = 34{,}650$$

Combinations

In a permutation we are interested in the order of arrangement of the objects. For example, *abc* is a different permutation from *bca*. In many problems, however, we are interested only in selecting or choosing objects without regard to order. Such selections are called *combinations*. For example, *abc* and *bca* are the same combination.

The total number of combinations of r objects selected from n (also called the *combinations of n things taken r at a time*) is denoted by $_nC_r$ or $\binom{n}{r}$. We have (see Problem 1.27)

$$\binom{n}{r} = {}_nC_r = \frac{n!}{r!(n-r)!} \tag{29}$$

It can also be written

$$\binom{n}{r} = \frac{n(n-1) \cdots (n-r+1)}{r!} = \frac{_nP_r}{r!} \tag{30}$$

It is easy to show that

$$\binom{n}{r} = \binom{n}{n-r} \qquad \text{or} \qquad {}_nC_r = {}_nC_{n-r} \tag{31}$$

EXAMPLE 1.18 The number of ways in which 3 cards can be chosen or selected from a total of 8 different cards is

$$_8C_3 = \binom{8}{3} = \frac{8 \cdot 7 \cdot 6}{3!} = 56$$

Binomial Coefficient

The numbers (29) are often called *binomial coefficients* because they arise in the *binomial expansion*

$$(x + y)^n = x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \cdots + \binom{n}{n}y^n \tag{32}$$

They have many interesting properties.

EXAMPLE 1.19

$$\begin{aligned}(x + y)^4 &= x^4 + \binom{4}{1}x^3y + \binom{4}{2}x^2y^2 + \binom{4}{3}xy^3 + \binom{4}{4}y^4 \\ &= x^4 + 4x^3y + 6x^2y^2 + 4xy^3 + y^4\end{aligned}$$

Stirling's Approximation to *n*!

When n is large, a direct evaluation of $n!$ may be impractical. In such cases use can be made of the approximate formula

$$n! \sim \sqrt{2\pi n}\, n^n e^{-n} \tag{33}$$

where $e = 2.71828\ldots$, which is the base of natural logarithms. The symbol $\sim$ in (33) means that the ratio of the left side to the right side approaches 1 as $n \to \infty$.

Computing technology has largely eclipsed the value of Stirling's formula for numerical computations, but the approximation remains valuable for theoretical estimates (see Appendix A).

SOLVED PROBLEMS

Random experiments, sample spaces, and events

1.1. A card is drawn at random from an ordinary deck of 52 playing cards. Describe the sample space if consideration of suits (a) is not, (b) is, taken into account.

(a) If we do not take into account the suits, the sample space consists of ace, two, . . . , ten, jack, queen, king, and it can be indicated as $\{1, 2, \ldots, 13\}$.

(b) If we do take into account the suits, the sample space consists of ace of hearts, spades, diamonds, and clubs; . . . ; king of hearts, spades, diamonds, and clubs. Denoting hearts, spades, diamonds, and clubs, respectively, by 1, 2, 3, 4, for example, we can indicate a jack of spades by (11, 2). The sample space then consists of the 52 points shown in Fig. 1-5.

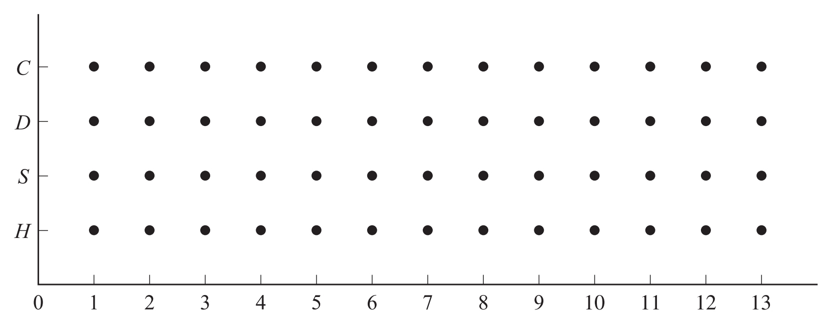


Fig. 1-5

1.2. Referring to the experiment of Problem 1.1, let A be the event {king is drawn} or simply {king} and B the event {club is drawn} or simply {club}. Describe the events (a) $A \cup B$, (b) $A \cap B$, (c) $A \cup B'$, (d) $A' \cup B'$, (e) $A - B$, (f) $A' - B'$, (g) $(A \cap B) \cup (A \cap B')$.

(a) $A \cup B$ = {either king or club (or both, i.e., king of clubs)}.

(b) $A \cap B$ = {both king and club} = {king of clubs}.

(c) Since B = {club}, B' = {not club} = {heart, diamond, spade}.

Then $A \cup B'$ = {king or heart or diamond or spade}.

(d) $A' \cup B'$ = {not king or not club} = {not king of clubs} = {any card but king of clubs}.

This can also be seen by noting that $A' \cup B' = (A \cap B)'$ and using (b).

(e) $A - B$ = {king but not club}.

This is the same as $A \cap B'$ = {king and not club}.

(f) $A' - B'$ = {not king and not "not club"} = {not king and club} = {any club except king}.

This can also be seen by noting that $A' - B' = A' \cap (B')' = A' \cap B$.

(g) $(A \cap B) \cup (A \cap B')$ = {(king and club) or (king and not club)} = {king}.

This can also be seen by noting that $(A \cap B) \cup (A \cap B') = A$.

1.3. Use Fig. 1-5 to describe the events (a) $A \cup B$, (b) $A' \cap B'$.

The required events are indicated in Fig. 1-6. In a similar manner, all the events of Problem 1.2 can also be indicated by such diagrams. It should be observed from Fig. 1-6 that $A' \cap B'$ is the complement of $A \cup B$.

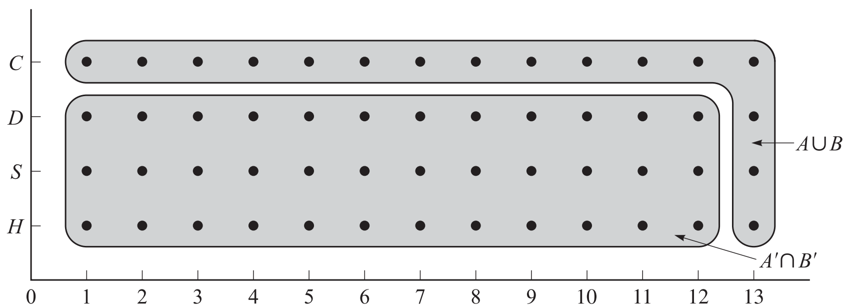


Fig. 1-6

Theorems on probability

1.4. Prove (a) Theorem 1-1, (b) Theorem 1-2, (c) Theorem 1-3, page 5.

(a) We have $A_2 = A_1 \cup (A_2 - A_1)$ where A_1 and $A_2 - A_1$ are mutually exclusive. Then by Axiom 3, page 5:

$$P(A_2) = P(A_1) + P(A_2 - A_1)$$

so that
$$P(A_2 - A_1) = P(A_2) - P(A_1)$$

Since $P(A_2 - A_1) \geq 0$ by Axiom 1, page 5, it also follows that $P(A_2) \geq P(A_1)$.

(b) We already know that $P(A) \geq 0$ by Axiom 1. To prove that $P(A) \leq 1$, we first note that $A \subset S$. Therefore, by Theorem 1-1 [part (a)] and Axiom 2,

$$P(A) \leq P(S) = 1$$

(c) We have $S = S \cup \varnothing$. Since $S \cap \varnothing = \varnothing$, it follows from Axiom 3 that

$$P(S) = P(S) + P(\varnothing) \qquad \text{or} \qquad P(\varnothing) = 0$$

1.5. Prove (a) Theorem 1-4, (b) Theorem 1-6.

(a) We have $A \cup A' = S$. Then since $A \cap A' = \varnothing$, we have

$$P(A \cup A') = P(S) \qquad \text{or} \qquad P(A) + P(A') = 1$$

i.e.,

$$P(A') = 1 - P(A)$$

(b) We have from the Venn diagram of Fig. 1-7,

$$(1) \qquad A \cup B = A \cup [B - (A \cap B)]$$

Then since the sets A and $B - (A \cap B)$ are mutually exclusive, we have, using Axiom 3 and Theorem 1-1,

$$\begin{aligned} P(A \cup B) &= P(A) + P[B - (A \cap B)] \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

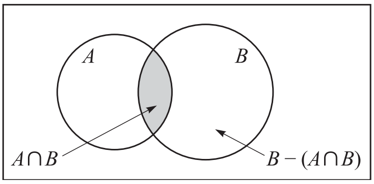


Fig. 1-7

Calculation of probabilities

1.6. A card is drawn at random from an ordinary deck of 52 playing cards. Find the probability that it is (a) an ace, (b) a jack of hearts, (c) a three of clubs or a six of diamonds, (d) a heart, (e) any suit except hearts, (f) a ten or a spade, (g) neither a four nor a club.

Let us use for brevity H, S, D, C to indicate heart, spade, diamond, club, respectively, and $1, 2, \ldots, 13$ for ace, two, . . . , king. Then $3 \cap H$ means three of hearts, while $3 \cup H$ means three or heart. Let us use the sample space of Problem 1.1(b), assigning equal probabilities of $1/52$ to each sample point. For example, $P(6 \cap C) = 1/52$.

(a)
$$\begin{aligned} P(1) &= P(1 \cap H \text{ or } 1 \cap S \text{ or } 1 \cap D \text{ or } 1 \cap C) \\ &= P(1 \cap H) + P(1 \cap S) + P(1 \cap D) + P(1 \cap C) \\ &= \frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{1}{13} \end{aligned}$$

This could also have been achieved from the sample space of Problem 1.1(a) where each sample point, in particular ace, has probability $1/13$. It could also have been arrived at by simply reasoning that there are 13 numbers and so each has probability $1/13$ of being drawn.

(b) $P(11 \cap H) = \dfrac{1}{52}$

(c) $P(3 \cap C \text{ or } 6 \cap D) = P(3 \cap C) + P(6 \cap D) = \dfrac{1}{52} + \dfrac{1}{52} = \dfrac{1}{26}$

(d) $P(H) = P(1 \cap H \text{ or } 2 \cap H \text{ or} \ldots 13 \cap H) = \dfrac{1}{52} + \dfrac{1}{52} + \cdots + \dfrac{1}{52} = \dfrac{13}{52} = \dfrac{1}{4}$

This could also have been arrived at by noting that there are four suits and each has equal probability$1/2$ of being drawn.

(e) $P(H') = 1 - P(H) = 1 - \dfrac{1}{4} = \dfrac{3}{4}$ using part (d) and Theorem 1-4, page 6.

(f) Since 10 and S are not mutually exclusive, we have, from Theorem 1-6,

$$P(10 \cup S) = P(10) + P(S) - P(10 \cap S) = \frac{1}{13} + \frac{1}{4} - \frac{1}{52} = \frac{4}{13}$$

(g) The probability of neither four nor club can be denoted by $P(4' \cap C')$. But $4' \cap C' = (4 \cup C)'$.

Therefore,

$$P(4' \cap C') = P[(4 \cup C)'] = 1 - P(4 \cup C)$$
$$= 1 - [P(4) + P(C) - P(4 \cap C)]$$
$$= 1 - \left[\frac{1}{13} + \frac{1}{4} - \frac{1}{52}\right] = \frac{9}{13}$$

We could also get this by noting that the diagram favorable to this event is the complement of the event shown circled in Fig. 1-8. Since this complement has $52 - 16 = 36$ sample points in it and each sample point is assigned probability $1/52$, the required probability is $36/52 = 9/13$.

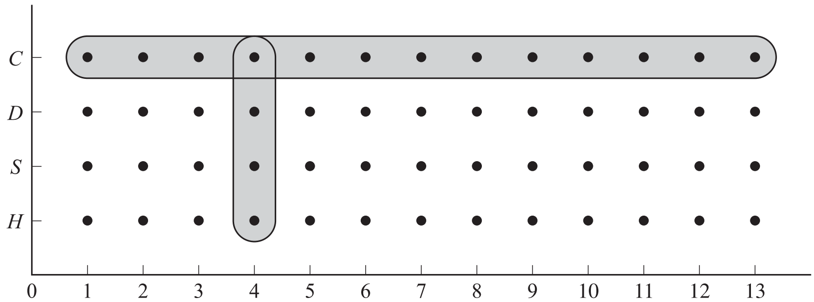


Fig. 1-8

1.7. A ball is drawn at random from a box containing 6 red balls, 4 white balls, and 5 blue balls. Determine the probability that it is (a) red, (b) white, (c) blue, (d) not red, (e) red or white.

(a) **Method 1**

Let R, W, and B denote the events of drawing a red ball, white ball, and blue ball, respectively. Then

$$P(R) = \frac{\text{ways of choosing a red ball}}{\text{total ways of choosing a ball}} = \frac{6}{6 + 4 + 5} = \frac{6}{15} = \frac{2}{5}$$

Method 2

Our sample space consists of $6 + 4 + 5 = 15$ sample points. Then if we assign equal probabilities $1/15$ to each sample point, we see that $P(R) = 6/15 = 2/5$, since there are 6 sample points corresponding to "red ball."

(b) $P(W) = \dfrac{4}{6 + 4 + 5} = \dfrac{4}{15}$

(c) $P(B) = \dfrac{5}{6 + 4 + 5} = \dfrac{5}{15} = \dfrac{1}{3}$

(d) $P(\text{not red}) = P(R') = 1 - P(R) = 1 - \dfrac{2}{5} = \dfrac{3}{5}$ by part (a).

(e) **Method 1**

$$P(\text{red or white}) = P(R \cup W) = \frac{\text{ways of choosing a red or white ball}}{\text{total ways of choosing a ball}}$$
$$= \frac{6 + 4}{6 + 4 + 5} = \frac{10}{15} = \frac{2}{3}$$

This can also be worked using the sample space as in part (a).

Method 2

$$P(R \cup W) = P(B') = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3} \text{ by part (c).}$$

Method 3

Since events R and W are mutually exclusive, it follows from (4), page 5, that

$$P(R \cup W) = P(R) + P(W) = \frac{2}{5} + \frac{4}{15} = \frac{2}{3}$$

Conditional probability and independent events

1.8. A fair die is tossed twice. Find the probability of getting a 4, 5, or 6 on the first toss and a 1, 2, 3, or 4 on the second toss.

Let A_1 be the event "4, 5, or 6 on first toss," and A_2 be the event "1, 2, 3, or 4 on second toss." Then we are looking for $P(A_1 \cap A_2)$.

Method 1

$$P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1) = P(A_1)P(A_2) = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = \frac{1}{3}$$

We have used here the fact that the result of the second toss is *independent* of the first so that $P(A_2 \mid A_1) = P(A_2)$. Also we have used $P(A_1) = 3/6$ (since 4, 5, or 6 are 3 out of 6 equally likely possibilities) and $P(A_2) = 4/6$ (since 1, 2, 3, or 4 are 4 out of 6 equally likely possibilities).

Method 2

Each of the 6 ways in which a die can fall on the first toss can be associated with each of the 6 ways in which it can fall on the second toss, a total of $6 \cdot 6 = 36$ ways, all equally likely.

Each of the 3 ways in which A_1 can occur can be associated with each of the 4 ways in which A_2 can occur to give $3 \cdot 4 = 12$ ways in which both A_1 and A_2 can occur. Then

$$P(A_1 \cap A_2) = \frac{12}{36} = \frac{1}{3}$$

This shows directly that A_1 and A_2 are independent since

$$P(A_1 \cap A_2) = \frac{1}{3} = \left(\frac{3}{6}\right)\left(\frac{4}{6}\right) = P(A_1)P(A_2)$$

1.9. Find the probability of not getting a 7 or 11 total on either of two tosses of a pair of fair dice.

The sample space for each toss of the dice is shown in Fig. 1-9. For example, (5, 2) means that 5 comes up on the first die and 2 on the second. Since the dice are fair and there are 36 sample points, we assign probability $1/36$ to each.

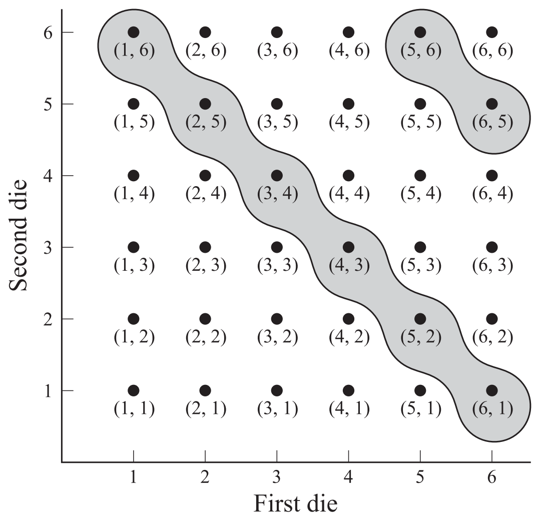


Fig. 1-9

If we let A be the event "7 or 11," then A is indicated by the circled portion in Fig. 1-9. Since 8 points are included, we have $P(A) = 8/36 = 2/9$. It follows that the probability of no 7 or 11 is given by

$$P(A') = 1 - P(A) = 1 - \frac{2}{9} = \frac{7}{9}$$

Using subscripts 1, 2 to denote 1st and 2nd tosses of the dice, we see that the probability of no 7 or 11 on either the first or second tosses is given by

$$P(A_1')P(A_2' \mid A_1') = P(A_1')P(A_2') = \left(\frac{7}{9}\right)\left(\frac{7}{9}\right) = \frac{49}{81},$$

using the fact that the tosses are independent.

1.10. Two cards are drawn from a well-shuffled ordinary deck of 52 cards. Find the probability that they are both aces if the first card is (a) replaced, (b) not replaced.

Method 1

Let A_1 = event "ace on first draw" and A_2 = event "ace on second draw." Then we are looking for $P(A_1 \cap A_2) = P(A_1)\,P(A_2 \mid A_1)$.

(a) Since for the first drawing there are 4 aces in 52 cards, $P(A_1) = 4/52$. Also, if the card is replaced for the second drawing, then $P(A_2 \mid A_1) = 4/52$, since there are also 4 aces out of 52 cards for the second drawing. Then

$$P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1) = \left(\frac{4}{52}\right)\left(\frac{4}{52}\right) = \frac{1}{169}$$

(b) As in part (a), $P(A_1) = 4/52$. However, if an ace occurs on the first drawing, there will be only 3 aces left in the remaining 51 cards, so that $P(A_2 \mid A_1) = 3/51$. Then

$$P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1) = \left(\frac{4}{52}\right)\left(\frac{3}{51}\right) = \frac{1}{221}$$

Method 2

(a) The first card can be drawn in any one of 52 ways, and since there is replacement, the second card can also be drawn in any one of 52 ways. Then both cards can be drawn in (52)(52) ways, all equally likely.

In such a case there are 4 ways of choosing an ace on the first draw and 4 ways of choosing an ace on the second draw so that the number of ways of choosing aces on the first and second draws is (4)(4). Then the required probability is

$$\frac{(4)(4)}{(52)(52)} = \frac{1}{169}$$

(b) The first card can be drawn in any one of 52 ways, and since there is no replacement, the second card can be drawn in any one of 51 ways. Then both cards can be drawn in (52)(51) ways, all equally likely.

In such a case there are 4 ways of choosing an ace on the first draw and 3 ways of choosing an ace on the second draw so that the number of ways of choosing aces on the first and second draws is (4)(3). Then the required probability is

$$\frac{(4)(3)}{(52)(51)} = \frac{1}{221}$$

1.11. Three balls are drawn successively from the box of Problem 1.7. Find the probability that they are drawn in the order red, white, and blue if each ball is (a) replaced, (b) not replaced.

Let R_1 = event "red on first draw," W_2 = event "white on second draw," B_3 = event "blue on third draw." We require $P(R_1 \cap W_2 \cap B_3)$.

(a) If each ball is replaced, then the events are independent and

$$\begin{aligned} P(R_1 \cap W_2 \cap B_3) &= P(R_1)P(W_2 \mid R_1)P(B_3 \mid R_2 \cap W_2) \\ &= P(R_1)P(W_2)P(B_3) \\ &= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{6+4+5}\right)\left(\frac{5}{6+4+5}\right) = \frac{8}{225} \end{aligned}$$

(b) If each ball is not replaced, then the events are dependent and

$$P(R_1 \cap W_2 \cap B_3) = P(R_1)P(W_2 \mid R_1)P(B_3 \mid R_1 \cap W_2)$$

$$= \left(\frac{6}{6+4+5}\right)\left(\frac{4}{5+4+5}\right)\left(\frac{5}{5+3+5}\right) = \frac{4}{91}$$

1.12. Find the probability of a 4 turning up at least once in two tosses of a fair die.

Let A_1 = event "4 on first toss" and A_2 = event "4 on second toss." Then

$$A_1 \cup A_2 = \text{event "4 on first toss or 4 on second toss or both"}$$
$$= \text{event "at least one 4 turns up,"}$$

and we require $P(A_1 \cup A_2)$.

Method 1

Events A_1 and A_2 are not mutually exclusive, but they are independent. Hence, by (*10*) and (*21*),

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$
$$= P(A_1) + P(A_2) - P(A_1)P(A_2)$$
$$= \frac{1}{6} + \frac{1}{6} - \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{11}{36}$$

Method 2

$$P(\text{at least one 4 comes up}) + P(\text{no 4 comes up}) = 1$$

Then

$$P(\text{at least one 4 comes up}) = 1 - P(\text{no 4 comes up})$$
$$= 1 - P(\text{no 4 on 1st toss and no 4 on 2nd toss})$$
$$= 1 - P(A_1' \cap A_2') = 1 - P(A_1')P(A_2')$$
$$= 1 - \left(\frac{5}{6}\right)\left(\frac{5}{6}\right) = \frac{11}{36}$$

Method 3

Total number of equally likely ways in which both dice can fall $= 6 \cdot 6 = 36$.

Also Number of ways in which A_1 occurs but not $A_2 = 5$
Number of ways in which A_2 occurs but not $A_1 = 5$
Number of ways in which both A_1 and A_2 occur $= 1$

Then the number of ways in which at least one of the events A_1 or A_2 occurs $= 5 + 5 + 1 = 11$. Therefore, $P(A_1 \cup A_2) = 11/36$.

1.13. One bag contains 4 white balls and 2 black balls; another contains 3 white balls and 5 black balls. If one ball is drawn from each bag, find the probability that (a) both are white, (b) both are black, (c) one is white and one is black.

Let W_1 = event "white ball from first bag," W_2 = event "white ball from second bag."

(a) $P(W_1 \cap W_2) = P(W_1)P(W_2 \mid W_1) = P(W_1)P(W_2) = \left(\dfrac{4}{4+2}\right)\left(\dfrac{3}{3+5}\right) = \dfrac{1}{4}$

(b) $P(W_1' \cap W_2') = P(W_1')P(W_2' \mid W_1') = P(W_1')P(W_2') = \left(\dfrac{2}{4+2}\right)\left(\dfrac{5}{3+5}\right) = \dfrac{5}{24}$

(c) The required probability is

$$1 - P(W_1 \cap W_2) - P(W_1' \cap W_2') = 1 - \frac{1}{4} - \frac{5}{24} = \frac{13}{24}$$

1.14. Prove Theorem 1-10, page 7.

We prove the theorem for the case $n = 2$. Extensions to larger values of n are easily made. If event A must result in one of the two mutually exclusive events A_1, A_2, then

$$A = (A \cap A_1) \cup (A \cap A_2)$$

But $A \cap A_1$ and $A \cap A_2$ are mutually exclusive since A_1 and A_2 are. Therefore, by Axiom 3,

$$\begin{aligned} P(A) &= P(A \cap A_1) + P(A \cap A_2) \\ &= P(A_1)\, P(A \mid A_1) + P(A_2)\, P(A \mid A_2) \end{aligned}$$

using (18), page 7.

1.15. Box *I* contains 3 red and 2 blue marbles while Box *II* contains 2 red and 8 blue marbles. A fair coin is tossed. If the coin turns up heads, a marble is chosen from Box *I*; if it turns up tails, a marble is chosen from Box *II*. Find the probability that a red marble is chosen.

Let R denote the event "a red marble is chosen" while I and II denote the events that Box I and Box II are chosen, respectively. Since a red marble can result by choosing either Box I or II, we can use the results of Problem 1.14 with $A = R$, $A_1 = I$, $A_2 = II$. Therefore, the probability of choosing a red marble is

$$P(R) = P(I)P(R \mid I) + P(II)P(R \mid II) = \left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{2+8}\right) = \frac{2}{5}$$

Bayes' theorem

1.16. Prove Bayes' theorem (Theorem 1-11, page 8).

Since A results in one of the mutually exclusive events $A_1, A_2, \ldots, A_n$, we have by Theorem 1-10 (Problem 1.14),

$$P(A) = P(A_1)P(A \mid A_1) + \cdots + P(A_n)P(A \mid A_n) = \sum_{j=1}^{n} P(A_j)P(A \mid A_j)$$

Therefore,
$$P(A_k \mid A) = \frac{P(A_k \cap A)}{P(A)} = \frac{P(A_k)P(A \mid A_k)}{\sum_{j=1}^{n} P(A_j)P(A \mid A_j)}$$

1.17. Suppose in Problem 1.15 that the one who tosses the coin does not reveal whether it has turned up heads or tails (so that the box from which a marble was chosen is not revealed) but does reveal that a red marble was chosen. What is the probability that Box *I* was chosen (i.e., the coin turned up heads)?

Let us use the same terminology as in Problem 1.15, i.e., $A = R$, $A_1 = I$, $A_2 = II$. We seek the probability that Box I was chosen given that a red marble is known to have been chosen. Using Bayes' rule with $n = 2$, this probability is given by

$$P(I \mid R) = \frac{P(I)P(R \mid I)}{P(I)P(R \mid I) + P(II)P(R \mid II)} = \frac{\left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right)}{\left(\frac{1}{2}\right)\left(\frac{3}{3+2}\right) + \left(\frac{1}{2}\right)\left(\frac{2}{2+8}\right)} = \frac{3}{4}$$

Combinational analysis, counting, and tree diagrams

1.18. A committee of 3 members is to be formed consisting of one representative each from labor, management, and the public. If there are 3 possible representatives from labor, 2 from management, and 4 from the public, determine how many different committees can be formed using (a) the fundamental principle of counting and (b) a tree diagram.

(a) We can choose a labor representative in 3 different ways, and after this a management representative in 2 different ways. Then there are $3 \cdot 2 = 6$ different ways of choosing a labor and management representative. With each of these ways we can choose a public representative in 4 different ways. Therefore, the number of different committees that can be formed is $3 \cdot 2 \cdot 4 = 24$.

(b) Denote the 3 labor representatives by L_1, L_2, L_3; the management representatives by M_1, M_2; and the public representatives by P_1, P_2, P_3, P_4. Then the tree diagram of Fig. 1-10 shows that there are 24 different committees in all. From this tree diagram we can list all these different committees, e.g., $L_1M_1P_1$, $L_1M_1P_2$, etc.

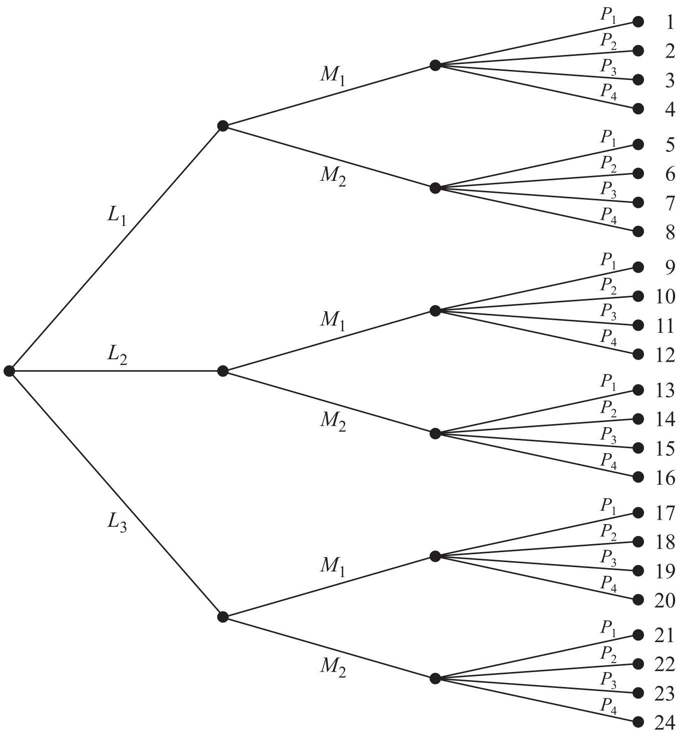


Fig. 1-10

Permutations

1.19. In how many ways can 5 differently colored marbles be arranged in a row?

We must arrange the 5 marbles in 5 positions thus: — — — — —. The first position can be occupied by any one of 5 marbles, i.e., there are 5 ways of filling the first position. When this has been done, there are 4 ways of filling the second position. Then there are 3 ways of filling the third position, 2 ways of filling the fourth position, and finally only 1 way of filling the last position. Therefore:

$$\text{Number of arrangements of 5 marbles in a row} = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! = 120$$

In general,

$$\text{Number of arrangements of } n \text{ different objects in a row} = n(n-1)(n-2) \cdots 1 = n!$$

This is also called the *number of permutations of n different objects taken n at a time* and is denoted by ${}_nP_n$.

1.20. In how many ways can 10 people be seated on a bench if only 4 seats are available?

The first seat can be filled in any one of 10 ways, and when this has been done, there are 9 ways of filling the second seat, 8 ways of filling the third seat, and 7 ways of filling the fourth seat. Therefore:

$$\text{Number of arrangements of 10 people taken 4 at a time} = 10 \cdot 9 \cdot 8 \cdot 7 = 5040$$

In general,

$$\text{Number of arrangements of } n \text{ different objects taken } r \text{ at a time} = n(n-1) \cdots (n-r+1)$$

This is also called the *number of permutations of n different objects taken r at a time* and is denoted by ${}_nP_r$. Note that when $r = n$, ${}_nP_n = n!$ as in Problem 1.19.

1.21. Evaluate (a) ${}_8P_3$, (b) ${}_6P_4$, (c) ${}_{15}P_1$, (d) ${}_3P_3$.

(a) ${}_8P_3 = 8 \cdot 7 \cdot 6 = 336$ (b) ${}_6P_4 = 6 \cdot 5 \cdot 4 \cdot 3 = 360$ (c) ${}_{15}P_1 = 15$ (d) ${}_3P_3 = 3 \cdot 2 \cdot 1 = 6$

1.22. It is required to seat 5 men and 4 women in a row so that the women occupy the even places. How many such arrangements are possible?

The men may be seated in ${}_5P_5$ ways, and the women in ${}_4P_4$ ways. Each arrangement of the men may be associated with each arrangement of the women. Hence,

$$\text{Number of arrangements} = {}_5P_5 \cdot {}_4P_4 = 5!\,4! = (120)(24) = 2880$$

1.23. How many 4-digit numbers can be formed with the 10 digits 0, 1, 2, 3, . . . , 9 if (a) repetitions are allowed, (b) repetitions are not allowed, (c) the last digit must be zero and repetitions are not allowed?

(a) The first digit can be any one of 9 (since 0 is not allowed). The second, third, and fourth digits can be any one of 10. Then $9 \cdot 10 \cdot 10 \cdot 10 = 9000$ numbers can be formed.

(b) The first digit can be any one of 9 (any one but 0).

The second digit can be any one of 9 (any but that used for the first digit).

The third digit can be any one of 8 (any but those used for the first two digits).

The fourth digit can be any one of 7 (any but those used for the first three digits).

Then $9 \cdot 9 \cdot 8 \cdot 7 = 4536$ numbers can be formed.

Another method

The first digit can be any one of 9, and the remaining three can be chosen in ${}_9P_3$ ways. Then $9 \cdot 9P_3 = 9 \cdot 9 \cdot 8 \cdot 7 = 4536$ numbers can be formed.

(c) The first digit can be chosen in 9 ways, the second in 8 ways, and the third in 7 ways. Then $9 \cdot 8 \cdot 7 = 504$ numbers can be formed.

Another method

The first digit can be chosen in 9 ways, and the next two digits in ${}_8P_2$ ways. Then $9 \cdot {}_8P_2 = 9 \cdot 8 \cdot 7 = 504$ numbers can be formed.

1.24. Four different mathematics books, six different physics books, and two different chemistry books are to be arranged on a shelf. How many different arrangements are possible if (a) the books in each particular subject must all stand together, (b) only the mathematics books must stand together?

(a) The mathematics books can be arranged among themselves in ${}_4P_4 = 4!$ ways, the physics books in ${}_6P_6 = 6!$ ways, the chemistry books in ${}_2P_2 = 2!$ ways, and the three groups in ${}_3P_3 = 3!$ ways. Therefore,

$$\text{Number of arrangements} = 4!6!2!3! = 207{,}360.$$

(b) Consider the four mathematics books as one big book. Then we have 9 books which can be arranged in ${}_9P_9 = 9!$ ways. In all of these ways the mathematics books are together. But the mathematics books can be arranged among themselves in ${}_4P_4 = 4!$ ways. Hence,

$$\text{Number of arrangements} = 9!4! = 8{,}709{,}120$$

1.25. Five red marbles, two white marbles, and three blue marbles are arranged in a row. If all the marbles of the same color are not distinguishable from each other, how many different arrangements are possible?

Assume that there are N different arrangements. Multiplying N by the numbers of ways of arranging (a) the five red marbles among themselves, (b) the two white marbles among themselves, and (c) the three blue marbles among themselves (i.e., multiplying N by 5!2!3!), we obtain the number of ways of arranging the 10 marbles if they were all distinguishable, i.e., 10!.

Then $$(5!2!3!)N = 10! \quad \text{and} \quad N = 10!/(5!2!3!)$$

In general, the number of different arrangements of n objects of which n_1 are alike, n_2 are alike, . . . , n_k are alike is $\dfrac{n!}{n_1!n_2! \cdots n_k!}$ where $n_1 + n_2 + \cdots + n_k = n$.

1.26. In how many ways can 7 people be seated at a round table if (a) they can sit anywhere, (b) 2 particular people must not sit next to each other?

(a) Let 1 of them be seated anywhere. Then the remaining 6 people can be seated in $6! = 720$ ways, which is the total number of ways of arranging the 7 people in a circle.

(b) Consider the 2 particular people as 1 person. Then there are 6 people altogether and they can be arranged in 5! ways. But the 2 people considered as 1 can be arranged in 2! ways. Therefore, the number of ways of arranging 7 people at a round table with 2 particular people sitting together $= 5!2! = 240$.

Then using (a), the total number of ways in which 7 people can be seated at a round table so that the 2 particular people do not sit together $= 730 - 240 = 480$ ways.

Combinations

1.27. In how many ways can 10 objects be split into two groups containing 4 and 6 objects, respectively?

This is the same as the number of arrangements of 10 objects of which 4 objects are alike and 6 other objects are alike. By Problem 1.25, this is $\dfrac{10!}{4!6!} = \dfrac{10 \cdot 9 \cdot 8 \cdot 7}{4!} = 210$.

The problem is equivalent to finding the number of selections of 4 out of 10 objects (or 6 out of 10 objects), the order of selection being immaterial. In general, the number of selections of r out of n objects, called the *number of combinations of n things taken r at a time,* is denoted by ${}_nC_r$ or $\dbinom{n}{r}$ and is given by

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{{}_nP_r}{r!} $$

1.28. Evaluate (a) ${}_7C_4$, (b) ${}_6C_5$, (c) ${}_4C_4$.

(a) ${}_7C_4 = \dfrac{7!}{4!3!} = \dfrac{7 \cdot 6 \cdot 5 \cdot 4}{4!} = \dfrac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 35.$

(b) ${}_6C_5 = \dfrac{6!}{5!1!} = \dfrac{6 \cdot 5 \cdot 4 \cdot 3 \cdot 2}{5!} = 6,$ or ${}_6C_5 = {}_6C_1 = 6.$

(c) ${}_4C_4$ is the number of selections of 4 objects taken 4 at a time, and there is only one such selection. Then ${}_4C_4 = 1$. Note that formally

$$ {}_4C_4 = \frac{4!}{4!0!} = 1 \quad \text{if we } \textit{define } 0! = 1. $$

1.29. In how many ways can a committee of 5 people be chosen out of 9 people?

$$ \binom{9}{5} = {}_9C_5 = \frac{9!}{5!4!} = \frac{9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}{5!} = 126 $$

1.30. Out of 5 mathematicians and 7 physicists, a committee consisting of 2 mathematicians and 3 physicists is to be formed. In how many ways can this be done if (a) any mathematician and any physicist can be included, (b) one particular physicist must be on the committee, (c) two particular mathematicians cannot be on the committee?

(a) 2 mathematicians out of 5 can be selected in ${}_5C_2$ ways.
3 physicists out of 7 can be selected in ${}_7C_3$ ways.

$$ \text{Total number of possible selections} = {}_5C_2 \cdot {}_7C_3 = 10 \cdot 35 = 350 $$

(b) 2 mathematicians out of 5 can be selected in ${}_5C_2$ ways.
2 physicists out of 6 can be selected in ${}_6C_2$ ways.

$$ \text{Total number of possible selections} = {}_5C_2 \cdot {}_6C_2 = 10 \cdot 15 = 150 $$

(c) 2 mathematicians out of 3 can be selected in ${}_3C_2$ ways.
3 physicists out of 7 can be selected in ${}_7C_3$ ways.

$$ \text{Total number of possible selections} = {}_3C_2 \cdot {}_7C_3 = 3 \cdot 35 = 105 $$

1.31. How many different salads can be made from lettuce, escarole, endive, watercress, and chicory?

Each green can be dealt with in 2 ways, as it can be chosen or not chosen. Since each of the 2 ways of dealing with a green is associated with 2 ways of dealing with each of the other greens, the number of ways of dealing with the 5 greens = 2^5 ways. But 2^5 ways includes the case in which no greens is chosen. Hence,

$$\text{Number of salads} = 2^5 - 1 = 31$$

Another method

One can select either 1 out of 5 greens, 2 out of 5 greens, . . . , 5 out of 5 greens. Then the required number of salads is

$${}_5C_1 + {}_5C_2 + {}_5C_3 + {}_5C_4 + {}_5C_5 = 5 + 10 + 10 + 5 + 1 = 31$$

In general, for any positive integer n, ${}_nC_1 + {}_nC_2 + {}_nC_3 + \cdots + {}_nC_n = 2^n - 1$.

1.32. From 7 consonants and 5 vowels, how many words can be formed consisting of 4 different consonants and 3 different vowels? The words need not have meaning.

The 4 different consonants can be selected in ${}_7C_4$ ways, the 3 different vowels can be selected in ${}_5C_3$ ways, and the resulting 7 different letters (4 consonants, 3 vowels) can then be arranged among themselves in ${}_7P_7 = 7!$ ways. Then

$$\text{Number of words} = {}_7C_4 \cdot {}_5C_3 \cdot 7! = 35 \cdot 10 \cdot 5040 = 1{,}764{,}000$$

The Binomial Coefficients

1.33. Prove that $\binom{n}{r} = \binom{n-1}{r} + \binom{n-1}{r-1}$.

We have

$$\begin{aligned}\binom{n}{r} = \frac{n!}{r!(n-r)!} &= \frac{n(n-1)!}{r!(n-r)!} = \frac{(n-r+r)(n-1)!}{r!(n-r)!} \\ &= \frac{(n-r)(n-1)!}{r!(n-r)!} + \frac{r(n-1)!}{r!(n-r)!} \\ &= \frac{(n-1)!}{r!(n-r-1)!} + \frac{(n-1)!}{(r-1)!(n-r)!} \\ &= \binom{n-1}{r} + \binom{n-1}{r-1}\end{aligned}$$

The result has the following interesting application. If we write out the coefficients in the binomial expansion of $(x + y)^n$ for $n = 0, 1, 2, \ldots$, we obtain the following arrangement, called *Pascal's triangle*:

$$\begin{array}{lccccccccccccc}
n = 0 & & & & & & & 1 & & & & & & \\
n = 1 & & & & & & 1 & & 1 & & & & & \\
n = 2 & & & & & 1 & & 2 & & 1 & & & & \\
n = 3 & & & & 1 & & 3 & & 3 & & 1 & & & \\
n = 4 & & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
n = 5 & & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
n = 6 & 1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1 \\
\text{etc.} & & & & & & & & & & & & &
\end{array}$$

An entry in any line can be obtained by adding the two entries in the preceding line that are to its immediate left and right. Therefore, $10 = 4 + 6$, $15 = 10 + 5$, etc.

1.34. Find the constant term in the expansion of $\left(x^2 + \frac{1}{x}\right)^{12}$.

According to the binomial theorem,

$$\left(x^2 + \frac{1}{x}\right)^{12} = \sum_{k=0}^{12} \binom{12}{k} (x^2)^k \left(\frac{1}{x}\right)^{12-k} = \sum_{k=0}^{12} \binom{12}{k} x^{3k-12}.$$

The constant term corresponds to the one for which $3k - 12 = 0$, i.e., $k = 4$, and is therefore given by

$$\binom{12}{4} = \frac{12 \cdot 11 \cdot 10 \cdot 9}{4 \cdot 3 \cdot 2 \cdot 1} = 495$$

Probability using combinational analysis

1.35. A box contains 8 red, 3 white, and 9 blue balls. If 3 balls are drawn at random without replacement, determine the probability that (a) all 3 are red, (b) all 3 are white, (c) 2 are red and 1 is white, (d) at least 1 is white, (e) 1 of each color is drawn, (f) the balls are drawn in the order red, white, blue.

(a) **Method 1**

Let R_1, R_2, R_3 denote the events, "red ball on 1st draw," "red ball on 2nd draw," "red ball on 3rd draw," respectively. Then $R_1 \cap R_2 \cap R_3$ denotes the event "all 3 balls drawn are red." We therefore have

$$\begin{aligned} P(R_1 \cap R_2 \cap R_3) &= P(R_1)P(R_2 \mid R_1)P(R_3 \mid R_1 \cap R_2) \\ &= \left(\frac{8}{20}\right)\left(\frac{7}{19}\right)\left(\frac{6}{18}\right) = \frac{14}{285} \end{aligned}$$

Method 2

$$\text{Required probability} = \frac{\text{number of selections of 3 out of 8 red balls}}{\text{number of selections of 3 out of 20 balls}} = \frac{{}_8C_3}{{}_{20}C_3} = \frac{14}{285}$$

(b) Using the second method indicated in part (a),

$$P(\text{all 3 are white}) = \frac{{}_3C_3}{{}_{20}C_3} = \frac{1}{1140}$$

The first method indicated in part (a) can also be used.

(c) $P(\text{2 are red and 1 is white})$

$$\begin{aligned} &= \frac{(\text{selections of 2 out of 8 red balls})(\text{selections of 1 out of 3 white balls})}{\text{number of selections of 3 out of 20 balls}} \\ &= \frac{({}_8C_2)({}_3C_1)}{{}_{20}C_3} = \frac{7}{95} \end{aligned}$$

(d) $P(\text{none is white}) = \frac{{}_{17}C_3}{{}_{20}C_3} = \frac{34}{57}$. Then

$$P(\text{at least 1 is white}) = 1 - \frac{34}{57} = \frac{23}{57}$$

(e) $P(\text{l of each color is drawn}) = \frac{({}_8C_1)({}_3C_1)({}_9C_1)}{{}_{20}C_3} = \frac{18}{95}$

(f) $P(\text{balls drawn in order red, white, blue}) = \frac{1}{3!} P(\text{l of each color is drawn})$

$$= \frac{1}{6}\left(\frac{18}{95}\right) = \frac{3}{95}, \text{ using (e)}$$

Another method

$$\begin{aligned} P(R_1 \cap W_2 \cap B_3) &= P(R_1)P(W_2 \mid R_1)P(B_3 \mid R_1 \cap W_2) \\ &= \left(\frac{8}{20}\right)\left(\frac{3}{19}\right)\left(\frac{9}{18}\right) = \frac{3}{95} \end{aligned}$$

1.36. In the game of *poker* 5 cards are drawn from a pack of 52 well-shuffled cards. Find the probability that (a) 4 are aces, (b) 4 are aces and 1 is a king, (c) 3 are tens and 2 are jacks, (d) a nine, ten, jack, queen, king are obtained in any order, (e) 3 are of any one suit and 2 are of another, (f) at least 1 ace is obtained.

(a) $P(4 \text{ aces}) = \dfrac{({}_4C_4)({}_{48}C_1)}{{}_{52}C_5} = \dfrac{1}{54{,}145}.$

(b) $P(4 \text{ aces and 1 king}) = \dfrac{({}_4C_4)({}_4C_1)}{{}_{52}C_5} = \dfrac{1}{649{,}740}.$

(c) $P(3 \text{ are tens and 2 are jacks}) = \dfrac{({}_4C_3)({}_4C_2)}{{}_{52}C_5} = \dfrac{1}{108{,}290}.$

(d) $P(\text{nine, ten, jack, queen, king in any order}) = \dfrac{({}_4C_1)({}_4C_1)({}_4C_1)({}_4C_1)({}_4C_1)}{{}_{52}C_5} = \dfrac{64}{162{,}435}.$

(e) $P(3 \text{ of any one suit, 2 of another}) = \dfrac{(4 \cdot {}_{13}C_3)(3 \cdot {}_{13}C_2)}{{}_{52}C_5} = \dfrac{429}{4165},$

since there are 4 ways of choosing the first suit and 3 ways of choosing the second suit.

(f) $P(\text{no ace}) = \dfrac{{}_{48}C_5}{{}_{52}C_5} = \dfrac{35{,}673}{54{,}145}.$ Then $P(\text{at least one ace}) = 1 - \dfrac{35{,}673}{54{,}145} = \dfrac{18{,}472}{54{,}145}.$

1.37. Determine the probability of three 6s in 5 tosses of a fair die.

Let the tosses of the die be represented by the 5 spaces – – – – –. In each space we will have the events 6 or not 6 (6′). For example, three 6s and two not 6s can occur as 6 6 6′6 6′ or 6 6′6 6′6, etc.

Now the probability of the outcome 6 6 6′ 6 6′ is

$$P(6\,6\,6'\,6\,6') = P(6)\,P(6)\,P(6')\,P(6)\,P(6') = \frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\cdot\frac{5}{6} = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2$$

since we assume independence. Similarly,

$$P = \left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2$$

for all other outcomes in which three 6s and two not 6s occur. But there are ${}_5C_3 = 10$ such outcomes, and these are mutually exclusive. Hence, the required probability is

$$P(6\,6\,6'6\,6' \text{ or } 6\,6'6\,6'6 \text{ or} \ldots) = {}_5C_3\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = \frac{5!}{3!2!}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^2 = \frac{125}{3888}$$

In general, if $p = P(A)$ and $q = 1 - p = P(A')$, then by using the same reasoning as given above, the probability of getting exactly x A's in n independent trials is

$${}_nC_x p^x q^{n-x} = \binom{n}{x} p^x q^{n-x}$$

1.38. A shelf has 6 mathematics books and 4 physics books. Find the probability that 3 particular mathematics books will be together.

All the books can be arranged among themselves in ${}_{10}P_{10} = 10!$ ways. Let us assume that the 3 particular mathematics books actually are replaced by 1 book. Then we have a total of 8 books that can be arranged among themselves in ${}_8P_8 = 8!$ ways. But the 3 mathematics books themselves can be arranged in ${}_3P_3 = 3!$ ways. The required probability is thus given by

$$\frac{8!\,3!}{10!} = \frac{1}{15}$$

Miscellaneous problems

1.39. A and B play 12 games of chess of which 6 are won by A, 4 are won by B, and 2 end in a draw. They agree to play a tournament consisting of 3 games. Find the probability that (a) A wins all 3 games, (b) 2 games end in a draw, (c) A and B win alternately, (d) B wins at least 1 game.

Let A_1, A_2, A_3 denote the events "A wins" in 1st, 2nd, and 3rd games, respectively, B_1, B_2, B_3 denote the events "B wins" in 1st, 2nd, and 3rd games, respectively. On the basis of their past performance (empirical probability), we shall assume that

$$P(\text{A wins any one game}) = \frac{6}{12} = \frac{1}{2}, \qquad P(\text{B wins any one game}) = \frac{4}{12} = \frac{1}{3}$$

(a) $$P(\text{A wins all 3 games}) = P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2)\,P(A_3) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{8}$$

assuming that the results of each game are independent of the results of any others. (This assumption would not be justifiable if either player were *psychologically influenced* by the other one's winning or losing.)

(b) In any one game the probability of a nondraw (i.e., either A or B wins) is $q = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$ and the probability of a draw is $p = 1 - q = \frac{1}{6}$. Then the probability of 2 draws in 3 trials is (see Problem 1.37)

$$\binom{3}{2} p^2 q^{3-2} = 3\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right) = \frac{5}{72}$$

(c)
$$\begin{aligned} P(\text{A and B win alternately}) &= P(\text{A wins then B wins then A wins} \\ &\qquad \textit{or } \text{B wins then A wins then B wins}) \\ &= P(A_1 \cap B_2 \cap A_3) + P(B_1 \cap A_2 \cap B_3) \\ &= P(A_1)P(B_2)P(A_3) + P(B_1)P(A_2)P(B_3) \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{3}\right)\left(\frac{1}{2}\right)\left(\frac{1}{3}\right) = \frac{5}{36} \end{aligned}$$

(d)
$$\begin{aligned} P(\text{B wins at least one game}) &= 1 - P(\text{B wins no game}) \\ &= 1 - P(B_1' \cap B_2' \cap B_3') \\ &= 1 - P(B_1')\,P(B_2')\,P(B_3') \\ &= 1 - \left(\frac{2}{3}\right)\left(\frac{2}{3}\right)\left(\frac{2}{3}\right) = \frac{19}{27} \end{aligned}$$

1.40. A and B play a game in which they alternately toss a pair of dice. The one who is first to get a total of 7 wins the game. Find the probability that (a) the one who tosses first will win the game, (b) the one who tosses second will win the game.

(a) The probability of getting a 7 on a single toss of a pair of dice, assumed fair, is $1/6$ as seen from Problem 1.9 and Fig. 1-9. If we suppose that A is the first to toss, then A will win in any of the following mutually exclusive cases with indicated associated probabilities:

(1) A wins on 1st toss. Probability $= \dfrac{1}{6}$.

(2) A loses on 1st toss, B then loses, A then wins. Probability $= \left(\dfrac{5}{6}\right)\left(\dfrac{5}{6}\right)\left(\dfrac{1}{6}\right)$.

(3) A loses on 1st toss, B loses, A loses, B loses, A wins. Probability $= \left(\dfrac{5}{6}\right)\left(\dfrac{5}{6}\right)\left(\dfrac{5}{6}\right)\left(\dfrac{5}{6}\right)\left(\dfrac{1}{6}\right)$.

. . . .

Then the probability that A wins is

$$\begin{aligned} &\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \cdots \\ &\qquad = \frac{1}{6}\left[1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \cdots\right] = \frac{1/6}{1 - (5/6)^2} = \frac{6}{11} \end{aligned}$$

where we have used the result 6 of Appendix A with $x = (5/6)^2$.

(b) The probability that B wins the game is similarly

$$\begin{aligned} \left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{5}{6}\right)\left(\frac{1}{6}\right) + \cdots &= \left(\frac{5}{6}\right)\left(\frac{1}{6}\right)\left[1 + \left(\frac{5}{6}\right)^2 + \left(\frac{5}{6}\right)^4 + \cdots\right] \\ &= \frac{5/36}{1 - (5/6)^2} = \frac{5}{11} \end{aligned}$$

Therefore, we would give 6 to 5 odds that the first one to toss will win. Note that since

$$\frac{6}{11} + \frac{5}{11} = 1$$

the probability of a tie is zero. This would not be true if the game was limited. See Problem 1.100.

1.41. A machine produces a total of 12,000 bolts a day, which are on the average 3% defective. Find the probability that out of 600 bolts chosen at random, 12 will be defective.

Of the 12,000 bolts, 3%, or 360, are defective and 11,640 are not. Then:

$$\text{Required probability} = \frac{{}_{360}C_{12}\ {}_{11,640}C_{588}}{{}_{12,000}C_{600}}$$

1.42. A box contains 5 red and 4 white marbles. Two marbles are drawn successively from the box without replacement, and it is noted that the second one is white. What is the probability that the first is also white?

Method 1

If W_1, W_2 are the events "white on 1st draw," "white on 2nd draw," respectivley, we are looking for $P(W_1 \mid W_2)$. This is given by

$$P(W_1 \mid W_2) = \frac{P(W_1 \cap W_2)}{P(W_2)} = \frac{(4/9)(3/8)}{4/9} = \frac{3}{8}$$

Method 2

Since the second is known to be white, there are only 3 ways out of the remaining 8 in which the first can be white, so that the probability is 3/8.

1.43. The probabilities that a husband and wife will be alive 20 years from now are given by 0.8 and 0.9, respectively. Find the probability that in 20 years (a) both, (b) neither, (c) at least one, will be alive.

Let H, W be the events that the husband and wife, respectively, will be alive in 20 years. Then $P(H) = 0.8$, $P(W) = 0.9$. We suppose that H and W are independent events, which may or may not be reasonable.

(a) $P(\text{both will be alive}) = P(H \cap W) = P(H)P(W) = (0.8)(0.9) = 0.72$.

(b) $P(\text{neither will be alive}) = P(H' \cap W') = P(H')\,P(W') = (0.2)(0.1) = 0.02$.

(c) $P(\text{at least one will be alive}) = 1 - P(\text{neither will be alive}) = 1 - 0.02 = 0.98$.

1.44. An inefficient secretary places n different letters into n differently addressed envelopes at random. Find the probability that at least one of the letters will arrive at the proper destination.

Let $A_1, A_2, \ldots A_n$ denote the events that the 1st, 2nd, . . . , nth letter is in the correct envelope. Then the event that at least one letter is in the correct envelope is $A_1 \cup A_2 \cup \cdots \cup A_n$, and we want to find $P(A_1 \cup A_2 \cup \cdots \cup A_n)$. From a generalization of the results (*10*) and (*11*), page 6, we have

$$(1) \qquad P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum P(A_k) - \sum P(A_j \cap A_k) + \sum P(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \cdots \cap A_n)$$

where $\sum P(A_k)$ the sum of the probabilities of A_k from 1 to n, $\sum P(A_j \cap A_k)$ is the sum of the probabilities of $A_j \cap A_k$ with j and k from 1 to n and $k > j$, etc. We have, for example, the following:

$$(2) \qquad P(A_1) = \frac{1}{n} \quad \text{and similarly} \quad P(A_k) = \frac{1}{n}$$

since, of the n envelopes, only 1 will have the proper address. Also

$$(3) \qquad P(A_1 \cap A_2) = P(A_1)P(A_2 \mid A_1) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)$$

since, if the 1st letter is in the proper envelope, then only 1 of the remaining $n - 1$ envelopes will be proper. In a similar way we find

$$(4) \qquad P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right)$$

etc., and finally

(5) $$P(A_1 \cap A_2 \cap \cdots \cap A_n) = \left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right) \cdots \left(\frac{1}{1}\right) = \frac{1}{n!}$$

Now in the sum $\sum P(A_j \cap A_k)$ there are $\binom{n}{2} = {}_nC_2$ terms all having the value given by (3). Similarly in $\sum P(A_i \cap A_j \cap A_k)$, there are $\binom{n}{3} = {}_nC_3$ terms all having the value given by (4). Therefore, the required probability is

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \binom{n}{1}\left(\frac{1}{n}\right) - \binom{n}{2}\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right) + \binom{n}{3}\left(\frac{1}{n}\right)\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right)$$

$$- \cdots + (-1)^{n-1}\binom{n}{n}\left(\frac{1}{n!}\right)$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\frac{1}{n!}$$

From calculus we know that (see Appendix A)

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

so that for $x = -1$

$$e^{-1} = 1 - \left(1 - \frac{1}{2!} + \frac{1}{3!} - \cdots\right)$$

or $$1 - \frac{1}{2!} + \frac{1}{3!} - \cdots = 1 - e^{-1}$$

It follows that if n is large, the required probability is very nearly $1 - e^{-1} = 0.6321$. This means that there is a good chance of at least 1 letter arriving at the proper destination. The result is remarkable in that the probability remains practically constant for all $n > 10$. Therefore, the probability that at least 1 letter will arrive at its proper destination is practically the same whether n is 10 or 10,000.

1.45. Find the probability that n people ($n \leq 365$) selected at random will have n different birthdays.

We assume that there are only 365 days in a year and that all birthdays are equally probable, assumptions which are not quite met in reality.

The first of the n people has of course some birthday with probability $365/365 = 1$. Then, if the second is to have a different birthday, it must occur on one of the other 364 days. Therefore, the probability that the second person has a birthday different from the first is $364/365$. Similarly the probability that the third person has a birthday different from the first two is $363/365$. Finally, the probability that the nth person has a birthday different from the others is $(365 - n + 1)/365$. We therefore have

$$P(\text{all } n \text{ birthdays are different}) = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - n + 1}{365}$$

$$= \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right) \cdots \left(1 - \frac{n-1}{365}\right)$$

1.46. Determine how many people are required in Problem 1.45 to make the probability of distinct birthdays less than $1/2$.

Denoting the given probability by p and taking natural logarithms, we find

(1) $$\ln p = \ln\left(1 - \frac{1}{365}\right) + \ln\left(1 - \frac{2}{365}\right) + \cdots + \ln\left(1 - \frac{n-1}{365}\right)$$

But we know from calculus (Appendix A, formula 7) that

(2) $$\ln(1 - x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots$$

so that (1) can be written

$$\text{(3)} \qquad \ln p = -\left[\frac{1 + 2 + \cdots + (n-1)}{365}\right] - \frac{1}{2}\left[\frac{1^2 + 2^2 + \cdots + (n-1)^2}{(365)^2}\right] - \cdots$$

Using the facts that for $n = 2, 3, \ldots$ (Appendix A, formulas 1 and 2)

$$\text{(4)} \qquad 1 + 2 + \cdots + (n-1) = \frac{n(n-1)}{2}, \qquad 1^2 + 2^2 + \cdots + (n-1)^2 = \frac{n(n-1)(2n-1)}{6}$$

we obtain for (3)

$$\text{(5)} \qquad \ln p = -\frac{n(n-1)}{730} - \frac{n(n-1)(2n-1)}{12(365)^2} - \cdots$$

For n small compared to 365, say, $n < 30$, the second and higher terms on the right of (5) are negligible compared to the first term, so that a good approximation in this case is

$$\text{(6)} \qquad \text{In } p = \frac{n(n-1)}{730}$$

For $p = \frac{1}{2}$, $\ln p = -\ln 2 = -0.693$. Therefore, we have

$$\text{(7)} \qquad \frac{n(n-1)}{730} = 0.693 \qquad \text{or} \qquad n^2 - n - 506 = 0 \qquad \text{or} \qquad (n-23)(n+22) = 0$$

so that $n = 23$. Our conclusion therefore is that, if n is larger than 23, we can give better than even odds that at least 2 people will have the same birthday.

SUPPLEMENTARY PROBLEMS

Calculation of probabilities

1.47. Determine the probability p, or an estimate of it, for each of the following events:

(a) A king, ace, jack of clubs, or queen of diamonds appears in drawing a single card from a well-shuffled ordinary deck of cards.

(b) The sum 8 appears in a single toss of a pair of fair dice.

(c) A nondefective bolt will be found next if out of 600 bolts already examined, 12 were defective.

(d) A 7 or 11 comes up in a single toss of a pair of fair dice.

(e) At least 1 head appears in 3 tosses of a fair coin.

1.48. An experiment consists of drawing 3 cards in succession from a well-shuffled ordinary deck of cards. Let A_1 be the event "king on first draw," A_2 the event "king on second draw," and A_3 the event "king on third draw." State in words the meaning of each of the following:

(a) $P(A_1 \cap A_2')$, (b) $P(A_1 \cup A_2)$, (c) $P(A_1' \cup A_2')$, (d) $P(A_1' \cap A_2' \cap A_3')$, (e) $P[(A_1 \cap A_2) \cup (A_2' \cap A_3)]$.

1.49. A marble is drawn at random from a box containing 10 red, 30 white, 20 blue, and 15 orange marbles. Find the probability that it is (a) orange or red, (b) not red or blue, (c) not blue, (d) white, (e) red, white, or blue.

1.50. Two marbles are drawn in succession from the box of Problem 1.49, replacement being made after each drawing. Find the probability that (a) both are white, (b) the first is red and the second is white, (c) neither is orange, (d) they are either red or white or both (red and white), (e) the second is not blue, (f) the first is orange, (g) at least one is blue, (h) at most one is red, (i) the first is white but the second is not, (j) only one is red.

1.51. Work Problem 1.50 with no replacement after each drawing.

Conditional probability and independent events

1.52. A box contains 2 red and 3 blue marbles. Find the probability that if two marbles are drawn at random (without replacement), (a) both are blue, (b) both are red, (c) one is red and one is blue.

1.53. Find the probability of drawing 3 aces at random from a deck of 52 ordinary cards if the cards are (a) replaced, (b) not replaced.

1.54. If at least one child in a family with 2 children is a boy, what is the probability that both children are boys?

1.55. Box *I* contains 3 red and 5 white balls, while Box *II* contains 4 red and 2 white balls. A ball is chosen at random from the first box and placed in the second box without observing its color. Then a ball is drawn from the second box. Find the probability that it is white.

Bayes' theorem or rule

1.56. A box contains 3 blue and 2 red marbles while another box contains 2 blue and 5 red marbles. A marble drawn at random from one of the boxes turns out to be blue. What is the probability that it came from the first box?

1.57. Each of three identical jewelry boxes has two drawers. In each drawer of the first box there is a gold watch. In each drawer of the second box there is a silver watch. In one drawer of the third box there is a gold watch while in the other there is a silver watch. If we select a box at random, open one of the drawers and find it to contain a silver watch, what is the probability that the other drawer has the gold watch?

1.58. Urn *I* has 2 white and 3 black balls; Urn *II*, 4 white and 1 black; and Urn *III*, 3 white and 4 black. An urn is selected at random and a ball drawn at random is found to be white. Find the probability that Urn *I* was selected.

Combinatorial analysis, counting, and tree diagrams

1.59. A coin is tossed 3 times. Use a tree diagram to determine the various possibilities that can arise.

1.60. Three cards are drawn at random (without replacement) from an ordinary deck of 52 cards. Find the number of ways in which one can draw (a) a diamond and a club and a heart in succession, (b) two hearts and then a club or a spade.

1.61. In how many ways can 3 different coins be placed in 2 different purses?

Permutations

1.62. Evaluate (a) ${}_4P_2$, (b) ${}_7P_5$, (c) ${}_{10}P_3$.

1.63. For what value of n is ${}_{n+1}P_3 = {}_nP_4$?

1.64. In how many ways can 5 people be seated on a sofa if there are only 3 seats available?

1.65. In how many ways can 7 books be arranged on a shelf if (a) any arrangement is possible, (b) 3 particular books must always stand together, (c) two particular books must occupy the ends?

1.66. How many numbers consisting of five different digits each can be made from the digits 1, 2, 3, . . . , 9 if (a) the numbers must be odd, (b) the first two digits of each number are even?

1.67. Solve Problem 1.66 if repetitions of the digits are allowed.

1.68. How many different three-digit numbers can be made with 3 fours, 4 twos, and 2 threes?

1.69. In how many ways can 3 men and 3 women be seated at a round table if (a) no restriction is imposed, (b) 2 particular women must not sit together, (c) each woman is to be between 2 men?

Combinations

1.70. Evaluate (a) ${}_5C_3$, (b) ${}_8C_4$, (c) ${}_{10}C_8$.

1.71. For what value of n is $3 \cdot {}_{n+1}C_3 = 7 \cdot {}_nC_2$?

1.72. In how many ways can 6 questions be selected out of 10?

1.73. How many different committees of 3 men and 4 women can be formed from 8 men and 6 women?

1.74. In how many ways can 2 men, 4 women, 3 boys, and 3 girls be selected from 6 men, 8 women, 4 boys and 5 girls if (a) no restrictions are imposed, (b) a particular man and woman must be selected?

1.75. In how many ways can a group of 10 people be divided into (a) two groups consisting of 7 and 3 people, (b) three groups consisting of 5, 3, and 2 people?

1.76. From 5 statisticians and 6 economists, a committee consisting of 3 statisticians and 2 economists is to be formed. How many different committees can be formed if (a) no restrictions are imposed, (b) 2 particular statisticians must be on the committee, (c) 1 particular economist cannot be on the committee?

1.77. Find the number of (a) combinations and (b) permutations of 4 letters each that can be made from the letters of the word *Tennessee*.

Binomial coefficients

1.78. Calculate (a) ${}_6C_3$, (b) $\binom{11}{4}$, (c) $({}_8C_2)({}_4C_3)/{}_{12}C_5$.

1.79. Expand (a) $(x + y)^6$, (b) $(x - y)^4$, (c) $(x - x^{-1})^5$, (d) $(x^2 + 2)^4$.

1.80. Find the coefficient of x in $\left(x + \frac{2}{x}\right)^9$.

Probability using combinatorial analysis

1.81. Find the probability of scoring a total of 7 points (a) once, (b) at least once, (c) twice, in 2 tosses of a pair of fair dice.

1.82. Two cards are drawn successively from an ordinary deck of 52 well-shuffled cards. Find the probability that (a) the first card is not a ten of clubs or an ace; (b) the first card is an ace but the second is not; (c) at least one card is a diamond; (d) the cards are not of the same suit; (e) not more than 1 card is a picture card (jack, queen, king); (f) the second card is not a picture card; (g) the second card is not a picture card given that the first was a picture card; (h) the cards are picture cards or spades or both.

1.83. A box contains 9 tickets numbered from 1 to 9, inclusive. If 3 tickets are drawn from the box 1 at a time, find the probability that they are alternately either odd, even, odd or even, odd, even.

1.84. The odds in favor of A winning a game of chess against B are 3:2. If 3 games are to be played, what are the odds (a) in favor of A winning at least 2 games out of the 3, (b) against A losing the first 2 games to B?

1.85. In the game of *bridge*, each of 4 players is dealt 13 cards from an ordinary well-shuffled deck of 52 cards. Find the probability that one of the players (say, the eldest) gets (a) 7 diamonds, 2 clubs, 3 hearts, and 1 spade; (b) a complete suit.

1.86. An urn contains 6 red and 8 blue marbles. Five marbles are drawn at random from it without replacement. Find the probability that 3 are red and 2 are blue.

1.87. (a) Find the probability of getting the sum 7 on at least 1 of 3 tosses of a pair of fair dice, (b) How many tosses are needed in order that the probability in (a) be greater than 0.95?

1.88. Three cards are drawn from an ordinary deck of 52 cards. Find the probability that (a) all cards are of one suit, (b) at least 2 aces are drawn.

1.89. Find the probability that a bridge player is given 13 cards of which 9 cards are of one suit.

Miscellaneous problems

1.90. A sample space consists of 3 sample points with associated probabilities given by $2p$, p^2, and $4p - 1$. Find the value of p.

1.91. How many words can be made from 5 letters if (a) all letters are different, (b) 2 letters are identical, (c) all letters are different but 2 particular letters cannot be adjacent?

1.92. Four integers are chosen at random between 0 and 9, inclusive. Find the probability that (a) they are all different, (b) not more than 2 are the same.

1.93. A pair of dice is tossed repeatedly. Find the probability that an 11 occurs for the first time on the 6th toss.

1.94. What is the least number of tosses needed in Problem 1.93 so that the probability of getting an 11 will be greater than (a) 0.5, (b) 0.95?

1.95. In a game of poker find the probability of getting (a) a *royal flush*, which consists of the ten, jack, queen, king, and ace of a single suit; (b) a *full house*, which consists of 3 cards of one face value and 2 of another (such as 3 tens and 2 jacks); (c) all different cards; (d) 4 aces.

1.96. The probability that a man will hit a target is $\frac{2}{3}$. If he shoots at the target until he hits it for the first time, find the probability that it will take him 5 shots to hit the target.

1.97. (a) A shelf contains 6 separate compartments. In how many ways can 4 indistinguishable marbles be placed in the compartments? (b) Work the problem if there are n compartments and r marbles. This type of problem arises in physics in connection with *Bose-Einstein statistics*.

1.98. (a) A shelf contains 6 separate compartments. In how many ways can 12 indistinguishable marbles be placed in the compartments so that no compartment is empty? (b) Work the problem if there are n compartments and r marbles where $r > n$. This type of problem arises in physics in connection with *Fermi-Dirac statistics*.

1.99. A poker player has cards 2, 3, 4, 6, 8. He wishes to discard the 8 and replace it by another card which he hopes will be a 5 (in which case he gets an "inside straight"). What is the probability that he will succeed assuming that the other three players together have (a) one 5, (b) two 5s, (c) three 5s, (d) no 5? Can the problem be worked if the number of 5s in the other players' hands is unknown? Explain.

1.100. Work Problem 1.40 if the game is limited to 3 tosses.

1.101. Find the probability that in a game of bridge (a) 2, (b) 3, (c) all 4 players have a complete suit.

ANSWERS TO SUPPLEMENTARY PROBLEMS

1.47. (a) 5/26 (b) 5/36 (c) 0.98 (d) 2/9 (e) 7/8

1.48. (a) Probability of king on first draw and no king on second draw.

(b) Probability of either a king on first draw or a king on second draw or both.

(c) No king on first draw or no king on second draw or both (no king on first and second draws).

(d) No king on first, second, and third draws.

(e) Probability of either king on first draw and king on second draw or no king on second draw and king on third draw.

1.49. (a) 1/3 (b) 3/5 (c) 11/15 (d) 2/5 (e) 4/5

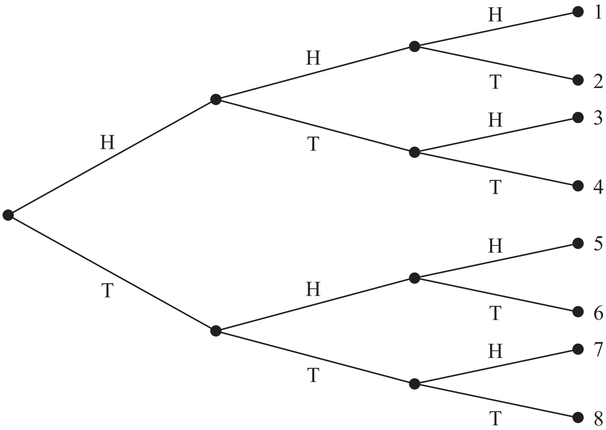
1.50. (a) 4/25 (c) 16/25 (e) 11/15 (g) 104/225 (i) 6/25
(b) 4/75 (d) 64/225 (f) 1/5 (h) 221/225 (j) 52/225

1.51. (a) 29/185 (c) 118/185 (e) 11/15 (g) 86/185 (i) 9/37
(b) 2/37 (d) 52/185 (f) 1/5 (h) 182/185 (j) 26/111

1.52. (a) 3/10 (b) 1/10 (c) 3/5 **1.53.** (a) 1/2197 (b) 1/17,576

1.54. 1/3 **1.55.** 21/56 **1.56.** 21/31 **1.57.** 1/3 **1.58.** 14/57

1.59.



1.60. (a) $13 \times 13 \times 13$ (b) $13 \times 12 \times 26$ **1.61.** 8 **1.62.** (a) 12 (b) 2520 (c) 720

1.63. $n = 5$ **1.64.** 60 **1.65.** (a) 5040 (b) 720 (c) 240 **1.66.** (a) 8400 (b) 2520

1.67. (a) 32,805 (b) 11,664 **1.68.** 26 **1.69.** (a) 120 (b) 72 (c) 12

1.70. (a) 10 (b) 70 (c) 45 **1.71.** $n = 6$ **1.72.** 210 **1.73.** 840

1.74. (a) 42,000 (b) 7000 **1.75.** (a) 120 (b) 2520 **1.76.** (a) 150 (b) 45 (c) 100

1.77. (a) 17 (b) 163 **1.78.** (a) 20 (b) 330 (c) 14/99

1.79. (a) $x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^3 + 6xy^5 + y^6$

(b) $x^4 - 4x^3y + 6x^2y^2 - 4xy^3 + y^4$

(c) $x^5 - 5x^3 + 10x - 10x^{-1} + 5x^{-3} - x^{-5}$

(d) $x^8 + 8x^6 + 24x^4 + 32x^2 + 16$

1.80. 2016 **1.81.** (a) 5/18 (b) 11/36 (c) 1/36

1.82. (a) 47/52 (b) 16/221 (c) 15/34 (d) 13/17 (e) 210/221 (f) 10/13 (g) 40/51 (h) 77/442

1.83. 5/18 **1.84.** (a) 81 : 44 (b) 21 : 4

1.85. (a) $({}_{13}C_7)({}_{13}C_2)({}_{13}C_3)({}_{13}C_1)/{}_{52}C_{13}$ (b) $4/{}_{52}C_{13}$ **1.86.** $({}_6C_3)({}_8C_2)/{}_{14}C_5$

1.87. (a) 91/216 (b) at least 17 **1.88.** (a) $4 \cdot {}_{13}C_{3//52}C_3$ (b) $({}_4C_2 \cdot {}_{48}C_1 + {}_4C_3)/{}_{52}C_3$

1.89. $4({}_{13}C_9)({}_{39}C_4)/{}_{52}C_{13}$ **1.90.** $\sqrt{11} - 3$ **1.91.** (a) 120 (b) 60 (c) 72

1.92. (a) 63/125 (b) 963/1000 **1.93.** 1,419,857/34,012,224 **1.94.** (a) 13 (b) 53

1.95. (a) $4/{}_{52}C_5$ (b) $(13)(2)(4)(6)/{}_{52}C_5$ (c) $4^5\,({}_{13}C_5)/{}_{52}C_5$ (d) (5)(4)(3)(2)/(52)(51)(50)(49)

1.96. 2/243 **1.97.** (a) 126 (b) ${}_{n+r-1}C_{n-1}$ **1.98.** (a) 462 (b) ${}_{r-1}C_{n-1}$

1.99. (a) 3/32 (b) 1/16 (c) 1/32 (d) 1/8

1.100. prob. A wins = 61/216, prob. B wins = 5/36, prob. of tie = 125/216

1.101. (a) $12/({}_{52}C_{13})({}_{39}C_{13})$ (b) $24/({}_{52}C_{13})({}_{39}C_{13})({}_{26}C_{13})$

CHAPTER 2

Random Variables and Probability Distributions

Random Variables

Suppose that to each point of a sample space we assign a number. We then have a *function* defined on the sample space. This function is called a *random variable* (or *stochastic variable*) or more precisely a *random function* (*stochastic function*). It is usually denoted by a capital letter such as X or Y. In general, a random variable has some specified physical, geometrical, or other significance.

EXAMPLE 2.1 Suppose that a coin is tossed twice so that the sample space is $S = \{HH, HT, TH, TT\}$. Let X represent the number of heads that can come up. With each sample point we can associate a number for X as shown in Table 2-1. Thus, for example, in the case of HH (i.e., 2 heads), $X = 2$ while for TH (1 head), $X = 1$. It follows that X is a random variable.

Table 2-1

Sample Point	HH	HT	TH	TT
X	2	1	1	0

It should be noted that many other random variables could also be defined on this sample space, for example, the square of the number of heads or the number of heads minus the number of tails.

A random variable that takes on a finite or countably infinite number of values (see page 4) is called a *discrete random variable* while one which takes on a noncountably infinite number of values is called a *nondiscrete random variable*.

Discrete Probability Distributions

Let X be a discrete random variable, and suppose that the possible values that it can assume are given by x_1, x_2, x_3, . . . , arranged in some order. Suppose also that these values are assumed with probabilities given by

$$P(X = x_k) = f(x_k) \qquad k = 1, 2, \ldots \tag{1}$$

It is convenient to introduce the *probability function*, also referred to as *probability distribution*, given by

$$P(X = x) = f(x) \tag{2}$$

For $x = x_k$, this reduces to (1) while for other values of x, $f(x) = 0$.

In general, $f(x)$ is a probability function if

1. $f(x) \geq 0$
2. $\sum_x f(x) = 1$

where the sum in 2 is taken over all possible values of x.

EXAMPLE 2.2 Find the probability function corresponding to the random variable X of Example 2.1. Assuming that the coin is fair, we have

$$P(HH) = \frac{1}{4} \qquad P(HT) = \frac{1}{4} \qquad P(TH) = \frac{1}{4} \qquad P(TT) = \frac{1}{4}$$

Then

$$P(X = 0) = P(TT) = \frac{1}{4}$$

$$P(X = 1) = P(HT \cup TH) = P(HT) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X = 2) = P(HH) = \frac{1}{4}$$

The probability function is thus given by Table 2-2.

Table 2-2

x	0	1	2
$f(x)$	1/4	1/2	1/4

Distribution Functions for Random Variables

The *cumulative distribution function*, or briefly the *distribution function*, for a random variable X is defined by

$$F(x) = P(X \leq x) \tag{3}$$

where x is any real number, i.e., $-\infty < x < \infty$.

The distribution function $F(x)$ has the following properties:

1. $F(x)$ is nondecreasing [i.e., $F(x) \leq F(y)$ if $x \leq y$].
2. $\lim_{x \to -\infty} F(x) = 0$; $\lim_{x \to \infty} F(x) = 1$.
3. $F(x)$ is continuous from the right [i.e., $\lim_{h \to 0^+} F(x + h) = F(x)$ for all x].

Distribution Functions for Discrete Random Variables

The distribution function for a discrete random variable X can be obtained from its probability function by noting that, for all x in $(-\infty, \infty)$,

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u) \tag{4}$$

where the sum is taken over all values u taken on by X for which $u \leq x$.

If X takes on only a finite number of values $x_1, x_2, \ldots, x_n$, then the distribution function is given by

$$F(x) = \begin{cases} 0 & -\infty < x < x_1 \\ f(x_1) & x_1 \leq x < x_2 \\ f(x_1) + f(x_2) & x_2 \leq x < x_3 \\ \vdots & \vdots \\ f(x_1) + \cdots + f(x_n) & x_n \leq x < \infty \end{cases} \tag{5}$$

EXAMPLE 2.3 (a) Find the distribution function for the random variable X of Example 2.2. (b) Obtain its graph.

(a) The distribution function is

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ \frac{1}{4} & 0 \leq x < 1 \\ \frac{3}{4} & 1 \leq x < 2 \\ 1 & 2 \leq x < \infty \end{cases}$$

(b) The graph of $F(x)$ is shown in Fig. 2-1.

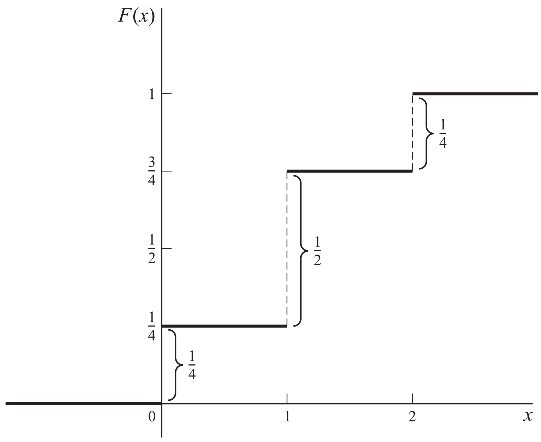


Fig. 2-1

The following things about the above distribution function, which are true in general, should be noted.

1. The magnitudes of the jumps at 0, 1, 2 are $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$ which are precisely the probabilities in Table 2-2. This fact enables one to obtain the probability function from the distribution function.
2. Because of the appearance of the graph of Fig. 2-1, it is often called a *staircase function* or *step function*. The value of the function at an integer is obtained from the higher step; thus the value at 1 is $\frac{3}{4}$ and not $\frac{1}{4}$. This is expressed mathematically by stating that the distribution function is *continuous from the right* at 0, 1, 2.
3. As we proceed from left to right (i.e. going *upstairs*), the distribution function either remains the same or increases, taking on values from 0 to 1. Because of this, it is said to be a *monotonically increasing function*.

It is clear from the above remarks and the properties of distribution functions that the probability function of a discrete random variable can be obtained from the distribution function by noting that

$$f(x) = F(x) - \lim_{u \to x^-} F(u). \tag{6}$$

Continuous Random Variables

A nondiscrete random variable X is said to be *absolutely continuous*, or simply *continuous*, if its distribution function may be represented as

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du \qquad (-\infty < x < \infty) \tag{7}$$

where the function $f(x)$ has the properties

1. $f(x) \geq 0$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

It follows from the above that if X is a continuous random variable, then the probability that X takes on any one particular value is zero, whereas the *interval probability* that X lies *between two different values*, say, a and b, is given by

$$P(a < X < b) = \int_{a}^{b} f(x)\,dx \tag{8}$$

EXAMPLE 2.4 If an individual is selected at random from a large group of adult males, the probability that his height X is precisely 68 inches (i.e., 68.000 . . . inches) would be zero. However, there is a probability greater than zero than X is between 67.000 . . . inches and 68.500 . . . inches, for example.

A function $f(x)$ that satisfies the above requirements is called a *probability function* or *probability distribution* for a continuous random variable, but it is more often called a *probability density function* or simply *density function*. Any function $f(x)$ satisfying Properties 1 and 2 above will automatically be a density function, and required probabilities can then be obtained from (8).

EXAMPLE 2.5 (a) Find the constant c such that the function

$$f(x) = \begin{cases} cx^2 & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

is a density function, and (b) compute $P(1 < X < 2)$.

(a) Since $f(x)$ satisfies Property 1 if $c \geq 0$, it must satisfy Property 2 in order to be a density function. Now

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^3 cx^2\,dx = \frac{cx^3}{3}\bigg|_0^3 = 9c$$

and since this must equal 1, we have $c = 1/9$.

(b) $$P(1 < X < 2) = \int_1^2 \frac{1}{9}x^2\,dx = \frac{x^3}{27}\bigg|_1^2 = \frac{8}{27} - \frac{1}{27} = \frac{7}{27}$$

In case $f(x)$ is continuous, which we shall assume unless otherwise stated, the probability that X is equal to any particular value is zero. In such case we can replace either or both of the signs $<$ in (8) by $\leq$. Thus, in Example 2.5,

$$P(1 \leq X \leq 2) = P(1 \leq X < 2) = P(1 < X \leq 2) = P(1 < X < 2) = \frac{7}{27}$$

EXAMPLE 2.6 (a) Find the distribution function for the random variable of Example 2.5. (b) Use the result of (a) to find $P(1 < x \leq 2)$.

(a) We have

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

If $x < 0$, then $F(x) = 0$. If $0 \leq x < 3$, then

$$F(x) = \int_0^x f(u)\,du = \int_0^x \frac{1}{9}u^2\,du = \frac{x^3}{27}$$

If $x \geq 3$, then

$$F(x) = \int_0^3 f(u)\,du + \int_3^x f(u)\,du = \int_0^3 \frac{1}{9}u^2\,du + \int_3^x 0\,du = 1$$

Thus the required distribution function is

$$F(x) = \begin{cases} 0 & x < 0 \\ x^3/27 & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Note that $F(x)$ increases monotonically from 0 to 1 as is required for a distribution function. It should also be noted that $F(x)$ in this case is continuous.

(b) We have

$$\begin{aligned} P(1 < X \le 2) &= P(X \le 2) - P(X \le 1) \\ &= F(2) - F(1) \\ &= \frac{2^3}{27} - \frac{1^3}{27} = \frac{7}{27} \end{aligned}$$

as in Example 2.5.

The probability that X is between x and $x + \Delta x$ is given by

$$P(x \le X \le x + \Delta x) = \int_x^{x+\Delta x} f(u)\,du \tag{9}$$

so that if Δx is small, we have approximately

$$P(x \le X \le x + \Delta x) = f(x)\Delta x \tag{10}$$

We also see from (7) on differentiating both sides that

$$\frac{dF(x)}{dx} = f(x) \tag{11}$$

at all points where $f(x)$ is continuous; i.e., the derivative of the distribution function is the density function.

It should be pointed out that random variables exist that are neither discrete nor continuous. It can be shown that the random variable X with the following distribution function is an example.

$$F(x) = \begin{cases} 0 & x < 1 \\ \dfrac{x}{2} & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

In order to obtain (11), we used the basic property

$$\frac{d}{dx}\int_a^x f(u)\,du = f(x) \tag{12}$$

which is one version of the Fundamental Theorem of Calculus.

Graphical Interpretations

If $f(x)$ is the density function for a random variable X, then we can represent $y = f(x)$ graphically by a curve as in Fig. 2-2. Since $f(x) \ge 0$, the curve cannot fall below the x axis. The entire area bounded by the curve and the x axis must be 1 because of Property 2 on page 36. Geometrically the probability that X is between a and b, i.e., $P(a < X < b)$, is then represented by the area shown shaded, in Fig. 2-2.

The distribution function $F(x) = P(X \le x)$ is a monotonically increasing function which increases from 0 to 1 and is represented by a curve as in Fig. 2-3.

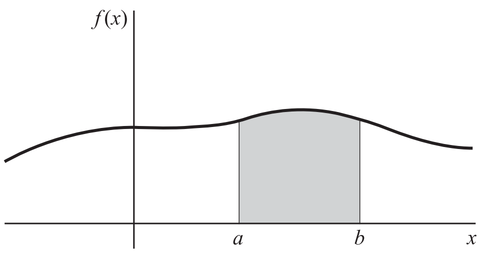


Fig. 2-2

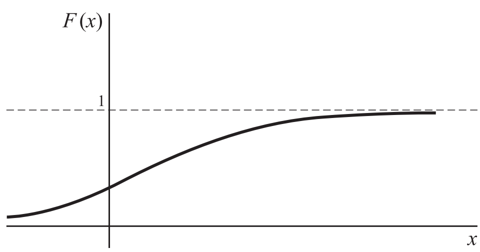


Fig. 2-3

Joint Distributions

The above ideas are easily generalized to two or more random variables. We consider the typical case of two random variables that are either both discrete or both continuous. In cases where one variable is discrete and the other continuous, appropriate modifications are easily made. Generalizations to more than two variables can also be made.

1. **DISCRETE CASE.** If X and Y are two discrete random variables, we define the *joint probability function* of X and Y by

$$P(X = x, Y = y) = f(x, y) \tag{13}$$

where 1. $f(x, y) \geq 0$

2. $\sum_x \sum_y f(x, y) = 1$

i.e., the sum over all values of x and y is 1.

Suppose that X can assume any one of m values $x_1, x_2, \ldots, x_m$ and Y can assume any one of n values $y_1, y_2, \ldots, y_n$. Then the probability of the event that $X = x_j$ and $Y = y_k$ is given by

$$P(X = x_j, Y = y_k) = f(x_j, y_k) \tag{14}$$

A joint probability function for X and Y can be represented by a *joint probability table* as in Table 2-3. The probability that $X = x_j$ is obtained by adding all entries in the row corresponding to x_i and is given by

$$P(X = x_j) = f_1(x_j) = \sum_{k=1}^{n} f(x_j, y_k) \tag{15}$$

Table 2-3

X \ Y	y_1	y_2	$\ldots$	y_n	Totals ↓	
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	$\ldots$	$f(x_1, y_n)$	$f_1(x_1)$	
x_2	$f(x_2, y_1)$	$f(x_2, y_2)$	$\ldots$	$f(x_2, y_n)$	$f_1(x_2)$	
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	
x_m	$f(x_m, y_1)$	$f(x_m, y_2)$	$\ldots$	$f(x_m, y_n)$	$f_1(x_m)$	
Totals →	$f_2(y_1)$	$f_2(y_2)$	$\ldots$	$f_2(y_n)$	1	← Grand Total

For $j = 1, 2, \ldots, m$, these are indicated by the entry totals in the extreme right-hand column or margin of Table 2-3. Similarly the probability that $Y = y_k$ is obtained by adding all entries in the column corresponding to y_k and is given by

$$P(Y = y_k) = f_2(y_k) = \sum_{j=1}^{m} f(x_j, y_k) \tag{16}$$

For $k = 1, 2, \ldots, n$, these are indicated by the entry totals in the bottom row or margin of Table 2-3.

Because the probabilities (15) and (16) are obtained from the margins of the table, we often refer to $f_1(x_j)$ and $f_2(y_k)$ [or simply $f_1(x)$ and $f_2(y)$] as the *marginal probability functions* of X and Y, respectively.

It should also be noted that

$$\sum_{j=1}^{m} f_1(x_j) = 1 \quad \sum_{k=1}^{n} f_2(y_k) = 1 \tag{17}$$

which can be written

$$\sum_{j=1}^{m}\sum_{k=1}^{n} f(x_j, y_k) = 1 \tag{18}$$

This is simply the statement that the total probability of all entries is 1. The *grand total* of 1 is indicated in the lower right-hand corner of the table.

The *joint distribution function* of X and Y is defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x}\sum_{v \leq y} f(u, v) \tag{19}$$

In Table 2-3, $F(x, y)$ is the sum of all entries for which $x_j \leq x$ and $y_k \leq y$.

2. CONTINUOUS CASE. The case where both variables are continuous is obtained easily by analogy with the discrete case on replacing sums by integrals. Thus the *joint probability function* for the random variables X and Y (or, as it is more commonly called, the *joint density function* of X and Y) is defined by

1. $f(x, y) \geq 0$
2. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$

Graphically $z = f(x, y)$ represents a surface, called the *probability surface*, as indicated in Fig. 2-4. The total volume bounded by this surface and the xy plane is equal to 1 in accordance with Property 2 above. The probability that X lies between a and b while Y lies between c and d is given graphically by the shaded volume of Fig. 2-4 and mathematically by

$$P(a < X < b, c < Y < d) = \int_{x=a}^{b}\int_{y=c}^{d} f(x, y)\,dx\,dy \tag{20}$$

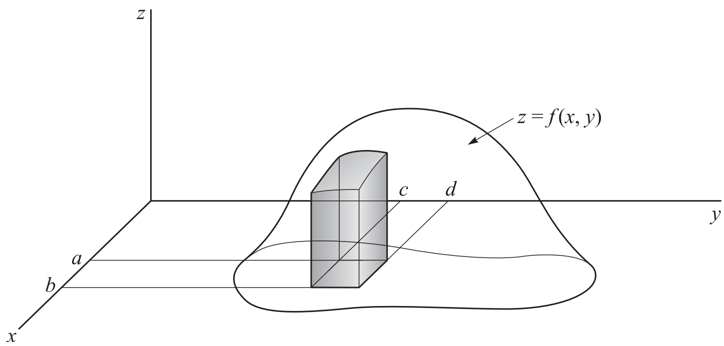


Fig. 2-4

More generally, if A represents any event, there will be a region $\mathcal{R}_A$ of the xy plane that corresponds to it. In such case we can find the probability of A by performing the integration over $\mathcal{R}_A$, i.e.,

$$P(A) = \iint_{\mathcal{R}_A} f(x, y)\,dx\,dy \tag{21}$$

The *joint distribution function* of X and Y in this case is defined by

$$F(x, y) = P(X \leq x, Y \leq y) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{y} f(u, v)\,du\,dv \tag{22}$$

It follows in analogy with (11), page 38, that

$$\frac{\partial^2 F}{\partial x\,\partial y} = f(x, y) \tag{23}$$

i.e., the density function is obtained by differentiating the distribution function with respect to x and y.

From (22) we obtain

$$P(X \le x) = F_1(x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv \tag{24}$$

$$P(Y \le y) = F_2(y) = \int_{u=-\infty}^{\infty}\int_{v=-\infty}^{y} f(u, v)\,du\,dv \tag{25}$$

We call (24) and (25) the *marginal distribution functions*, or simply the *distribution functions*, of X and Y, respectively. The derivatives of (24) and (25) with respect to x and y are then called the *marginal density functions,* or simply the *density functions*, of X and Y and are given by

$$f_1(x) = \int_{v=-\infty}^{\infty} f(x, v)\,dv \qquad f_2(y) = \int_{u=-\infty}^{\infty} f(u, y)\,du \tag{26}$$

Independent Random Variables

Suppose that X and Y are discrete random variables. If the events $X = x$ and $Y = y$ are independent events for all x and y, then we say that X and Y are *independent random variables*. In such case,

$$P(X = x, Y = y) = P(X = x)P(Y = y) \tag{27}$$

or equivalently

$$f(x, y) = f_1(x)f_2(y) \tag{28}$$

Conversely, if for all x and y the joint probability function $f(x, y)$ can be expressed as the product of a function of x alone and a function of y alone (which are then the marginal probability functions of X and Y), X and Y are independent. If, however, $f(x, y)$ cannot be so expressed, then X and Y are *dependent*.

If X and Y are continuous random variables, we say that they are *independent random variables* if the events $X \le x$ and $Y \le y$ are independent events for all x and y. In such case we can write

$$P(X \le x, Y \le y) = P(X \le x)P(Y \le y) \tag{29}$$

or equivalently

$$F(x, y) = F_1(x)F_2(y) \tag{30}$$

where $F_1(z)$ and $F_2(y)$ are the (marginal) distribution functions of X and Y, respectively. Conversely, X and Y are independent random variables if for all x and y, their joint distribution function $F(x, y)$ can be expressed as a product of a function of x alone and a function of y alone (which are the marginal distributions of X and Y, respectively). If, however, $F(x, y)$ cannot be so expressed, then X and Y are dependent.

For continuous independent random variables, it is also true that the joint density function $f(x, y)$ is the product of a function of x alone, $f_1(x)$, and a function of y alone, $f_2(y)$, and these are the (marginal) density functions of X and Y, respectively.

Change of Variables

Given the probability distributions of one or more random variables, we are often interested in finding distributions of other random variables that depend on them in some specified manner. Procedures for obtaining these distributions are presented in the following theorems for the case of discrete and continuous variables.

1. DISCRETE VARIABLES

Theorem 2-1 Let X be a discrete random variable whose probability function is $f(x)$. Suppose that a discrete random variable U is defined in terms of X by $U = \phi(X)$, where to each value of X there corresponds one and only one value of U and conversely, so that $X = \psi(U)$. Then the probability function for U is given by

$$g(u) = f[\psi(u)] \tag{31}$$

Theorem 2-2 Let X and Y be discrete random variables having joint probability function $f(x, y)$. Suppose that two discrete random variables U and V are defined in terms of X and Y by $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$, where to each pair of values of X and Y there corresponds one and only one pair of values of U and V and conversely, so that $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$. Then the joint probability function of U and V is given by

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)] \tag{32}$$

2. CONTINUOUS VARIABLES

Theorem 2-3 Let X be a continuous random variable with probability density $f(x)$. Let us define $U = \phi(X)$ where $X = \psi(U)$ as in Theorem 2-1. Then the probability density of U is given by $g(u)$ where

$$g(u)|du| = f(x)|dx| \tag{33}$$

or

$$g(u) = f(x)\left|\frac{dx}{du}\right| = f[\psi(u)]\,|\psi'(u)| \tag{34}$$

Theorem 2-4 Let X and Y be continuous random variables having joint density function $f(x, y)$. Let us define $U = \phi_1(X, Y)$, $V = \phi_2(X, Y)$ where $X = \psi_1(U, V)$, $Y = \psi_2(U, V)$ as in Theorem 2-2. Then the joint density function of U and V is given by $g(u, v)$ where

$$g(u, v)|du\, dv| = f(x, y)|dx\, dy| \tag{35}$$

or

$$g(u, v) = f(x, y)\left|\frac{\partial(x, y)}{\partial(u, v)}\right| = f[\psi_1(u, v), \psi_2(u, v)]\,|J| \tag{36}$$

In (36) the *Jacobian determinant*, or briefly *Jacobian*, is given by

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} \tag{37}$$

Probability Distributions of Functions of Random Variables

Theorems 2-2 and 2-4 specifically involve joint probability functions of two random variables. In practice one often needs to find the probability distribution of some specified function of several random variables. Either of the following theorems is often useful for this purpose.

Theorem 2-5 Let X and Y be continuous random variables and let $U = \phi_1(X, Y)$, $V = X$ (the second choice is arbitrary). Then the density function for U is the marginal density obtained from the joint density of U and V as found in Theorem 2-4. A similar result holds for probability functions of discrete variables.

Theorem 2-6 Let $f(x, y)$ be the joint density function of X and Y. Then the density function $g(u)$ of the random variable $U = \phi_1(X, Y)$ is found by differentiating with respect to u the distribution function given by

$$G(u) = P[\phi_1(X, Y) \le u] = \iint_{\mathcal{R}} f(x, y)\,dx\,dy \tag{38}$$

Where $\mathcal{R}$ is the region for which $\phi_1(x, y) \le u$.

Convolutions

As a particular consequence of the above theorems, we can show (see Problem 2.23) that the density function of the sum of two continuous random variables X and Y, i.e., of $U = X + Y$, having joint density function $f(x, y)$ is given by

$$g(u) = \int_{-\infty}^{\infty} f(x, u - x)\,dx \tag{39}$$

In the special case where X and Y are independent, $f(x, y) = f_1(x)f_2(y)$, and (39) reduces to

$$g(u) = \int_{-\infty}^{\infty} f_1(x)\,f_2(u - x)\,dx \tag{40}$$

which is called the *convolution* of f_1 and f_2, abbreviated, $f_1 * f_2$.

The following are some important properties of the convolution:

1. $f_1 * f_2 = f_2 * f_1$
2. $f_1 *(f_2 * f_3) = (f_1 * f_2) * f_3$
3. $f_1 *(f_2 + f_3) = f_1 * f_2 + f_1 * f_3$

These results show that f_1, f_2, f_3 obey the *commutative*, *associative*, and *distributive laws* of algebra with respect to the operation of convolution.

Conditional Distributions

We already know that if $P(A) > 0$,

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{41}$$

If X and Y are discrete random variables and we have the events (A: $X = x$), (B: $Y = y$), then (41) becomes

$$P(Y = y \mid X = x) = \frac{f(x, y)}{f_1(x)} \tag{42}$$

where $f(x, y) = P(X = x, Y = y)$ is the joint probability function and $f_1(x)$ is the marginal probability function for X. We define

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{43}$$

and call it the *conditional probability function of Y given X*. Similarly, the conditional probability function of X given Y is

$$f(x \mid y) \equiv \frac{f(x, y)}{f_2(y)} \tag{44}$$

We shall sometimes denote $f(x \mid y)$ and $f(y \mid x)$ by $f_1(x \mid y)$ and $f_2(y \mid x)$, respectively.

These ideas are easily extended to the case where X, Y are continuous random variables. For example, the *conditional density function of Y given X* is

$$f(y \mid x) \equiv \frac{f(x, y)}{f_1(x)} \tag{45}$$

where $f(x, y)$ is the joint density function of X and Y, and $f_1(x)$ is the marginal density function of X. Using (45) we can, for example, find that the probability of Y being between c and d given that $x < X < x + dx$ is

$$P(c < Y < d \mid x < X < x + dx) = \int_c^d f(y \mid x)\, dy \tag{46}$$

Generalizations of these results are also available.

Applications to Geometric Probability

Various problems in probability arise from geometric considerations or have geometric interpretations. For example, suppose that we have a target in the form of a plane region of area K and a portion of it with area K_1, as in Fig. 2-5. Then it is reasonable to suppose that the probability of hitting the region of area K_1 is proportional to K_1. We thus define

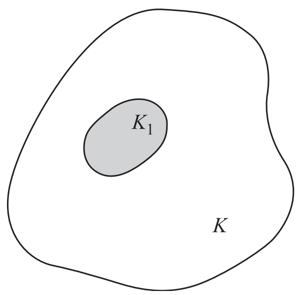


Fig. 2-5

$$P(\text{hitting region of area } K_1) = \frac{K_1}{K} \tag{47}$$

where it is assumed that the probability of hitting the target is 1. Other assumptions can of course be made. For example, there could be less probability of hitting outer areas. The type of assumption used defines the probability distribution function.

SOLVED PROBLEMS

Discrete random variables and probability distributions

2.1. Suppose that a pair of fair dice are to be tossed, and let the random variable X denote the sum of the points. Obtain the probability distribution for X.

The sample points for tosses of a pair of dice are given in Fig. 1-9, page 14. The random variable X is the sum of the coordinates for each point. Thus for (3, 2) we have $X = 5$. Using the fact that all 36 sample points are equally probable, so that each sample point has probability 1/36, we obtain Table 2-4. For example, corresponding to $X = 5$, we have the sample points (1, 4), (2, 3), (3, 2), (4, 1), so that the associated probability is 4/36.

Table 2-4

x	2	3	4	5	6	7	8	9	10	11	12
$f(x)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

2.2. Find the probability distribution of boys and girls in families with 3 children, assuming equal probabilities for boys and girls.

Problem 1.37 treated the case of n mutually independent trials, where each trial had just two possible outcomes, A and A', with respective probabilities p and $q = 1 - p$. It was found that the probability of getting exactly x A's in the n trials is ${}_nC_x p^x q^{n-x}$. This result applies to the present problem, under the assumption that successive births (the "trials") are independent as far as the sex of the child is concerned. Thus, with A being the event "a boy," $n = 3$, and $p = q = \frac{1}{2}$, we have

$$P(\text{exactly } x \text{ boys}) = P(X = x) = {}_3C_x\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{3-x} = {}_3C_x\left(\frac{1}{2}\right)^3$$

where the random variable X represents the number of boys in the family. (Note that X is defined on the sample space of 3 trials.) The probability function for X,

$$f(x) = {}_3C_x\left(\frac{1}{2}\right)^3$$

is displayed in Table 2-5.

Table 2-5

x	0	1	2	3
$f(x)$	1/8	3/8	3/8	1/8

Discrete distribution functions

2.3. (a) Find the distribution function $F(x)$ for the random variable X of Problem 2.1, and (b) graph this distribution function.

(a) We have $F(x) = P(X \le x) = \sum_{u \le x} f(u)$. Then from the results of Problem 2.1, we find

$$F(x) = \begin{cases} 0 & -\infty < x < 2 \\ 1/36 & 2 \le x < 3 \\ 3/36 & 3 \le x < 4 \\ 6/36 & 4 \le x < 5 \\ \vdots & \vdots \\ 35/36 & 11 \le x < 12 \\ 1 & 12 \le x < \infty \end{cases}$$

(b) See Fig. 2-6.

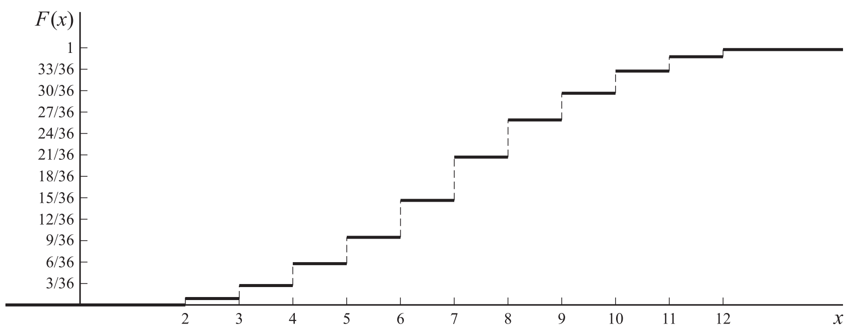


Fig. 2-6

2.4. (a) Find the distribution function $F(x)$ for the random variable X of Problem 2.2, and (b) graph this distribution function.

(a) Using Table 2-5 from Problem 2.2, we obtain

$$F(x) = \begin{cases} 0 & -\infty < x < 0 \\ 1/8 & 0 \le x < 1 \\ 1/2 & 1 \le x < 2 \\ 7/8 & 2 \le x < 3 \\ 1 & 3 \le x < \infty \end{cases}$$

(b) The graph of the distribution function of (a) is shown in Fig. 2-7.

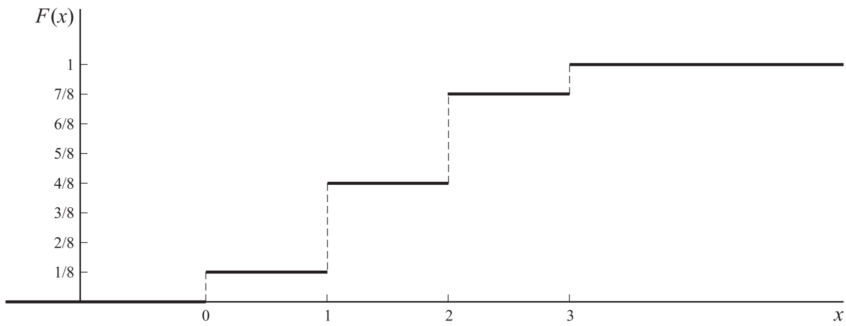


Fig. 2-7

Continuous random variables and probability distributions

2.5. A random variable X has the density function $f(x) = c/(x^2 + 1)$, where $-\infty < x < \infty$. (*a*) Find the value of the constant c. (b) Find the probability that X^2 lies between 1/3 and 1.

(a) We must have $\int_{-\infty}^{\infty} f(x)\,dx = 1$, i.e.,

$$\int_{-\infty}^{\infty} \frac{c\,dx}{x^2 + 1} = c \tan^{-1} x \Big|_{-\infty}^{\infty} = c\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = 1$$

so that $c = 1/\pi$.

(b) If $\frac{1}{3} \le X^2 \le 1$, then either $\frac{\sqrt{3}}{3} \le X \le 1$ or $-1 \le X \le -\frac{\sqrt{3}}{3}$. Thus the required probability is

$$\frac{1}{\pi}\int_{-1}^{-\sqrt{3}/3} \frac{dx}{x^2 + 1} + \frac{1}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1} = \frac{2}{\pi}\int_{\sqrt{3}/3}^{1} \frac{dx}{x^2 + 1}$$

$$= \frac{2}{\pi}\left[\tan^{-1}(1) - \tan^{-1}\left(\frac{\sqrt{3}}{3}\right)\right]$$

$$= \frac{2}{\pi}\left(\frac{\pi}{4} - \frac{\pi}{6}\right) = \frac{1}{6}$$

2.6. Find the distribution function corresponding to the density function of Problem 2.5.

$$F(x) = \int_{-\infty}^{x} f(u)\,du = \frac{1}{\pi}\int_{-\infty}^{x} \frac{du}{u^2 + 1} = \frac{1}{\pi}\left[\tan^{-1} u \Big|_{-\infty}^{x}\right]$$

$$= \frac{1}{\pi}[\tan^{-1} x - \tan^{-1}(-\infty)] = \frac{1}{\pi}\left[\tan^{-1} x + \frac{\pi}{2}\right]$$

$$= \frac{1}{2} + \frac{1}{\pi}\tan^{-1} x$$

2.7. The distribution function for a random variable X is

$$F(x) = \begin{cases} 1 - e^{-2x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Find (a) the density function, (b) the probability that $X > 2$, and (c) the probability that $-3 < X \leq 4$.

(a)
$$f(x) = \frac{d}{dx}F(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x < 0 \end{cases}$$

(b)
$$P(X > 2) = \int_2^{\infty} 2e^{-2u}du = -e^{-2u}\Big|_2^{\infty} = e^{-4}$$

Another method

By definition, $P(X \leq 2) = F(2) = 1 - e^{-4}$. Hence,

$$P(X > 2) = 1 - (1 - e^{-4}) = e^{-4}$$

(c)
$$P(-3 < X \leq 4) = \int_{-3}^{4} f(u)\,du = \int_{-3}^{0} 0\,du + \int_0^4 2e^{-2u}\,du$$
$$= -e^{-2u}\Big|_0^4 = 1 - e^{-8}$$

Another method

$$P(-3 < X \leq 4) = P(X \leq 4) - P(X \leq -3)$$
$$= F(4) - F(-3)$$
$$= (1 - e^{-8}) - (0) = 1 - e^{-8}$$

Joint distributions and independent variables

2.8. The joint probability function of two discrete random variables X and Y is given by $f(x, y) = c(2x + y)$, where x and y can assume all integers such that $0 \leq x \leq 2$, $0 \leq y \leq 3$, and $f(x, y) = 0$ otherwise.

(a) Find the value of the constant c.
(b) Find $P(X = 2, Y = 1)$.
(c) Find $P(X \geq 1, Y \leq 2)$.

(a) The sample points (x, y) for which probabilities are different from zero are indicated in Fig. 2-8. The probabilities associated with these points, given by $c(2x + y)$, are shown in Table 2-6. Since the grand total, $42c$, must equal 1, we have $c = 1/42$.

Table 2-6

X \ Y	0	1	2	3	Totals ↓
0	0	c	$2c$	$3c$	$6c$
1	$2c$	$3c$	$4c$	$5c$	$14c$
2	$4c$	$5c$	$6c$	$7c$	$22c$
Totals →	$6c$	$9c$	$12c$	$15c$	$42c$

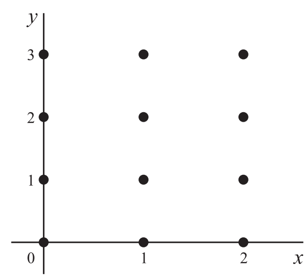


Fig. 2-8

(b) From Table 2-6 we see that

$$P(X = 2, Y = 1) = 5c + \frac{5}{42}$$

(c) From Table 2-6 we see that

$$P(X \geq 1, Y \leq 2) = \sum_{x \geq 1}\sum_{y \leq 2} f(x, y)$$
$$= (2c + 3c + 4c)(4c + 5c + 6c)$$
$$= 24c = \frac{24}{42} = \frac{4}{7}$$

as indicated by the entries shown shaded in the table.

2.9. Find the marginal probability functions (a) of X and (b) of Y for the random variables of Problem 2.8.

(a) The marginal probability function for X is given by $P(X = x) = f_1(x)$ and can be obtained from the margin totals in the right-hand column of Table 2-6. From these we see that

$$P(X = x) = f_1(x) = \begin{cases} 6c = 1/7 & x = 0 \\ 14c = 1/3 & x = 1 \\ 22c = 11/21 & x = 2 \end{cases}$$

Check: $\frac{1}{7} + \frac{1}{3} + \frac{11}{21} = 1$

(b) The marginal probability function for Y is given by $P(Y = y) = f_2(y)$ and can be obtained from the margin totals in the last row of Table 2-6. From these we see that

$$P(Y = y) = f_2(y) = \begin{cases} 6c = 1/7 & y = 0 \\ 9c = 3/14 & y = 1 \\ 12c = 2/7 & y = 2 \\ 15c = 5/14 & y = 3 \end{cases}$$

Check: $\frac{1}{7} + \frac{3}{14} + \frac{2}{7} + \frac{5}{14} = 1$

2.10. Show that the random variables X and Y of Problem 2.8 are dependent.

If the random variables X and Y are independent, then we must have, for all x and y,

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

But, as seen from Problems 2.8(b) and 2.9,

$$P(X = 2, Y = 1) = \frac{5}{42} \qquad P(X = 2) = \frac{11}{21} \qquad P(Y = 1) = \frac{3}{14}$$

so that
$$P(X = 2, Y = 1) \neq P(X = 2)P(Y = 1)$$

The result also follows from the fact that the joint probability function $(2x + y)/42$ cannot be expressed as a function of x alone times a function of y alone.

2.11. The joint density function of two continuous random variables X and Y is

$$f(x, y) = \begin{cases} cxy & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(a) Find the value of the constant c.

(b) Find $P(1 < X < 2, 2 < Y < 3)$.

(c) Find $P(X \geq 3, Y \leq 2)$.

(a) We must have the total probability equal to 1, i.e.,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\,dx\,dy = 1$$

Using the definition of $f(x, y)$, the integral has the value

$$\int_{x=0}^{4}\int_{y=1}^{5} cxy\,dx\,dy = c\int_{x=0}^{4}\left[\int_{y=1}^{5} xy\,dy\right]dx$$

$$= c\int_{z=0}^{4} \frac{xy^2}{2}\bigg|_{y=1}^{5} dx = c\int_{x=0}^{4}\left(\frac{25x}{2} - \frac{x}{2}\right)dx$$

$$= c\int_{x=0}^{4} 12x\,dx = c(6x^2)\bigg|_{x=0}^{4} = 96c$$

Then $96c = 1$ and $c = 1/96$.

(b) Using the value of c found in (a), we have

$$P(1 < X < 2, 2 < Y < 3) = \int_{x=1}^{2}\int_{y=2}^{3} \frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=1}^{2}\left[\int_{y=2}^{3} xy\,dy\right]dx = \frac{1}{96}\int_{x=1}^{2} \frac{xy^2}{2}\bigg|_{y=2}^{3} dx$$

$$= \frac{1}{96}\int_{x=1}^{2} \frac{5x}{2}\,dx = \frac{5}{192}\left(\frac{x^2}{2}\right)\bigg|_{1}^{2} = \frac{5}{128}$$

(c)
$$P(X \geq 3, Y \leq 2) = \int_{x=3}^{4}\int_{y=1}^{2} \frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=3}^{4}\left[\int_{y=1}^{2} xy\,dy\right]dx = \frac{1}{96}\int_{x=3}^{4} \frac{xy^2}{2}\bigg|_{y=1}^{2} dx$$

$$= \frac{1}{96}\int_{x=3}^{4} \frac{3x}{2}\,dx = \frac{7}{128}$$

2.12. Find the marginal distribution functions (a) of X and (b) of Y for Problem 2.11.

(a) The marginal distribution function for X if $0 \leq x < 4$ is

$$F_1(x) = P(X \leq x) = \int_{u=-\infty}^{x}\int_{v=-\infty}^{\infty} f(u, v)\,du\,dv$$

$$= \int_{u=0}^{x}\int_{v=1}^{5} \frac{uv}{96}\,du\,dv$$

$$= \frac{1}{96}\int_{u=0}^{x}\left[\int_{v=1}^{5} uv\,dv\right]du = \frac{x^2}{16}$$

For $x \geq 4$, $F_1(x) = 1$; for $x < 0$, $F_1(x) = 0$. Thus

$$F_1(x) = \begin{cases} 0 & x < 0 \\ x^2/16 & 0 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

As $F_1(x)$ is continuous at $x = 0$ and $x = 4$, we could replace $<$ by $\leq$ in the above expression.

(b) The marginal distribution function for Y if $1 \le y < 5$ is

$$F_2(y) = P(Y \le y) = \int_{u=-\infty}^{\infty} \int_{v=1}^{y} f(u, v)\,du\,dv$$

$$= \int_{u=0}^{4} \int_{v=1}^{y} \frac{uv}{96}\,du\,dv = \frac{y^2 - 1}{24}$$

For $y \ge 5$, $F_2(y) = 1$. For $y < 1$, $F_2(y) = 0$. Thus

$$F_2(y) = \begin{cases} 0 & y < 1 \\ (y^2 - 1)/24 & 1 \le y < 5 \\ 1 & y \ge 5 \end{cases}$$

As $F_2(y)$ is continuous at $y = 1$ and $y = 5$, we could replace $<$ by $\le$ in the above expression.

2.13. Find the joint distribution function for the random variables X, Y of Problem 2.11.

From Problem 2.11 it is seen that the joint density function for X and Y can be written as the product of a function of x alone and a function of y alone. In fact, $f(x, y) = f_1(x)f_2(y)$, where

$$f_1(x) = \begin{cases} c_1 x & 0 < x < 4 \\ 0 & \text{otherwise} \end{cases} \qquad f_2(y) = \begin{cases} c_2 y & 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

and $c_1 c_2 = c = 1/96$. It follows that X and Y are independent, so that their joint distribution function is given by $F(x, y) = F_1(x)F_2(y)$. The marginal distributions $F_1(x)$ and $F_2(y)$ were determined in Problem 2.12, and Fig. 2-9 shows the resulting piecewise definition of $F(x, y)$.

2.14. In Problem 2.11 find $P(X + Y < 3)$.

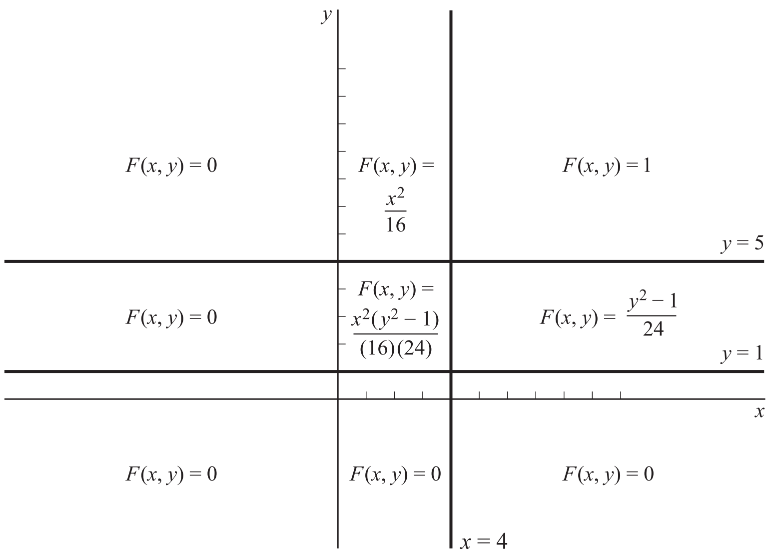


Fig. 2-9

In Fig. 2-10 we have indicated the square region $0 < x < 4$, $1 < y < 5$ within which the joint density function of X and Y is different from zero. The required probability is given by

$$P(X + Y < 3) = \iint_{\mathcal{R}} f(x, y)\,dx\,dy$$

where $\mathscr{R}$ is the part of the square over which $x + y < 3$, shown shaded in Fig. 2-10. Since $f(x, y) = xy/96$ over $\mathscr{R}$, this probability is given by

$$\int_{x=0}^{2}\int_{y=1}^{3-x}\frac{xy}{96}\,dx\,dy$$

$$= \frac{1}{96}\int_{x=0}^{2}\left[\int_{y=1}^{3-x} xy\,dy\right]dx$$

$$= \frac{1}{96}\int_{x=0}^{2}\frac{xy^2}{2}\bigg|_{y=1}^{3-x}dx = \frac{1}{192}\int_{x=0}^{2}[x(3-x)^2 - x] = \frac{1}{48}$$

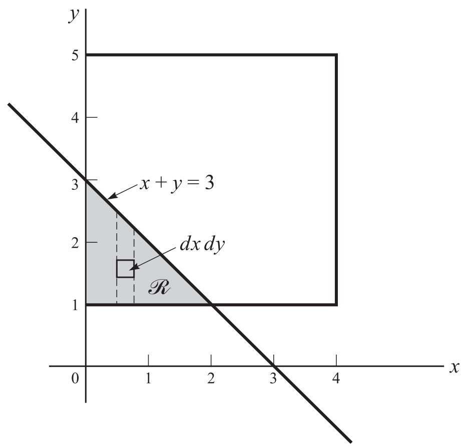


Fig. 2-10

Change of variables

2.15. Prove Theorem 2-1, page 42.

The probability function for U is given by

$$g(u) = P(U = u) = P[\phi(X) = u] = P[X = \psi(u)] = f[\psi(u)]$$

In a similar manner Theorem 2-2, page 42, can be proved.

2.16. Prove Theorem 2-3, page 42.

Consider first the case where $u = \phi(x)$ or $x = \psi(u)$ is an increasing function, i.e., u increases as x increases (Fig. 2-11). There, as is clear from the figure, we have

(1) $$P(u_1 < U < u_2) = P(x_1 < X < x_2)$$

or

(2) $$\int_{u_1}^{u_2} g(u)\,du = \int_{x_1}^{x_2} f(x)\,dx$$

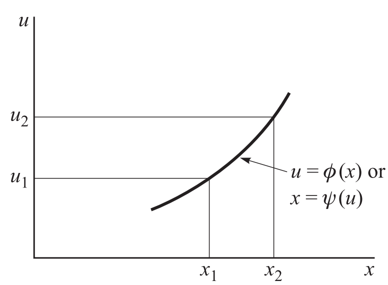


Fig. 2-11

Letting $x = \psi(u)$ in the integral on the right, (2) can be written

$$\int_{u_1}^{u_2} g(u)\,du = \int_{u_1}^{u_2} f[\psi(u)]\,\psi'(u)\,du$$

This can hold for all u_1 and u_2 only if the integrands are identical, i.e.,

$$g(u) = f[\psi(u)]\psi'(u)$$

This is a special case of (34), page 42, where $\psi'(u) > 0$ (i.e., the slope is positive). For the case where $\psi'(u) \leq 0$, i.e., u is a decreasing function of x, we can also show that (34) holds (see Problem 2.67). The theorem can also be proved if $\psi'(u) \geq 0$ or $\psi'(u) < 0$.

2.17. Prove Theorem 2-4, page 42.

We suppose first that as x and y increase, u and v also increase. As in Problem 2.16 we can then show that

$$P(u_1 < U < u_2, v_1 < V < v_2) = P(x_1 < X < x_2, y_1 < Y < y_2)$$

or

$$\int_{v_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{x_1}^{x_2}\int_{y_1}^{y_2} f(x, y)\,dx\,dy$$

Letting $x = \psi_1(u, v)$, $y = \psi_2(u, v)$ in the integral on the right, we have, by a theorem of advanced calculus,

$$\int_{v_1}^{u_2}\int_{v_1}^{v_2} g(u, v)\,du\,dv = \int_{u_1}^{u_2}\int_{v_1}^{v_2} f[\psi_1(u, v), \psi_2(u, v)]J\,du\,dv$$

where

$$J = \frac{\partial(x, y)}{\partial(u, v)}$$

is the *Jacobian*. Thus

$$g(u, v) = f[\psi_1(u, v), \psi_2(u, v)]J$$

which is (36), page 42, in the case where $J > 0$. Similarly, we can prove (36) for the case where $J < 0$.

2.18. The probability function of a random variable X is

$$f(x) = \begin{cases} 2^{-x} & x = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}$$

Find the probability function for the random variable $U = X^4 + 1$.

Since $U = X^4 + 1$, the relationship between the values u and x of the random variables U and X is given by $u = x^4 + 1$ or $x = \sqrt[4]{u - 1}$, where $u = 2, 17, 82, \ldots$ and the real positive root is taken. Then the required probability function for U is given by

$$g(u) = \begin{cases} 2^{-\sqrt[4]{u-1}} & u = 2, 17, 82, \ldots \\ 0 & \text{otherwise} \end{cases}$$

using Theorem 2-1, page 42, or Problem 2.15.

2.19. The probability function of a random variable X is given by

$$f(x) = \begin{cases} x^2/81 & -3 < x < 6 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability density for the random variable $U = \frac{1}{3}(12 - X)$.

We have $u = \frac{1}{3}(12 - x)$ or $x = 12 - 3u$. Thus to each value of x there is one and only one value of u and conversely. The values of u corresponding to $x = -3$ and $x = 6$ are $u = 5$ and $u = 2$, respectively. Since $\psi'(u) = dx/du = -3$, it follows by Theorem 2-3, page 42, or Problem 2.16 that the density function for U *is*

$$g(u) = \begin{cases} (12 - 3u)^2/27 & 2 < u < 5 \\ 0 & \text{otherwise} \end{cases}$$

Check: $$\int_2^5 \frac{(12 - 3u)^2}{27}\,du = -\frac{(12 - 3u)^3}{243}\bigg|_2^5 = 1$$

2.20. Find the probability density of the random variable $U = X^2$ where X is the random variable of Problem 2.19.

We have $u = x^2$ or $x = \pm\sqrt{u}$. Thus to each value of x there corresponds one and only one value of u, but to each value of $u \neq 0$ there correspond *two* values of x. The values of x for which $-3 < x < 6$ correspond to values of u for which $0 \le u < 36$ as shown in Fig. 2-12.

As seen in this figure, the interval $-3 < x \le 3$ corresponds to $0 \le u \le 9$ while $3 < x < 6$ corresponds to $9 < u < 36$. In this case we cannot use Theorem 2-3 directly but can proceed as follows. The distribution function for U is

$$G(u) = P(U \le u)$$

Now if $0 \le u \le 9$, we have

$$G(u) = P(U \le u) = P(X^2 \le u) = P(-\sqrt{u} \le X \le \sqrt{u})$$

$$= \int_{-\sqrt{u}}^{\sqrt{u}} f(x)\,dx$$

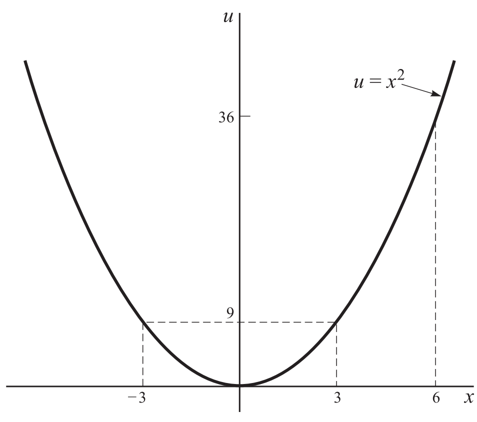


Fig. 2-12

But if $9 < u < 36$, we have

$$G(u) = P(U \le u) = P(-3 < X < \sqrt{u}) = \int_{-3}^{\sqrt{u}} f(x)\,dx$$

Since the density function $g(u)$ is the derivative of $G(u)$, we have, using (12),

$$g(u) = \begin{cases} \dfrac{f(\sqrt{u}) + f(-\sqrt{u})}{2\sqrt{u}} & 0 \le u \le 9 \\ \dfrac{f(\sqrt{u})}{2\sqrt{u}} & 9 < u < 36 \\ 0 & \text{otherwise} \end{cases}$$

Using the given definition of $f(x)$, this becomes

$$g(u) = \begin{cases} \sqrt{u}/81 & 0 \le u \le 9 \\ \sqrt{u}/162 & 9 < u < 36 \\ 0 & \text{otherwise} \end{cases}$$

Check:

$$\int_0^9 \frac{\sqrt{u}}{81}\,du + \int_9^{36} \frac{\sqrt{u}}{162}\,du = \frac{2u^{3/2}}{243}\bigg|_0^9 + \frac{u^{3/2}}{243}\bigg|_9^{36} = 1$$

2.21. If the random variables X and Y have joint density function

$$f(x, y) = \begin{cases} xy/96 & 0 < x < 4,\ 1 < y < 5 \\ 0 & \text{otherwise} \end{cases}$$

(see Problem 2.11), find the density function of $U = X + 2Y$.

Method 1

Let $u = x + 2y$, $v = x$, the second relation being chosen arbitrarily. Then simultaneous solution yields $x = v$, $y = \frac{1}{2}(u - v)$. Thus the region $0 < x < 4$, $1 < y < 5$ corresponds to the region $0 < v < 4$, $2 < u - v < 10$ shown shaded in Fig. 2-13.

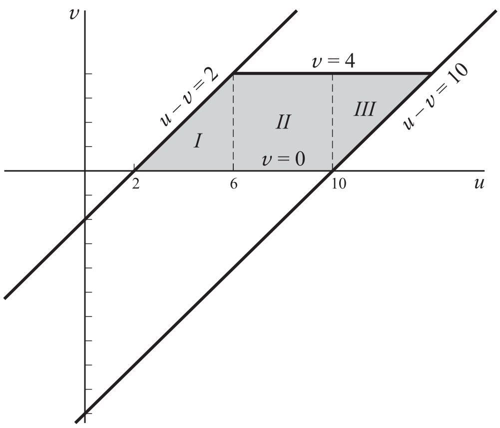


Fig. 2-13

The Jacobian is given by

$$J = \begin{vmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{2} & -\dfrac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

Then by Theorem 2-4 the joint density function of U and V is

$$g(u, v) = \begin{cases} v(u - v)/384 & 2 < u - v < 10, 0 < v < 4 \\ 0 & \text{otherwise} \end{cases}$$

The marginal density function of U is given by

$$g_1(u) = \begin{cases} \displaystyle\int_{v=0}^{u-2} \frac{v(u - v)}{384}\, dv & 2 < u < 6 \\ \displaystyle\int_{v=0}^{4} \frac{v(u - v)}{384}\, dv & 6 < u < 10 \\ \displaystyle\int_{v=u-10}^{4} \frac{v(u - v)}{384}\, dv & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

as seen by referring to the shaded regions *I*, *II*, *III* of Fig. 2-13. Carrying out the integrations, we find

$$g_1(u) = \begin{cases} (u - 2)^2(u + 4)/2304 & 2 < u < 6 \\ (3u - 8)/144 & 6 < u < 10 \\ (348u - u^3 - 2128)/2304 & 10 < u < 14 \\ 0 & \text{otherwise} \end{cases}$$

A check can be achieved by showing that the integral of $g_1(u)$ is equal to 1.

Method 2

The distribution function of the random variable $X + 2Y$ is given by

$$P(X + 2Y \leq u) = \iint\limits_{x+2y\leq u} f(x, y)\,dx\,dy = \iint\limits_{\substack{x+2y\leq u \\ 0<x<4 \\ 1<y<5}} \frac{xy}{96}\,dx\,dy$$

For $2 < u < 6$, we see by referring to Fig. 2-14, that the last integral equals

$$\int_{x=0}^{u-2}\int_{y=1}^{(u-x)/2} \frac{xy}{96}\,dx\,dy = \int_{x=0}^{u-2}\left[\frac{x(u - x)^2}{768} - \frac{x}{192}\right]dx$$

The derivative of this with respect to u is found to be $(u - 2)^2(u + 4)/2304$. In a similar manner we can obtain the result of Method 1 for $6 < u < 10$, etc.

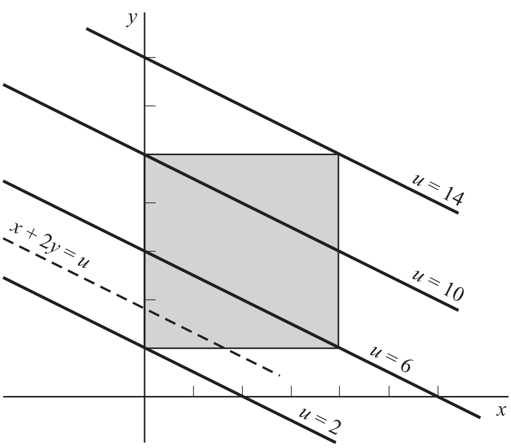


Fig. 2-14

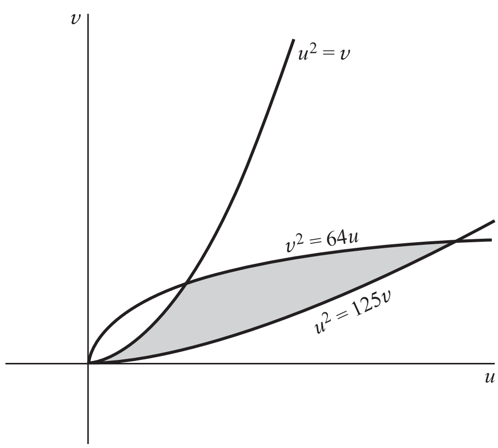


Fig. 2-15