Chapter 1

PROBABILITY

1.1 Introduction

The theory of probability had its origin in gambling and games of chance in mid-eighteenth-century. The theory thus developed for "heads or tails" or "red or black" soon found applications in situations where the outcomes were "boy or girl", " life or death ", "pass or fail" and scholars began to apply probability theory to actuarial problems and some aspects of the social sciences. Later, probability and statistics were introduced into physics by L. Boltzmann, J. Gibbs, and J. Maxwell, and in this century they have found applications in all phases of human endeavor which in some way involve an element of uncertainty or risk. The names which are connected most prominently with the growth of probability and mathematical statistics in the first half of this century are those A. N. Kolmogorov, R. A. Fisher, J. Neyman, E. S. Pearson, and A. Wald.

The concept of probability is frequently encountered in everyday communication. For example, we may hear a physician say that a patient has a 50-50 chance of surviving a certain operation. Another physician may say that he is 95 percent certain that a patient has a particular disease. Thus we are accustomed to measuring the probability of the occurrence of some event by a number between zero and one. The more likely the event, the closer the number is to one; and the more unlikely the event, the closer the number is to zero. An event that cannot occur has a probability of zero, and an event that is certain to occur has a probability of one.

1.2 Basic Definitions

We begin our development of probability by defining some basic key words.

(A) Experiment

The term **experiment** refers to the process of obtaining an observed result of some phenomenon, and a performance of an experiment, is called a **trial** of the experiment. An observed result, on a trial of the experiment, is called an **outcome**. This terminology is rather general; our interest will be in situations, called **random experiments**, where there is uncertainty about which outcome will occur when the experiment is performed. By a random experiment we will mean any procedure that:

1- all possible outcomes can be completely defined in advance.

2- can be repeated, theoretically, any number of times under identical conditions.

3- any performance of the experiment results in an outcome that is not known to certain occur in advance.

For example, a coin is tossed, assuming that the coin does not land on the side, there are two possible outcomes of the experiment: heads (denoted by **H**) and tails (denoted by **T**). On any performance of this experiment one does not know what the outcome will be. The coin can be tossed as many times as desired.

(B) Sample Space

The Sample Space, denoted by S, associated with a random experiment is the set of all possible outcomes of that experiment. Note that one and only one of the possible outcomes will occur on any given trial of the experiment.

Example 1.1

An experiment consists of tossing two coins, and the observed face of each coin is of interest. The set of all possible outcomes may be represented by the sample space;

$$S = \{ \text{ HH , HT , TH , TT } \}$$

An alternate way of representing such a sample space is to list all possible ordered pairs of the numbers 1 and 0,

$$S = \{ (1,1) , (1,0) , (0,1) , (0,0) \}$$

where, for example, (1,0) indicates that the first coin landed heads up and the second coin landed tails up. If we are interested in the total number of heads obtained from the two coins, an appropriate sample space could then be written as

$$S = \{ 0 , 1 , 2 \}$$

Thus, different sample spaces may be appropriated for the same experiment, depending on the characteristic of interest.

In tossing the coin three times (or 3 coins), the sample space consists of 8 outcomes,

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

Note that, in tossing the coin n times the sample space consists of 2^n outcomes.

First Toss	Second Toss	Third Toss	
		H	HHH
	H	T	HHT
		H	HTH
H	T	T	HTT
		H	THH
	H	T	THT
T	T	H	TTH
		T	TTT
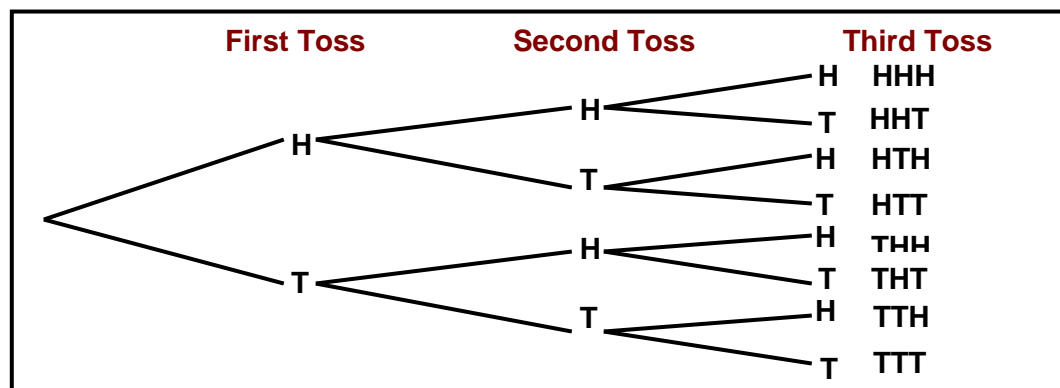
Example 1.2

If a coin is tossed repeatedly until a head occurs, then the natural sample space is S = {H, TH, TTH ...}. If one is interested in the number of tosses required to obtain a head, then a possible sample space for this experiment would be the set of all positive integers,

$$S^* = \{1, 2, 3 \, ...\},$$

and the outcomes would correspond directly to the number of tosses required to obtain the first head.

Example 1.3

A light bulb is placed in service and the time of operation until it burns out is measured. At least conceptually, the sample space for this experiment can be taken to be the set of nonnegative real numbers,

$$S = \{\, t : 0 \leq t < \infty\}$$

Note that if the actual failure time could be measured only to the nearest hour, then the sample space for the actual observed failure time would be the set of nonnegative integers,

$$S^* = \{\, 0, 1, 2, 3, \, ... \, \}.$$

A sample S space is said to be finite if it consists of a **finite** number of outcomes, say

$$S = \{e_1, e_2, ..., e_N\}$$

and it is said to be **countably infinite** if its outcomes can be put into a one correspondence with the positive integers, say

$$S = \{e_1, e_2, ... \}.$$

Definition

If a sample space S is either finite or countably infinite, then it is called a ***discrete sample space***.

A set that is either finite or countably infinite is also said to be **countable**. This is the case in the first two examples. It is also true for the last example when failure times are recorded to the nearest hour, but not for the conceptual sample space. Since the conceptual space involves outcomes that may assume any value in some interval of real numbers (i.e., the set of nonnegative real numbers), it could be termed a ***continuous sample space***, and it provides an example where a discrete sample space is not an appropriate model.

If the sample space consists of a continuum, such as all the points of a line segment or all the points in a plane (i.e. an uncountable infinite number of outcomes), it is

said to be ***continuous***. Continuous sample spaces arise in practice whenever the outcomes of experiments are measurements of physical properties such as temperature, speed, pressure, length, that are measured on continuous scales.

(C) Event

An event A is a collection of some of the possible outcomes of the random experiment. In the language of the set theory, A is a ***subset*** of the sample space S i.e.

$$A \subset S$$

In particular, we may regard each individual outcome as an event; since each outcome is a subset of S. (These are sometimes called **elementary events**). Also, the null set ϕ (i.e. the set which consists of no outcomes) and the complete sample space S are each special cases of subsets of S, and thus both ϕ and **S** may be regarded as events. ϕ is known as the **impossible event** and **S** as the **certain event**.
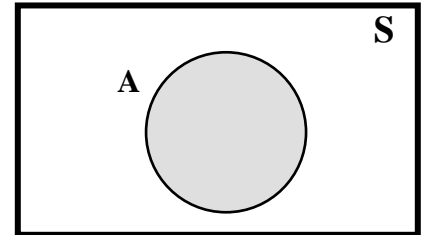
For example, in tossing a die, we might consider the following as typical events;

A_1 : Score is 1, i.e. {1}

A_2 : Score is even, i.e. {2,4,6}

A_3 : Score is less than 5, i.e. {1,2,3,4}.

(D) Occurrence of an event

We say that an event A had occurred if the outcome of the experiment belongs to the set A. For example, if the outcome in throwing a die was a score 6, then, as described above, A_2 occurred, but A_1 did not occur and A_3 did not occur.
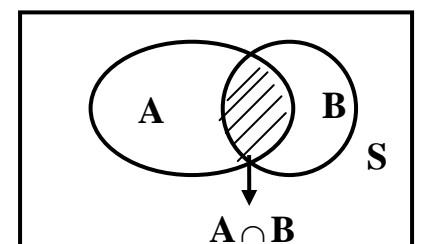
(E) Union of Events.

Given two events A & B, the union of A & B, written $A \cup B$ is defined as the set of outcomes which belong to either A or B or both. Thus, the event (A $\cup$ B) occurs whenever A or B occurs or both occur.
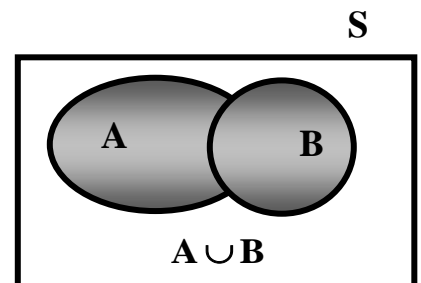
(F) Intersection of events

The intersection of A & B written as $A \cap B$ is defined as the set of outcomes which belong to both A and B. The set (A $\cap$ B) occurs when both A & B occur.
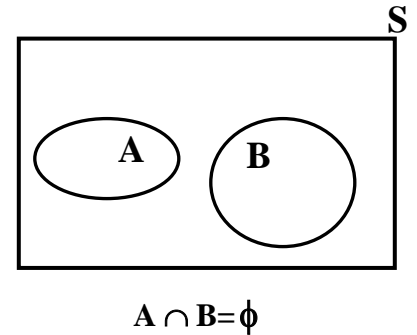
.

(G) Mutually exclusive Events

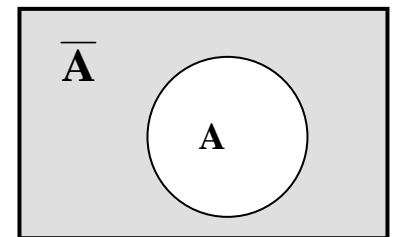Two events A and B are called mutually exclusive if both A and B cannot occur in the same time i.e. if

$$A \cap B = \phi$$

so that A and B are disjoint.

$A \cap B = \phi$

(H) Complimentary event

Given an event A, the complimentary, $\overline{A}$ (or A'), is defined as that subset of S which do not belong to A. Thus the event $\overline{A}$ occurs whenever A does not occur and vice versa.

(I) Equally Likely Events

If no outcome of a random experiment is any more likely to occur than any other, the outcomes are said to be equally likely.

1.3 Classical Probability

Until fairly recently, probability was thought of by statisticians and mathematicians only as an **objective phenomenon** derived from objective processes. The concept of *objective probability* may be categorized further under the headings of:

1- Classical, or a **priori**, **probability** and

2- The **relative frequency**, or a **posteriori**, concept of probability.

We may define probability in the classical sense as follows:

If a random experiment can result in **N** mutually exclusive and equally likely outcomes, **n(A)** of which corresponds to the occurrence of some event A, then the **probability** that the event A will occur, denoted by **P(A)**, will defined as the ratio **n(A)/N**, symbolically.

$$P(A) = \frac{n(A)}{N} = \frac{\text{number of outcomes belong to A}}{\text{total number of outcomes belong to S}} \tag{1.1}$$

The relative frequency approach to probability depends on the repeatability of some process and the ability to count the number of repetitions, as well as the number of times that some event of interest occurs. We may define the probability of observing some characteristic, **A**, of an event as follows:

If some process is repeated a large number of times, **N**, and if some resulting event with the characteristic, **A**, occurs n times, the relative frequency of occurrence of A, **n/N**, will be approximately equal to the probability of A.

For instance, consider the fair die. If our experiment consists of throwing a fair (unbiased) die, then there are 6 equally likely outcomes. We decide that the outcomes are equally likely by reasoning that if the die fair and a face (a number) is turned up at random there is certainly no reason to expect any face to be more likely to be turned up than any other. On any throw of a die, just one face and not other will turned up, so we conclude that the outcomes are mutually exclusive.

Example 1.4

Throw an unbiased coin three times and observe the sequence of heads and tails. Here the sample space is the collection of all possible sequences,

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

Since the outcomes are equally likely and mutually exclusive then the probability of each outcome is 1/8. Let A be the event that two or more heads appear consecutively, and B that all the tosses are the same. Then

$$A = \{ \text{ HHH, HHT, THH } \} \quad \text{and} \quad B = \{ \text{ HHH, TTT } \}$$

Therefore

$$P(A) = \frac{n(A)}{N} = \frac{3}{8} \quad , \quad P(B) = \frac{n(B)}{N} = \frac{2}{8} = \frac{1}{4}$$

$$P(A \cap B) = \frac{n(A \cap B)}{N} = P(\{ \text{HHH} \}) = \frac{1}{8} \text{ , and}$$

$$P(A \cup B) = \frac{n(A \cup B)}{N} = P(\{ \text{HHH, HHT, THH, TTT} \}) = \frac{4}{8} = \frac{1}{4}$$

Example 1.5

There are 15 balls, numbered from 1 to 15 in a bag. If a person selects one at random, what is the probability that the number printed on the ball will be

  i- a prime number greater than 5.

  ii- an odd number less than 9.

Solution

Let A_1 = { a prime number greater than 5 }

  A_2 = { an odd number less than 9 }

then

  A_1 = { 7, 11, 13 }, A_2 = { 1, 3, 5, 7 }, N = 15 , n(A_1) = 3 , n(A_2) = 4

$$\therefore \ P(A_1) = \frac{n(A_1)}{N} = \frac{3}{15} = \frac{1}{5} \quad \text{and} \quad P(A_2) = \frac{n(A_2)}{N} = \frac{4}{15}$$

Find $P(A_1 \cup A_2)$ and $P(A_1 \cap A_2)$?

Example 1.6

A fair die is tossed twice. What is the probability that the sum of the upturned faces is 9?

Solution

The sample space for this experiment is

$$S = \{ (i , j) : i = 1, 2, ...,6 ; j = 1, 2, ...,6 \}$$

Since the die is fair (unbiased), each of the 36 possible outcomes would be equally likely to occur. If A represents the event that the sum of the upturned faces is 9 then,

$$A = \{(3, 6) , (4, 5) , (5, 4) , (6, 3)\}$$

hence,

$$P(A) = 4/36 = 1/9$$

1.4 Axiomatic Approach To Probability

For a given experiment, S denotes the sample space, a set real function that associates a real value **P(A)** with each event **A** is called a probability function (or a probability measure), and **P(A)** is called the probability of **A**, if the following properties are satisfied:

  Axiom I. $P(A) \geq 0$, for every event $A \subset S$.

  Axiom II. $P(S) = 1$,

Axiom III. **If A_1, A_2, A_3, ... ,** is a finite or infinite sequence of mutually exclusive event of S, then

$$P(A_1 \cup A_2 \cup A_3 \cup...) = P(A_1) + P(A_2) + P(A_3) + ...$$

These axioms all seem to agree with our intuitive concept of probability and these few axioms are sufficient to allow a mathematical structure to be developed. Note, if A and B are two mutually exclusive events of S, then

$$P(A \cup B) = P(A) + P(B)$$

Example 1.7

Three horses A, B and C are in a race. A is twice as likely to win as B and B is twice as likely to win as C. What is their respective probabilities of winning, i.e. P(A), P(B) and P(C)?

Solution

Let $P(C) = p$,

since B is twice as likely to win as C, then

$P(B) = 2p$,

and since A is twice as likely to win as B, then

$P(C) = 4p$

Now the sum of the probabilities must be one (by property II), hence

$$P(A) + P(B) + P(C) = 1,$$

i.e.

$$4p + 2p + p = 1$$

Therefore, p=1/7, and accordingly

$P(A) = 4p = 4/7$, $P(B) = 2p = 2/7$ and $P(C) = p = 1/7$.

Probability in Discrete Spaces

If A is an event in a discrete sample space S, then P(A) equals the sum of the probabilities of the individual outcomes comprising A.

In particular if the sample space is given by

$$S = \{ e_1, e_2, e_3, ...\}$$

and we suppose that to each elementary event $\{e_i\}$ we assign a real number p_i, so that $P(\{e_i\}) = p_i$ such that;

$$p_i \geq 0 \text{ for all } i; \text{ and } \sum_i p_i = 1 \qquad \qquad (1.2)$$

Then the probability of an event A is given by

$$P(A) = \sum_{i: e_i \in A} p_i$$

It is understood, with this notation, that the summation is taken over all indices i such that e_i is an outcome in A.

Example 1.8

A die is loaded in such a way that each odd number is twice as likely to occur as each even number. Find P(G), where G is the event that a number greater than 3 occurs on a single roll of the die.

Solution

The sample space is S = {1, 2, 3, 4, 5, 6}. Hence, if we assign probability p to each even number and probability 2p to each odd number, we find that

$$2p + p + 2p + p + 2p + p = 9\,p = 1$$

in accordance with axiom II. It follows that $p = \dfrac{1}{9}$

$$P(G) = \frac{1}{9} + \frac{2}{9} + \frac{1}{9} = \frac{4}{9}$$

If a sample space is accountably infinite, probabilities will have to be assigned to the individual outcomes by means of a mathematical rule, preferably by means of a formula or equation.

1.5 Some Rules of Probability

Based on the three axioms of probability, we can derive many other rules which have important applications. Among them, the next four theorems are immediate consequences of the axioms.

Theorem 1.1

For any sample space S, $p(\phi) = 0$

Proof

Since S and φ are mutually exclusive and $S \cup \phi = S$ in accordance with definition of the empty set φ, it follows that

$$P(\ S\ ) = P(\ S \cup \phi\ )$$
$$= P(\ S\ ) + P(\ \phi\ ) \quad (\ \text{by Axiom III}\ )$$

and, hence, that $P(\phi) = 0$.

Theorem 1.2

If A and $\overline{A}$ are complementary events in a sample space S, then
$$P(\overline{A}) = 1 - P(A)$$

Proof

In the second and third steps of the proof that follows, we make use of the definition of a complement, according to which A and A' are mutually exclusive and $A \cup A' = S$. Thus, we write

$$1 = P(S) \quad (\ \text{by Axiom II}\ )$$
$$= P(\ A \cup \overline{A}\ )$$
$$= P(\ A\ ) + P(\ \overline{A}\ ) \quad (\ \text{by Axiom III}\ )$$

and it follows that $P(\overline{A}) = 1 - P(A)$.

Theorem 1.3

If A and B are events in S and $\mathbf{A} \subset \mathbf{B}$, then $P(A) \le P(B)$.

Proof

Since $A \subset B$, we can write
$$B = A \cup (\ \overline{A} \cap B\ )$$

as can easily be verified by means of a Venn diagram. Then, since $\mathbf{A}$ and $\mathbf{A' \cap B}$ are mutually exclusive, we get

$$P(\ B\ ) = P(\ A\ ) + P(\ \overline{A} \cup B\ ) \quad (\text{by Axiom III}\ )$$
$$\ge P(\ A\ ) \quad\quad\quad\quad (\text{by Axiom I}\ )$$

In words, this theorem states that if A is a subset of B, then P(A) cannot be greater than P(B).

Theorem 1.4

For any event A in S, $\mathbf{0 \le P(A) \le 1}$.

Proof

Using Theorem 1.3 and the fact that
$$\phi \subset A \subset S$$

for any event A in S, we have

$$P(\phi) \le P(A) \le P(S)$$

Then, $P(\phi) = 0$ and $P(S) = 1$ leads to the result that

$$0 \le P(A) \le 1$$

Theorem 1.5

If A and B are any two events in S, then the ***probability of occurrence of A and non-occurrence of B*** is given by

$$P(A \cap \overline{B}) = P(A \mid B) = P(A) - P(A \cap B)$$

Proof

The approach will be to express the event A as unions of mutually exclusive events. From set properties we have

$$A = (A \cap \overline{B}) \cup (A \cap B)$$

It follows that $A \cap \overline{B}$ and $A \cap B$ are mutually exclusive, so that

$$P(A) = P(A \cap \overline{B}) + P(A \cap B)$$

Hence the result.

Theorem 1.6

If A and B are any two events in a sample space S, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Proof

Expressing the event $A \cup B$ as union of mutually exclusive events as

$$A \cup B = (A \cap \overline{B}) \cup B$$

It follows that $A \cap \overline{B}$ and B are mutually exclusive, so that

$$P(A \cup B) = P(A \cap \overline{B}) + P(B)$$
$$= [P(A) - P(A \cap B)] + P(B)$$
$$= P(A) + P(B) - P(A \cap B)$$

Repeatedly using the formula of theorem 1.6, we can generalize this addition rule so that it will apply to any number of events. For instance, for three events we get,

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(A \cap C) + P(A \cap B \cap C)$$

Example 1.9

The probability that Jaillan Gabr passes mathematics is 0.75, and passes English is 0.85. If the probability of passing at least one course is 0.9, what is the probability that Jaillan will pass both courses?

Solution

If M is the event "passing mathematics" and E the event "passing English", then by transposing the terms in the Additive Rule,) we have

$$P(M \cup E) = P(M) + P(E) - P(M \cap E)$$
$$0.9 = 0.75 + 0.85 - P(M \cap E)$$

From which the probability that Jaillan will pass both courses is

$$P(M \cap E) = 0.7$$

<u>Questions</u>: what is the probability that Jaillan will fail both courses?

what is the probability that Jaillan will pass one and only one course?

1.6 Conditional Probability

A major objective of probability modeling is to determine how likely it is that an event A will occur when a certain experiment is performed. However, there are numerous cases in which the probability assigned to A will be affected by knowledge of the occurrence or nonoccurrence of another event B. In such case we will use the terminology "**conditional probability of A given B**", and the notation **P(A|B)** will be used to distinguish between this new concept and ordinary probability **P(A)**.

As an illustration, consider the event B of getting a 4 when a fair die is tossed. Based on the sample space S = {1,2,3,4,5,6} the probability of B occurring is $\frac{1}{6}$. Now suppose that it is known that the toss of the die resulted in a number greater than 3. We are now dealing with a reduced sample space A= {4,5,6}, which is a subset of S. Relative to the sample space A, we find that the probability that B occurs is 1/3. This example illustrates that events may have different probabilities when considered relative different sample spaces, and we would write

$$P(B|A) = \frac{1}{3},$$

which can also be written

$$P(B|A) = \frac{1}{3} = \frac{1/6}{3/6} = \frac{P(A \cap B)}{P(A)}$$

where P(A ∩ B) and P(A) are found from the original sample space S. In other words, a conditional probability relative to a subspace A of S may be calculated directly from the sample space S itself.

Definition

If A and B are any two events in S and **P(A) > 0**, the conditional probability of B given A is

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \tag{1.3}$$

Example 1.10

A box contains 100 microchips, some of which were produced by factory 1 and the rest by factory 2. Some of the microchips are defective and some are good (non-defective). An experiment consists of choosing one microchip at random from the box and testing whether it is good or defective. Let

	B	$\overline{B}$	Totals
A	15	5	20
$\overline{A}$	45	35	80
Totals	60	40	100

Table 1.1 Numbers of defective and nondefective microchips from two different factories

A be the event "obtaining a defective microchip"; consequently, and

B be the event "the microchip was produced by factory 1" , then

$\overline{A}$ is the event "obtaining a good microchip".

$\overline{B}$ is the event "the microchip was produced by factory 2".

Table 1.1 gives the number of microchips in each category.

The probability of obtaining a defective microchip is

$$P(A) = \frac{n(A)}{n(S)} = \frac{20}{100} = 0.20$$

Now suppose that each microchip has a number stamped on it that identifies which factory produced it. Thus, prior to testing whether it is defective, it can be determined

whether B has occurred (produced by factory 1) or $\overline{B}$ has occurred affects likelihood that a defective microchip is selected, and the use of conditional probability is appropriate. For example, if the event B has occurred, then the only microchips we should consider are those in the first column of Table 1.1, and the total number is n(B) = 60. Furthermore, the only defective chips to consider are those in both the first column and the first row, and the total number is n(A∪B) = 15. Thus, the conditional probability of A given B is

$$P(A|B) = \frac{n(A \cap B)}{n(B)} = \frac{15}{60} = 0.25$$

Notice that if we divide both the numerator and denominator by n(S) = 100, we can express conditional probability in terms of some ordinary unconditional probabilities,

$$P(A|B) = \frac{n(A \cap B)/n(S)}{n(B)/n(S)} = \frac{P(A \cap B)}{P(B)}$$

Example 1.11

The probability that a student, selected at random from certain college, will pass economics is 0.8 and will pass in both economics and religion is 0.5. What is the probability that he will pass religion if it is known that he had passed economics?

Solution

If E is the event "passing economics" and R the event "passing religion", then

$$P(R|E) = \frac{P(E \cap R)}{P(E)} = \frac{0.5}{0.8} = \frac{5}{8} = 0.625$$

Multiplication Rule

If we multiply the expressions on both sides of (1.3) by P(A), we obtain the following multiplication rule

Theorem 1.7

If A and B are two events in S, then

$$P(A \cap B) = P(A)P(B|A) = P(B)P(A|B) \qquad (1.4)$$

Example 1.12

If we randomly pick two television tubes in succession from a shipment of 240 television tubes of which 15 are defective, what is the probability that they will both defective?

Solution

If we assume equal probabilities for each selection (which is what we mean by "randomly" picking the tubes), the probability that the first tube will be defective is

$$\frac{15}{240}$$

and the probability that the second tube will be defective given that the first tube is defective is

$$\frac{14}{239}$$

Thus, the probability that both tubes will be defective is

$$\frac{15}{240} \cdot \frac{14}{239} = \frac{7}{1,912}$$

This assumes that we are sampling without replacement, namely, that the first tube is not replaced before the second tube is selected.

Theorem 1.7 can easily be generalized so that it applies to more than two events; for instance, for three events we have

Corollary:

If A, B, and C are any three events in a sample space S such that $P(A \cap B) > 0$, then

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

Proof

Writing $A \cap B \cap C$ as $(A \cap B) \cap C$ and using the formula of Theorem 2.7 twice, we get

$$P(A \cap B \cap C) = P[(A \cap B) \cap C]$$
$$= P(A \cap B) P(C|A \cap B]$$
$$= P(A) P(B|A) P(C|A \cap B)$$

Example 1.13

A box of fuses contains 20 fuses, of which 5 are defective. If 3 of the fuses are selected at random and removed from the box in succession without replacement, what is the probability that all three fuses are defective?

Solution

If D_1 is the event that the first fuse is defective, D_2 is the event that the second fuse is defective, and D_3 is the event that the third fuse is defective, then

$$P(D_1) = \frac{5}{20}, \qquad P(D_2|D_1) = \frac{4}{19}, \qquad P(D_3|D_2 \cap D_1) = \frac{3}{18}$$

and substitution into the formula yields

$$P(D_1 \cap D_2 \cap D_3) = \frac{5}{20} \cdot \frac{4}{19} \cdot \frac{3}{18} = \frac{1}{114}$$

Further generalization of Theorem 1.7 and its corollary to k events is straightforward, and the resulting formula can be proved by mathematical induction.

1.7 Total Probability and Bayes' Theorem

There are many situations where the outcome of an experiment depends on what happens in various intermediate stages. Suppose, in general the intermediate stage permits **k** different alternatives (whose occurrence is denoted by $B_1, B_2, ..., B_k$). The events $B_1, B_2, ..., B_k$, in this case, constitute a partition of the sample space **S** (i.e. the B's are pairwise mutually exclusive and their union equals S). To find the probability of occurrence of an event **A** that can occurs with one of the B's events, the following theorem is required, sometimes called the law or the rule of total probability.

Theorem 1.8 : Total Probability

If the events $B_1, B_2, ..., B_k$ constitute a partition of the sample space S and $P(B_i) > 0$ for i = 1, 2, ..., k, then for any event A in S

$$P(A) = \sum_{i=1}^{k} P(B_i) P(A \mid B_i) \qquad (1.5)$$

Proof

Since the events $B_1, B_2, ..., B_k$ are exhaustive (i.e their union equals S), then

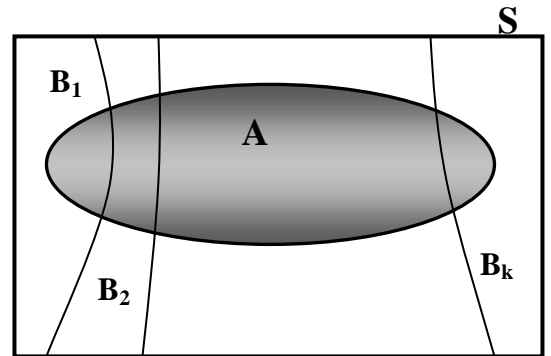$$A = A \cap S = (A \cap B_1) \cup (A \cap B_2) \cup ... \cup (A \cap B_k)$$

Since the events $B_1, B_2, ..., B_k$ are mutually exclusive, then $A \cap B_1$, $A \cap B_2$, ..., $A \cap B_k$ are also mutually exclusive. It follows that;

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i)$$

and the theorem results from applying (1.4) to each term in this summation.

Example 1.14

The members of a consulting firm rent cars from three rental agencies: 60% from

agency 1, 30 % from agency 2, and 10 % from agency 3. If 9% of the cars from agency 1 need a tune-up, 20 % of the cars from agency 2 need a tune-up, and 6% of the cars from agency 3 need a tune-up, what is the probability that a rental car delivered to the firm will need a tune-up?

Solution

If A is the event that the car needs a tune-up, and B_1, B_2, and B_3 are the events that the car comes from rental agencies 1, 2, or 3, we have

$$P(B_1) = 0.60, \qquad P(B_2) = 0.30, \qquad\qquad P(B_3) = 0.10, \qquad \text{and}$$
$$P(A \mid B_1) = 0.09, \quad P(A \mid B_2) = 0.20, \quad \text{and} \qquad P(A \mid B_3) = 0.06.$$

Substituting these values into the formula of Theorem 2.8, we get

$$P(A) = (0.60)(0.09) + (0.30)(0.20) + (0.10)(0.06) = 0.12$$

Thus, 12% of all the rental cars delivered to this firm will need a tune-up.

Tree diagram

It is sometimes helpful to illustrate this result with a **tree diagram** given in figure 1.1. The probability associated with branch B_i is $P(B_i)$, and the probability associated with each branch labeled **A** is a conditional probability $P(A| B_i)$, which may be different depending on which branch, B_i, it follows. In order for **A** to occur, it must occur jointly with one and only
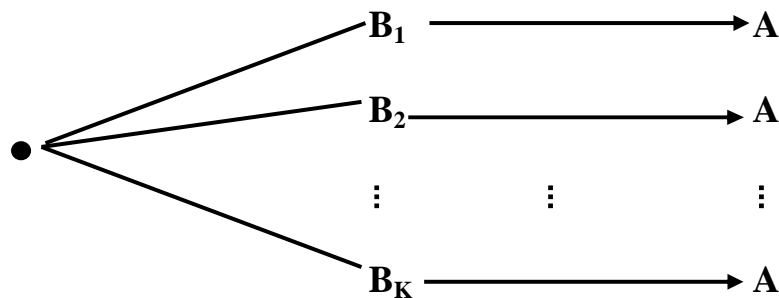
one of the events B_i (i.e their union equals S). With reference to the preceding example, suppose that we are interested in the following question: If a rental car delivered to the consulting firm needs a tune-up, what is the probability that it came from rental agency 2? To answer questions of this kind, we need the following theorem, called Bayes' theorem:

Theorem 1.9 : Bayes' Formula

If B_1, B_2, ..., B_k constitute a partition of the sample space S and $P(B_i) > 0$ for i = 1, 2, ..., k, then for any event A in S such that $P(A) > 0$,

$$P(B_r/A) = \frac{P(B_r) \cdot P(A|B_r)}{\sum\limits_{i=1}^{k} P(B_i) P(A|B_i)} \quad \text{for} \quad r = 1, 2, ..., k \qquad \text{(1.6)}$$

In words, the probability that event A was reached via the rah branch of the tree diagram of Figure 1.1, given that it was reached via one of its k branches, is the ratio of the probability associated with the rah branch to the sum of the probabilities associated with all k branches of the tree.

Proof

Writing in accordance with the definition of conditional probability, we have only to substitute $P(B_r) \cdot P(A|B_r)$ for $P(A \cap B_r)$ and the formula (1.5) for $P(A)$.

$$P(B_r|A) = \frac{P(A \cap B_r)}{P(A)}$$

Example 1.15

With reference to Example 1.14, if a rental car delivered to the consulting firm needs a tune-up, what is the probability that it came from rental agency 2?

Solution

Substituting the probabilities given in Example 1.14 into the formula of Theorem 1.9, we get

$$P(B_2|A) = \frac{(0.30)(0.20)}{(0.60)(0.09) + (0.30)(0.20) + (0.10)(0.06)}$$

$$= \frac{0.060}{0.120} = 0.5$$

Observe that although only 30% of the cars delivered to the firm come from agency 2, 50% of those requiring a tune-up come from that agency.

Example 1.16

We are given three similar boxes of microchips as follows:

Box B_1 contains 20 microchips of which 5 are defective,

Box B_2 contains 35 microchips of which 7 are defective, and

Box B_3 contains 40 microchips of which 5 are defective.

A box selected at random, then a microchip is selected at random from the box. If the component obtained is defective, find the probability that it came from box 2.

Solution

Defining the events B_1, B_2 and B_3 to be respectively choosing "box 1", "box 2" and

"box 3". Let A be the event "obtaining a defective microchip". Thus, we have

$$P(B_1) = P(B_2) = P(B_3) = 1/3$$

Box   **Microchip**
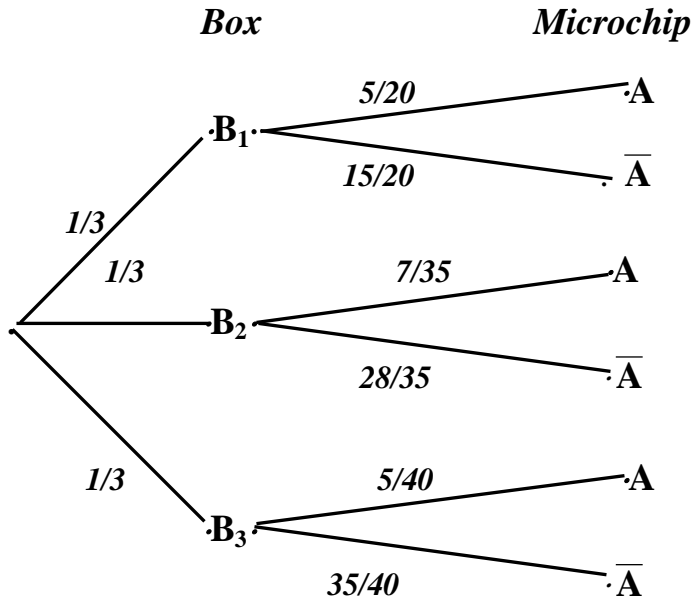
Therefore;

$$P(A) = \left(\frac{1}{3}\right)\left(\frac{5}{20}\right) + \left(\frac{1}{3}\right)\left(\frac{7}{35}\right) + \left(\frac{1}{3}\right)\left(\frac{5}{40}\right) = \frac{23}{120}$$

Now applying Bayes' formula (1.6), with k=3 and r=2 to find the probability that the component came from box 2 given that it is defective as follows

$$P(B_2 | A) = \frac{P(B_2)\,P(A | B_2)}{P(A)} = \frac{\left(\frac{1}{3}\right)\left(\frac{7}{35}\right)}{\frac{23}{120}} = \frac{8}{23} = 0.348$$

The tree diagram, given above, describes this process and gives the probability of each branch of the tree.

1.8 Independent Events

Informally speaking, two events A and B are said to be **independent** if the occurrence or nonoccurrence of either one does not affect the probability of the occurrence of the other.

Symbolically, two events A and B are independent if

$$P(B | A) = P(B) \qquad \text{and} \qquad P(A | B) = P(A)$$

and it can be shown that either of these equalities implies the other when both of the conditional probabilities exist, namely, when neither P(A) nor P(B) equals zero.

Now, if we substitute P(B) for P(B|A) into formula (1.3), we get

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

and this is used as the formal definition of independence.

Definition

Two events **A** and **B** are *independent* iff

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

otherwise, A and B are *dependent*.

It can be shown that if A and B are events such that $P(A) > 0$ and $P(B) > 0$, then A and B are independent iff either of the following holds

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

Example 1.17

A fair coin is tossed three times. If A is the event that a head occurs on each of the first two tosses, B is the event that a tail occurs on the third toss, and C is the event that exactly two tails occur in the tree tosses, show that

a- events A and B are independent;

b- events B and C are dependent.

Solution

Since the coin is fair then the eight possible outcomes, HHH, HHT, HTH, THH, HTT, THT, TTH, and TTT, are equally likely, and

A = {HHH, HHT}, B = {HHT, HTT, THT, TTT}, C = {HTT, THT, TTH}

A ∩ B = {HHT} and B ∩ C = {HTT, THT}

The assumption that the eight possible outcomes are all equiprobable yields

$P(A) = 1/4$, $P(B) = 1/2$, $P(C) = 3/8$, $P(A \cap B) = 1/8$ and $P(B \cap C) = 1/4$.

a- Since P(A). P(B) = 1/4.1/2 = P(A ∩ B), then the events A and B are independent.

b- Since P(B). P(C) = 1/2 .3/8 ≠ P(B ∩ C), the events B and C are not independent.

In connection with the definition of independence, given above, it can be shown that if A and B are independent, then so are A and $\overline{B}$, $\overline{A}$ and B, and $\overline{A}$ and $\overline{B}$. For instance:

To extend the concept of independence to more than two events, let us make the following definition.

Definition

Events A_1, A_2, ..., A_k are **independent** iff the probability of the intersection of any 2, 3, ... or k of these events equals the product of their respective probabilities.

For three events A, B, and C, for example, independence requires that;

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

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

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

and

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

It is of interest to note that three or more events can be **pairwise independent** without being independent.

A common example of dependent events occurs in connection with repeated **sampling without replacement**. In example 1.12 we considered the results of drawing fuses in succession. Suppose that, instead, the outcome of the first fuse is tested and then the fuse is replaced in the box before the second draw is made and the second fuse is replaced in the box before the third draw is made. This type of sampling is referred to as **sampling with replacement,** and it would reasonable to assume that the draws are independent trials. In this case the probability that all three fuses are defective is

$$P(A \cap B \cap C) = P(A) . P(B) . P(C) = \frac{5}{20} . \frac{5}{20} . \frac{5}{20} = \frac{1}{64}$$
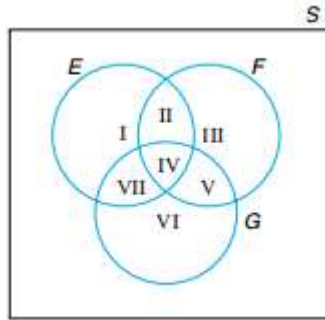
<+><+><+><+><+><+><+><+>

EXERCISES

[1] Let A, B and C be three arbitrary events. Find expressions for the following events that of A, B and C.

a- Only A occurs,   **b-** All three events occur,   **c-** Non occurs,

d- At least one occurs,   **e-** Not more than 2 occur,   **g-** One and only one occurs.

[2] For the following Venn diagram, describe in terms of E, F, and G the events denoted in the diagram by the Roman numerals I through VII.

$$I = E \cap \bar{F} \cap \bar{G} \text{ or } E \cap \overline{(F \cup G)} \qquad II = E \cap F \cap \bar{G}$$
$$III = \bar{E} \cap F \cap \bar{G} \qquad\qquad IV = E \cap F \cap G$$
$$V = \bar{E} \cap F \cap G \qquad\qquad VI = \bar{E} \cap \bar{F} \cap G$$
$$VII = E \cap \bar{F} \cap G$$

[3] Let $S = \{e_1, e_2, e_3, e_4\}$ be a sample space and let P be probability function defined on S,

a- Find $P(e_1)$ and $P(e_2)$ if $P(e_3) = P(e_4) = \frac{1}{4}$ and $P(e_1) = 2\,P(e_2)$.

b- Find $P(e_1)$ if $P(\{e_2, e_3\}) = 2/3$, $P(\{e_3, e_4\}) = \frac{1}{2}$ and $P(e_2) = 1/3$.

[4] An integer between 1 and 100 is selected at random, find the probability of getting a perfect square if

a- all integers are equally likely to be selected,

b- all integers between 1 and 50 are twice as likely to occur as the rest.

[5] Three students A, B and C are in a swimming race. A and B have the same probability of winning and each is twice as likely to win as C.

a- What is the probability that A does not win?

b- What is the probability that B or C wins?

[6] In a certain population of women 4% have had breast cancer, 20% are smokers and 3% are smokers and have had breast cancer. A woman is selected at random from the population. What is the probability that:

a- she has had breast cancer or smokes?

b- she has had breast cancer and she is not a smoker ?

[7] From a box containing 5 black balls and 3 green balls, 2 balls are drawn in succession, the first ball being replaced in the box before the second draw is made.

    a- What is the probability that both balls are the same color?

    b- What is the probability that each color is represented?

[8] Solve problem [6] in case of drawing the second ball without replacing the first one.

[9] The probability that a student, selected at random from a certain college, passes mathematics is 0.8, the probability that passes English is 0.6 and the probability of passing at least one of them is 0.9.

    a- What is the probability that he failed both Math. and English ?

    b- What is the probability that he failed Math only?

    c- If the student passed Math., what is the probability that he also passes English ?

    d- If the student failed Math., what is the probability that he also failed English ?

[10] If the three event A, B and C are independent with $P(A)=1/4$, $P(B) = 1/3$ and $P(C) = 1/5$, find the probability that at least one of the events *A, B* and *C* occurs if they are:
  (a) mutually exclusive;
  (b) independent.

[11] In a high school graduating class of 100 students, 54 studied mathematics, 69 studied English and 35 studied both mathematics and English. If one of these students is selected at random, find the probability that:

  a- the student took mathematics or English,

  b- the student did not take either of these subjects,

  c- the student took English but not mathematics.

[12] A regional telephone company operates three identical relay stations at different locations. During a one year period, the number of malfunctions reported by each station and the causes are shown below.

Station	A	B	C
Problems with electricity supplied	2	4	5
Computer malfunction	7	1	3
Malfunctioning electrical equipment	4	7	1
Caused by other human errors	2	2	5

Suppose that a malfunction was reported and it was found to be caused by other human errors. What is the probability that it came from station C?

[13] A factory has three machines M_1, M_2 and M_3 produce plastic gears. The output of machine M_1 is twice as the output of machine M_2 and the output of machine M_2 is triple as the output of machine M_3. The percentages of defective output of these machines are respectively 4%, 2% and 5%. A gear is selected at random and is found to be defective, find the probability that it was produced by machine M_2.

[14] A large industrial firm uses 3 local motels to provide overnight accommodations for its clients. From past experience it is known that 20%: of the clients are assigned rooms at the Ramada Inn, 50% at the Sheraton, and 30% at the Hilton. If the plumbing is faulty in 5% of the rooms at the Ramada Inn, in 4% of the rooms at the Sheraton, and in 8% of the rooms at the Hilton, what, is the probability that
(a) a client will be assigned a room with faulty plumbing?
(b) a person with a room having faulty plumbing was assigned accommodations at the Hilton?

[15] Fifty-two percent of the students at a certain college are females. Five percent of the students in this college are majoring in computer science. Two percent of the students are women majoring in computer science. If a student is selected at random, find the conditional probability that
(a) This student is female, given that the student is majoring in computer science;
(b) This student is majoring in computer science, given that the student is female.

[16] Police plan to enforce speed limits by using radar traps at four different locations within the city limits. The radar traps at each of the locations $L1$, $L2$, $L3$, and $L4$ will be operated 40%, 30%, 20%, and 10% of the time. If a person who is speeding on her way to work has probabilities of 0.2, 0.1, 0.5, and 0.2, respectively, of passing through these locations, what is the probability that she will receive a speeding ticket? If the person received a speeding ticket on her way to work, what is the probability that she passed through the radar trap located at $L2$?

[17] Let A and B events with $P(A) = ½$, $P(B) = 1/3$ and $P(A \cup B) = 7/12$. Find $P(A|B)$, $P(A|\overline{B})$, $P(\overline{A}|B)$ and $P(\overline{A}|\overline{B})$.

[18] Given $\mathbf{P(A)} = 0.5$ and $\mathbf{P(AUB)} = 0.8$, find $\mathbf{P(B)}$ if:
    a) A and **B** are mutually exclusive   **b) A** and **B** are independent
    c) P(A|B)=0.4.   **d) A** is a subset of **B**.

[19] At an electronic plant, it is known from past experience that the probability is 0.84 that

a new worker who has attended the company's training program will meet the production quota, and that the corresponding probability is 0.49 for a new worker who has not attended the company's training program. If 40% of all new workers attend the training program, what is the probability that a new worker will meet the production quota?

[20] When an experiment is performed, one and only one of the events A_1, A_2 or A_3 will occur. Find $P(A_1)$, $P(A_2)$, and $P(A_3)$ under each of the following assumptions:

    a- $P(A_1) = P(A_2) = P(A_3)$.   **b-** $P(A_1) = P(A_2)$ and $P(A_3) = 1/2$.

    c- $P(A_1) = 2\,P(A_2) = 3\,P(A_3)$.

[21] Multiple choices:

1. If the knowledge that an event A has occurred implies that a second event B cannot occur, the events A and B are said to be

    a- independent   **b-** mutually exclusive   **c-** A contains B   **d-** non of the above.

2. If P(A)=0.4, P(B)=0.5 and A & B are disjoint, then the probability of occurrence of both A and B is:

    a- 0.0   **b-** 0.1   **c-** 0.2   **d-** 0.9

3. If P(A)=0.4, P(B)=0.45 and $P(A \cap \overline{B}) = 0.25$ then $P(A \cup B)$ is :

    a- 0.7   **b-** 0.8   **c-** 0.9   **d-** non of the above.

4. If $P(A) = \frac{1}{2}$, $P(B) = 1/3$ and $P(A \cup B) = 2/3$ then the events A and B are:

    a- independent   **b-** mutually exclusive   **c-** dependent.   **d-** non of the above.

5. If $P(\overline{A}) = 0.6$ and $P(\overline{A} \cap B) = 0.5$ then the probability of neither A nor B occurs is:

    a- 5/6   **b-** 0.1   **c-** 0.05   **d-** non of the above.

6. If $P(A) = P(B) = 0.5$ and $P(A \cup B) = 0.75$, then A and B are:

    a. Independent   **b.** Mutually exclusive   **c.** dependent   **d.** $A \subset B$

7. If P(A|B)=1, then A & B are:

    a. independent   **b.** $B \subset A$ **c.** $A \subset B$ **d.** non of the above.

8. If $P(\overline{A}|B) = 1$, then A & B are:

    a. independent   **b.** $B \subset A$ **c.** $A \subset B$ **d.** Mutually exclusive

9. If $P(B|\overline{A}) = 0$, then A & B are:

    a. independent   **b.** $B \subset A$ **c.** $A \subset B$ **d.** Mutually exclusive

<center><+><+><+><+><+><+><+><+></center>