4 Introduction To Probability

CHAPTER CONTENTS

LEARNING OBJECTIVES

After reading this chapter and doing the exercises, you should be able to:

1 Appreciate the role probability information plays in the decision-making process.

2 Understand probability as a numerical measure of the likelihood of occurrence.

3 Appreciate the three methods commonly used for assigning probabilities and understand when they should be used.

4 Use the laws that are available for computing the probabilities of events.

5 Understand how new information can be used to revise initial (prior) probability estimates using Bayes' theorem.

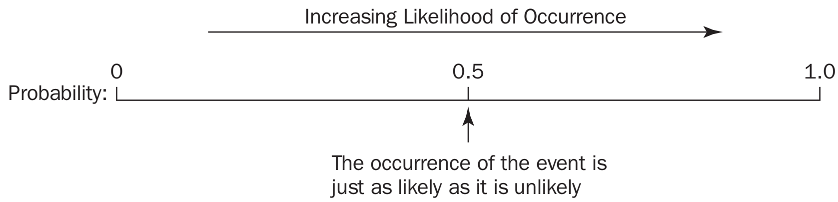
Managers often base their decisions on an analysis of uncertainties such as the following:

1. What are the chances that sales will decrease if we increase prices?
2. What is the likelihood a new assembly method will increase productivity?
3. How likely is it that the project will be finished on time?
4. What is the chance that a new investment will be profitable?

Probability is a numerical measure of the likelihood that an event will occur. Thus, probabilities can be used as measures of the degree of uncertainty associated with the four events previously listed. If probabilities are available, we can determine the likelihood of each event occurring.

Probability values are always assigned on a scale from 0 to 1. A probability near zero indicates an event is unlikely to occur; a probability near 1 indicates an event is almost certain to occur. Other probabilities between 0 and 1 represent degrees of likelihood that an event will occur. For example, if we consider the event ‘rain tomorrow’, we understand that when the weather report indicates ‘a near-zero probability of rain’, it means almost no chance of rain. However, if a 0.90 probability of rain is reported, we know that rain is likely to occur. A 0.50 probability indicates that rain is just as likely to occur as not. Figure 4.1 depicts the view of probability as a numerical measure of the likelihood of an event occurring.

FIGURE 4.1
Probability as a numerical measure of the likelihood of an event occurring



STATISTICS IN PRACTICE
Combating junk email

Junk email remains a major Internet scourge. In April 2012 it was estimated 77.2 per cent of electronic mail worldwide was spam (unsolicited commercial email).[1] In 2011, India, Russia and Vietnam accounted for more than 30 per cent of it.[2] In the past, spam has been inextricably linked to the spread of malware on the Web and indeed a significant proportion of spam was botnet-generated. Spam is often associated with porn – the notorious Facebook attack in 2011 being entirely in keeping with this growing phenomenon.[3] Spam is time-consuming to deal with and an increasing brake on further email take-up and usage.

Various initiatives have been undertaken to help counter the problem. However, determining which messages are ‘good’ and which are ‘spam’ is difficult to establish even with the most sophisticated spam filters (spam-busters). One of the earliest and most effective techniques for dealing with spam is the adaptive Naïve Bayes’ method which exploits the probability relationship

$$P(\text{spam} \mid \text{message}) = \frac{P(\text{message} \mid \text{spam}) \;\; P(\text{spam})}{P(\text{message})}$$

$$\text{where} \quad P(\text{message}) = P(\text{message} \mid \text{spam}) \;\; P(\text{spam}) + P(\text{message} \mid \text{good}) \;\; P(\text{good})$$

Here:

P(spam) is the prior probability a message is spam based on past experience,

P(message | spam) is estimated from a training corpus (a set of messages known to be good or spam) on the (naïve) assumption that every word in the message is independent of every other so that:

$$P(\text{message} \mid \text{spam}) = P(\text{first word} \mid \text{spam}) \; P(\text{second word} \mid \text{spam}) \ldots \; P(\text{last word} \mid \text{spam})$$

Similarly:

$$P(\text{message} \mid \text{good}) = P(\text{first word} \mid \text{good}) \; P(\text{second word} \mid \text{good}) \ldots \; P(\text{last word} \mid \text{good})$$

Advantages of Naïve Bayes are its simplicity and ease of implementation. Indeed it is often found to be very effective – even compared to methods based on more complex modelling procedures.

1 www.kaspersky.co.uk/about/news/spam/2012/Spam_in_April_2012_Junk_Mail_Gathers_Pace_in_the_US
2 www.cisco.com/en/US/prod/collateral/vpndevc/security_annual_report_2011.pdf
3 http://mashable.com/2011/11/15/facebook-spam-porn/

4.1 EXPERIMENTS, COUNTING RULES AND ASSIGNING PROBABILITIES

We define an **experiment** as a process that generates well-defined outcomes. On any single repetition of an experiment, one and only one of the possible experimental outcomes will occur. Several examples of experiments and their associated outcomes follow.

Experiment	*Experimental outcomes*
Toss a coin	Head, tail
Select a part for inspection	Defective, non-defective
Conduct a sales call	Purchase, no purchase
Role a die	1, 2, 3, 4, 5, 6
Play a football game	Win, lose, draw

By specifying all possible experimental outcomes, we identify the **sample space** for an experiment.

Sample space

The sample space for an experiment is the set of all experimental outcomes.

An experimental outcome is also called a **sample point** to identify it as an element of the sample space.

Consider the first experiment in the preceding table – tossing a coin. The upward face of the coin – a head or a tail – determines the experimental outcomes (sample points). If we let S denote the sample space, we can use the following notation to describe the sample space.

$$S = \{\text{Head, Tail}\}$$

The sample space for the second experiment in the table – selecting a part for inspection – can be described as follows

$$S = \{\text{Defective, Non-defective}\}$$

Both of the experiments just described have two experimental outcomes (sample points). However, suppose we consider the fourth experiment listed in the table – rolling a die. The possible experimental outcomes, defined as the number of dots appearing on the upward face of the die, are the six points in the sample space for this experiment.

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

Counting rules, combinations and permutations

Being able to identify and count the experimental outcomes is a necessary step in assigning probabilities. We now discuss three useful counting rules.

Multiple-step experiments

The first counting rule applies to multiple-step experiments. Consider the experiment of tossing two coins. Let the experimental outcomes be defined in terms of the pattern of heads and tails appearing on the upward faces of the two coins. How many experimental outcomes are possible for this experiment?

The experiment of tossing two coins can be thought of as a two-step experiment in which step 1 is the tossing of the first coin and step 2 is the tossing of the second coin. If we use H to denote a head and T to denote a tail, (H, H) indicates the experimental outcome with a head on the first coin and a head on the second coin. Continuing this notation, we can describe the sample space (S) for this coin-tossing experiment as follows:

$$S = \{(H, H), (H, T), (T, H), (T, T)\}$$

Thus, we see that four experimental outcomes are possible. In this case, we can easily list all of the experimental outcomes.

The counting rule for multiple-step experiments makes it possible to determine the number of experimental outcomes without listing them.

A counting rule for multiple-step experiments

If an experiment can be described as a sequence of k steps with n_1 possible outcomes on the first step, n_2 possible outcomes on the second step and so on, then the total number of experimental outcomes is given by:

$$n_1 \times n_2 \times \ \dots \ \times n_k$$

Viewing the experiment of tossing two coins as a sequence of first tossing one coin ($n_1 = 2$) and then tossing the other coin ($n_2 = 2$), we can see from the counting rule that there are $2 \times 2 = 4$ distinct experimental outcomes. They are $S = \{(H, H), (H, T), (T, H), (T, T)\}$. The number of experimental outcomes in an experiment involving tossing six coins is $2 \times 2 \times 2 \times 2 \times 2 \times 2 = 64$.

A **tree diagram** is a graphical representation that helps in visualizing a multiple-step experiment. Figure 4.2 shows a tree diagram for the experiment of tossing two coins. The sequence of steps moves from left to right through the tree. Step 1 corresponds to tossing the first coin, and step 2 corresponds to tossing the second coin. For each step, the two possible outcomes are head or tail. Note that for each possible outcome at step 1 two branches correspond to the two possible outcomes at step 2. Each of the points on the right end of the tree corresponds to an experimental outcome. Each path through the tree from the leftmost node to one of the nodes at the right side of the tree corresponds to a unique sequence of outcomes.

Let us now see how the counting rule for multiple-step experiments can be used in the analysis of a capacity expansion project for Kristof Projects Limited (KPL). KPL is starting a project designed to increase the generating capacity of one of its plants in southern Norway. The project is divided into two sequential stages or steps: stage 1 (design) and stage 2 (construction). Even though each stage will be scheduled and controlled as closely as possible, management cannot predict beforehand the exact time required to complete each stage of the project. An analysis of similar construction projects revealed possible completion times for the design stage of two, three or four months and possible completion times for the construction stage of six, seven or eight months.

FIGURE 4.2
Tree diagram for the experiment of tossing two coins

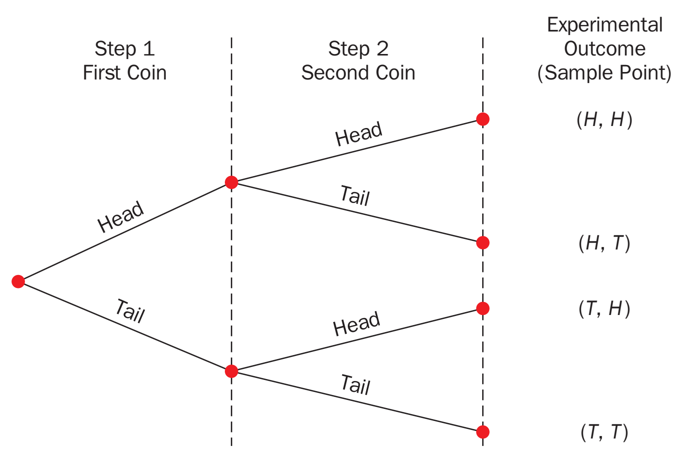


TABLE 4.1 Experimental outcomes (sample points) for the KPL project

Completion time (months)			
Stage 1 Design	**Stage 2 Construction**	**Notation for experimental outcome**	**Total project completion time (months)**
2	6	(2, 6)	8
2	7	(2, 7)	9
2	8	(2, 8)	10
3	6	(3, 6)	9
3	7	(3, 7)	10
3	8	(3, 8)	11
4	6	(4, 6)	10
4	7	(4, 7)	11
4	8	(4, 8)	12

In addition, because of the critical need for additional electrical power, management set a goal of ten months for the completion of the entire project.

Because this project has three possible completion times for the design stage (step 1) and three possible completion times for the construction stage (step 2), the counting rule for multiple-step experiments can be applied here to determine a total of $3 \times 3 = 9$ experimental outcomes. To describe the experimental outcomes, we use a two-number notation; for instance, (2, 6) indicates that the design stage is completed in two months and the construction stage is completed in six months. This experimental outcome results in a total of $2 + 6 = 8$ months to complete the entire project. Table 4.1 summarizes the nine experimental outcomes for the KPL problem. The tree diagram in Figure 4.3 shows how the nine outcomes (sample points) occur.

The counting rule and tree diagram help the project manager identify the experimental outcomes and determine the possible project completion times. We see that the project will be completed in 8 to 12 months, with six of the nine experimental outcomes providing the desired completion time of ten months or less. Even though identifying the experimental outcomes may be helpful, we need to consider how probability values can be assigned to the experimental outcomes before making an assessment of the probability that the project will be completed within the desired ten months.

Combinations

A second useful counting rule allows one to count the number of experimental outcomes when the experiment involves selecting n objects from a (usually larger) set of N objects. It is called the counting rule for combinations.

Counting rule for combinations

The number of combinations of N objects taken n at a time is:

$$^{N}C_{n} = \binom{N}{n} = \frac{N!}{n!(N-n)!}$$

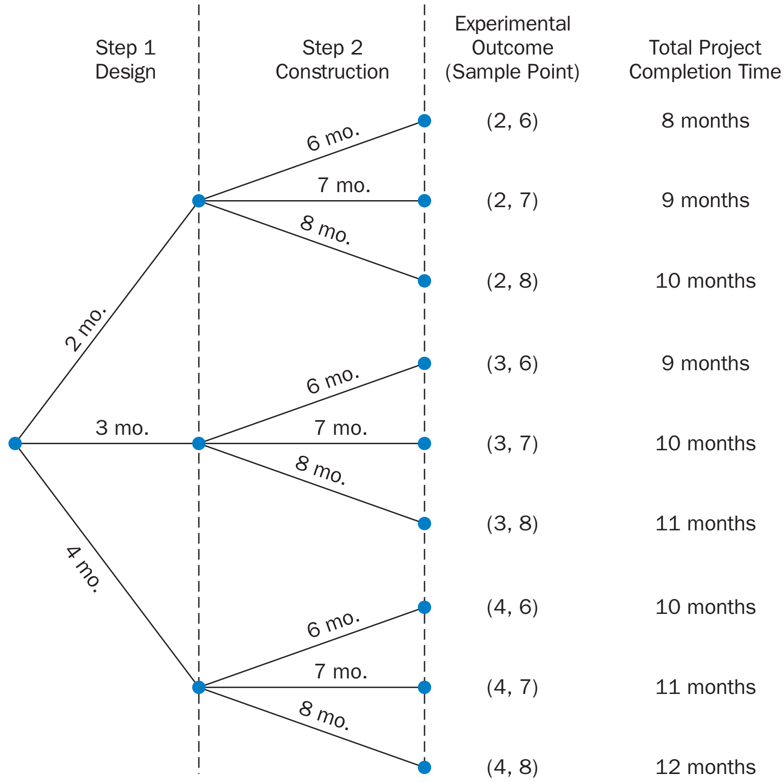
where:

$$N! = N \times (N-1) \times (N-2) \times \ldots \times (2) \times (1)$$
$$n! = n \times (n-1) \times (n-2) \times \ldots \times (2) \times (1)$$

and, by definition:

$$0! = 1$$

FIGURE 4.3
Tree diagram for the KPL project



The notation ! means *factorial*: for example, 5 factorial is $5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$.

Consider a quality control procedure in which an inspector randomly selects two of five parts to test for defects. In a group of five parts, how many combinations of two parts can be selected? The counting rule in equation (4.1) shows that with $N = 5$ and $n = 2$, we have:

$$^{5}C_{2} = \binom{5}{2} = \frac{5 \times 4 \times 3 \times 2 \times 1}{(2 \times 1) \times (3 \times 2 \times 1)} = \frac{120}{12} = 10 \qquad \textbf{(4.1)}$$

Thus, ten outcomes are possible for the experiment of randomly selecting two parts from a group of five. If we label the five parts as A, B, C, D and E, the ten combinations or experimental outcomes can be identified as AB, AC, AD, AE, BC, BD, BE, CD, CE and DE.

As another example, consider that the Spanish Lotto 6–49 system uses the random selection of six integers from a group of 49 to determine the weekly lottery winner. The counting rule for combinations, equation (4.1), can be used to determine the number of ways six different integers can be selected from a group of 49.

$$\binom{49}{6} = \frac{49!}{6!(49-6)!} = \frac{49!}{6!43!} = \frac{49 \times 48 \times 47 \times 46 \times 45 \times 44}{6 \times 5 \times 4 \times 3 \times 2 \times 1} = 13\,983\,816$$

The counting rule for combinations tells us that more than 13 million experimental outcomes are possible in the lottery drawing. An individual who buys a lottery ticket has one chance in 13 983 816 of winning.

Permutations

A third counting rule that is sometimes useful is the counting rule for permutations. It allows one to compute the number of experimental outcomes when n objects are to be selected from a set of N objects where the order of selection is important. The same n objects selected in a different order is considered a different experimental outcome.

Counting rule for permutations

The number of permutations of N objects taken at n is given by:

$$^{N}P_{n} = n!\binom{N}{n} = \frac{N!}{(N-n)!}$$

The counting rule for permutations closely relates to the one for combinations; however, an experiment results in more permutations than combinations for the same number of objects because every selection of n objects can be ordered in $n!$ different ways.

As an example, consider again the quality control process in which an inspector selects two of five parts to inspect for defects. How many permutations may be selected? The counting rule in equation (4.2) shows that with $N = 5$ and $n = 2$, we have:

$$^{5}P_{2} = \frac{5!}{(5-2)!} = \frac{5!}{3!} = \frac{5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1} = \frac{120}{6} = 20 \quad \textbf{(4.2)}$$

Thus, 20 outcomes are possible for the experiment of randomly selecting two parts from a group of five when the order of selection must be taken into account. If we label the parts A, B, C, D and E, the 20 permutations are AB, BA, AC, CA, AD, DA, AE, EA, BC, CB, BD, DB, BE, EB, CD, DC, CE, EC, DE and ED.

Assigning probabilities

Now let us see how probabilities can be assigned to experimental outcomes. The three approaches most frequently used are the classical, relative frequency and subjective methods. Regardless of the method used, two **basic requirements for assigning probabilities** must be met.

Basic requirements for assigning probabilities

1. The probability assigned to each experimental outcome must be between 0 and 1, inclusively. If we let E_i denote the ith experimental outcome and $P(E_i)$ its probability, then this requirement can be written as:

$$0 \leq P(E_i) \leq 1) \text{ for all } i \quad \textbf{(4.3)}$$

2. The sum of the probabilities for all the experimental outcomes must equal 1.0. For n experimental outcomes, this requirement can be written as:

$$P(E_1) + P(E_2) + \ldots + P(E_n) = 1 \quad \textbf{(4.4)}$$

The **classical method** of assigning probabilities is appropriate when all the experimental outcomes are equally likely. If n experimental outcomes are possible, a probability of $1/n$ is assigned to each experimental outcome. When using this approach, the two basic requirements for assigning probabilities are automatically satisfied.

For example, consider the experiment of tossing a fair coin: the two experimental outcomes – head and tail – are equally likely. Because one of the two equally likely outcomes is a head, the probability of observing a head is 1/2 or 0.50. Similarly, the probability of observing a tail is also 1/2 or 0.50.

As another example, consider the experiment of rolling a die. It would seem reasonable to conclude that the six possible outcomes are equally likely, and hence each outcome is assigned a probability of 1/6. If $P(1)$ denotes the probability that one dot appears on the upward face of the die, then $P(1) = 1/6$. Similarly, $P(2) = 1/6$, $P(3) = 1/6$, $P(4) = 1/6$, $P(5) = 1/6$ and $P(6) = 1/6$. Note that these probabilities satisfy the two basic requirements of equations (4.3) and (4.4) because each of the probabilities is greater than or equal to zero and they sum to 1.0.

The **relative frequency method** of assigning probabilities is appropriate when data are available to estimate the proportion of the time the experimental outcome will occur if the experiment is repeated a large number of times. As an example, consider a study of waiting times in the X-ray department for a local hospital. A clerk recorded the number of patients waiting for service at 9:00 a.m. on 20 successive days, and obtained the following results.

Number waiting	*Number of days outcome occurred*
0	2
1	5
2	6
3	4
4	3
	Total = 20

These data show that on two of the 20 days, zero patients were waiting for service; on five of the days, one patient was waiting for service and so on. Using the relative frequency method, we would assign a probability of $2/20 = 0.10$ to the experimental outcome of zero patients waiting for service, $5/20 = 0.25$ to the experimental outcome of one patient waiting, $6/20 = 0.30$ to two patients waiting, $4/20 = 0.20$ to three patients waiting and $3/20 = 0.15$ to four patients waiting. As with the classical method, using the relative frequency method automatically satisfies the two basic requirements of equations (4.3) and (4.4).

The **subjective method** of assigning probabilities is most appropriate when one cannot realistically assume that the experimental outcomes are equally likely and when little relevant data are available. When the subjective method is used to assign probabilities to the experimental outcomes, we may use any information available, such as our experience or intuition. After considering all available information, a probability value that expresses our *degree of belief* (on a scale from 0 to 1) that the experimental outcome will occur, is specified. Because subjective probability expresses a person's degree of belief, it is personal. Using the subjective method, different people can be expected to assign different probabilities to the same experimental outcome.

The subjective method requires extra care to ensure that the two basic requirements of equations (4.3) and (4.4) are satisfied. Regardless of a person's degree of belief, the probability value assigned to each experimental outcome must be between 0 and 1, inclusive, and the sum of all the probabilities for the experimental outcomes must equal 1.0.

Consider the case in which Tomas and Margit Elsbernd make an offer to purchase a house. Two outcomes are possible:

$$E_1 = \textit{their offer is accepted}$$
$$E_2 = \textit{their offer is rejected}$$

Margit believes that the probability their offer will be accepted is 0.8; thus, Margit would set $P(E_1) = 0.8$ and $P(E_2) = 0.2$. Tomas, however, believes that the probability that their offer will be accepted is 0.6; hence, Tomas would set $P(E_1) = 0.6$ and $P(E_2) = 0.4$. Note that Tomas' probability estimate for E_1 reflects a greater pessimism that their offer will be accepted.

Both Margit and Tomas assigned probabilities that satisfy the two basic requirements. The fact that their probability estimates are different emphasizes the personal nature of the subjective method.

Even in business situations where either the classical or the relative frequency approach can be applied, managers may want to provide subjective probability estimates. In such cases, the best probability estimates often are obtained by combining the estimates from the classical or relative frequency approach with subjective probability estimates.

Probabilities for the KPL project

To perform further analysis on the KPL project, we must develop probabilities for each of the nine experimental outcomes listed in Table 4.1. On the basis of experience and judgement, management concluded that the experimental outcomes were not equally likely. Hence, the classical method of assigning probabilities could not be used. Management then decided to conduct a study of the completion times for similar projects undertaken by KPL over the past three years. The results of a study of 40 similar projects are summarized in Table 4.2.

After reviewing the results of the study, management decided to employ the relative frequency method of assigning probabilities. Management could have provided subjective probability estimates, but felt that the current project was quite similar to the 40 previous projects. Thus, the relative frequency method was judged best.

In using the data in Table 4.2 to compute probabilities, we note that outcome (2, 6) – stage 1 completed in two months and stage 2 completed in six months – occurred six times in the 40 projects. We can use the relative frequency method to assign a probability of $6/40 = 0.15$ to this outcome. Similarly, outcome (2, 7) also occurred in six of the 40 projects, providing a $6/40 = 0.15$ probability. Continuing in this manner, we obtain the probability assignments for the sample points of the KPL project shown in Table 4.3.

TABLE 4.2 Completion results for 40 KPL projects

Completion times (months)			
Stage 1 Design	**Stage 2 Construction**	**Sample point**	**Number of past projects having these completion times**
2	6	(2, 6)	6
2	7	(2, 7)	6
2	8	(2, 8)	2
3	9	(3, 6)	4
3	7	(3, 7)	8
3	8	(3, 8)	2
4	6	(4, 6)	2
4	7	(4, 7)	4
4	8	(4, 8)	6
			Total = 40

TABLE 4.3 Probability assignments for the KPL project based on the relative frequency method

Sample point	Project completion time	Probability of sample point
(2, 6)	8 months	$P(2, 6) = 6/40 = 0.15$
(2, 7)	9 months	$P(2, 7) = 6/40 = 0.15$
(2, 8)	10 months	$P(2, 8) = 2/40 = 0.05$
(3, 6)	9 months	$P(3, 6) = 4/40 = 0.10$
(3, 7)	10 months	$P(3, 7) = 8/40 = 0.20$
(3, 8)	11 months	$P(3, 8) = 2/40 = 0.05$
(4, 6)	10 months	$P(4, 6) = 2/40 = 0.05$
(4, 7)	11 months	$P(4, 7) = 4/40 = 0.10$
(4, 8)	12 months	$P(4, 8) = 6/40 = 0.15$
		Total 1.00

Note that $P(2, 6)$ represents the probability of the sample point (2, 6), $P(2, 7)$ represents the probability of the sample point (2, 7) and so on.

EXERCISES

Methods

1. An experiment has three steps with three outcomes possible for the first step, two outcomes possible for the second step and four outcomes possible for the third step. How many experimental outcomes exist for the entire experiment?

COMPLETE SOLUTIONS

2. How many ways can three items be selected from a group of six items? Use the letters A, B, C, D, E and F to identify the items, and list each of the different combinations of three items.

3. How many permutations of three items can be selected from a group of six? Use the letters A, B, C, D, E and F to identify the items, and list each of the permutations of items B, D and F.

COMPLETE SOLUTIONS

4. Consider the experiment of tossing a coin three times.

a. Develop a tree diagram for the experiment.
b. List the experimental outcomes.
c. What is the probability for each experimental outcome?

5. Suppose an experiment has five equally likely outcomes: E_1, E_2, E_3, E_4, E_5. Assign probabilities to each outcome and show that the requirements in equations (4.3) and (4.4) are satisfied. What method did you use?

6. An experiment with three outcomes has been repeated 50 times, and it was learned that E_1 occurred 20 times, E_2 occurred 13 times and E_3 occurred 17 times. Assign probabilities to the outcomes. What method did you use?

COMPLETE SOLUTIONS

7. A decision-maker subjectively assigned the following probabilities to the four outcomes of an experiment: $P(E_1) = 0.10$, $P(E_2) = 0.15$, $P(E_3) = 0.40$ and $P(E_4) = 0.20$. Are these probability assignments valid? Explain.

8. Applications for zoning changes in a large metropolitan city go through a two-step process: a review by the planning commission and a final decision by the city council. At step 1 the planning commission reviews the zoning change request and makes a positive or negative recommendation concerning the change. At step 2 the city council reviews the planning commission's recommendation and then votes to approve or to disapprove the zoning change. Suppose the developer of an apartment complex submits an application for a zoning change. Consider the application process as an experiment.

a. How many sample points are there for this experiment? List the sample points.
b. Construct a tree diagram for the experiment.

9. A total of 11 Management students, four International Management and American Business Studies (IMABS) and eight International Management and French Studies (IMF) students have volunteered to take part in an inter-university tournament.

a. How many different ways can a team consisting of eight Management students, two IMABS and five IMF students be selected?
b. If after the team has been selected, one Management, one IMABS and two IMF students are found to be suffering from glandular fever and are unable to play, what is the probability that the team will not have to be changed?

COMPLETE SOLUTIONS

10. A company that franchises coffee houses conducted taste tests for a new coffee product. Four blends were prepared, then randomly chosen individuals were asked to taste the blends and state which one they liked best. Results of the taste test for 100 individuals are given.

Blend	*Number choosing*
1	20
2	30
3	35
4	15

a. Define the experiment being conducted. How many times was it repeated?
b. Prior to conducting the experiment, it is reasonable to assume preferences for the four blends are equal. What probabilities would you assign to the experimental outcomes prior to conducting the taste test? What method did you use?
c. After conducting the taste test, what probabilities would you assign to the experimental outcomes? What method did you use?

11. A company that manufactures toothpaste is studying five different package designs. Assuming that one design is just as likely to be selected by a consumer as any other design, what selection probability would you assign to each of the package designs? In an actual experiment, 100 consumers were asked to pick the design they preferred. The following data were obtained. Do the data confirm the belief that one design is just as likely to be selected as another? Explain.

Design times	*Number of preferred*
1	5
2	15
3	30
4	40
5	10

4.2 EVENTS AND THEIR PROBABILITIES

In the introduction to this chapter we used the term *event* much as it would be used in everyday language. Then, in Section 4.1 we introduced the concept of an experiment and its associated experimental outcomes or sample points. Sample points and events provide the foundation for the study of probability. We must now introduce the formal definition of an **event** as it relates to sample points. Doing so will provide the basis for determining the probability of an event.

Event

An event is a collection of sample points.

For example, let us return to the KPL project and assume that the project manager is interested in the event that the entire project can be completed in ten months or less. Referring to Table 4.3, we see that six sample points – (2, 6), (2, 7), (2, 8), (3, 6), (3, 7) and (4, 6) – provide a project completion time of ten months or less. Let C denote the event that the project is completed in ten months or less; we write

$$C = \{(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6)\}$$

Event C is said to occur if *any one* of these six sample points appears as the experimental outcome.

Other events that might be of interest to KPL management include the following:

$L =$ The event that the project is completed in *less* than ten months

$M =$ The event that the project is completed in *more* than ten months

Using the information in Table 4.3, we see that these events consist of the following sample points:

$$L = \{(2, 6), (2, 7), (3, 6)\}$$

$$M = \{3, 8), (4, 7), (4, 8)\}$$

A variety of additional events can be defined for the KPL project, but in each case the event must be identified as a collection of sample points for the experiment.

Given the probabilities of the sample points shown in Table 4.3, we can use the following definition to compute the probability of any event that KPL management might want to consider.

Probability of an event

The probability of any event is equal to the sum of the probabilities of the sample points for the event.

Using this definition, we calculate the probability of a particular event by adding the probabilities of the sample points (experimental outcomes) that make up the event. We can now compute the probability that the project will take ten months or less to complete. Because this event is given by $C = \{(2, 6), (2, 7), (2, 8), (3, 6), (3, 7), (4, 6)\}$, the probability of event C, denoted $P(C)$, is given by:

$$\begin{aligned} P(C) &= P(2, 6) + P(2, 7) + P(2, 8) + P(3, 6) + P(3, 7) + P(4, 6) \\ &= 0.15 + 0.15 + 0.05 + 0.10 + 0.20 + 0.05 = 0.70 \end{aligned}$$

Similarly, because the event that the project is completed in less than ten months is given by $L = \{(2, 6), (2, 7), (3, 6)\}$, the probability of this event is given by:

$$\begin{aligned} P(L) &= P(2, 6) + P(2, 7) + (3, 6) \\ &= 0.15 + 0.15 + 0.10 = 0.40 \end{aligned}$$

Finally, for the event that the project is completed in more than ten months, we have $M = \{(3, 8), (4, 7), (4, 8)\}$ and thus:

$$\begin{aligned} P(M) &= P(3, 8) + P(4, 7) + (4, 8) \\ &= 0.05 + 0.10 + 0.15 = 0.30 \end{aligned}$$

Using these probability results, we can now tell KPL management that there is a 0.70 probability that the project will be completed in ten months or less, a 0.40 probability that the project will be completed in less than ten months, and a 0.30 probability that the project will be completed in more than ten months. This procedure of computing event probabilities can be repeated for any event of interest to the KPL management.

Any time that we can identify all the sample points of an experiment and assign probabilities to each, we can compute the probability of an event using the definition. However, in many experiments the large number of sample points makes the identification of the sample points, as well as the determination of their associated probabilities, extremely cumbersome, if not impossible. In the remaining sections of this chapter, we present some basic probability relationships that can be used to compute the probability of an event without knowledge of all the sample point probabilities.

EXERCISES

Methods

12. An experiment has four equally likely outcomes: E_1, E_2, E_3 and E_4.

- **a.** What is the probability that E_2 occurs?
- **b.** What is the probability that any two of the outcomes occur (e.g. E_1 or E_3)?
- **c.** What is the probability that any three of the outcomes occur (e.g. E_1 or E_2 or E_4)?

13. Consider the experiment of selecting a playing card from a deck of 52 playing cards. Each card corresponds to a sample point with a 1/52 probability.

- **a.** List the sample points in the event an ace is selected.
- **b.** List the sample points in the event a club is selected.
- **c.** List the sample points in the event a face card (jack, queen or king) is selected.
- **d.** Find the probabilities associated with each of the events in parts (a), (b) and (c).

COMPLETE SOLUTIONS

14. Consider the experiment of rolling a pair of dice. Suppose that we are interested in the sum of the face values showing on the dice.

- **a.** How many sample points are possible? (*Hint:* Use the counting rule for multiple-step experiments.)
- **b.** List the sample points.
- **c.** What is the probability of obtaining a value of 7?
- **d.** What is the probability of obtaining a value of 9 or greater?
- **e.** Because each roll has six possible even values (2, 4, 6, 8, 10 and 12) and only five possible odd values (3, 5, 7, 9 and 11), the dice should show even values more often than odd values. Do you agree with this statement? Explain.
- **f.** What method did you use to assign the probabilities requested?

Applications

15. Refer to the KPL sample points and sample point probabilities in Tables 4.2 and 4.3.

- **a.** The design stage (stage 1) will run over budget if it takes four months to complete. List the sample points in the event the design stage is over budget.
- **b.** What is the probability that the design stage is over budget?
- **c.** The construction stage (stage 2) will run over budget if it takes eight months to complete. List the sample points in the event the construction stage is over budget.
- **d.** What is the probability that the construction stage is over budget?
- **e.** What is the probability that both stages are over budget?

COMPLETE SOLUTIONS

16. Suppose that a manager of a large apartment complex provides the following subjective probability estimates about the number of vacancies that will exist next month.

Vacancies	*Probability*
0	0.10
1	0.15
2	0.30
3	0.20
4	0.15
5	0.10

Provide the probability of each of the following events.

- **a.** No vacancies.
- **b.** At least four vacancies.
- **c.** Two or fewer vacancies.

17. When three marksmen take part in a shooting contest, their chances of hitting the target are 1/2, 1/3 and 1/4 respectively. If all three marksmen fire at it simultaneously

a. What is the chance that one and only one bullet will hit the target?
b. What is the chance that two marksmen will hit the target (and therefore one will not)?
c. What is the chance that all three marksmen will hit the target?

4.3 SOME BASIC RELATIONSHIPS OF PROBABILITY

Complement of an event

Given an event A, the **complement of A** is defined to be the event consisting of all sample points that are *not* in A. The complement of A is denoted by $\bar{A}$. Figure 4.4 is a diagram, known as a **Venn diagram**, which illustrates the concept of a complement. The rectangular area represents the sample space for the experiment and as such contains all possible sample points. The circle represents event A and contains only the sample points that belong to A. The shaded region of the rectangle contains all sample points not in event A, and is by definition the complement of A.

In any probability application, either event A or its complement $\bar{A}$ must occur. Therefore, we have:

$$P(A) + P(\bar{A}) = 1$$

Solving for $P(A)$, we obtain the following result.

Computing probability using the complement

$$P(A) = 1 - P(\bar{A}) \tag{4.5}$$

Equation (4.5) shows that the probability of an event A can be computed easily if the probability of its complement, $P(\bar{A})$, is known.

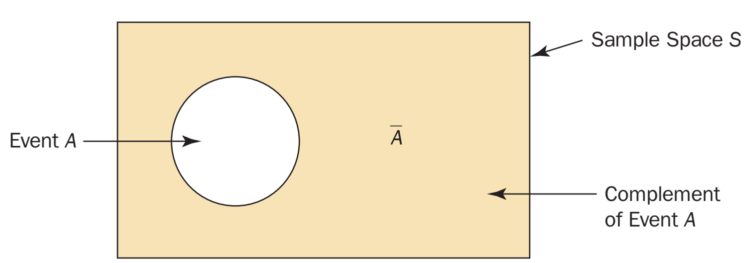
As an example, consider the case of a sales manager who, after reviewing sales reports, states that 80 per cent of new customer contacts result in no sale. By allowing A to denote the event of a sale and $\bar{A}$ to denote the event of no sale, the manager is stating that $P(\bar{A}) = 0.80$. Using equation (4.5), we see that:

$$P(A) = 1 - P(\bar{A}) = 1 - 0.80 = 0.20$$

We can conclude that a new customer contact has a 0.20 probability of resulting in a sale.

In another example, a purchasing agent states a 0.90 probability that a supplier will send a shipment that is free of defective parts. Using the complement, we can conclude that there is a $1 - 0.90 = 0.10$ probability that the shipment will contain defective parts.

FIGURE 4.4
Complement of event A is shaded



Addition law

The addition law is helpful when we are interested in knowing the probability that at least one of two events occurs. That is, with events A and B we are interested in knowing the probability that event A or event B or both occur.

Before we present the addition law, we need to discuss two concepts related to the combination of events: the *union* of events and the *intersection* of events. Given two events A and B, the **union of A and B** is defined as follows.

Union of two events

The *union* of A and B is the event containing *all* sample points belonging to A *or* B *or both*. The union is denoted by $A \cup B$.

The Venn diagram in Figure 4.5 depicts the union of events A and B. Note that the two circles contain all the sample points in event A as well as all the sample points in event B.

The fact that the circles overlap indicates that some sample points are contained in both A and B.

The definition of the **intersection of A and B** follows.

Intersection of two events

Given two events A and B, the *intersection* of A and B is the event containing the sample points belonging to *both A and B*. The intersection is denoted by $A \cap B$.

The Venn diagram depicting the intersection of events A and B is shown in Figure 4.6. The area where the two circles overlap is the intersection; it contains the sample points that are in both A and B.

FIGURE 4.5
Union of events A and B is shaded

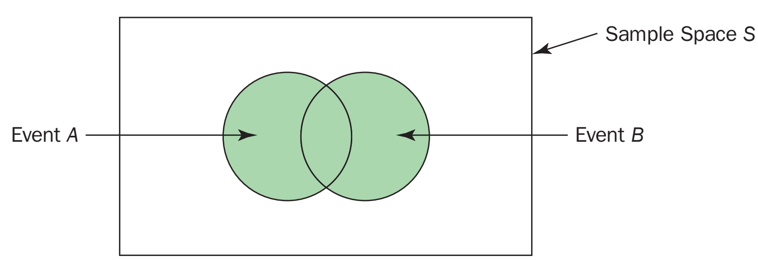
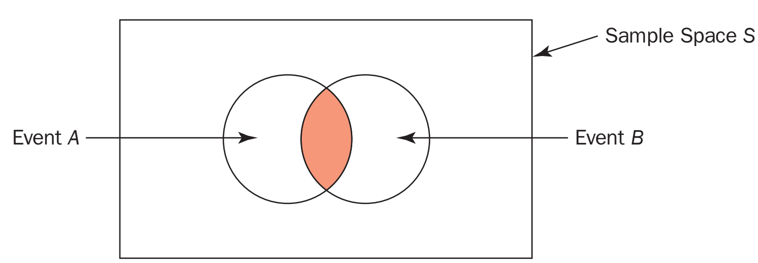


FIGURE 4.6
Intersection of events A and B is shaded



The **addition law** provides a way to compute the probability that event A or event B or both occur. In other words, the addition law is used to compute the probability of the union of two events. The addition law is written as follows in equation (4.6).

Addition law

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{4.6}$$

To understand the addition law intuitively, note that the first two terms in the addition law, $P(A) + P(B)$, account for all the sample points in $A \cup B$. However, because the sample points in the intersection $A \cap B$ are in both A and B, when we compute $P(A) + P(B)$, we are in effect counting each of the sample points in $A \cup B$ twice. We correct for this over-counting by subtracting $P(A \cap B)$.

As an example of an application of the addition law, consider the case of a small assembly plant with 50 employees. Each worker is expected to complete work assignments on time and in such a way that the assembled product will pass a final inspection. On occasion, some of the workers fail to meet the performance standards by completing work late or assembling a defective product. At the end of a performance evaluation period, the production manager found that five of the 50 workers completed work late, six of the 50 workers assembled a defective product and two of the 50 workers both completed work late *and* assembled a defective product.

Let:

$$L = \text{the event that the work is completed}$$
$$D = \text{the event that the assembled product is defective}$$

The relative frequency information leads to the following probabilities:

$$P(L) = \frac{5}{50} = 0.10$$
$$P(D) = \frac{6}{50} = 0.12$$
$$P(L \cap D) = \frac{2}{50} = 0.04$$

After reviewing the performance data, the production manager decided to assign a poor performance rating to any employee whose work was either late or defective; thus the event of interest is $L \cup D$. What is the probability that the production manager assigned an employee a poor performance rating?

Using equation (4.6), we have:

$$\begin{aligned} P(L \cup D) &= P(L) + P(D) - P(L \cap D) \\ &= 0.10 + 0.12 - 0.04 = 0.18 \end{aligned}$$

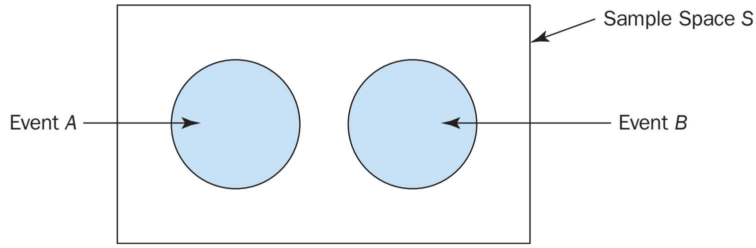
This calculation tells us that there is a 0.18 probability that a randomly selected employee received a poor performance rating.

As another example of the addition law, consider a recent study conducted by the personnel manager of a major computer software company. The study showed that 30 per cent of the employees who left the firm within two years did so primarily because they were dissatisfied with their salary, 20 per cent left because they were dissatisfied with their work assignments and 12 per cent of the former employees indicated dissatisfaction with *both* their salary and their work assignments. What is the probability that an employee who leaves within two years does so because of dissatisfaction with salary, dissatisfaction with the work assignment or both?

Let:

$$S = \text{the event that the employee leaves because of salary}$$
$$W = \text{the event that the employee leaves because of work assignment}$$

FIGURE 4.7
Mutually exclusive events



We have $P(S) = 0.30$, $P(W) = 0.20$ and $P(S \cap W) = 0.12$. Using equation (4.6) we have

$$P(S) + P(W) - P(S \cap W) = 0.30 + 0.20 - 0.12 = 0.38$$

We find a 0.38 probability that an employee leaves for salary or work assignment reasons.

Before we conclude our discussion of the addition law, let us consider a special case that arises for **mutually exclusive events**.

Mutually exclusive events

Two events are said to be mutually exclusive if the events have no sample points in common.

Events A and B are mutually exclusive if, when one event occurs, the other cannot occur. Thus, a requirement for A and B to be mutually exclusive is that their intersection must contain no sample points. The Venn diagram depicting two mutually exclusive events A and B is shown in Figure 4.7. In this case $P(A \cap B) = 0$ and the addition law can be written as follows.

Addition law for mutually exclusive events

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

EXERCISES

Methods

18. Suppose that we have a sample space with five equally likely experimental outcomes: E_1, E_2, E_3, E_4, E_5. Let:

$$A = \{E_1, E_2\}$$
$$B = \{E_3, E_4\}$$
$$C = \{E_2, E_3, E_5\}$$

a. Find $P(A)$, $P(B)$ and $P(C)$.
b. Find $P(A \cup B)$. Are A and B mutually exclusive?
c. Find $\bar{A}$, $\bar{C}$, $P(\bar{A})$ and $P(\bar{C})$.
d. Find $A \cup \bar{B}$ and $P(A \cup \bar{B})$.
e. Find $P(B \cup C)$.

19. Suppose that we have a sample space $S = \{E_1, E_2, E_3, E_4, E_5, E_6, E_7\}$, where $E_1, E_2, \ldots, E_7$ denote the sample points. The following probability assignments apply: $P(E_1) = 0.05$, $P(E_2) = 0.20$, $P(E_3) = 0.20$, $P(E_4) = 0.25$, $P(E_5) = 0.15$, $P(E_6) = 0.10$, and $P(E_7) = 0.05$. Let:

$$A = \{E_1, E_2\}$$
$$B = \{E_3, E_4\}$$
$$C = \{E_2, E_3, E_5\}$$

a. Find $P(A)$, $P(B)$, and $P(C)$.
b. Find $A \cup B$ and $P(A \cup B)$.
c. Find $A \cap B$ and $P(A \cap B)$.
d. Are events A and C mutually exclusive?
e. Find $\overline{B}$ and $P(\overline{B})$.

Applications

20. A survey of magazine subscribers showed that 45.8 per cent rented a car during the past 12 months for business reasons, 54 per cent rented a car during the past 12 months for personal reasons and 30 per cent rented a car during the past 12 months for both business and personal reasons.

a. What is the probability that a subscriber rented a car during the past 12 months for business or personal reasons?

b. What is the probability that a subscriber did not rent a car during the past 12 months for either business or personal reasons?

4.4 CONDITIONAL PROBABILITY

Often, the probability of an event is influenced by whether a related event already occurred. Suppose we have an event A with probability $P(A)$. If we obtain new information and learn that a related event, denoted by B, already occurred, we will want to take advantage of this information by calculating a new probability for event A. This new probability of event A is called a **conditional probability** and is written $P(A \mid B)$. We use the notation $\mid$ to indicate that we are considering the probability of event A *given* the condition that event B has occurred. Hence, the notation $P(A \mid B)$ reads 'the probability of A given B'.

Consider the situation of the promotion status of male and female police officers of a regional police force in France. The police force consists of 1200 officers: 960 men and 240 women. Over the past two years, 324 officers on the police force received promotions. The specific breakdown of promotions for male and female officers is shown in Table 4.4.

After reviewing the promotion record, a committee of female officers raised a discrimination case on the basis that 288 male officers had received promotions but only 36 female officers had received promotions.

TABLE 4.4 Promotion status of police officers over the past two years

	Men	Women	Total
Promoted	288	36	324
Not promoted	672	204	876
Totals	960	240	1200

The police administration argued that the relatively low number of promotions for female officers was due not to discrimination, but to the fact that relatively few females are members of the police force. Let us show how conditional probability could be used to analyze the discrimination charge.

Let:

$$M = \text{event an officer is a man}$$
$$W = \text{event an officer is a woman}$$
$$A = \text{event an officer is promoted}$$
$$\bar{A} = \text{event an officer is not promoted}$$

Dividing the data values in Table 4.4 by the total of 1200 officers enables us to summarize the available information with the following probability values.

$P(M \cap A) = 288/1200 = 0.24 =$ probability that a randomly selected officer is a man *and* is promoted

$P(M \cap \bar{A}) = 672/1200 = 0.56 =$ probability that a randomly selected officer is a man *and* not promoted

$P(W \cap A) = 36/1200 = 0.03 =$ probability that a randomly selected officer is a woman *and* is promoted

$P(W \cap \bar{A}) = 204/1200 = 0.17 =$ probability that a randomly selected officer is a woman *and* is not promoted

Because each of these values gives the probability of the intersection of two events, the probabilities are called **joint probabilities**. Table 4.5 is referred to as a *joint probability table*.

The values in the margins of the joint probability table provide the probabilities of each event separately. That is, $P(M) = 0.80$, $P(W) = 0.20$, $P(A) = 0.27$ and $P(\bar{A}) = 0.73$. These probabilities are referred to as **marginal probabilities** because of their location in the margins of the joint probability table. We note that the marginal probabilities are found by summing the joint probabilities in the corresponding row or column of the joint probability table. For instance, the marginal probability of being promoted is $P(A) = P(M \cap A) + P(W \cap A) = 0.24 + 0.03 = 0.27$. From the marginal probabilities, we see that 80 per cent of the force is male, 20 per cent of the force is female, 27 per cent of all officers received promotions and 73 per cent were not promoted.

Consider the probability that an officer is promoted given that the officer is a man. In conditional probability notation, we are attempting to determine $P(A \mid M)$. By definition, $P(A \mid M)$ tells us that we are concerned only with the promotion status of the 960 male officers. Because 288 of the 960 male officers received promotions, the probability of being promoted given that the officer is a man is $288/960 = 0.30$. In other words, given that an officer is a man, that officer has a 30 per cent chance of receiving a promotion over the past two years.

TABLE 4.5 Joint probability table for promotions

Joint probabilities appear in the body of the table

	Men (M)	**Women (W)**	**Totals**
Promoted (A)	0.24	0.03	0.27
Not Promoted ($\bar{A}$)	0.56	0.17	0.73
Totals	0.80	0.20	1.00

Marginal probabilities appear in the margins of the table.

This procedure was easy to apply because the values in Table 4.4 show the number of officers in each category. We now want to demonstrate how conditional probabilities such as $P(A \mid M)$ can be computed directly from related event probabilities rather than the frequency data of Table 4.4.

We have shown that $P(A \mid M) = 288/960 = 0.30$. Let us now divide both the numerator and denominator of this fraction by 1200, the total number of officers in the study.

$$P(A|M) = \frac{288}{960} = \frac{288/1200}{960/1200} = \frac{0.24}{0.80} = 0.30$$

We now see that the conditional probability $P(A \mid M)$ can be computed as 0.24/0.80. Refer to the joint probability table (Table 4.5). Note in particular that 0.24 is the joint probability of A and M; that is, $P(A \cap M) = 0.24$. Also note that 0.80 is the marginal probability that a randomly selected officer is a man; that is, $P(M) = 0.80$. Thus, the conditional probability $P(A \mid M)$ can be computed as the ratio of the joint probability $P(A \cap M)$ to the marginal probability $P(M)$.

$$P(A \mid M) = \frac{P(A \cap M)}{P(M)} = \frac{0.24}{0.80} = 0.30$$

The fact that conditional probabilities can be computed as the ratio of a joint probability to a marginal probability provides the following general formula (equations (4.7) and (4.8)) for conditional probability calculations for two events A and B.

Conditional probability

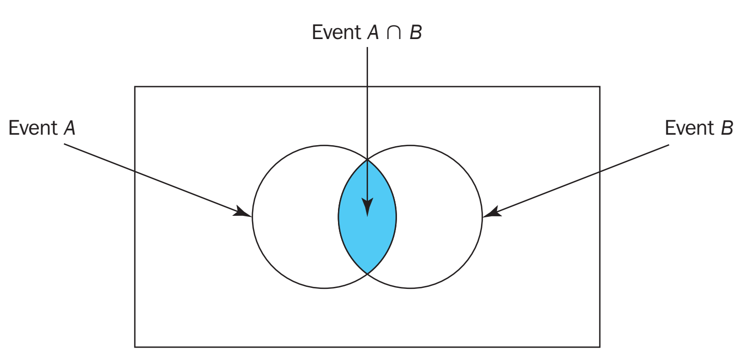
$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{4.7}$$

or

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{4.8}$$

The Venn diagram in Figure 4.8 is helpful in obtaining an intuitive understanding of conditional probability. The circle on the right shows that event B has occurred; the portion of the circle that overlaps with event A denotes the event $(A \cap B)$. We know that, once event B has occurred, the only way that we can also observe event A is for the event $(A \cap B)$ to occur. Thus, the ratio $P(A \cap B)/P(B)$ provides the conditional probability that we will observe event A given that event B has already occurred.

FIGURE 4.8
Conditional probability



Let us return to the issue of discrimination against the female officers. The marginal probability in row 1 of Table 4.5 shows that the probability of promotion of an officer is $P(A) = 0.27$ (regardless of whether that officer is male or female). However, the critical issue in the discrimination case involves the two conditional probabilities $P(A \mid M)$ and $P(A \mid W)$. That is, what is the probability of a promotion *given* that the officer is a man, and what is the probability of a promotion *given* that the officer is a woman? If these two probabilities are equal, a discrimination argument has no basis because the chances of a promotion are the same for male and female officers. However, a difference in the two conditional probabilities will support the position that male and female officers are treated differently in promotion decisions.

We already determined that $P(A \mid M) = 0.30$. Let us now use the probability values in Table 4.5 and the basic relationship of conditional probability in equation (4.7) to compute the probability that an officer is promoted given that the officer is a woman; that is, $P(A \mid W)$. Using equation (4.7), with W replacing B, we obtain:

$$P(A|W) = \frac{P(A \cap W)}{P(W)} = \frac{0.03}{0.20} = 0.15$$

What conclusion do you draw? The probability of a promotion given that the officer is a man is 0.30, twice the 0.15 probability of a promotion given that the officer is a woman. Although the use of conditional probability does not in itself prove that discrimination exists in this case, the conditional probability values support the argument presented by the female officers.

Independent events

In the preceding illustration, $P(A) = 0.27$, $P(A \mid M) = 0.30$ and $P(A \mid W) = 0.15$. We see that the probability of a promotion (event A) is affected or influenced by whether the officer is a man or a woman. Particularly, because $P(A \mid M) \neq P(A)$, we would say that events A and M are dependent events. That is, the probability of event A (promotion) is altered or affected by knowing that event M (the officer is a man) exists. Similarly, with $P(A \mid W) \neq P(A)$, we would say that events A and W are *dependent events*. However, if the probability of event A is not changed by the existence of event M – that is, $P(A \mid M) = P(A)$ – we would say that events A and M are **independent events**. This situation leads to the following definition of the independence of two events.

Independent events

Two events A and B are independent if

$$P(A \mid B) = P(A) \tag{4.9}$$

or

$$P(B \mid A) = P(B) \tag{4.10}$$

Multiplication law

Whereas the addition law of probability is used to compute the probability of a union of two events, the multiplication law is used to compute the probability of the intersection of two events. The multiplication law is based on the definition of conditional probability. Using equations (4.7) and (4.8) and solving for $P(A \cap B)$, we obtain the **multiplication law**, as in equations (4.11) and (4.12).

Multiplication law

$$P(A \cap B) = P(A)P(B \mid A) \quad \textbf{(4.11)}$$

or

$$P(A \cap B) = P(B)P(A \mid B) \quad \textbf{(4.12)}$$

To illustrate the use of the multiplication law, consider a newspaper circulation department where it is known that 84 per cent of the households in a particular neighbourhood subscribe to the daily edition of the paper. If we let D denote the event that a household subscribes to the daily edition, $P(D) = 0.84$. In addition, it is known that the probability that a household that already holds a daily subscription also subscribes to the Sunday edition (event S) is 0.75; that is, $P(S \mid D) = 0.75$.

What is the probability that a household subscribes to both the Sunday and daily editions of the newspaper? Using the multiplication law, we compute the desired $P(S \cap D)$ as

$$P(S \cap D) = P(D)P(S \cap D) = 0.84 \times 0.75 = 0.63$$

We now know that 63 per cent of the households subscribe to both the Sunday and daily editions.

Before concluding this section, let us consider the special case of the multiplication law when the events involved are independent. Recall that events A and B are independent whenever $P(A \mid B) = P(A)$ or $P(B \mid A) = P(B)$. Hence, using equations (4.11) and (4.12) for the special case of independent events, we obtain the following multiplication law (equation (4.13)).

Multiplication law for independent events

$$P(A \cap B) = P(A)P(B) \quad \textbf{(4.13)}$$

To compute the probability of the intersection of two independent events, we simply multiply the corresponding probabilities. Note that the multiplication law for independent events provides another way to determine whether A and B are independent. That is, if $P(A \cap B) = P(A)P(B)$, then A and B are independent; if $P(A \cap B) \neq P(A)P(B)$, then A and B are dependent.

As an application of the multiplication law for independent events, consider the situation of a service station manager who knows from past experience that 80 per cent of the customers use a credit card when they purchase petrol. What is the probability that the next two customers purchasing petrol will each use a credit card? If we let

A = the event that the first customer uses a credit card

B = the event that the second customer uses a credit card

then the event of interest is $A \cap B$. Given no other information, we can reasonably assume that A and B are independent events. Thus

$$P(A \cap B) = P(A)P(B) = 0.80 \times 0.80 = 0.64$$

To summarize this section, we note that our interest in conditional probability is motivated by the fact that events are often related. In such cases, we say the events are dependent and the conditional probability formulae in equations (4.7) and (4.8) must be used to compute the event probabilities. If two events are not related, they are independent; in this case neither event's probability is affected by whether the other event occurred.

EXERCISES

Methods

21. Suppose that we have two events, A and B, with $P(A) = 0.50$, $P(B) = 0.60$ and $P(A \cap B) = 0.40$.

a. Find $P(A \mid B)$.
b. Find $P(B \mid A)$.
c. Are A and B independent? Why or why not?

22. Assume that we have two events, A and B, that are mutually exclusive. Assume further that we know $P(A) = 0.30$ and $P(B) = 0.40$.

a. What is $P(A \cap B)$?
b. What is $P(A \mid B)$?
c. A student in statistics argues that the concepts of mutually exclusive events and independent events are really the same, and that if events are mutually exclusive they must be independent. Do you agree with this statement? Use the probability information in this problem to justify your answer.
d. What general conclusion would you make about mutually exclusive and independent events given the results of this problem?

Applications

23. A Paris nightclub obtains the following data on the age and marital status of 140 customers.

	Marital status	
Age	*Single*	*Married*
Under 30	77	14
30 or over	28	21

a. Develop a joint probability table for these data.
b. Use the marginal probabilities to comment on the age of customers attending the club.
c. Use the marginal probabilities to comment on the marital status of customers attending the club.
d. What is the probability of finding a customer who is single and under the age of 30?
e. If a customer is under 30, what is the probability that he or she is single?
f. Is marital status independent of age? Explain, using probabilities.

24. A slot machine in Melbourne has a hold facility. A gambler experiments with this to see if their success rate is higher when they use 'hold' compared to when they do not.

The results from 120 plays can be summarized as follows.

	Win	*Lose*
Hold	14	36
Not hold	10	60

What is the probability that the gambler:

a. Holds?
b. Wins?
c. Wins given that they held?
d. Held and lost?
e. Held given that they won?

25. A sample of convictions and compensation orders issued at a number of Manx courts was followed up to see whether the offender had paid the compensation to the victim. Details by gender of offender are as follows:

Offender gender	*Payment outcome*		
	Paid in full	*Part paid*	*Nothing paid*
Male	754	62	61
Female	157	7	6

a. What is the probability that no compensation was paid?
b. What is the probability that the offender was not male given that compensation was part paid?

26. A purchasing agent in Haifa placed rush orders for a particular raw material with two different suppliers, *A* and *B*. If neither order arrives in four days, the production process must be shut down until at least one of the orders arrives. The probability that supplier *A* can deliver the material in four days is 0.55. The probability that supplier *B* can deliver the material in four days is 0.35.

a. What is the probability that both suppliers will deliver the material in four days? Because two separate suppliers are involved, we are willing to assume independence.
b. What is the probability that at least one supplier will deliver the material in four days?
c. What is the probability that the production process will be shut down in four days because of a shortage of raw material (that is, both orders are late)?

COMPLETE SOLUTIONS

4.5 BAYES' THEOREM

In the discussion of conditional probability, we indicated that revising probabilities when new information is obtained is an important phase of probability analysis. Often, we begin the analysis with initial or **prior probability** estimates for specific events of interest. Then, from sources such as a sample, a special report or a product test, we obtain additional information about the events. Given this new information, we update the prior probability values by calculating revised probabilities, referred to as **posterior probabilities**. **Bayes' theorem** provides a means for making these probability calculations. The steps in this probability revision process are shown in Figure 4.9.

As an application of Bayes' theorem, consider a manufacturing firm that receives shipments of parts from two different suppliers. Let A_1 denote the event that a part is from supplier 1 and A_2 denote the event that a part is from supplier 2. Currently, 65 per cent of the parts purchased by the company are from supplier 1 and the remaining 35 per cent are from supplier 2. Hence, if a part is selected at random, we would assign the prior probabilities $P(A_1) = 0.65$ and $P(A_2) = 0.35$.

The quality of the purchased parts varies with the source of supply. Historical data suggest that the quality ratings of the two suppliers are as shown in Table 4.6.

FIGURE 4.9
Probability revision using Bayes' theorem

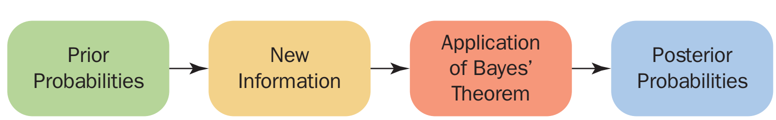


TABLE 4.6 Historical quality levels of two suppliers

	Percentage good parts	Percentage bad parts
Supplier 1	98	2
Supplier 2	95	5

If we let G denote the event that a part is good and B denote the event that a part is bad, the information in Table 4.6 provides the following conditional probability values:

$$P(G \mid A_1) = 0.98 \qquad P(B \mid A_1) = 0.02$$
$$P(G \mid A_2) = 0.95 \qquad P(B \mid A_2) = 0.05$$

The tree diagram in Figure 4.10 depicts the process of the firm receiving a part from one of the two suppliers and then discovering that the part is good or bad as a two-step experiment. We see that four experimental outcomes are possible: two correspond to the part being good and two correspond to the part being bad.

Each of the experimental outcomes is the intersection of two events, so we can use the multiplication rule to compute the probabilities. For instance:

$$P(A_1, G) = P(A_1 \cap G) = P(A_1)P(G \mid A_1) = 0.05$$

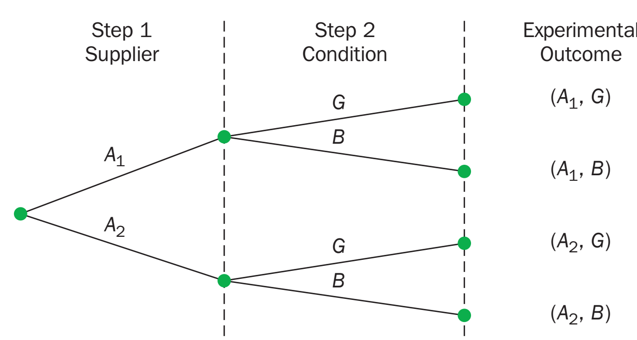
The process of computing these joint probabilities can be depicted in what is called a probability tree (see Figure 4.11). From left to right through the tree, the probabilities for each branch at step 1 are prior probabilities and the probabilities for each branch at step 2 are conditional probabilities. To find the probabilities of each experimental outcome, we simply multiply the probabilities on the branches leading to the outcome. Each of these joint probabilities is shown in Figure 4.11 along with the known probabilities for each branch.

Suppose now that the parts from the two suppliers are used in the firm's manufacturing process and that a machine breaks down because it attempts to process a bad part. Given the information that the part is bad, what is the probability that it came from supplier 1 and what is the probability that it came from supplier 2? With the information in the probability tree (Figure 4.11), Bayes' theorem can be used to answer these questions.

Letting B denote the event that the part is bad, we are looking for the posterior probabilities $P(A_1 \mid B)$ and $P(A_2 \mid B)$. From the law of conditional probability, we know that:

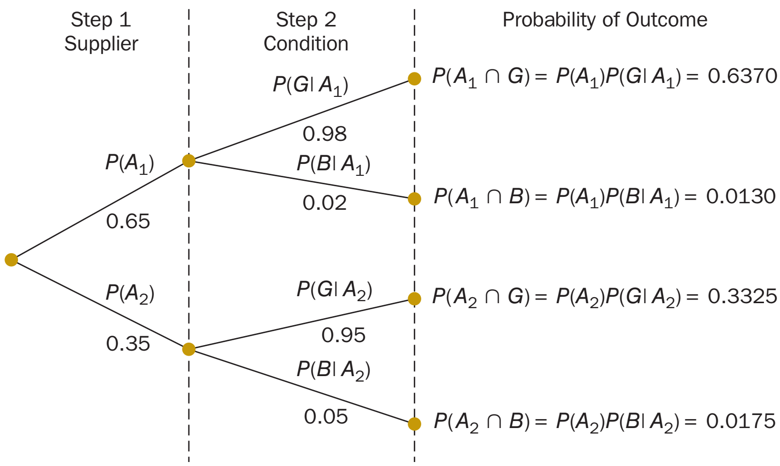
$$P(A_1 \mid B) = \frac{P(A_1 \cap B)}{P(B)} \qquad \textbf{(4.14)}$$

FIGURE 4.10
Tree diagram for two-supplier example



Note: Step 1 shows that the part comes from one of two suppliers, and Step 2 shows whether the part is good or bad.

FIGURE 4.11
Probability tree for two-supplier example



Referring to the probability tree, we see that:

$$P(A_1 \cap B) = P(A_1)P(B \mid A_1) \tag{4.15}$$

To find $P(B)$, we note that event B can occur in only two ways: $(A_1 \cap B)$ and $(A_2 \cap B)$. Therefore, we have:

$$\begin{aligned} P(B) &= P(A_1 \cap B) + P(A_2 \cap B) \\ &= P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) \end{aligned} \tag{4.16}$$

Substituting from equations (4.15) and (4.16) into equation (4.14) and writing a similar result for $P(A_2 \mid B)$, we obtain Bayes' theorem for the case of two events.

Bayes' theorem (two-event case)

$$P(A_1 \mid B) = \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \tag{4.17}$$

$$P(A_2 \mid B) = \frac{P(A_2)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \tag{4.18}$$

Using equation (4.17) and the probability values provided in the example, we have

$$\begin{aligned} P(A_1 \mid B) &= \frac{P(A_1)P(B \mid A_1)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2)} \\ &= \frac{0.65 \times 0.02}{0.65 \times 0.02 + 0.35 \times 0.05} = \frac{0.0130}{0.0130 + 0.0175} \\ &= \frac{0.0130}{0.0305} = 0.4262 \end{aligned}$$

In addition, using equation (4.18), we find $P(A_2 \mid B)$.

$$\begin{aligned} P(A_2 \mid B) &= \frac{0.35 \times 0.05}{0.65 \times 0.02 + 0.35 \times 0.05} \\ &= \frac{.0175}{0.0130 + 0.0175} = \frac{0.0175}{0.0305} = 0.5738 \end{aligned}$$

Note that in this application we started with a probability of 0.65 that a part selected at random was from supplier 1. However, given information that the part is bad, the probability that the part is from supplier 1 drops to 0.4262. In fact, if the part is bad, it has better than a 50–50 chance that it came from supplier 2; that is, $P(A_2 \mid B) = 0.5738$.

Bayes' theorem is applicable when the events for which we want to compute posterior probabilities are mutually exclusive and their union is the entire sample space.* For the case of n mutually exclusive events $A_1, A_2, \ldots, A_n$, whose union is the entire sample space, Bayes' theorem can be used to compute any posterior probability $P(A_i \mid B)$ as shown in equation (4.19).

Bayes' theorem

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \ldots + P(A_n)P(B \mid A_n)} \tag{4.19}$$

With prior probabilities $P(A_1), P(A_2), \ldots, P(A_n)$ and the appropriate conditional probabilities $P(B \mid A_1), P(B \mid A_2), \ldots, P(B \mid A_n)$, equation (4.19) can be used to compute the posterior probability of the events $A_1, A_2, \ldots, A_n$.

Tabular approach

A tabular approach is helpful in conducting the Bayes' theorem calculations. Such an approach is shown in Table 4.7 for the parts supplier problem. The computations shown there are done in the following steps.

Step 1 Prepare the following three columns:
Column 1 – The mutually exclusive events A_i for which posterior probabilities are desired.
Column 2 – The prior probabilities $P(A_i)$ for the events.
Column 3 – The conditional probabilities $P(B \mid A_i)$ of the new information B given each event.

Step 2 In column 4, compute the joint probabilities $P(A_i \cap B)$ for each event and the new information B by using the multiplication law. These joint probabilities are found by multiplying the prior probabilities in column 2 by the corresponding conditional probabilities in column 3: that is, $P(A_i \cap B) = P(A_i)P(B \mid A_i)$.

TABLE 4.7 Tabular approach to Bayes' theorem calculations for the two-supplier problem

(1) Events A_i	(2) Prior probabilities $P(A_i)$	(3) Conditional probabilities $P(B \mid A_i)$	(4) Joint probabilities $P(A_i \cap B)$	(5) Posterior probabilities $P(A_i \mid B)$
A_1	0.65	0.02	0.0130	0.0130/0.0305 = 0.4262
A_2	0.35	0.05	0.0175	0.0175/0.0305 = 0.5738
			$P(B)$ = 0.0305	1.0000

* If the union of events is the entire sample space, the events are said to be collectively exhaustive.

Step 3 Sum the joint probabilities in column 4. The sum is the probability of the new information, $P(B)$. Thus we see in Table 4.7 that there is a 0.0130 probability that the part came from supplier 1 and is bad and a 0.0175 probability that the part came from supplier 2 and is bad. Because these are the only two ways in which a bad part can be obtained, the sum $0.0130 + 0.0175$ shows an overall probability of 0.0305 of finding a bad part from the combined shipments of the two suppliers.

Step 4 In column 5, compute the posterior probabilities using the basic relationship of conditional probability.

$$P(A_i \mid B) = \frac{P(A_i \cap B)}{P(B)}$$

Note that the joint probabilities $P(A_i \cap B)$ are in column (4) and the probability $P(B)$ is the sum of column (4).

EXERCISES

Methods

27. The prior probabilities for events A_1 and A_2 are $P(A_1) = 0.40$ and $P(A_2) = 0.60$. It is also known that $P(A_1 \cap A_2) = 0$. Suppose $P(B \mid A_1) = 0.20$ and $P(B \mid A_2) = 0.05$.

- **a.** Are A_1 and A_2 mutually exclusive? Explain.
- **b.** Compute $P(A_1 \cap B)$ and $P(A_2 \cap B)$.
- **c.** Compute $P(B)$.
- **d.** Apply Bayes' theorem to compute $P(A_1 \mid B)$ and $P(A_2 \mid B)$.

28. The prior probabilities for events A_1, A_2 and A_3 are $P(A_1) = 0.20$, $P(A_2) = 0.50$ and $P(A_3) = 0.30$. The conditional probabilities of event B given A_1, A_2 and A_3 are $P(B \mid A_1) = 0.50$, $P(B \mid A_2) = 0.40$ and $P(B \mid A_3) = 0.30$.

- **a.** Compute $P(B \cap A_1)$, $P(B \cap A_2)$ and $P(B \cap A_3)$.
- **b.** Apply Bayes' theorem, equation (4.19), to compute the posterior probability $P(A_2 \mid B)$.
- **c.** Use the tabular approach to applying Bayes' theorem to compute $P(A_1 \mid B)$, $P(A_2 \mid B)$ and $P(A_3 \mid B)$.

Applications

29. Records show that for every 100 items produced in a factory during the day shift, two are defective and for every 100 items produced during the night shift, four are defective. What is the prior probability of the bid being successful (that is, prior to the request for additional information)?

- **a.** If during a 24-hour period, 2000 items are produced during the day and 800 at night, what is the probability that an item picked at random from the output over 24 hours came from the night shift if it was defective?

30. A company is about to sell to a new client. It knows from past experience that there is a real possibility that the client may default on payment. As a precaution the company checks with a consultant on the likelihood of the client defaulting in this case and is given an estimate of 20 per cent. Sometimes the consultant gets it wrong. Your own experience of the consultant is that he is correct 70 per cent of the time when he predicts that the client will default but that 20 per cent of clients who he believes will not default actually do.

COMPLETE SOLUTIONS

- **a.** What is the probability that the new client will not default?

31. In 2011, there were 1901 fatalities recorded on Britain's roads, 60 of which were for children (Department of Transport, 2012). Correspondingly, serious injuries totalled 23 122 of which 20 770 were for adults.

a. What is the probability of a serious injury given the victim was a child?
b. What is the probability that the victim was an adult given a fatality occurred?

32. The following cross-tabulation shows industry type and price/earnings (P/E) ratio for 100 companies in the consumer products and banking industries.

	P/E ratio					
Industry	*5–9*	*10–14*	*15–19*	*20–24*	*25–29*	*Total*
Consumer	4	10	18	10	8	50
Banking	14	14	12	6	4	50
Total	18	24	30	16	12	100

a. What is the probability that a company had a P/E greater than 9 and belonged to the consumer industry?
b. What is the probability that a company with a P/E in the range 15–19 belonged to the banking industry?

33. A large investment advisory service has a number of analysts who prepare detailed studies of individual companies. On the basis of these studies the analysts make 'buy' or 'sell' recommendations on the companies' shares. The company classes an excellent analyst as one who will be correct 80 per cent of the time, a good analyst as who will be correct 60 per cent of the time and a poor analyst who will be correct 40 per cent of the time.

Two years ago, the advisory service hired Mr Smith who came with considerable experience from the research department of another firm. At the time he was hired it was thought that the probability was 0.90 that he was an excellent analyst, 0.09 that he was a good analyst and 0.01 that he was a poor analyst. In the past two years he has made ten recommendations of which only three have been correct.

Assuming that each recommendation is an independent event what probability would you assign to Mr Smith being:

a. An excellent analyst?
b. A good analyst?
c. A poor analyst?

34. An electronic component is produced by four production lines in a manufacturing operation. The components are costly, are quite reliable and are shipped to suppliers in 50-component lots. Because testing is destructive, most buyers of the components test only a small number before deciding to accept or reject lots of incoming components. All four production lines usually only produce 1 per cent defective components which are randomly dispersed in the output. Unfortunately, production line 1 suffered mechanical difficulty and produced 10 per cent defectives during the month of April. This situation became known to the manufacturer after the components had been shipped. A customer received a lot in April and tested five components. Two failed. What is the probability that this lot came from production line 1?

ONLINE RESOURCES

For the data files, additional online summary, questions and answers for Chapter 4, visit the online platform.

SUMMARY

In this chapter we introduced basic probability concepts and illustrated how probability analysis can be used to provide helpful information for decision-making. We described how probability can be interpreted as a numerical measure of the likelihood that an event will occur and reviewed classical, relative frequency and subjective methods for deriving probabilities. In addition, we saw that the probability of an event can be computed either by summing the probabilities of the experimental outcomes (sample points) comprising the event or by using the relationships established by the addition, conditional probability, and multiplication laws of probability. For cases in which new information is available, we showed how Bayes' theorem can be used to obtain revised or posterior probabilities.

KEY TERMS

Addition law
Basic requirements for assigning probabilities
Bayes' theorem
Classical method
Complement of *A*
Conditional probability
Event
Experiment
Independent events
Intersection of *A* and *B*
Joint probability
Marginal probability
Multiplication law
Mutually exclusive events
Posterior probabilities
Prior probabilities
Probability
Relative frequency method
Sample point
Sample space
Subjective method
Tree diagram
Union of *A* and *B*
Venn diagram

KEY FORMULAE

Counting rule for combinations

$$^{N}C_{n} = \binom{N}{n} = \frac{N!}{n!(N-n)!} \qquad \textbf{(4.1)}$$

Counting rule for permutations

$$^{N}P_{n} = n!\binom{N}{n} = \frac{N!}{(N-n)!} \qquad \textbf{(4.2)}$$

Computing probability using the complement

$$P(A) = 1 - P(\bar{A}) \qquad \textbf{(4.5)}$$

Addition law

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \qquad \textbf{(4.6)}$$

Conditional probability

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \tag{4.7}$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \tag{4.8}$$

Multiplication law

$$P(A \cap B) = P(B)P(A \mid B) \tag{4.11}$$

$$P(A \cap B) = P(A)P(B \mid A) \tag{4.12}$$

Multiplication law for independent events

$$P(A \cap B) = P(A)P(B) \tag{4.13}$$

Bayes' theorem

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \ldots + P(A_n)P(B \mid A_n)} \tag{4.19}$$

CASE PROBLEM

BAC and the Alcohol Test

In 2005, 6.7 per cent of accidents with injuries in Austria were caused by drunk drivers. The police in Wachau, Austria, a region which is famous for its wine production, is interested in buying equipment for testing drivers' blood alcohol levels. The law in Austria requires that the driver's licence be withdrawn if the driver is found to have more than 0.05 per cent BAC (blood alcohol concentration).

Due to the large number of factors that come into play regarding the consumption and reduction (burn off) rates of different people, there is no blood alcohol calculator that is 100 per cent accurate. Factors include the sex (male/female) of the drinker, differing metabolism rates, various health issues and the combination of medications being taken, drinking frequency, amount of food in the stomach and small intestine and when it was eaten, elapsed time and many others. The best that can be done is a rough estimate of the BAC level based on known inputs.

There are three types of equipment available with the following conditions:

1. The Saliva Screen is a disposable strip which can be used once – this is the cheapest method.
2. The Alcometer™ is an instrument attached to a container into which the driver breathes, with the Alcometer™ then measuring the BAC concentration through an analysis of the driver's breath. The draw-back to the Alcometer™ is that it can only detect the alcohol level correctly if it is used within two hours of alcohol consumption. It is less effective if used beyond this two-hour period.

Type	*False positive*	*False negative*
Saliva Screen	0.020	0.03
Alcometer™	0.015	0.02
Intoximeter	0.020	0.01

3. The Intoximeter is the most expensive of the three and it works through a blood sample of the driver. The advantage for this is that it can test the BAC up to 12 hours after alcohol consumption. False positive is the situation where the test indicates a high BAC level in a driver that actually does not have such a level. The false negative is when the test indicates a low level of BAC when the driver is actually highly intoxicated.

Police records show that the percentage of drivers (late night) that drink heavily and drive, ranges between 6 per cent on weekdays and 10 per cent on the weekend.

Managerial report

Carry out an appropriate probability analysis of this information on behalf of the police and advise them accordingly. (Note that it would be particularly helpful if you could assess the effectiveness of the different equipment types separately for weekdays and

Case problem provided by Dr Ibrahim Wazir, Webster University, Vienna