Uncertainty in Forensic Science

1.1 INTRODUCTION

The purpose of this book is to discuss the statistical and probabilistic evaluation of scientific evidence for forensic scientists. For the most part the evidence to be evaluated will be so-called transfer or trace evidence.

There is a well-known principle in forensic science known as Locard’s principle which states that every contact leaves a trace:

\[ \text{tantôt le malfaiteur a laissé sur les lieux les marques de son passage, tantôt, par une action inverse, il a emporté sur son corps ou sur ses vêtements, les indices de son séjour ou de son geste. (Locard, 1920)} \]

Iman and Rudin (2001) translate this as follows:

either the wrong-doer has left signs at the scene of the crime, or, on the other hand, has taken away with him – on his person (body) or clothes – indications of where he has been or what he has done.

The principle was reiterated using different words in 1929:

\[ \text{Les débris microscopiques qui recouvrent nos habits et notre corps sont les témoins muets, assurés et fidèles de chacun de nos gestes et de chacun de nos rencontres. (Locard, 1929)} \]

This may be translated as follows:

Traces which are present on our clothes or our person are silent, sure and faithful witnesses of every action we do and of every meeting we have.

Transfer evidence and Locard’s principle may be illustrated as follows. Suppose a person gains entry to a house by breaking a window and assaults the man of the house, during which assault blood is spilt by both victim and assailant. The criminal may leave traces of his presence at the crime scene in the form of bloodstains from the assault and fibres from his clothing. This evidence is said to be transferred from the criminal to the scene of the crime. The criminal may also take traces of the crime scene away with him. These could include
Uncertainty in forensic science

bloodstains from the assault victim, fibres of their clothes and fragments of glass from the broken window. Such evidence is said to be transferred to the criminal from the crime scene. A suspect is soon identified, at a time at which he will not have had the opportunity to change his clothing. The forensic scientists examining the suspect’s clothing find similarities amongst all the different types of evidence: blood, fibres and glass fragments. They wish to evaluate the strength of this evidence. It is hoped that this book will enable them so to do.

Quantitative issues relating to the distribution of characteristics of interest will be discussed. However, there will also be discussion of qualitative issues such as the choice of a suitable population against which variability in the measurements of the characteristics of interest may be compared. Also, a brief history of statistical aspects of the evaluation of evidence is given in Chapter 4.

1.2 STATISTICS AND THE LAW

The book does not focus on the use of statistics and probabilistic thinking for legal decision making, other than by occasional reference. Also, neither the role of statistical experts as expert witnesses presenting statistical assessments of data nor their role as consultants preparing analyses for counsel is discussed. There is a distinction between these two issues (Fienberg, 1989; Tribe, 1971). The main focus of this book is on the assessment of evidence for forensic scientists, in particular for identification purposes. The process of addressing the issue of whether or not a particular item came from a particular source is most properly termed individualisation. ‘Criminalistics is the science of individualisation’ (Kirk, 1963), but established forensic and judicial practices have led to it being termed identification. The latter terminology will be used throughout this book. An identification, however, is more correctly defined as ‘the determination of some set to which an object belongs or the determination as to whether an object belongs to a given set’ (Kingston, 1965a). Further discussion is given in Kwan (1977) and Evett et al. (1998a).

For example, in a case involving a broken window, similarities may be found between the refractive indices of fragments of glass found on the clothing of a suspect and the refractive indices of fragments of glass from the broken window. The assessment of this evidence, in associating the suspect with the scene of the crime, is part of the focus of this book (and is discussed in particular in Section 10.4.2).

For those interested in the issues of statistics and the law beyond those of forensic science, in the sense used in this book, there are several books available and some of these are discussed briefly.

‘The Evolving Role of Statistical Assessments as Evidence in the Courts’ is the title of a report, edited by Fienberg (1989), by the Panel on Statistical Assessments as Evidence in the Courts formed by the Committee on National Statistics and the Committee on Research on Law Enforcement and the Administration of Justice.
of the USA, and funded by the National Science Foundation. Through the use of case studies the report reviews the use of statistics in selected areas of litigation, such as employment discrimination, antitrust litigation and environment law. One case study is concerned with identification in a criminal case. Such a matter is the concern of this book and the ideas relevant to this case study, which involves the evidential worth of similarities amongst human head hair samples, will be discussed in greater detail later (Sections 4.5.2 and 4.5.5). The report makes various recommendations, relating to the role of the expert witness, pre-trial discovery, the provision of statistical resources, the role of court-appointed experts, the enhancement of the capability of the fact-finder and statistical education for lawyers.

Two books which take the form of textbooks on statistics for lawyers are Vito and Latessa (1989) and Finkelstein and Levin (2001). The former focuses on the presentation of statistical concepts commonly used in criminal justice research. It provides criminological examples to demonstrate the calculation of basic statistics. The latter introduces rather more advanced statistical techniques and again uses case studies to illustrate such techniques.

The area of discrimination litigation is covered by a set of papers edited by Kaye and Aickin (1986). This starts by outlining the legal doctrines that underlie discrimination litigation. In particular, there is a fundamental issue relating to discrimination in hiring. The definition of the relevant market from which an employer hires has to be made very clear. For example, consider the case of a man who applies, but is rejected for, a secretarial position. Is the relevant population the general population, the representation of men amongst secretaries in the local labour force or the percentage of male applicants? The choice of a suitable reference population is also one with which the forensic scientist has to be concerned. This is discussed at several points in this book.

Another textbook, which comes in two volumes, is Gastwirth (1988a, b). The book is concerned with civil cases and ‘is designed to introduce statistical concepts and their proper use to lawyers and interested policymakers’ (1988a, p. xvii). Two areas are stressed which are usually given less emphasis in most statistical textbooks. The first area is concerned with measures of relative or comparative inequality. These are important because many legal cases are concerned with issues of fairness or equal treatment. The second area is concerned with the combination of results of several related statistical studies. This is important because existing administrative records or currently available studies often have to be used to make legal decisions and public policy; it is not possible to undertake further research. Gastwirth (2000) has also edited a collection of essays on statistical science in the courtroom, some of which are directly relevant to this current book and will be referred to as appropriate.

A collection of papers on Statistics and Public Policy has been edited by Fairley and Mosteller (1977). One issue in the book which relates to a particularly infamous case, the Collins case, is discussed in detail later (Section 4.4). Other articles concern policy issues and decision making.
Uncertainty in forensic science

The remit of this book is one which is not covered by these others in great detail. The use of statistics in forensic science in general is discussed in a collection of essays edited by Aitken and Stoney (1991). The remit of this book is to describe statistical procedures for the evaluation of evidence for forensic scientists. This will be done primarily through a modern Bayesian approach. This approach has its origins in the work of I.J. Good and A.M. Turing as code-breakers at Bletchley Park during World War II. A brief review of the history is given in Good (1991). An essay on the topic of probability and the weighing of evidence is Good (1950). This also refers to entropy (Shannon, 1948), the expected amount of information from an experiment, and Good remarks that the expected weight of evidence in favour of a hypothesis $H$ and against its complement $ar{H}$ is equal to the difference of the entropies assuming $H$ and $ar{H}$, respectively. A brief discussion of a frequentist approach and the problems associated with it is given in Section 4.6.

It is of interest to note that a high proportion of situations involving the formal objective presentation of statistical evidence uses the frequentist approach with tests of significance (Fienberg and Schervish, 1986). However, Fienberg and Schervish go on to say that the majority of examples cited for the use of the Bayesian approach are in the area of identification evidence. It is this area which is the main focus of this book and it is Bayesian analyses which will form the basis for the evaluation of evidence as discussed here. Examples of the applications of such analyses to legal matters include Cullison (1969), Fairley (1973), Finkelstein and Fairley (1970, 1971), Lempert (1977), Lindley (1977a, b), Fienberg and Kadane (1983) and Anderson and Twining (1998).

Another approach which will not be discussed here is that of Shafer (1976, 1982). This concerns so-called belief functions (see Section 4.1). The theory of belief functions is a very sophisticated theory for assessing uncertainty which endeavours to answer criticisms of both the frequentist and Bayesian approaches to inference. Belief functions are non-additive in the sense that belief in an event $A$ (denoted Bel($A$)) and belief in the opposite of $A$ (denoted Bel($\bar{A}$)) do not sum to 1. See also Shafer (1978) for a historical discussion of non-additivity. Further discussion is beyond the scope of this book. Practical applications are few. One such, however, is to the evaluation of evidence concerning the refractive index of glass (Shafer, 1982).

It is very tempting when assessing evidence to try to determine a value for the probability of the so-called probandum of interest (or the ultimate issue) such as the guilt of a suspect, or a value for the odds in favour of guilt, and perhaps even to reach a decision regarding the suspect’s guilt. However, this is the role of the jury and/or judge. It is not the role of the forensic scientist or statistical expert witness to give an opinion on this (Evett, 1983). It is permissible for the scientist to say that the evidence is 1000 times more likely, say, if the suspect is guilty than if he is innocent. It is not permissible to interpret this to say that, because of the evidence, it is 1000 times more likely that the suspect is guilty than innocent. Some of the difficulties associated with assessments of probabilities are discussed...
by Tversky and Kahneman (1974) and by Kahneman et al. (1982) and are further described in Section 3.3. An appropriate representation of probabilities is useful because it fits the analytic device most used by lawyers, namely the creation of a story. This is a narration of events ‘abstracted from the evidence and arranged in a sequence to persuade the fact-finder that the story told is the most plausible account of “what really happened” that can be constructed from the evidence that has been or will be presented’ (Anderson and Twining, 1998, p. 166). Also of relevance is Kadane and Schum (1996), which provides a Bayesian analysis of evidence in the Sacco–Vanzetti case (Sacco, 1969) based on subjectively determined probabilities and assumed relationships amongst evidential events. A similar approach is presented in Chapter 14.

1.3 UNCERTAINTY IN SCIENTIFIC EVIDENCE

Scientific evidence requires considerable care in its interpretation. Emphasis needs to be put on the importance of asking the question ‘what do the results mean in this particular case?’ (Jackson, 2000). Scientists and jurists have to abandon the idea of absolute certainty in order to approach the identification process in a fully objective manner. If it can be accepted that nothing is absolutely certain then it becomes logical to determine the degree of confidence that may be assigned to a particular belief (Kirk and Kingston, 1964).

There are various kinds of problems concerned with the random variation naturally associated with scientific observations. There are problems concerned with the definition of a suitable reference population against which concepts of rarity or commonality may be assessed. There are problems concerned with the choice of a measure of the value of the evidence.

The effect of the random variation can be assessed with the appropriate use of probabilistic and statistical ideas. There is variability associated with scientific observations. Variability is a phenomenon which occurs in many places. People are of different sexes, determination of which is made at conception. People are of different heights, weights and intellectual abilities, for example. The variation in height and weight is dependent on a person’s sex. In general, females tend to be lighter and shorter than males. However, variation is such that there can be tall, heavy females and short, light males. At birth, it is uncertain how tall or how heavy the baby will be as an adult. However, at birth, it is known whether the baby is a boy or a girl. This knowledge affects the uncertainty associated with the predictions of adult height and weight.

People are of different blood groups. A person’s blood group does not depend on the age or sex of the person but does depend on the person’s ethnicity. The refractive index of glass varies within and between windows. Observation of glass as to whether it is window or bottle glass will affect the uncertainty associated with the prediction of its refractive index and that of other pieces of glass which may be thought to be from the same origin.
It may be thought that, because there is variation in scientific observations, it is not possible to make quantitative judgements regarding any comparisons between two sets of observations. The two sets are either different or they are not, and there is no more to be said. However, this is not so. There are many phenomena which vary, but they vary in certain specific ways. It is possible to represent these specific ways mathematically. Several such ways, including various probability distributions, are introduced in Chapter 2. It is then possible to assess differences quantitatively and to provide a measure of uncertainty associated with such assessments.

It is useful to recognise the distinction between statistics and probability. Probability is a deductive process which argues from the general to the particular. Consider a fair coin, that is, one for which, when tossed, the probability of a head landing uppermost equals the probability of a tail landing uppermost equals 1/2. A fair coin is tossed ten times. Probability theory enables a determination to be made of the probability that there are three heads and seven tails, say. The general concept of a fair coin is used to determine something about the outcome of the particular case in which it was tossed ten times.

On the other hand, statistics is an inductive process which argues from the particular to the general. Consider a coin which is tossed ten times and there are three heads and seven tails. Statistics enables the question as to whether the coin is fair or not to be addressed. The particular outcome of three heads and seven tails in ten tosses is used to determine something about the general case of whether the coin was fair or not.

Fundamental to both statistics and probability is uncertainty. Given a fair coin, the numbers of heads and tails in ten tosses is uncertain. The probability associated with each outcome may be determined but the actual outcome itself cannot be predicted with certainty. Given the outcome of a particular sequence of ten tosses, information is then available about the fairness or otherwise of the coin. For example, if the outcome were ten heads and no tails, one may believe that the coin is double-headed but it is not certain that this is the case. There is still a non-zero probability (1/1024) that ten tosses of a fair coin will result in ten heads. Indeed, this has occurred in the first author’s experience. A class of some 130 students were asked to each toss a coin ten times. One student tossed ten consecutive heads from what it is safe to assume was a fair coin. The probability of this happening is $1 - \left(1 - \frac{1}{1024}\right)^{130} = 0.12$.

### 1.3.1 The frequentist method

Consider a consignment of compact discs containing \(N\) discs. The consignment is said to be of size \(N\). It is desired to make inferences about the proportion \(\theta\) (\(0 \leq \theta \leq 1\)) of the consignment which is pirated. It is not practical to inspect the whole consignment so a sample of size \(n\), where \(n < N\), is inspected.

The frequentist method assumes that the proportion \(\theta\) of the consignment which is pirated is unknown but fixed. The data, the number of discs in the
sample which are pirated, are variable. A so-called confidence interval is calculated. The name confidence is used since no probability can be attached to the uncertain event that the interval contains \( \theta \). These ideas are discussed further in Chapter 5.

The frequentist approach derives its name from the relative frequency definition of probability. The probability that a particular event, \( A \), occurs is defined as the relative frequency of the number of occurrences of event \( A \) compared with the total number of occurrences of all possible events, over a long run of observations, conducted under identical conditions of all possible events.

For example, consider tossing a coin \( n \) times. It is not known if the coin is fair. The outcomes of the \( n \) tosses are to be used to determine the probability of a head occurring on an individual toss. There are two possible outcomes, heads (\( H \)) and tails (\( T \)). Let \( n(H) \) be the number of \( H \) and \( n(T) \) be the number of \( T \) such that \( n(H) + n(T) = n \). Then the probability of tossing a head on an individual toss of the coin is defined as the limit as \( n \to \infty \) of the fraction \( n(H)/n \). The frequentist approach relies on a belief in the long-run repetition of trials under identical conditions. This is an idealised situation, seldom, if ever, realised in practice. More discussion on the interpretation of such a result is given in Section 4.6.

The way in which statistics and probability may be used to evaluate evidence is the theme of this book. Care is required. Statisticians are familiar with variation, as are forensic scientists who observe it in the course of their work. Lawyers, however, prefer certainties. A defendant is found guilty or not guilty (or also, in Scotland, not proven). The scientist’s role is to testify as to the worth of the evidence, the role of the statistician and this book is to provide the scientist with a quantitative measure of this worth. It is shown that there are few forms of evidence that are so definite that statistical treatment is neither needed nor desirable. It is up to other people (the judge and/or the jury) to use this information as an aid to their deliberations. It is for neither the statistician nor the scientist to pass judgement. The scientist’s role in court is restricted to giving evidence in terms of what has been called cognition (Kind, 1994): the problem of whether or not the evidence from two places (e.g., the scene of the crime and the suspect) has the same origin.

The use of these ideas in forensic science is best introduced through the discussion of several examples. These examples will provide a constant theme throughout the book. Populations from which the criminal may be thought to have come, to which reference is made below, are considered in detail in Section 8.5, where they are called relevant populations. The value of evidence is measured by a statistic known as the likelihood ratio, and its logarithm. These are introduced in Sections 3.4 and 3.5.

1.3.2 Stains of body fluids

Example 1.1 A crime is committed. A bloodstain is found at the scene of the crime. All innocent explanations for the presence of the stain are eliminated.
A suspect is found. His DNA profile is identified and found to match that of the crime stain. What is the evidential value of this match? This is a very common situation, yet the answer to the question provides plenty of opportunity for discussion of the theme of this book.

Certain other questions need to be addressed before this particular one can be answered. Where was the crime committed, for example? Does it matter? Does the value of the evidence of the bloodstain change depending on where the crime was committed?

Apart from his DNA profile, what else is known about the criminal? In particular, is there any information, such as ethnicity, which may be related to his DNA profile? What is the population from which the criminal may be thought to have come? Could he be another member of the suspect’s family?

Questions such as these and their effect on the interpretation and evaluation of evidence will be discussed in greater detail. First, consider only the evidence of the DNA profile in isolation and one particular locus, \( \text{LDLR} \). Assume the crime was committed in Chicago and that there is eyewitness evidence that the criminal was a Caucasian. Information is available to the investigating officer about the genotypic distribution for the \( \text{LDLR} \) locus in Caucasians in Chicago and is given in Table 1.1. The information about the location of the crime and the ethnicity of the criminal is relevant. Genotypic frequencies vary across locations and among ethnic groups. A suspect is identified. For locus \( \text{LDLR} \) the genotype of the crime stain and that of the suspect match. The investigating officer knows a little about probability and works out that the probability of two people chosen at random and unrelated to the suspect having matching alleles, given the figures in Table 1.1

\[
0.188^2 + 0.321^2 + 0.491^2 = 0.379
\]

(see Section 4.5). He is not too sure what this result means. Is it high, and is a high value incriminating for the suspect? Is it low, and is a low value incriminating? In fact, a low value is more incriminating than a high value.

He thinks a little more and remembers that the genotypes not only match but also are both of type \( BB \). The frequencies of genotypes \( AA \) and \( AB \) are not relevant. He works out the probability that two people chosen at random both have genotype \( BB \) as

\[
0.321^2 = 0.103
\]

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( AA )</td>
<td>18.8</td>
</tr>
<tr>
<td>( BB )</td>
<td>32.1</td>
</tr>
<tr>
<td>( AB )</td>
<td>49.1</td>
</tr>
</tbody>
</table>

Table 1.1 Genotypic frequencies for locus \( \text{LDLR} \) amongst Caucasians in Chicago based on a sample of size 200 (from Johnson and Peterson, 1999)
Uncertainty in scientific evidence

(see Section 4.5). He is still not too sure what this means but feels that it is more representative of the information available to him than the previous probability, since it takes account of the actual genotypes of the crime stain and the suspect.

The genotype of the crime stain for locus LDLR is BB. The genotype of the suspect is also BB (if it were not he would not be a suspect). What is the value of this evidence? The discussion above suggests various possible answers.

1. The probability that two people chosen at random have the same genotype for locus LDLR. This is 0.379.
2. The probability that two people chosen at random have the same, pre-specified, genotype. For genotype BB this is 0.103.
3. The probability that one person chosen at random has the same genotype as the crime stain. If the crime stain is of group BB, this probability is 0.321, from Table 1.1.

The phrase at random is taken to include the caveat that the people chosen are unrelated to the suspect.

The relative merits of these answers will be discussed in Section 4.5 for (1) and (2) and Section 9.2 for (3).

1.3.3 Glass fragments

The previous section discussed an example of the interpretation of the evidence of DNA profiling. Consider now an example concerning glass fragments and the measurement of the refractive index of these.

Example 1.2 As before, consider the investigation of a crime. A window has been broken during the commission of the crime. A suspect is found with fragments of glass on his clothing, similar in refractive index to the broken window. Several fragments are taken for investigation and their refractive index measurements taken.

Note that there is a difference here from Example 1.1 where it was assumed that the crime stain had come from the criminal and been transferred to the crime scene. In Example 1.2 glass is transferred from the crime scene to the criminal. Glass on the suspect need not have come from the scene of the crime; it may have come from elsewhere and by perfectly innocent means. This is an asymmetry associated with this kind of evidence. The evidence is known as transfer evidence, as discussed in Section 1.1, because evidence (e.g., blood or glass fragments) has been transferred from the criminal to the scene or vice versa. Transfer from the criminal to the scene has to be considered differently from evidence transferred from the scene to the criminal. A full discussion of this is given in Chapter 8.
Comparison in Example 1.2 has to be made between the two sets of fragments on the basis of their refractive index measurements. The evidential value of the outcome of this comparison has to be assessed. Notice that it is assumed that none of the fragments has any distinctive features and comparison is based only on the refractive index measurements.

Methods for evaluating such evidence were discussed in many papers in the late 1970s and early 1980s (Evett, 1977, 1978; Evett and Lambert, 1982, 1984, 1985; Grove, 1981, 1984; Lindley, 1977a; Seheult, 1978; Shafer, 1982). These methods will be described as appropriate in later chapters. Knowledge-based computer systems such as CAGE (Computer Assistance for Glass Evidence) and CAGE 2000 have been developed (Curran et al., 2000; Hicks, 2004). See also Curran et al. (2000) for a review of current practice in the forensic interpretation of glass evidence.

Evett (1977) gave an example of the sort of problem which may be considered and developed a procedure for evaluating the evidence which mimicked the interpretative thinking of the forensic scientist of the time. The case is an imaginary one. Five fragments from a suspect are to be compared with ten fragments from a window broken at the scene of a crime. The values of the refractive index measurements are given in Table 1.2. The rather arbitrary and hybrid procedure developed by Evett is a two-stage one. It is described here briefly; more details are given in Chapter 4. While it follows the thinking of the forensic scientist, there are interpretative problems, which are described here, in attempting to provide due weight to the evidence. An alternative approach which overcomes these problems is described in Chapter 10.

The first stage is known as the comparison stage. The two sets of measurements are compared. The comparison takes the form of the calculation of a statistic, $D$ say. This statistic provides a measure of the difference, known as a standardised difference, between the two sets of measurements which takes account of the natural variation there is in the refractive index measurements of glass fragments from within the same window. If the absolute value of $D$ is less than (or equal to) some pre-specified value, known as a threshold value, then the two sets of fragments are deemed to be similar and the second stage is implemented. If the absolute value of $D$ is greater than the threshold value then the two sets of fragments are deemed to be dissimilar. The two sets of fragments are then deemed to have come from different sources and the second stage is not implemented. (Note the use here of the word statistic, which in this context

<table>
<thead>
<tr>
<th>Measurements from the window</th>
<th>1.51844</th>
<th>1.51848</th>
<th>1.51844</th>
<th>1.51850</th>
<th>1.51840</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurements from the suspect</td>
<td>1.51848</td>
<td>1.51846</td>
<td>1.51846</td>
<td>1.51844</td>
<td>1.51848</td>
</tr>
</tbody>
</table>
Uncertainty in scientific evidence

can be thought of simply as a function of the observations.) A classical example of such an approach is the use of the Student t-test or the modified Welch test for the comparison of means (Welch, 1937; Walsh et al., 1996; Curran et al., 2000); more details are given in Chapter 10.

The second stage is known as the significance stage. This stage attempts to determine the significance of the finding from the first stage that the two sets of fragments were similar. The significance is determined by calculating the probability of the result that the two sets of fragments were found to be similar, under the assumption that the two sets had come from different sources. If this probability is very low then this assumption is deemed to be false. The fragments are then assumed to come from the same source, an assumption which places the suspect at the crime scene.

The procedure can be criticised on two points. First, in the comparison stage the threshold provides a qualitative step which may provide very different outcomes for two different pairs of observations. One pair of sets of fragments may provide a value of $D$ which is just below the threshold, whereas the other pair may provide a value of $D$ just above the threshold. The first pair will proceed to the significance stage, the second stage will not. Yet, the two pairs may have measurements which are close together. The difference in the consequences is greater than the difference in the measurements' merits (such an approach is called a fall-off-the-cliff effect; see Robertson and Vignaux, 1995a). A better approach, which is described in Chapter 10, provides a measure of the value of the evidence which decreases as the distance between the two sets of measurements increases, subject, as explained later, to the rarity or otherwise of the measurements.

The second criticism is that the result is difficult to interpret. Because of the effect of the comparison stage, the result is not simply the probability of the evidence, assuming the two sets of fragments came from different sources. A reasonable interpretation, as will be explained in Section 3.5, of the value of the evidence is the effect that it has on the odds in favour of guilt of the suspect. In the two-stage approach this effect is difficult to measure. The first stage discards certain sets of measurements which may have come from the same source and may not discard other sets of measurements which have come from different sources. The second stage calculates a probability, not of the evidence but of that part of the evidence for which $D$ was not greater than the threshold value, assuming the two sets came from different sources. It is necessary to compare this probability with the probability of the same result, assuming the two sets came from the same source. There is also an implication in the determination of the probability in the significance stage that a small probability for the evidence, assuming the two sets came from different sources, means that there is a large probability that the two sets came from the same source. This implication is unfounded; see Section 3.3.1.

A review of the two-stage approach and the development of a Bayesian approach is provided by Evett (1986).
Uncertainty in forensic science

As with DNA profiling, there are problems associated with the definition of a suitable population from which probability distributions for refractive measurements may be obtained; see, for example, Walsh and Buckleton (1986).

These examples have been introduced to provide a framework within which the evaluation of evidence may be considered. In order to evaluate evidence, something about which there is much uncertainty, it is necessary to establish a suitable terminology and to have some method for the assessment of uncertainty. First, some terminology will be introduced, followed by the method for the assessment of uncertainty. This method is probability. The role of uncertainty, as represented by probability, in the assessment of the value of scientific evidence will form the basis of the rest of this chapter. A commentary on so-called knowledge management, of which this is one part, has been given by Evett (1993b).

1.4 TERMINOLOGY

It is necessary to have clear definitions of certain terms. The crime scene and suspect materials have fundamentally different roles. Determination of the probability of a match between two randomly chosen sets of materials is not the important issue. One set of materials, crime scene or suspect, can be assumed to have a known source. It is then required to assess the probability of the corresponding item, suspect or crime scene, matching in some sense, the known set of materials, under two competing propositions. Examples 1.1 and 1.2 serve to illustrate this.

Example 1.1 (continued) A crime is committed. A bloodstain is found at the scene of the crime. All innocent explanations for the presence of the stain are eliminated. A suspect is found. His DNA profile is found to match that of the crime stain. The crime scene material is the DNA profile of the crime stain. The suspect material is the DNA profile of the suspect.

Example 1.2 (continued) As before, consider the investigation of a crime. A window has been broken during the commission of the crime. Several fragments are taken for investigation and measurements made of their refractive indices. These fragments, as their origin is known, are sometimes known as control fragments and the corresponding measurements are known as control measurements. A suspect is found. Fragments of glass are found on his person and measurements of the refractive indices of these fragments are made. These fragments (measurements) are sometimes known as recovered fragments (measurements). Their origin is not known. They may have come from the window broken at the crime scene but need not necessarily have done so.

The crime scene material is the fragments of glass and the measurements of refractive index of these at the scene of the crime. The suspect material is the fragments of glass found on the suspect and their refractive index measurements.
Evidence such as this where the source is known and is of a bulk form will be known as *source* or *bulk form* evidence. These fragments of glass will be known as source or bulk fragments and the corresponding measurements will be known as source or bulk measurements, as their source is known and they have been taken from a bulk form of glass, namely a window (Stoney, 1991a). In general, only the term *source* will be used when referring to this type of evidence.

A suspect is found. Fragments of glass are found on his person and measurements of the refractive indices of these fragments are made. Evidence such as this where the evidence has been received and is in particulate form will be known as *receptor* or *transferred particle* evidence. These fragments (measurements) of glass in this example will be known as receptor or transferred particle fragments (measurements). Their origin is not known. They have been ‘received’ from somewhere by the suspect. They are particles which have been transferred to the suspect from somewhere. They may have come from the window broken at the crime scene but need not necessarily have done so.

There will also be occasion to refer to the location at which, or the person on which, the evidence was found. Evidence found at the scene of the crime will be referred to as *crime* evidence. Evidence found on the suspect’s clothing or in the suspect’s natural environment, such as his home, will be referred to as *suspect* evidence. Note that this does not mean that the evidence itself is of a suspect nature!

*Locard’s principle* (see Section 1.1) is that every contact leaves a trace. In the examples above the contact is that of the criminal with the crime scene. In Example 1.1, the trace is the bloodstain at the crime scene. In Example 1.2, the trace is the fragments of glass that would be removed from the crime scene by the criminal (and, later, hopefully, be found on his clothing).

The evidence in both examples is *transfer evidence* (see Section 1.1) or sometimes *trace evidence*. Material has been transferred between the criminal and the scene of the crime. In Example 1.1 blood has been transferred *from* the criminal to the scene of the crime. In Example 1.2 fragments of glass *may* have been transferred *from* the scene of the crime to the criminal. The direction of transfer in these two examples is different. Also, in the first example the blood at the crime scene has been identified as coming from the criminal. Transfer is known to have taken place. In the second example it is not known that glass has been transferred from the scene of the crime to the criminal. The suspect has glass fragments on his clothing but these need not necessarily have come from the scene of the crime. Indeed, if the suspect is innocent and has no connection with the crime scene the fragments will not have come from the crime scene.

Other terms have been suggested and these suggestions provide a potential source of confusion. For example, the term *control* has been used to indicate the material whose origin is known. This can be either the bulk (source) form of the material or the transferred particle form, however. Similarly, the term *recovered* has been used to indicate the material whose origin is unknown.
Again, this can be either the bulk (source) form or the transferred particle form, depending on which has been designated the control form. Alternatively, known has been used for ‘control’ and questioned has been used for ‘recovered’. See, for example, Brown and Cropp (1987). Also Kind et al. (1979) used crime for material known to be associated with a crime and questioned for material thought to be associated with a crime. All these terms are ambiguous. The need to distinguish the various objects or persons associated with a crime was pointed out by Stoney (1984a).

Consider the recovery of a jumper of unknown origin from a crime scene (Stoney, 1991a). A suspect is identified and fibres similar in composition to those on the jumper are found at his place of residence. The two parts of the transfer evidence are the jumper found at the crime scene and the fibres found at the suspect’s place of residence. The bulk (source) form of the material is the jumper. The transferred particle form of the material is the fibres. However, the jumper may not be the control evidence. It is of unknown origin. The fibres, the transferred particle form, could be considered the control as their source is known, in the sense that they have been found at the suspect’s place of residence. They are associated with the suspect in a way the jumper, by itself, is not. Similarly, the fibres have a known origin, and the jumper has a questioned origin.

Definitions given in the context of fibres evidence are provided by Champod and Taroni (1999). The object or person on which traces have been recovered is defined as the receptor and the object or person that could be the source (or one of the sources) of the traces, and which is the origin of the material defined as known material, is defined as the known source.

Material will be referred to as source form where appropriate and to receptor or transferred particle form where appropriate. This terminology conveys no information as to which of the two forms is of known origin. There are two possibilities for the origin of the material which is taken to be known: the scene of the crime and the suspect. One or other is taken to be known, the other to be unknown. The two sets of material are compared by determining two probabilities, both of which depend on what is assumed known and provide probabilities for what is assumed unknown. The two possibilities for the origin of the material which is taken to be known are called scene-anchored and suspect-anchored, where the word ‘anchored’ refers to that which is assumed known (Stoney, 1991a). The distinction between scene-anchoring and suspect-anchoring is important when determining so-called correspondence probabilities (Section 8.2); it is not so important in the determination of likelihood ratios. Reference to form (source or receptor, bulk or transferred particle) is a reference to one of the two parts of the evidence. Reference to anchoring (scene or suspect) is a reference to a perspective for the evaluation of the evidence.

Other terms such as control, known, recovered, questioned, will be avoided as far as possible. However, it is sometimes useful to refer to a sample of material found at the scene of a crime as the crime sample and to a sample of material
found on or about a suspect as the suspect sample. This terminology reflects
the site at which the material was found. It does not indicate the kind of
material (bulk or transferred particle form) or the perspective (scene- or suspect-
anchored) by which the evidence will be evaluated.

1.5 TYPES OF DATA

A generic name for observations which are made on items of interest, such as
bloodstains or refractive indices of glass, is data. There are different types of data
and some terminology is required to differentiate amongst them. For example,
consider the ABO blood grouping system. The observations of interest are the
blood groups of the crime stain and of the suspect. These are not quantifiable.
There is no numerical significance which may be attached to these. The blood
group is a qualitative characteristic. As such, it is an example of so-called
qualitative data. The observation of interest is a quality, the blood group, which
has no numerical significance. The different blood groups are sometime known
as categories. The assignation of a person to a particular category is called a
classification. A person may be said to be classified into one of several categories
(see the discussion on the definition of identification in Section 1.2).

It is not possible to order blood groups and say that one is larger or smaller
than another. However, there are other qualitative data which do have a
natural ordering, such as the level of burns on a body. There is not a numerical
measure of this but the level of burns may be classified as first, second or
third degree, for example. Qualitative data which have no natural ordering are
known as nominal data. Qualitative data to which a natural ordering may be
attached are known as ordinal data. An ordinal characteristic is one in which
there is an underlying order even though it is not quantifiable. Pain is one such
characteristic; level of trauma may be ordered as none, slight, mild, severe, very
severe. The simplest case of nominal data arises when an observation (e.g., of a
person’s blood group) may be classified into one of only two possible categories.
For example, consider the old methodology of blood typing, such as the Kell
genetic marker system where a person may be classified as either Kell+ or
Kell−. Such data are known as binary. Alternatively, the variable of interest,
here the Kell system, is known as dichotomous.

Other types of data are known as quantitative data. These may be either
counts (known as discrete data, since the counts take discrete, integer, values)
or measurements (known as continuous data, since the measurements may take
any value on a continuous interval).

A violent crime involving several people, victims and offenders, may result
in much blood being split and many stains from each of several DNA profiles
being identified. Then the numbers of stains for each of the different profiles are
examples of discrete, quantitative data.

The refractive indices and elemental concentrations of glass fragments are
examples of continuous measurements. In practice, variables are rarely truly
Uncertainty in forensic science

Continuous because of the limits imposed by the sensitivity of the measuring instruments. Refractive indices may be measured only to a certain accuracy.

Observations, or data, may thus be classified as qualitative or quantitative. Qualitative data may be classified further as nominal or ordinal and quantitative data may be classified further as discrete or continuous.

1.6 PROBABILITY

1.6.1 Introduction

The interpretation of scientific evidence may be thought of as the assessment of a comparison. This comparison is that between evidential material found at the scene of a crime (denote this by \( M_c \)) and evidential material found on a suspect, a suspect’s clothing or around his environment (denote this by \( M_s \)). Denote the combination by \( M = (M_c, M_s) \). As a first example, consider the bloodstains of Example 1.1. The crime stain is \( M_c \), the receptor or transferred particle form of the evidential material; \( M_s \) is the genotype of the suspect and is the source form of the material. From Example 1.2, suppose glass is broken during the commission of a crime. \( M_c \) would be the fragments of glass (the source form of the material) found at the crime scene, \( M_s \) would be fragments of glass (the receptor or transferred particle form of the material) found on the clothing of a suspect, and \( M \) would be the two sets of fragments.

Qualities (such as the genotypes) or measurements (such as the refractive indices of the glass fragments) are taken from \( M \). Comparisons are made of the source form and the receptor form. Denote these by \( E_c \) and \( E_s \), respectively, and let \( E = (E_c, E_s) \) denote the combined set. Comparison of \( E_c \) and \( E_s \) is to be made and the assessment of this comparison has to be quantified. The totality of the evidence is denoted \( E_v \) and is such that \( E_v = (M, E) \).

Statistics has developed as a subject, one of whose main concerns is the quantification of the assessments of comparisons. The performance of a new treatment, drug or fertiliser has to be compared with that of an old treatment, drug or fertiliser, for example. It seems natural that statistics and forensic science should come together. Two samples, the crime and the suspect sample, are to be compared. Yet, apart from the examples discussed in Chapter 4, it is only recently that this has happened. As discussed in Section 1.2, there have been several books describing the role of statistics in the law. Until the first edition of this book there had been none concerned with statistics and the evaluation of scientific evidence. Two factors may have been responsible for this.

First, there was a lack of suitable data from relevant populations. There was a consequential lack of a baseline against which measures of typicality of any characteristics of interest might be determined. One exception are the reference
data which have been available for many years on blood group frequencies among certain populations. Not only has it been possible to say that the suspect’s blood group matched that of a stain found at the scene of a crime but also that this group is only present in, say, 0.01% of the population. Now these have been superseded by databases of DNA profiles. Also, data collections exist for the refractive index of glass fragments found at random on clothing and for transfer and persistence parameters linked to glass evidence (Curran et al., 2000). Contributions towards estimating the frequency of fibre types have also been published (Grieve and Biermann, 1997; Grieve, 2000a, b; Grieve et al., 2001). There is also much information about the frequency of characteristics in DNA profiles. Announcements of population data are published regularly in peer-reviewed journals such as Forensic Science International and the Journal of Forensic Sciences.

Secondly, the approach adopted by forensic scientists in the assessment of their evidence has been difficult to model. The approach has been one of comparison and significance. Characteristics of the crime and suspect samples are compared. If the examining scientists believe them to be similar, the typicality, and hence the significance of the similarity, of the characteristics is then assessed. This approach is what has been modelled by the two-stage approach of Evett (1977), described briefly in Section 1.3.3 and in fuller detail in Chapter 4. However, interpretation of the results provided by this approach is difficult.

Then, in a classic paper, Lindley (1977a) described an approach which was easy to justify, to implement and to interpret. It combined the two parts of the two-stage approach into one statistic and is discussed in detail in Section 10.2. The approach compares two probabilities, the probability of the evidence, assuming one proposition about the suspect to be true (that he is guilty, for example) and the probability of the evidence, assuming another proposition about the suspect to be true (that he is innocent, for example). (Note: some people use the term hypothesis rather than proposition; the authors will endeavour to use the term proposition as they believe this reduces the risk of confusion of their ideas with the ideas of hypothesis testing associated with the alternative term.) This approach implies that it is not enough for a prosecutor to show that evidence is unlikely if a suspect is innocent. The evidence has also to be more likely if the suspect is guilty. Such an approach had a good historical pedigree (Good, 1950; see also Good, 1991, for a review) yet it had received very little attention in the forensic science literature, even though it was clearly proposed at the beginning of the twentieth century (Taroni et al., 1998). It is also capable of extension beyond the particular type of example discussed by Lindley, as will be seen by the discussion throughout this book, for example in Section 10.4.

However, in order to proceed it is necessary to have some idea about how uncertainty can be measured. This is best done through probability (Lindley, 1991, 1998).
1.6.2 A standard for uncertainty

An excellent description of probability and its role in forensic science has been given by Lindley (1991). Lindley’s description starts with the idea of a standard for uncertainty. He provides an analogy using the concept of balls in an urn. Initially, the balls are of two different colours, black and white. In all other respects – size, weight, texture etc. – they are identical. In particular, if one were to pick a ball from the urn, without looking at its colour, it would not be possible to tell what colour it was. The two colours of balls are in the urn in proportions $b$ and $w$ for black and white balls, respectively, such that $b + w = 1$. For example, if there were 10 balls in the urn of which 6 were black and 4 were white, then $b = 0.6$, $w = 0.4$ and $b + w = 0.6 + 0.4 = 1$.

The urn is shaken up and the balls thoroughly mixed. A ball is then drawn from the urn. Because of the shaking and mixing it is assumed that each ball, regardless of colour, is equally likely to be selected. Such a selection process, in which each ball is equally likely to be selected, is known as a random selection, and the chosen ball is said to have been chosen at random.

The ball, chosen at random, can be either black, an event which will be denoted $B$, or white, an event which will be denoted $W$. There are no other possibilities; one and only one of these two events has to occur. The uncertainty of the event $B$, the drawing of a black ball, is related to the proportion $b$ of black balls in the urn. If $b$ is small (close to zero), $B$ is unlikely. If $b$ is large (close to 1), $B$ is likely. A proportion $b$ close to 1/2 implies that $B$ and $W$ are about equally likely. The proportion $b$ is referred to as the probability of obtaining a black ball on a single random draw from the urn. In a similar way, the proportion $w$ is referred to as the probability of obtaining a white ball on a single random draw from the urn.

Notice that on this simple model probability is represented by a proportion. As such it can vary between 0 and 1. A value of $b = 0$ occurs if there are no black balls in the urn and it is, therefore, impossible to draw a black ball from the urn. The probability of obtaining a black ball on a single random draw from the urn is zero. A value of $b = 1$ occurs if all the balls in the urn are black. It is certain that a ball drawn at random from the urn will be black. The probability of obtaining a black ball on a single random draw from the urn is one. All values between these extremes of 0 and 1 are possible (by considering very large urns containing very large numbers of balls).

A ball has been drawn at random from the urn. What is the probability that the selected ball is black? The event $B$ is the selection of a black ball. Each ball has an equal chance of being selected. The colours black and white of the balls are in the proportions $b$ and $w$. The proportion, $b$, of black balls corresponds to the probability that a ball, drawn in the manner described (i.e., at random) from the urn is black. It is then said that the probability a black ball is drawn from the urn, when selection is made at random, is $b$. Some notation is needed to denote the probability of an event. The probability of $B$, the drawing of a
black ball, is denoted \( Pr(B) \), and similarly \( Pr(W) \) denotes the probability of the drawing of a white ball. Then it can be written that \( Pr(B) = b, Pr(W) = w \). Note that

\[
Pr(B) + Pr(W) = b + w = 1.
\]

This concept of balls in an urn can be used as a reference for considering uncertain events. Let \( R \) denote the uncertain event that the England football team will win the next European football championship. Let \( B \) denote the uncertain event that a black ball will be drawn from the urn. A choice has to be made between \( R \) and \( B \) and this choice has to be ethically neutral. If \( B \) is chosen and a black ball is indeed drawn from the urn then a prize is won. If \( R \) is chosen and England do win the championship the same prize is won. The proportion \( b \) of black balls in the urn is known in advance. Obviously, if \( b = 0 \) then \( R \) is the better choice, assuming, of course, that England do have some non-zero chance of winning the championship. If \( b = 1 \) then \( B \) is the better choice. Somewhere in the interval \([0, 1]\), there is a value of \( b, b_0 \) say, where the choice does not matter. One is indifferent as to whether \( R \) or \( B \) is chosen. If \( B \) is chosen, \( Pr(B) = b_0 \). Then it is said that \( Pr(R) = b_0 \), also. In this way the uncertainty in relation to any event can be measured by a probability \( b_0 \), where \( b_0 \) is the proportion of black balls which leads to indifference between the two choices, namely the choice of drawing a black ball from the urn and the choice of the uncertain event in whose probability one is interested.

Notice, though, that there is a difference between these two probabilities. By counting, the proportion of black balls in the urn can be determined precisely. Probabilities of other events such as the outcome of the toss of a coin or the roll of a die are also relatively straightforward to determine, based on assumed physical characteristics such as fair coins and fair dice. Let \( H \) denote the event that when a coin is tossed it lands head uppermost. Then, for a fair coin, in which the outcomes of a head \( H \) and a tail \( T \) at any one toss are equally likely, the probability the coin comes down head uppermost is 1/2. Let \( F \) denote the event that when a die is rolled it lands 4 uppermost. Then, for a fair die, in which the outcomes 1, 2, 3, 4, 5, 6 at any one roll are equally likely, the probability the die lands 4 uppermost is 1/6.

Probabilities relating to the outcomes of sporting events, such as football matches or championships or horse races, or to the outcome of a civil or criminal trial, are rather different in nature. It may be difficult to decide on a particular value for \( b_0 \). The value may change as evidence accumulates such as the results of particular matches and the fitness or otherwise of particular players, the fitness of horses, the identity of the jockey or the going of the race track. Also, different people may attach different values to the probability of a particular event. These kinds of probability are sometimes known as subjective or personal probabilities; see de Finetti (1931), Good (1959), Savage (1954) and DeGroot (1970). Another term is measure of belief, since the probability may be
thought to provide a measure of one’s belief in a particular event. Despite these difficulties, the arguments concerning probability still hold. Given an uncertain event \( R \), the probability of \( R \), \( Pr(R) \), is defined as the proportion of balls \( b_0 \) in the urn such that if one had to choose between \( B \) (the event that a black ball was chosen) where \( Pr(B) = b_0 \) and \( R \) then one would be indifferent as to which one was chosen. There are difficulties, but the point of importance is that a standard for probability exists. A comment on subjective probabilities is given in Section 9.5.5. A use of probability as a measure of belief is described in Section 9.5, where it is used to represent relevance. The differences and similarities in the two kinds of probability discussed above and their ability to be combined has been referred to as a duality (Hacking, 1975).

It is helpful also to consider two quotes concerning the relationship amongst probability, logic and consistency, both from Ramsey (1931).

We find, therefore, that a precise account of the nature of partial beliefs reveals that the laws of probability are laws of consistency, an extension to partial beliefs of formal logic, the logic of consistency. They do not depend for their meaning on any degree of belief in a proposition being uniquely determined as the rational one; they merely distinguish those sets of beliefs which obey them as consistent ones. (p. 182)

We do not regard it as belonging to formal logic to say what should be a man’s expectation of drawing a white or black ball from an urn; his original expectations may within the limits of consistency be any he likes; all we have to point out is that if he has certain expectations he is bound in consistency to have certain others. This is simply bringing probability into line with ordinary formal logic, which does not criticize premisses but merely declares that certain conclusions are the only ones consistent with them. (p. 189)

### 1.6.3 Events

The outcome of the drawing of a ball from the urn was called an event. If the ball was black, the event was denoted \( B \). If the ball was white, the event was denoted \( W \). It was not certain which of the two events would happen; would the ball be black, event \( B \), or white, event \( W \)? The degree of uncertainty of the event (\( B \) or \( W \)) was measured by the proportion of balls of the appropriate colour (\( B \) or \( W \)) in the urn and this proportion was called the probability of the event (\( B \) or \( W \)). In general, for an event \( R \), \( Pr(R) \) denotes the probability that \( R \) occurs.

Events can be events which may have happened (past events), events which may be relevant at the present time (present events) and events which may happen in the future (future events). There is uncertainty associated with each of these. In each case, a probability may be associated with the event.

- Past event: a crime is committed and a bloodstain of a particular type is found at the crime scene. A suspect is found. The event of interest is that the suspect left the stain at the crime scene. Though the suspect either did or
did not leave the stain, the knowledge of it is uncertain and hence the event can have a probability associated with it.

• Present event: a person is selected. The event of interest is that he is of blood group O. Again, before the result of a blood test is available, this knowledge is uncertain.

• Future event: The event of interest is that it will rain tomorrow.

All of these events are uncertain and have probabilities associated with them. Notice, in particular, that even if an event has happened, there can still be uncertainty associated with it. The probability the suspect left the stain at the crime scene requires consideration of many factors, including the possible location of the suspect at the crime scene and the properties of transfer of blood from a person to a site. With reference to the blood group, consideration has to be given to the proportion of people in some population of blood group O. Probabilistic predictions are common with weather forecasting. Thus, it may be said, for example, that the probability it will rain tomorrow is 0.8 (though it may not always be obvious what this means).

1.6.4 Subjective probability

In forensic science, it is often emphasised that there is a real paucity of numerical data, so that the calculation of likelihood ratios (Section 3.4.1) is sometimes very difficult. Examples of this difficulty are the numerical assessments of parameters such as transfer or persistence probabilities (see Chapter 8) or even the relevance of a piece of evidence (see Section 9.5). The Bayesian approach considers probabilities as measures of belief (also called subjective probabilities) since such probabilities may be thought of as measures of one’s belief in the occurrence of a particular event. The approach enables the combination of the objective probabilities, based on data, and subjective probabilities, for which the certified knowledge and experience of the forensic scientist may assist in the provision of estimates. Jurists are also interested in probability calculations using subjective probabilities, notably probabilities associated with the credibility of witnesses and the conclusions that might be drawn from their testimony.

From a formal point of view, the frequentist definition of probability involves a long sequence of repetitions of a given situation, under identical conditions. Consider a sequence of $N$ repetitions in which an event $E$ occurs $X$ times, where $X$ is some value greater than or equal to 0 and less than or equal to $N$. The relative frequency $X/N$ could vary in different sequences of $N$ repetitions, but it is supposed that, in a sequence where the number $N$ of repetitions grows indefinitely under identical conditions, the relative frequency tends to a definite limiting value. In a frequentist framework the probability of event $E$, denoted $Pr(E)$, is defined to be that limiting value. Note, as with the balls in the urn, that this is a value between 0 and 1. In reality, it is difficult, if not impossible,
Uncertainty in forensic science

to maintain identical conditions between trials. Therefore, in anything other than idealised situations, such a definition of probability proves unworkable. For example, consider prediction of the unemployment rate for the following year. It is not possible to use the frequentist definition to determine the probability that the rate will lie between 3% and 4% of the workforce, since it is not possible to consider unemployment as a sequence of repetitions under identical conditions. Unemployment in the following year is a unique, one-time event (Berger, 1985).

Frequentist probabilities are objective probabilities. They are objective in the sense that there is a well-defined set of circumstances for the long-run repetition of the trials, such that the corresponding probabilities are well defined and that one’s personal or subjective views will not alter the value of the probabilities. Each person considering these circumstances will provide the same values for the probabilities. The frequency model relates to a relative frequency obtained in a long sequence of trials, assumed to be performed in an identical manner, physically independent of each other. Such a circumstance has certain difficulties. If taken strictly, this point of view does not allow a statement of probability for any situation that does not happen to be embedded, at least conceptually, in a long sequence of events giving equally likely outcomes.

The underlying idea of subjective probability is that the probability an event happens reflects a measure of personal belief in the occurrence of the event. For example, a person may have a personal feeling that the unemployment rate will be between 3% and 4%, even though no frequency probability can be assigned to the event. There is nothing surprising about this. It is common to think in terms of personal probabilities all the time, such as when betting on the outcome of a football game or when stating the probability of rain tomorrow.

In many situations, law being a prime example, we cannot assume equally likely outcomes any more than we can count past occurrences of events to determine relative frequencies. The reason is that the events of interest, if they have occurred, have done so only once (Schum, 2000). So, a subjective probability is defined as ‘a degree of belief (as actually held by someone based on his whole knowledge, experience, information) regarding the truth of a statement, or event $E$ (a fully specified single event or statement whose truth or falsity is, for whatever reason, unknown to the person)’ (de Finetti, 1968, p. 45). There are three factors to consider for this probabilistic assessment. First, it depends on the available information. Second, it may change as the information changes. Third, it may vary amongst individuals because different individuals may have different information or assessment criteria. The only constraint in such an assessment is that it must be what is known as coherent. Coherence may be understood through consideration of subjective probability in terms of betting, in particular on a horse race. For the probabilities on winning for each horse in a race to be coherent, the sum of the probabilities over all the horses in the race must be 1 (Taroni et al., 2001).

Under either definition (frequentist or Bayesian), probability takes a value between 0 and 1. Events or parameters of interest, in a wide range of academic
fields (such as history, theology, law, forensic science), are usually not the result of repetitive or replicable processes. These events are singular, unique or one of a kind. It is not possible to repeat the events under identical conditions and tabulate the number of occasions on which some past event actually occurred. The use of subjective probabilities allows us to consider probability for events in situations such as these.

For a historical and philosophical definition of subjective probabilities and a commentary on the work of statisticians, de Finetti and Savage, working in this field in the middle of the twentieth century, see Lindley (1980) and Taroni et al. (2001).

1.6.5 Laws of probability

There are several laws of probability which describe the values which probability may take and how probabilities may be combined. These laws are given here, first for events which are not conditioned on any other information and then for events which are conditioned on other information.

The first law of probability has already been suggested implicitly in the context of proportions.

First law of probability Probability can take any value between 0 and 1, inclusive, and only those values. Let \( R \) be any event and let \( P(\text{R}) \) denote the probability that \( R \) occurs. Then \( 0 \leq P(\text{R}) \leq 1 \). For an event which is known to be impossible, the probability is zero. Thus if \( R \) is impossible, \( P(\text{R}) = 0 \). This law is sometimes known as the convexity rule (Lindley, 1991).

Consider the hypothetical example of the balls in the urn of which a proportion \( b \) are black and a proportion \( w \) white, with no other colours present, such that \( b + w = 1 \). Proportions lie between 0 and 1: hence \( 0 \leq b \leq 1 \), \( 0 \leq w \leq 1 \). For any event \( R \), \( 0 \leq P(\text{R}) \leq 1 \). Consider \( B \), the drawing of a black ball. If this event is impossible then there are no black balls in the urn and \( b = 0 \). This law is sometimes strengthened to say that a probability can only be 0 when the associated event is known to be impossible.

The first law concerns only one event. The next two laws, sometimes known as the second and third laws of probability, are concerned with combinations of events. Events combine in two ways. Let \( R \) and \( S \) be two events. One form of combination is to consider the event ‘\( R \) and \( S \)’, the event that occurs if and only if \( R \) and \( S \) both occur. This is known as the conjunction of \( R \) and \( S \).

Consider the roll of a six-sided fair die. Let \( R \) denote the throwing of an odd number. Let \( S \) denote the throwing of a number greater than 3 (i.e., a 4, 5 or 6). Then the event ‘\( R \) and \( S \)’ denotes the throwing of a 5.

Secondly, consider rolling two six-sided fair die. Let \( R \) denote the throwing of a six with the first die. Let \( S \) denote the throwing of a six with the second die. Then the event ‘\( R \) and \( S \)’ denotes the throwing of a double six.
The second form of combination is to consider the event ‘R or S’, the event that occurs if R or S (or both) occurs. This is known as the disjunction of R and S.

Consider again the roll of a single six-sided fair die. Let R, the throwing of an odd number (1, 3 or 5), and S, the throwing of a number greater than 3 (4, 5 or 6), be as before. Then ‘R or S’ denotes the throwing of any number other than a 2 (which is both even and less than 3).

Secondly, consider drawing a card from a well-shuffled pack of 52 playing cards, such that each card is equally likely to be drawn. Let R denote the event that the card drawn is a spade. Let S denote the event that the card drawn is a club. Then the event ‘R or S’ is the event that the card drawn is from a black suit.

The second law of probability concerns the disjunction ‘R or S’ of two events. Events are called mutually exclusive when the occurrence of one excludes the occurrence of the other. For such events, the conjunction ‘R and S’ is impossible. Thus \( Pr(R \text{ and } S) = 0 \).

The third law of probability concerns the conjunction of two events. Initially, it will be assumed that the two events are what is known as independent. By independence is meant that knowledge of the occurrence of one of the two events does not alter the probability of occurrence of the other event. Thus two events which are mutually exclusive cannot be independent. As a simple example of independence, consider the rolling of two six-sided fair dice, A and B say. The outcome of the throw of A does not affect the outcome of the throw of B. If A lands ‘6’ uppermost, this result does not alter the probability that B...
will land ‘6’ uppermost. The same argument applies if one die is rolled two or more times. Outcomes of earlier throws do not affect the outcomes of later throws.

**Third law of probability**  Let \( R \) and \( S \) be two independent events. Then

\[
Pr(\text{R and S}) = Pr(\text{R}) \times Pr(\text{S}).
\]

This relationship is sometimes used as the definition of independence. Thus, two events \( R \) and \( S \) are said to be independent if \( Pr(\text{R and S}) = Pr(\text{R}) \times Pr(\text{S}) \). There is symmetry in this definition. Event \( R \) is independent of \( S \) and \( S \) is independent of \( R \). This law may be generalised to more than two events. Consider \( n \) events \( S_1, S_2, \ldots, S_n \). If they are mutually independent, then

\[
Pr(S_1 \text{ and } S_2 \text{ and } \ldots \text{ and } S_n) = Pr(S_1) \times Pr(S_2) \times \cdots \times Pr(S_n) = \prod_{i=1}^{n} Pr(S_i).
\]

### 1.6.6 Dependent events and background information

Not all events are independent. Consider one roll of a fair die, with \( R \) the throwing of an odd number as before, and \( S \) the throwing of a number greater than 3 as before. Then, \( Pr(R) = 1/2, Pr(S) = 1/2, Pr(R) \times Pr(S) = 1/4 \) but \( Pr(R \text{ and } S) = Pr(\text{throwing a5}) = 1/6 \).

Events which are not independent are said to be dependent. The third law of probability for dependent events was first presented by Thomas Bayes (1763), see also Barnard (1958), Pearson and Kendall (1970) and Poincaré (1912). It is the general law for the conjunction of events. The law for independent events is a special case. Before the general statement of the third law is made, some discussion of dependence is helpful.

It is useful to consider that a probability assessment depends on two things: the event \( R \) whose probability is being considered and the information \( I \) available when \( R \) is being considered. The probability \( Pr(R \mid I) \) is referred to as a conditional probability, acknowledging that \( R \) is conditional or dependent on \( I \). Note the use of the vertical bar \( \mid \). Events listed to the left of it are events whose probability is of interest. Events listed to the right are events whose outcomes are known and which may affect the probability of the events listed to the left of the bar, the vertical bar having the meaning ‘given’ or ‘conditional on’.

Consider a defendant in a trial who may or may not be guilty. Denote the event that he is guilty by \( G \). The uncertainty associated with his guilt, the probability that he is guilty, may be denoted by \( Pr(G) \). It is a subjective probability. The uncertainty will fluctuate during the course of a trial. It will fluctuate as evidence is presented. It depends on the evidence. Yet neither the notation, \( Pr(G) \), nor the language, the probability of guilt, makes mention of
Uncertainty in forensic science

this dependence. The probability of guilt at any particular time depends on the knowledge (or information) available at that time. Denote this information by $I$. It is then possible to speak of the probability of guilt given, or conditional on, the information available at that time. This is written as $Pr(G \mid I)$. If additional evidence $E$ is presented this then becomes, along with $I$, part of what is known. What is taken as known is then ‘$E$ and $I$’, the conjunction of $E$ and $I$. The revised probability of guilt is $Pr(G \mid E$ and $I)$.

All probabilities may be thought of as conditional probabilities. Personal experience informs judgements made about events. For example, judgement concerning the probability of rain the following day is conditioned on personal experiences of rain following days with similar weather patterns to the current one. Similarly, judgement concerning the value of evidence or the guilt of a suspect is conditional on many factors. These include other evidence at the trial but may also include a factor to account for the reliability of the evidence. There may be eyewitness evidence that the suspect was seen at the scene of the crime, but this evidence may be felt to be unreliable. Its value will then be lessened.

The value of scientific evidence will be conditioned on the background data relevant to the type of evidence being assessed. Evidence concerning frequencies of different DNA profiles will be conditioned on information regarding ethnicity of the people concerned for the values of these frequencies. Evidence concerning distributions of the refractive indices of glass fragments will be conditioned on information regarding the type of glass from which the fragments have come (e.g., building window, car headlights). The existence of such conditioning events will not always be stated explicitly. However, they should not be forgotten. As stated above, all probabilities may be thought of as conditional probabilities. The first two laws of probability can be stated in the new notation, for events $R,S$ and information $I$ as follows:

First law

$$0 \leq Pr(R \mid I) \leq 1$$

and $Pr(\text{not } I \mid I) = 0$. If $I$ is known, the event ‘not $I$’ is impossible and thus $Pr(I \mid I) = 1$.

Second law

$$Pr(R \text{ or } S \mid I) = Pr(R \mid I) + Pr(S \mid I) - Pr(R \text{ and } S \mid I).$$

Third law for independent events

If $R$ and $S$ are independent, then, conditional on $I$,

$$Pr(R \text{ and } S \mid I) = Pr(R \mid I) \times Pr(S \mid I).$$
Notice that the event \( I \) appears as a conditioning event in all the probability expressions. The laws are the same as before but with this simple extension. The third law for dependent events is given later by (1.7).

As an example of the use of the ideas of independence, consider a diallelic system in genetics in which the alleles are denoted \( A \) and \( a \), with \( \Pr(A) = p \), \( \Pr(a) = q \) and \( \Pr(A) + \Pr(a) = p + q = 1 \). This gives rise to three genotypes that, assuming Hardy–Weinberg equilibrium to hold, are expected to have the following probabilities:

- \( p^2 \) (homozygotes for allele \( A \)),
- \( 2pq \) (heterozygotes),
- \( q^2 \) (homozygotes for allele \( a \)).

The genotype probabilities are calculated by simply multiplying the two allele probabilities together on the assumption that the allele inherited from one’s father is independent of the allele inherited from one’s mother. The factor 2 arises for heterozygotes case because two cases must be considered, that in which allele \( A \) was contributed by the mother and allele \( a \) by the father, and vice versa. Each of these cases has probability \( pq \) because of the assumption of independence (see Table 1.3). The particular locus under consideration is said to be in Hardy–Weinberg equilibrium when the two parental alleles are independent.

Suppose now that events \( R \) and \( S \) are dependent – that the knowledge that \( R \) has occurred affects the probability that \( S \) will occur, and vice versa. For example, let \( R \) be the outcome of drawing a card from a well-shuffled pack of 52 playing cards. This card is not replaced in the pack, so there are now only 51 cards in the pack. Let \( S \) be the draw of a card from this reduced pack of cards. Let \( R \) be the event ‘an ace is drawn’. Thus \( \Pr(R) = 4/52 = 1/13 \). (Note here the conditioning information \( I \) that the pack is well shuffled, with its implication that each of the 52 cards is equally likely to be drawn, has been omitted for simplicity of notation; explicit mention of \( I \) will be omitted in many cases but its existence should never be forgotten.) Let \( S \) be the event ‘an ace is drawn’ also. Then \( \Pr(S | R) \) is the probability that an ace was drawn at the second draw, given that an ace was drawn at the first draw (and given everything

### Table 1.3 Genotype probabilities, assuming Hardy–Weinberg equilibrium, for a diallelic system with allele probabilities \( p \) and \( q \)

<table>
<thead>
<tr>
<th>Allele from mother</th>
<th>Allele from father</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( A(p) )</td>
</tr>
<tr>
<td>( A(p) )</td>
<td>( p^2 )</td>
</tr>
<tr>
<td>( a(q) )</td>
<td>( pq )</td>
</tr>
</tbody>
</table>
else that is known, in particular that the first card was not replaced. There are 51 cards at the time of the second draw, of which 3 are aces. (Remember that an ace was drawn the first time, which is the information contained in \( R \).) Thus \( \Pr(S \mid R) = \frac{3}{51} \). It is now possible to formulate the third law of probability for dependent events:

**Third law of probability for dependent events**

\[
\Pr(R \text{ and } S \mid I) = \Pr(R \mid I) \times \Pr(S \mid R \text{ and } I). 
\]

Thus in the example of the drawing of the aces from the pack, the probability of drawing two aces is

\[
\Pr(R \text{ and } S \mid I) = \frac{4}{52} \times \frac{3}{51}.
\]

Note that if the first card had been replaced in the pack and the pack shuffled again, the two draws would have been independent and the probability of drawing two aces would have been \( \frac{4}{52} \times \frac{4}{52} \).

**Example 1.3** Consider the genetic markers Kell and Duffy. For both of these markers a person may be either positive (+) or negative (−). In a relevant population 60% of the people are Kell+ and 70% are Duffy+. An individual is selected at random from this population; that is, an individual is selected using a procedure such that each individual is equally likely to be selected.

Consider the Kell marker first in relation to the urn example. Let those people with Kell+ correspond to black balls and those with Kell− correspond to white balls. By analogy with the urn example, the probability a randomly selected individual (ball) is Kell+ (black) corresponds to the proportion of Kell+ people (black balls) in the population, namely 0.6. Let \( R \) be the event that a randomly selected individual is Kell+. Then \( \Pr(R) = 0.6 \). Similarly, let \( S \) be the event that a randomly selected individual is Duffy+. Then \( \Pr(S) = 0.7 \).

If Kell and Duffy markers were independent, it would be calculated that the probability that a person selected at random were both Kell+ and Duffy+, namely \( \Pr(R \text{ and } S) \), was given by

\[
\Pr(R \text{ and } S) = \Pr(R) \times \Pr(S) = 0.6 \times 0.7 = 0.42.
\]

However, it may not necessarily be the case that 42% of the population is both Kell+ and Duffy+. Nothing in the information currently available justifies the assumption of independence used to derive this result. It is perfectly feasible that, say, 34% of the population are both Kell+ and Duffy+, that is, \( \Pr(R \text{ and } S) = 0.34 \). In such a situation where \( \Pr(R \text{ and } S) \neq \Pr(R) \times \Pr(S) \) it can be said that the Kell and Duffy genetic markers are not independent. (See also Chapter 3.)
Table 1.4 The proportion of people in a population who fall into the four possible categories of genetic markers

<table>
<thead>
<tr>
<th></th>
<th>Duffy</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>34</td>
<td>60</td>
</tr>
<tr>
<td>−</td>
<td>36</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>70</td>
<td>100</td>
</tr>
</tbody>
</table>

The information about the Kell and Duffy genetic markers can be represented in a tabular form, known as a 2 × 2 (two-by-two) table as in Table 1.4, the rows of which refer to Kell (positive or negative) and the columns of which refer to Duffy (positive or negative). It is possible to verify the third law of probability for dependent events (1.7) using this table. Of the 60% of the population who are Kell+ (R), 34/60 are Duffy+ (S). Thus Pr(S | R) = 34/60 = 0.68. Also, Pr(R) = 60/100 and

\[ Pr(R \text{ and } S) = Pr(R) \times Pr(S | R) = \frac{60}{100} \times \frac{34}{60} = 0.34, \]

as may be derived from the table directly.

This example also illustrates the symmetry of the relationship between R and S as follows. Of the 70% of the population who are Duffy+ (S), 34/70 are Kell+ (R). Thus Pr(R | S) = 34/70. Also Pr(S) = 70/100 and

\[ Pr(R \text{ and } S) = Pr(S) \times Pr(R | S) = \frac{70}{100} \times \frac{34}{70} = 0.34. \]

Thus, for dependent events, R and S, the third law of probability, (1.7) may be written as

\[ Pr(R \text{ and } S) = Pr(S | R) \times Pr(R) = Pr(R | S) \times Pr(S). \quad (1.8) \]

where the conditioning on I has been omitted. The result for independent events follows as a special case with Pr(R | S) = Pr(R) and Pr(S | R) = Pr(S).

1.6.7 Law of total probability

This is sometimes known as the extension of the conversation (Lindley, 1991). Events S₁, S₂, . . . , Sₙ are said to be mutually exclusive and exhaustive if one of them has to be true and only one of them can be true; they exhaust the possibilities and the occurrence of one excludes the possibility of any other.
Alternatively, they are called a partition. The event \((S_1 \text{ or } \ldots \text{ or } S_n)\) formed from the conjunction of the individual events \(S_1, \ldots, S_n\) is certain to happen since the events are exclusive. Thus, it has probability 1 and

\[
Pr(S_1 \text{ or } \ldots \text{ or } S_n) = Pr(S_1) + \cdots + Pr(S_n) = 1.
\] (1.9)

a generalisation of the second law of probability (1.5) for exclusive events.

Consider allelic distributions at a locus, for example locus \textit{LDLR} for Caucasians in Chicago (Johnson and Peterson, 1999). The three alleles \(A, B\) and \(C\) are mutually exclusive and exhaustive.

Consider \(n = 2\). Let \(R\) be any other event. The events ‘\(R\) and \(S_1\)’ and ‘\(R\) and \(S_2\)’ are exclusive. They cannot both occur. The event ‘ \(R\) and \(S_1\)’ or ‘\(R\) and \(S_2\)’ is simply \(R\). Let \(S_1\) be male, \(S_2\) be female, \(R\) be left-handed. Then ‘\(R\) and \(S_1\)’ denotes a left-handed male, while ‘\(R\) and \(S_2\)’ denotes a left-handed female. The event ‘ \(R\) and \(S_1\)’ or ‘\(R\) and \(S_2\)’ is the event that a person is a left-handed male or a left-handed female, which implies the person is left-handed \((R)\). Thus,

\[
Pr(R) = Pr(R \text{ and } S_1) + Pr(R \text{ and } S_2)
\]

\[
= Pr(R \mid S_1)Pr(S_1) + Pr(R \mid S_2)Pr(S_2).
\]

The argument extends to any number of mutually exclusive and exhaustive events to give the law of total probability.

\[\text{Law of total probability} \quad \text{If } S_1, S_2, \ldots, S_n \text{ are } n \text{ mutually exclusive and exhaustive events,}
\]

\[
Pr(R) = Pr(R \mid S_1)Pr(S_1) + \cdots + Pr(R \mid S_n)Pr(S_n).
\] (1.10)

An example for blood types and paternity cases is given by Lindley (1991). Consider two possible groups, \(S_1\) (Rh\(−\)) and \(S_2\) (Rh\(+\)) for the father, so here \(n = 2\). Assume the relative frequencies of the two groups are \(p\) and \(1 - p\), respectively. The child is Rh\(−\) (event \(R\)) and the mother is also Rh\(−\) (event \(M\)). The probability of interest is the probability a Rh\(−\) mother will have a Rh\(−\) child, in symbols \(Pr(R \mid M)\). This probability is not easily derived directly but the derivation is fairly straightforward if the law of total probability is invoked to include the father:

\[
Pr(R \mid M) = Pr(R \mid M \text{ and } S_1)Pr(S_1 \mid M) + Pr(R \mid M \text{ and } S_2)Pr(S_2 \mid M).
\] (1.11)

This is a generalisation of the law to include information \(M\). If both parents are Rh\(−\), event ‘\(M\) and \(S_1\)’, then the child is Rh\(−\) with probability 1, so \(Pr(R \mid M \text{ and } S_1) = 1\). If the father is Rh\(+\) (the mother is still Rh\(−\)), event \(S_2\), then \(Pr(R \mid M \text{ and } S_2) = 1/2\). Assume that parents mate at random with respect
to the rhesus quality. Then $Pr(S_1 | M) = p$, the relative frequency of Rh− in the population, independent of $M$. Similarly, $Pr(S_2 | M) = 1 - p$, the relative frequency of Rh+ in the population. These probabilities can now be inserted in (1.11) to obtain

$$Pr(R | M) = 1(p) + \frac{1}{2}(1 - p) = (1 + p)/2$$

for the probability that a Rh− mother will have a Rh− child. This result is not intuitively obvious, unless one considers the approach based on the law of total probability.

An example using DNA profiles is given in Evett and Weir (1998). According to the 1991 census, the New Zealand population consists of 83.47% Caucasians, 12.19% Maoris and 4.34% Pacific Islanders; denote the event that a person chosen at random from the 1991 New Zealand population is Caucasian, Maori or Pacific Islander as $Ca$, $Ma$ and $Pa$, respectively. The probabilities of finding the same YNH24 genotype $g$ (event $G$), as in a crime sample, for a Caucasian, Maori or Pacific Islander are 0.012, 0.045 and 0.039, respectively. These values are the assessments for the following three conditional probabilities:

$$Pr(G | Ca), Pr(G | Ma), Pr(G | Pa).$$

Then the probability of finding the YNH24 genotype, $G$, in a person taken at random from the whole population of New Zealand is

$$Pr(G) = Pr(G | Ca)Pr(Ca) + Pr(G | Ma)Pr(Ma) + Pr(G | Pa)Pr(Pa)$$

$$= 0.012 \times 0.8347 + 0.045 \times 0.1219 + 0.039 \times 0.0434$$

$$= 0.017.$$

A further extension of this law to consider probabilities for combinations of genetic marker systems in a racially heterogeneous population has been given by Walsh and Buckleton (1988). Let $C$ and $D$ be two genetic marker systems. Let $S_1$ and $S_2$ be two mutually exclusive and exhaustive sub-populations such that a person from the population belongs to one and only one of $S_1$ and $S_2$. Let $Pr(S_1)$ and $Pr(S_2)$ be the probabilities that a person chosen at random from the population belongs to $S_1$ and to $S_2$, respectively. Then $Pr(S_1) + Pr(S_2) = 1$. Within each sub-population $C$ and $D$ are independent so that the probability an individual chosen at random from one of these sub-populations is of type $CD$ is simply the product of the individual probabilities. Thus

$$Pr(CD | S_1) = Pr(C | S_1) \times Pr(D | S_1).$$

$$Pr(CD | S_2) = Pr(C | S_2) \times Pr(D | S_2).$$
However, such a so-called conditional independence result does not imply unconditional independence (i.e., that \( \Pr(CD) = \Pr(C) \times \Pr(D) \)). The probability that an individual chosen at random from the population is \( CD \), without regard to his sub-population membership, may be written as

\[
\Pr(CD) = \Pr(CDS_1) + \Pr(CDS_2) \\
= \Pr(CD \mid S_1) \times \Pr(S_1) + \Pr(CD \mid S_2) \times \Pr(S_2) \\
= \Pr(C \mid S_1) \times \Pr(D \mid S_1) \times \Pr(S_1) + \Pr(C \mid S_2) \times \Pr(D \mid S_2) \times \Pr(S_2).
\]

This is not necessarily equal to \( \Pr(C) \times \Pr(D) \) as is illustrated in the following example. Let \( \Pr(C \mid S_1) = \gamma_1, \Pr(C \mid S_2) = \gamma_2, \Pr(D \mid S_1) = \delta_1, \Pr(D \mid S_2) = \delta_2, \Pr(S_1) = \theta \) and \( \Pr(S_2) = 1 - \theta \). Then

\[
\Pr(CD) = \gamma_1 \delta_1 \theta + \gamma_2 \delta_2 (1 - \theta), \\
\Pr(C) = \gamma_1 \theta + \gamma_2 (1 - \theta), \\
\Pr(D) = \delta_1 \theta + \delta_2 (1 - \theta).
\]

The product of \( \Pr(C) \) and \( \Pr(D) \) is not necessarily equal to \( \Pr(CD) \). Suppose, for example that \( \theta = 0.40, \gamma_1 = 0.10, \gamma_2 = 0.20, \delta_1 = 0.15 \) and \( \delta_2 = 0.05 \). Then

\[
\Pr(CD) = \gamma_1 \delta_1 \theta + \gamma_2 \delta_2 (1 - \theta) \\
= 0.10 \times 0.15 \times 0.4 + 0.20 \times 0.05 \times 0.6 \\
= 0.0120, \\
\Pr(C) = \gamma_1 \theta + \gamma_2 (1 - \theta) = 0.04 + 0.12 = 0.16, \\
\Pr(D) = \delta_1 \theta + \delta_2 (1 - \theta) = 0.06 + 0.03 = 0.09, \\
\Pr(C) \times \Pr(D) = 0.0144 \neq 0.0120 = \Pr(CD).
\]

### 1.6.8 Updating of probabilities

Notice that the probability of guilt is a subjective probability, as mentioned before (Section 1.6.4). Its value will change as evidence accumulates. Also, different people will have different values for it. The following examples, adapted from similar ones in DeGroot (1970), illustrate how probabilities may change with increasing information. The examples have several parts and each part has to be considered in turn without information from a later part.
Example 1.4
(a) Consider four events $S_1, S_2, S_3$, and $S_4$. Event $S_1$ is that the area of Lithuania is no more than 50,000 km$^2$, $S_2$ is the event that the area of Lithuania is greater than 50,000 km$^2$ but no more than 75,000 km$^2$, $S_3$ is the event that the area of Lithuania is greater than 75,000 km$^2$ but no more than 100,000 km$^2$, and $S_4$ is the event that the area of Lithuania is greater than 100,000 km$^2$. Assign probabilities to each of these four events. Remember that these are four mutually exclusive events and that the four probabilities should add up to 1. Which do you consider the most probable and what probability do you assign to it? Which do you consider the least probable and what probability do you assign to it?

(b) Now consider the information that Lithuania is the 25th largest country in Europe (excluding Russia). Use this information to reconsider your probabilities in part (a).

c) Consider the information that Estonia, which is the 30th largest country in Europe, has an area of 45,000 km$^2$ and use it to reconsider your probabilities from the previous part.

d) Consider the information that Austria, which is the 21st largest country in Europe, has an area of 84,000 km$^2$ and use it to reconsider your probabilities from the previous part.

The area of Lithuania is given at the end of the chapter.

Example 1.5
(a) Imagine you are on a jury. The trial is about to begin but no evidence has been led. Consider the two events: $S_1$, the defendant is guilty; $S_2$, the defendant is innocent. What are your probabilities for these two events?

(b) The defendant is a tall Caucasian male. An eyewitness says he saw a tall Caucasian male running from the scene of the crime. What are your probabilities now for $S_1$ and $S_2$?

c) A bloodstain at the scene of the crime was identified as coming from the criminal. A partial DNA profile has been obtained, with proportion 2% in the local Caucasian population. What are your probabilities now for $S_1$ and $S_2$?

d) A window was broken during the commission of the crime. Fragments of glass were found on the defendant’s clothing of a similar refractive index to that of the crime window. What are your probabilities now for $S_1$ and $S_2$?

e) The defendant works as a demolition worker near to the crime scene. Windows on the demolition site have refractive indices similar to the crime window. What are your probabilities now for $S_1$ and $S_2$?

This example is designed to mimic the presentation of evidence in a court case. Part (a) asks for a prior probability of guilt before the presentation of any evidence. It may be considered as a question concerning the understanding of
Uncertainty in forensic science

the dictum ‘innocent until proven guilty’. See Section 3.5.5 for further discussion of this with particular reference to the logical problem created if a prior probability of zero is assigned to the event that the suspect is guilty.

Part (b) involves two parts. First, the value of the similarity in physical characteristics between the defendant and the person running from the scene of the crime, assuming the eyewitness is reliable, has to be assessed. Secondly, the assumption that the eyewitness is reliable has to be assessed.

In part (c) it is necessary to check that the defendant has the same DNA profile. It is not stated that he has, but if he has not he should never have been a defendant. Secondly, is the local Caucasian population the correct population? The evaluation of evidence of the form in (c) is discussed in Chapter 9.

The evaluation of refractive index measurements mentioned in (d) is discussed in Chapter 10. Variation both within and between windows has to be considered. Finally, how information about the defendant’s lifestyle may be considered is discussed in Chapter 8.

It should be noted that the questions asked initially in Example 1.5 are questions which should be addressed by the judge and/or jury. The forensic scientist is concerned with the evaluation of his evidence, not with probabilities of guilt or innocence. These are the concern of the jury. The jury combines the evidence of the scientist with all other evidence and uses its judgement to reach a verdict. The theme of this book is the evaluation of evidence. Discussion of issues relating to guilt or otherwise of suspects will not be very detailed.

(As a tail piece to this chapter, the area of Lithuania is 65,301 km².)